UCL
Université catholique de Louvain

de Duve INSTITUTE

Identification of a novel cancer-germline transcript within the miRNA harboring *GABRA3* gene. Epigenetic alterations of the locus in tumors.

Aurélie VAN TONGELEN

Mai 2017

Thèse présentée en vue de l'obtention du grade
de docteur en sciences biomédicales et pharmaceutiques

Secteur des sciences de la santé

President of the jury

Professor Frederic Lemaigre
de Duve institute
Université catholique de Louvain

Jury members

Doctor Paola Arimodo
ETaC – Unité de pharmacochimie de la régulation épigénétique du cancer
CNRS - Laboratoires Pierre Fabre Centre de recherche & développement

Doctor Philippe Arnaud
GReD – Génétique Reproduction et Développement
CNRS - Clermont Université – INSERM

Professor Anabelle Decottignies
de Duve institute
Université catholique de Louvain

Professor Patrick Jacquemin
de Duve institute
Université catholique de Louvain

Professor Guido Bommer
de Duve institute
Université catholique de Louvain

Doctor Axelle Loriot
de Duve institute
Université catholique de Louvain

Promoter

Professor Charles De Smet
de Duve institute
Université catholique de Louvain

This thesis was supported by PhD fellowship from F.R.S – FNRS Télévie and UCL – FSR

Après avoir passé presque six années dans le laboratoire de Charles De Smet, j'en sors grandie d'expérience, d'ouverture d'esprit, de connaissance, d'organisation et d'indépendance. Mais ces six années n'ont pas seulement été faites de sciences, c'est aussi une grande expérience de vie. Je remercie toutes les personnes qui m'ont permis d'arriver où je suis aujourd'hui.

Je tiens tout d'abord et tout particulièrement à remercier le professeur Charles De Smet, sans qui cette thèse n'aurait pu voir le jour. En tant que promoteur de thèse, il m'a constamment guidée dans mes recherches et m'a sans cesse aidée à trouver des solutions pour avancer. J'ai énormément de gratitude à son égard pour tous ses efforts à garantir la continuité et l'aboutissement de ma thèse. Je le remercie pour tout ce temps et toute sa patience qu'il m'a consacré pour garantir la qualité de mon travail. Plusieurs pages auraient été nécessaires pour lister tout ce qu'il a fait pour que j'atteigne mon objectif.

Malgré son désordre inégalable, Axelle Loriot m'a apporté un enseignement pratique de grande qualité et très riche en trucs et astuces. Je la remercie pour sa générosité extraordinaire et son efficacité à toute épreuve. Je la remercie également pour tous ces moments privilégiés que nous avons partagés, pour toutes ces discussion très matinales, pour nos weekends relais qPCR durant la période de Noel, pour tous ces breaks « mentos » qui suivaient les expériences fructueuses (une chance qu'on a eu beaucoup) et pour son énergie phénoménale qui fait vivre le laboratoire.

Je remercie Florian Poulain, mon compagnon qui m'a encouragée et soutenue tout au long de ma thèse. Je le remercie très fortement pour son aide précieuse apportée lors des nuits blanches passées à la rédaction de ce manuscrit. Je le remercie également pour son écoute attentive, son réconfort et ses idées pertinentes. Sans lui, je ne serais pas là aujourd'hui.

Je remercie Julie Cannuyer pour sa gentillesse, pour notre complicité tissée durant près de cinq années et pour sa disciplinee olympienne.

Je souhaite remercier Jean Fain pour sa fabuleuse motivation qui m'a donné l'énergie d'approfondir mon dernier sujet de thèse. Sa présence a permis de rendre les journées plus atypiques et plus cocasses. Rien de tel que de subtiles petites blagues pour égailler les journées.

Je remercie Anna Diacofotakis, pour sa profonde gentillesse et toute sa sincérité. Je la remercie également pour son aide à la relecture et à la correction de ce manuscrit, pour ses délicieux gâteaux agrémentés de « cocoa » et pour le petit grain de Grèce qu'elle a apporté au laboratoire.

Enfin, je souhaite remercier ma famille et mes parents pour leur soutien et leur amour constant. Ce sont eux qui m'ont encouragée et permis d'entreprendre mes études et cette fabuleuse aventure qu'aura été le doctorat.

« Je dédie cette thèse à mon compagnon ainsi qu'à mes parents »

TABLE OF CONTENTS

SUMMARY

It is now well established that alterations in DNA methylation patterns contribute to tumor development. Both gains (DNA hypermethylation) and losses (DNA hypomethylation) of this repressive mark are observed in tumors. Our group demonstrated that DNA hypomethylation in tumors induces transcriptional activation of a defined group of genes, which normally show specific expression in the germline. These genes were grouped under the term "cancer-germline" genes (CG genes). The main goal of the laboratory is to characterize the epigenetic regulation and functional roles of CG genes. Our work focused on the *GABRA3* gene locus, where we discovered the existence of two overlapping transcripts: *BT-GABRA3*, which is expressed specifically in brain and testis; and *CT-GABRA3*, starting ~250 kb upstream, which is expressed exclusively in testis. *CT-GABRA3* (but not *BT-GABRA3*) exhibits typical features of a CG gene, as it shows promoter hypomethylation and transcriptional activation in various tumors. *CT-GABRA3* carries a clustered pair of miRNAs (miR-105 and miR-767), which show concurrent expression in tumors. Interestingly, an independent group identified miR-105 as a crucial promoter of cancer metastasis, due to its ability to weaken vascular endothelial barriers following exosomal secretion. On the other hand, we demonstrated that miR-767 inhibits expression of *TET1*, a gene with tumor suppressive functions involved in epigenetic processes of DNA demodification. Moreover, our studies revealed that *CT-GABRA3* hypomethylation/activation in tumors is correlated with hypermethylation of the downstream *BT-GABRA3* promoter. The mechanism underlying this interdependent epigenetic alteration appears to involve deposition of the histone mark H3K36me3 on the entire *CT-GABRA3* transcribed region, which includes the *BT-GABRA3* promoter. Finally, we observed that in tumor cells where the *BT-GABRA3* promoter is hypermethylated, the gene becomes sensitive to DNA demethylation, suggesting a process of epigenetic switch. Together our work revealed the existence of a novel miRNA-producing CG gene with oncogenic potential. It also uncovered a unusual mechanism of epigenetic alteration in tumors, whereby DNA hypomethylation and hypermethylation are linked via a process of transcriptional overlap.

ABREVIATIONS

DNA	*Deoxyribonucleic acid*
-MEL	*melanoma cell lines*
3'/5'-UTR	*3'/5'-untranslated region*
5-azadC	*5-aza-2'-deoxycytidine*
5'-RACE	*5'-Rapid amplification of cDNA ends*
5caC	*5-carboxylcytosine*
5fC	*5-formylcytosine*
5hmC	*5-hydroxymethylcytosines*
5mC	*5-methylcytosines*
BER	*Base excision repair*
Cas9	*CRISPR-associated protein 9*
CG	*Cancer-germline*
CGI	*CpG island*
CpG	*Cytosine-phosphate-guanine*
CRISPR	*Clustered, Regularly Interspaced, Short Palindromic Repeats*
CT	*Cancer-testis*
DNMT	*DNA methyltransferase*
DSB	*double-strand breaks*
DSS	*salmon sperm DNA*
ESC	*Embryonic stem cell*
FACS	*fluorescence-activated cell sorting*
FLP	*Flippase*
gRNA	*guide RNA*
H3K27ac	*Histone 3 lysine 27 acetylation*
H3K27me3	*Histone 3 lysine 27 trimethylation*
H3K36	*histone H3 at lysine 36 trimethylation*
H3K4me1/2/3	*Histone 3 lysine 4 mono/di/trimethylation*
H3K79me2	*Histone 3 lysine 79 dimethylation*
HDAC	*Histone deacetylases*
HDR	*homology-directed repair*
ICM	*Inner cell mass*
KDM2A/B	*Lysine-specific demethylase 2A/B*
LHA	*left homology arm*
LNA	*Locked nucleic acid*
MBD	*Methyl-CpG-binding domain*
MBP	*Methyl-CpG-binding proteins*
MeCP2	*methyl CpG binding protein 2*
miRNA	*micro RNA*
mRNA	*messenger RNA*
Neo	*neomycin resistance cassette*
NHEJ	*non-homologous end-joining*
NSCLC	*non-small cell lung carcinoma*
PAM	*Protospacer Adjacent Motif*
PGC	*Primordial Germ Cell*
PolII	*RNA polymerase II*
PRC1/2	*Polycomb repressive complexes 1/2*
PWWP	*Pro-Trp-Trp-Pro*
RHA	*right homology arm*
RT-PCR	*Reverse transcription polymerase chain reaction*
RT-qPCR	*Quantitative reverse transcription polymerase chain reaction*
SAM	*S-Adenosyl methionine*
sgRNAs	*single guide RNAs*
SRA	*SET and RING finger-associated domain*
TF	*Transcription factor*
tracrRNA	*trans-activating CRISPR RNA*
tRNA	*transfer RNA*
TSS	*transcription start site*
WT	*Wild type*

INTRODUCTION

DNA IS NOT DESTINY: THE NEW SCIENCE OF EPIGENETICS

The DNA structure was discovered in 1953 by Watson and Crick, and few years later, the genetic code was cracked by Khorana and Nirenberg, along with Robert Holley. By the year 2001, the entire human genome sequence was published. Three billion base pairs of the human genome have been decrypted with about 25.000 genes coding for proteins that have been precisely located on the 23 chromosomes pairs. By that time, it was thought that all that defines a human individual is irreversibly written in the DNA code, and that the sequence determines the way genes are expressed. Genetic polymorphisms and mutations were expected to explain differences in appearance, behavior and susceptibility for disease. However, DNA sequence is not destiny and is not enough to explain all traits.

Regulation of the genetic machinery is defined by additional information that can be transmitted through cell division. If all cells have the same genetic information the key to these differences might come from “epigenetics”. According to Mark Ptashne, epigenetics corresponds to “a change in the state of expression of a gene that does not involve a mutation, but that is nevertheless inherited in the absence of the signal or event that initiated the change” (Ptashne 2007). A healthy human body needs many different cells such as brain cells, blood cells, skin cells etc. all deriving from a single fertilized egg cell, the zygote. During cell differentiation, some genes get switched on, while others are switched off. These gene expression patterns are transmitted from one cell to its daughter. Epigenetics guides cells through development and confers an epigenetic “memory” to maintain cell type-specific transcriptional patterns in the adult (D'Urso and Brickner 2014). The reading of the genome depends therefore on its cellular history. Epigenetics involves multiple mechanisms inducing chromatin changes such as DNA modifications or chemical adjunctions on histones, which are proteins acting as coils around which DNA rolls. If epigenetic mechanisms are altered, genes can be turned on or off in the wrong cells or at the wrong time and lead to disease.

DNA METHYLATION: A FORM OF EPIGENETIC CONTROL OF GENE EXPRESSION

The first suggestion of the existence of modified DNA has been proposed in 1947, even before the double helix structure had been discovered (Vischer and Chargaff 1947). The predominant modified DNA mark in vertebrate genomes is the covalent addition of a methyl group on position five of the pyrimidine ring of a cytosine, which is called 5-methylcytosine (5mC) (Rollins et al. 2006; Schubeler 2015). Mammalian methylation occurs predominantly in a cytosine-phosphate-guanine (CpG) dinucleotide context (Gruenbaum et al. 1981) and is the most studied epigenetic mark. 5mC is extensively present in all human cells.

1 Genomic distribution of modified cytosines

In mammals, 70% to 80% of all CpG dinucleotides are methylated (Ehrlich et al. 1982; Robertson 2005). Surprisingly, CpG dinucleotides in mammalian genomes are only present at 21% of the expected frequency. This has been attributed to the hypermutability of methylated CpG, where methylated cytosines are converted to thymines by a deamination reaction (**Figure 1**). It is thought that this hypermutability has led to a depletion of CpG dinucleotides during evolution (Bird 1980; Holliday and Grigg 1993; Smallwood et al. 2011).

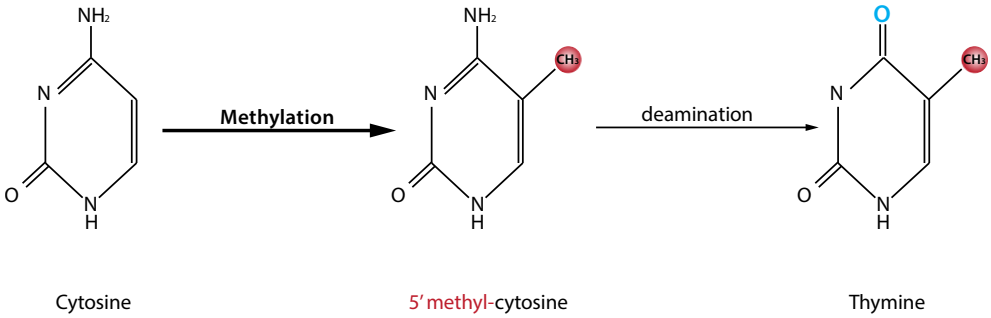


Figure 1: Cytosine methylation in a CpG context and 5-methylcytosine deamination

The distribution of the CpG dinucleotides in the genome is not homogenous. While the majority of the genome is nearly completely methylated, it is punctuated however by non-methylated and CpG-dense DNA sequences called "CpG islands" (CGIs) (**Figure 2**). These short regions have kept a high density of CpG sites (Gardiner-Garden and Frommer 1987) probably because they are kept unmethylated (Smallwood et al. 2011). A majority of the annotated promoters in the human genome has been reported to contain CGIs (Saxonov et al. 2006; Jiang et al. 2014), but CGIs can also be present in intra- and intergenic regions (**Figure 2**) (Illingworth et al. 2010).

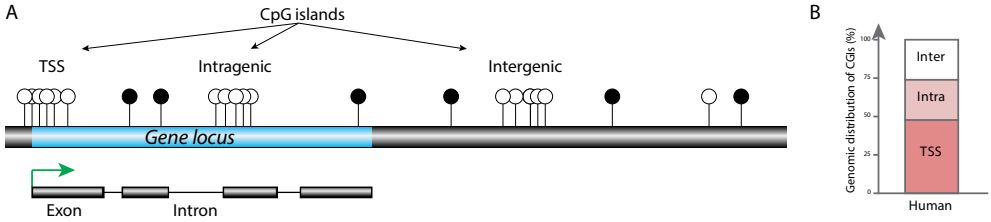


Figure 2: Genomic distribution of CGIs. (A) CGIs can be located in transcription start sites (TSS), within coding gene region (Intragenic), or between genes (Intergenic). (B) Proportion of CGIs through the three distinct genomic regions according to Illingworth (Illingworth et al. 2010). White lollipop: non-methylated CpG, black lollipop: methylated CpG..

2 DNA METHYLATION : a transcriptional regulation mechanism

What is the effect of DNA methylation on gene transcription?

Classically, DNA methylation induces gene transcription silencing (Razin and Riggs 1980; Jones and Takai 2001). This action is mediated by two main mechanisms (**Figure 3**). Firstly, the presence of 5mC in a promoter region can physically impede the binding of transcription factors (TF) to their targeted sequence (Wade 2001). Secondly, 5mC can be recognized by specialized proteins such as Methyl-CpG binding domain proteins (MBDs). In turn, these proteins enhance chromatin compaction by recruiting repressors that establish a repressive chromatin state suitable for stable gene silencing (Herman and Baylin 2003; Goll and Bestor 2005).

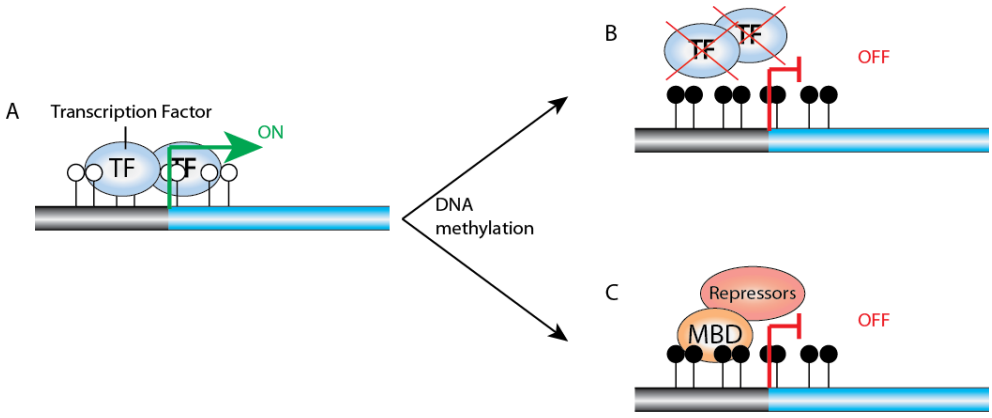


Figure 3: Transcriptional repression by DNA methylation. (A) Active genes with an unmethylated promoter enables transcription factors to bind. (B) DNA methylation masks the targeted sequences of transcription factors, which are therefore unable to initiate gene transcription. (C) DNA methylation is recognized by specific proteins like Methyl-CpG binding *proteins* (MBDs) that will subsequently recruit co-repressor complexes and inhibit gene transcription. White lollipop: non-methylated CpG, black lollipop: methylated CpG.

While active genes exhibit an absence of DNA methylation around their promoter regions, substantial amounts of DNA methylation are often found within their transcribed region (**Figure 4**) (Jones 2012). It is suggested that DNA methylation in the core region of a gene enhances transcriptional expression (Zhang et al. 2006). This suggestion is strengthened by the fact that methylation located immediately downstream of a transcription start site (TSS) antagonizes the binding of Polycomb repressive complexes (PRC) protein complexes that induces histone modifications involved in gene repression (Wu et al. 2010). Methylation in the gene bodies does not block transcription elongation and might even stimulate it (Yang et al. 2014). Methylation within gene bodies is also suggested to silence spurious transcription or alternative splicing as methylation of gene body CpGs appears to be associated with repression of intragenic promoters (Maunakea et al. 2010).

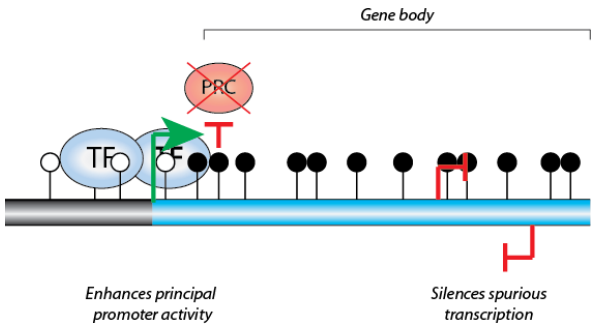


Figure 4: Illustration of the potential functions of CpG methylation within the gene body of a transcriptionally active gene. DNA methylation could enhance the principal promoter activity by inhibiting PRC repressor complex and could silence downstream spurious transcription. White lollipop: non-methylated CpG, black lollipop: methylated CpG.

What are the genes targeted by DNA methylation?

It was originally proposed that DNA methylation would be a general mechanism to control gene transcription during cellular differentiation (Riggs 1975). However, many tissue-specific genes show an inconstant relationship between promoter methylation and expression level (Walsh and Bestor 1999; Warnecke and Clark 1999). Many tissue-specific genes for instance, contain a CGI within their promoter that is always unmethylated even in non-expressing tissues. Moreover, experimental demethylation of tissue-specific gene promoters is generally not sufficient to induce their activation (Michalowsky and Jones 1989). Although DNA methylation is essential for normal development (Gaudet et al. 2003), its role determining tissue-specific expression patterns is less pivotal than previously suggested.

Even if the majority of CGIs are always unmethylated, some of them stand out as exceptions to this rule. These genes include imprinted genes, genes located on the inactivated X-chromosome and several genes with specific expression in germline cells (Li et al. 1993; De Smet et al. 1999; Plass and Soloway 2002; Sharp et al. 2011). Many genomic repeated sequences also contain CGIs that are generally methylated in most tissues (Ehrlich 2002).

Imprinted genes

Gene transcription occurs generally from both alleles. In mammals, a small proportion (<1%) of genes are however expressed from only one allele (Wilkinson et al. 2007). During gametogenesis, certain genes will receive an "imprint", which allows to distinguish the maternal origin from the paternal origin (Surani et al. 1984). Most imprinted genes possess differentially methylated regions (DMRs) whereby allelic methylation depends on the parent of origin (Reik and Walter 2001). Allele-specific DNA methylation profiles are acquired during the development of germ cells, and are maintained through fertilization and the development of the embryo.

X-inactivated chromosome

In female somatic cells, one of the two copies of the X-chromosome is inactivated. This phenomenon allows a balance between female cells, which have two copies of the X-chromosome, and male cells, which have only one copy (Lyon 1999). This process is initiated by the transcription of a long-non-coding RNA (lncRNA), *Xist*, which is specifically expressed from the X-chromosome that will be inactivated (Brockdorff et al. 1991; Penny et al. 1996). After local accumulation of this lncRNA, the whole X-chromosome is covered, and repressive chromatin modifications are brought onto the inactive chromosome (Avner and Heard 2001). Modifications include DNA methylation which extends within CGIs of gene promoters (Wutz 2011) (**Figure 5**). Even if DNA methylation is not the initial factor that represses these genes, it is essential to maintain their repression across cell divisions (Heard and Disteche 2006).

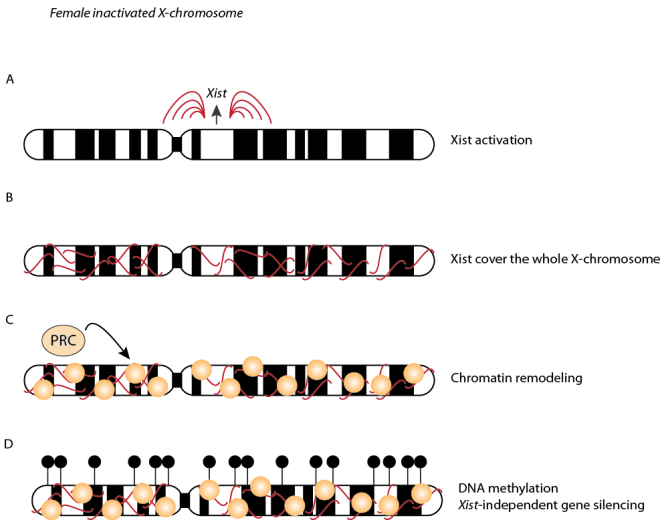


Figure 5: X-chromosome inactivation process. (A) Xist transcriptional activation. (B) Spreading of Xist all over the X-chromosome. (C) Factors such as Polycomb repressor complex (PRC) catalyze chromatin modification marks. (D) Methylation deposition maintains Xist-independent gene silencing from the X-inactivated chromosome. Black lollipop: methylated CpG.

Tissue-specific genes

DNA methylation is not the general mechanism regulating tissue-specific expression. Most genes indeed use other mechanisms, including histone modifications, to ensure tissue-specific expression. However, for a few genes, such as certain germline-specific genes, DNA methylation appears to be the primary mechanism of repression in non-expressing tissues (De Smet et al. 1999). As these genes hold a central place in my thesis, a separate chapter is devoted to them at page 41.

Genomic repeated sequences

Endogenous transposable elements, which constitute about 40% of mammalian genomes, may harbor strong transcriptional promoters that must be constitutively repressed to prevent uncontrolled transcription. This is largely obtained through the methylation of the CpG sites contained in these sequences. (Bourc'his and Bestor 2004; Hackett et al. 2012). It appears therefore that DNA methylation has an important role in maintaining stable silencing of

mobile transposable elements, thereby reducing their threat for genome integrity (Walsh et al. 1998; Reik 2007).

3 Global DNA methylation level is dynamic during the mammalian life cycle

The various profiles of DNA methylation observed in the different tissues are established during development. Once established, they are largely maintained during cell divisions in the adult. Two waves of DNA methylation reprogramming occur during embryo development (Reik et al. 2001) (***Figure 6***).

Early Embryo Reprogramming

A first wave of reprogramming is observed in the somatic cells of the very early embryo (***Figure 6***). Observations in mice revealed that following fertilization, the zygote undergoes a genome-wide demethylation that is completed by the stage of blastocyst. The paternal genome undergoes demethylation immediately after fertilization and is completed before DNA replication begins (Mayer et al. 2000; Oswald et al. 2000). In the maternal genome, demethylation occurs more slowly. Then, establishment of new DNA methylation patterns on the embryonic genome (both maternal and paternal genomes) starts in cells from the inner cell mass (ICM) at the time of implantation of the expanded blastocyst, and is completed prior to birth. This wave of DNA demethylation and remethylation is likely to play a role in the removal of parentally acquired epigenetic modifications (Reik et al. 2001). It has to be noticed that DMRs associated with imprinted genes escape this reprogramming process and keep the methylation patterns that correspond to their parental origin (Howell et al. 2001).

Germ Cells Reprogramming

The other wave of DNA methylation reprogramming (***Figure 6***) occurs early in the development of the primordial germ cells (PGCs) (Monk et al. 1987; Brandeis et al. 1993). In mice, DNA demethylation of PGC is completed by embryonic day 13 (Kafri et al. 1992). In the male germ line, remethylation begins

several days later to be completed at birth. In the female germ line, remethylation begins at birth to be completed at puberty (Kafri et al. 1992). This process is true for both single-copy gene sequences and DMRs regions associated with imprinted genes. The purpose of this reprogramming wave is the resetting of parental imprints and the removal of acquired epigenetic modifications (Reik et al. 2001).

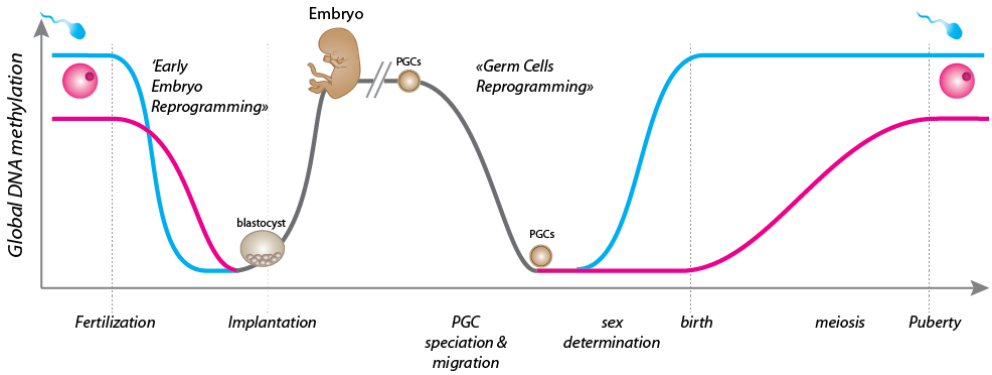


Figure 6: Schematic representation of the two waves of global DNA methylation reprogramming during development. One wave occurs after fertilization in somatic cells. It is the "early embryo reprogramming". The second wave occurs in primordial germ cells (PGC). It is the "germ cells reprogramming".

4 Histone marks associated with DNA methylation

DNA methylation is tightly interconnected with histone post-translational modifications to control chromatin structure and function (***Figure 7***). Histones are proteins around which DNA is tightly wrapped. They can undergo a panel of chemical modifications such as methylation, acetylation, phosphorylation, sumoylation and ubiquitinylation. Among these, the role of histone acetylation, histone 3 lysine 4 methylation (H3K4me), histone 3 lysine 27 trimethylation (H3K27me3), histone 3 lysine 9 trimethylation (H3K9me3) and histone 3 lysine 36 trimethylation (H3K36me3) in transcriptional regulation are well documented and were investigated in the context of this thesis. These histone modifications will therefore be discussed in the following paragraphs (***Figure 8***) & (**Table 1**).

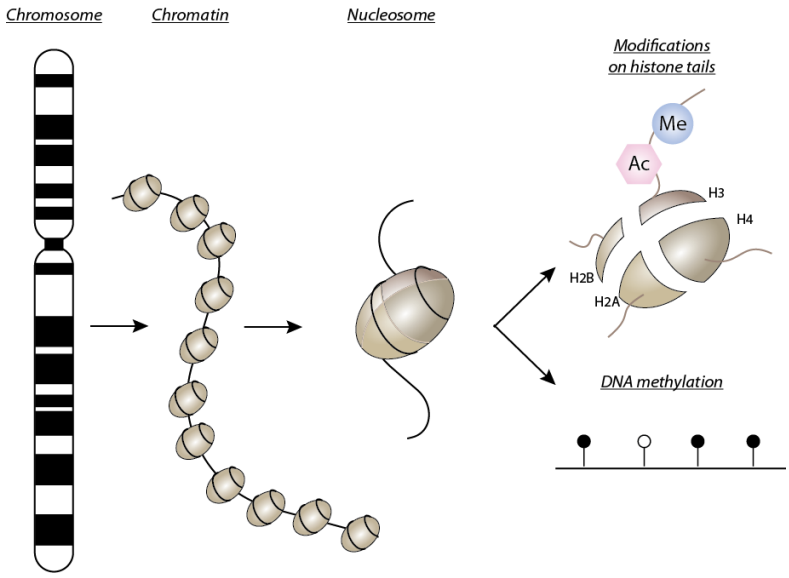


Figure 7: Schematic representation of chromatin organization. The basic unit of chromatin is the nucleosome. The latter contains the DNA which can be methylated and the histones can be methylated or acetylated. There are four different histones variants H2A, H2B, H3 and H4. Tails of the histones are accessible outside the nucleosome and undergo several post-translational modifications such as acetylation or methylation. White lollipop: non-methylated CpG, black lollipop: methylated CpG.

Acetylation

Histone acetylation corresponds to the addition of an acetyl group (CH_3CO) to a lysine located on a histone tail. This reaction is catalyzed by histone acetyltransferases (HAT) (Roth et al. 2001). Conversely, histone acetylation can be erased by enzymes named histone deacetylases (HDACs). Histone acetylation is associated with an active chromatin state, as only transcriptionally active regions harbor this modification. Acetylation of the lysine neutralizes the positive charge and hence decreases the interaction between the histones and the negatively charged phosphate groups of DNA. Histone acetylation also attracts chromatin remodeling complexes and general transcription factors that facilitate transcription locally (Yang and Seto 2003).

HAT and HDAC enzymes are usually found in large multiprotein complexes near euchromatic regions of the chromatin (Saha and Pahan 2006). They can be specifically recruited to promoters by direct interaction with certain transcription factors. The local balance between HAT and HDAC activities in a chromatin region directs the level of histone acetylation and the level of transcription.

Methylation

Histone methylation can be catalyzed on lysine or arginine residues, and plays a central role in transcriptional regulation. The residues can be mono-, di- or tri- methylated and can be associated with transcriptional repression or activation.

H3K4me: The methylation of lysine 4 of the histone3 (H3K4me) correlates positively with gene expression. H3K4 can be mono-, bi-, or trimethylated (-me1, -me2, -me3). H3K4me3 is mainly enriched in active transcription start sites (TSS) and H3K4me2 in transcribed regions of genes, while H3K4me1 is more associated with enhancer regions (Barski et al. 2007). H3K4me3 and –me2 facilitate transcription by promoting the binding of positive transcription factors and by attracting chromatin remodeling complexes that favor an open chromatin configuration (Nishioka et al. 2002; Lauberth et al. 2013).

H3K4me can be catalyzed by histone lysine methyl-transferases (HKMTs) such as MLL (mixed lineage leukemia) protein complexes and erased by histone lysine demethylases (HKDMs) such as LSD1 (lysine-specific demethylase 1).

H3K9me3: The trimethylation of lysine 9 of histone 3 (H3K9me3) is associated with silent genes (Barski et al. 2007). H3K9me3 is catalyzed by Suv-39H enzyme or G9a (Rice et al. 2003) and can be bound by the heterochromatin protein 1 (HP1) to form condensed chromatin (Lehnertz et al. 2003). H3K9me3 can be erased by HKDMs such as JMJD2 (jumanji domain 2).

H3K27me3: The trimethylation of lysine 27 of the histone 3 is tightly associated with inactive gene promoters. The major enzyme that catalyzes H3K27 trimethylation is the EZH2 (enhancer of zeste homolog 2) enzyme

(Kuzmichev et al. 2002). This enzyme is a subunit of the large complex PRC2 (Polycomb repressive complex 2) involved in the establishment and maintenance of repressive chromatin state during development and cell differentiation (Bracken et al. 2006). H3K27me3 is also an important mark of the inactive X chromosome (Rougeulle et al. 2004). H3K27 trimethylation can be erased by several HKDMs such as UTX (Ubiquitously transcribed tetratricopeptide repeat) or JMJD3 (jumanji domain 3) (Chen et al. 2012; Van der Meulen et al. 2014).

H3K36me3: The trimethylation of lysine 36 of the histone 3 is specifically enriched in the body of those genes (Barski et al. 2007). During transcription H3K36me3 is deposited on histones by the SETD2 (SET domain containing 2) enzyme (Edmunds et al. 2008), which is associated with the transcriptional machinery (including RNA PolII). H3K36me3 recruits HDACs to deacetylate the histones and induce a repressive chromatin context to prevent spurious transcription that may otherwise initiate within the gene body (Carrozza et al. 2005).

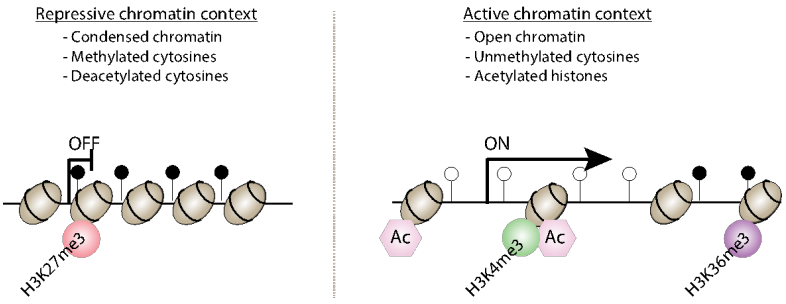


Figure 8: Histone modifications and their association with chromatin context that correlates with transcriptional activation or repression of a gene. White lollipop: non-methylated CpG, black lollipop: methylated CpG.

Table 1: Major histone modifications, effects on gene transcription and enzymes involved in their settings.

Modifications	Transcriptional association	« Writers »	« Erasers »
Acetylation	Activation	HATs	HDACs
H3K4me3	Activation	MLL	LSD1/KDM1a
H3K9me3	Repression	Suv39 / G9a/SetDB1	JMJD2
H3K27me3	Repression	EZH2	UTX / JMJD3
H3K36me3	Activation	SETD2	KDM4A

5 Regulation of DNA methylation

5.1 Writing DNA methylation

DNA methyltransferases (DNMTs) are the enzymes that catalyze the transfer of a methyl group from S-Adenosyl Methionine (SAM) to a cytosine in a CpG context. The family of DNMTs is composed of five members: DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L (***Figure 9***). Two different DNA methyltransferase activities can be distinguished. The establishment of DNA methylation profiles during embryogenesis implies a *de novo* DNA methyltransferase activity which induces methylation of sequences initially unmethylated, whereas preservation of DNA methylation patterns after replication implies a maintenance DNA methyltransferase activity (***Figure 10***).

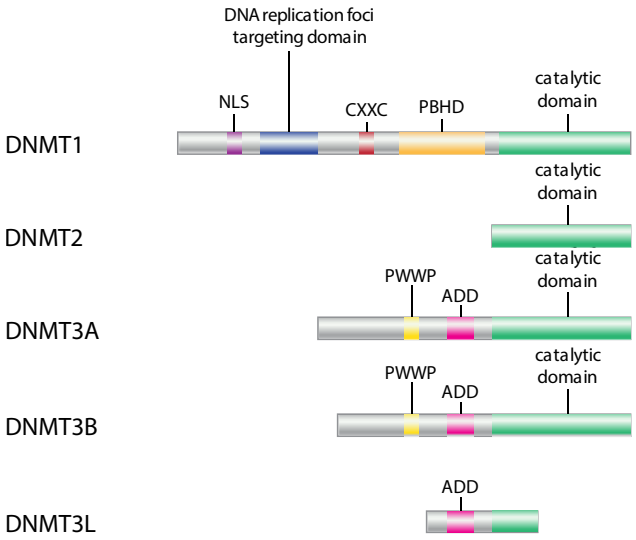


Figure 9: Schematic structure of the five members of DNMTs family enzymes with their main regulatory domains: ADD (ATRX-DNMT3-DNMT3L), CXXC (cysteine-rich ZN2+-binding domain), PBD (PCNA binding domain) PWWP (proline-tryptophan-tryptophan-proline), NLS (nuclear localization signal).

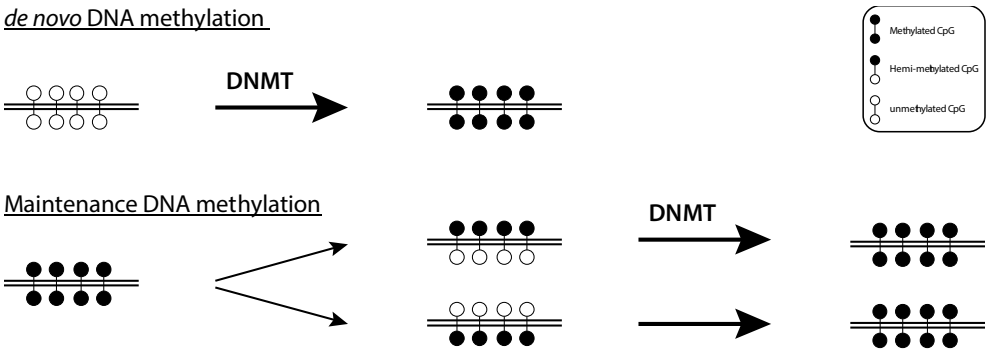


Figure 10: Illustration of DNA methylation de novo and maintenance mechanisms

de novo DNA methyltransferases

The so-called *de novo* DNA methylation process involves DNMT3A and DNMT3B. These enzymes are essential for the setting of DNA methylation profiles during embryo development as it was demonstrated by knockout experiments in mice. The inactivation of DNMT3A and DNMT3B in mouse blocks

de novo methylation in post-implantation embryos and causes embryonic lethality (Okano et al. 1999). These enzymes are partially redundant in their functions, even if certain sequences are specifically targeted by DNMT3A or DNMT3B. DNMT3A is for example essential for the establishment of imprinting, and DNMT3B is responsible for the methylation of pericentromeric repetitive regions (Bachman et al. 2001). These two enzymes play a very important role in development and are only expressed at low levels in adult somatic tissues (Okano et al. 1999).

The mechanisms which direct DNMT3s to their target sequences are not completely understood. Interconnections between DNMTs and histone marks seem to play a crucial role. DNMT3A and DNMT3B share an ADD (ATRX-DNMT3-DNMT3L) domain that recognizes the unmethylated lysine of histone H3 (H3K4) (Zhang et al. 2010b).

Interestingly, in gene bodies, several studies have suggested that the tri-methylation lysine 36 of histone H3 (H3K36me3) could target *de novo* DNA methylation to these regions (***Figure 11*****A**). Catalyzed by SETD2 enzyme, H3K36me3 is targeted to the body of actively transcribed genes (Sun et al. 2005). Within gene bodies, H3K36me3 is enriched in exons relative to introns (Kolasinska-Zwierz et al. 2009; Spies et al. 2009; Huff et al. 2010) and correlates with DNA methylation profiles (Lorincz et al. 2004; Brown et al. 2012). Furthermore, DNMT3A and DNMT3B contain in addition to their ADD domain, a PWWP (Pro-Trp-Trp-Pro) domain which appears to read H3K36me3 (Baubec et al. 2015). This interaction is believed to guide DNA methylation to transcriptionally active genic regions, which do not include promoters (Dhayalan et al. 2010). During germline development however, processes of read through transcription appear to target DNA methylation towards downstream imprinted DMRs and gene promoters (Smallwood and Kelsey 2012).

De novo methylation often involves the initial binding of protein complexes capable of recruiting histone methylases or deacetylylases (***Figure 11*** **B**). EZH2 was shown to recruit DNMT3s, which then methylate the underlying

DNA (Vire et al. 2006). It has also been reported that DNMT3A binds to HDAC using its ADD domain to silence transcription (Fuks et al. 2001). Hence, it appears that the chromatin context guides *de novo* methyltransferases.

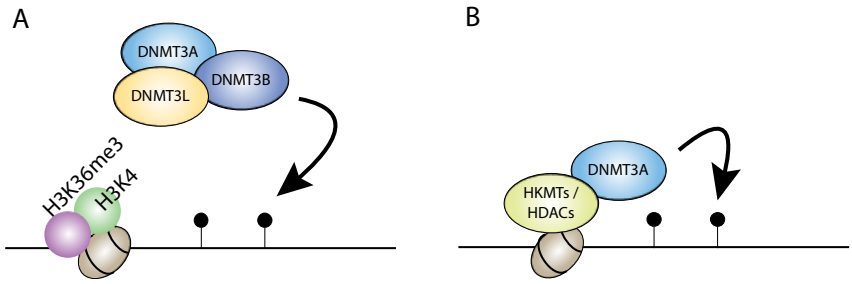


Figure 11: Mechanisms that induce de novo DNA methylation involving DNTM3s proteins. (A) DNMT3A and DNMT3L recognize DNA or chromatin by specific domains. The ADD domain of DNMT3A and DNMT3L was shown to interact with the unmodified Lys 4 of histone H3 (H3K4). In gene bodies, H3K36me3 mark recruits DNMT3s via their PWWP domains. (B) DNMTs are recruited to DNA through protein–protein interactions with chromatin-modifying enzymes, including KMTs and HDACs. Black lollipop: methylated CpG

Maintenance DNA methyltransferases

DNMT1, the maintenance cytosine methyltransferase, is ubiquitously expressed. DNMT1 knockout embryos die at late gastrulation stage and show extensive demethylation of many genomic sequences, including repetitive sequences and imprinted genes (Li et al. 1992). During DNA replication, DNMT1 is targeted to the replication fork and functions with the help of partner proteins including UHRF1 (Ubiquitin-like, containing PHD and RING finger domains, 1) and PCNA (Proliferating cell nuclear antigen) (Iida et al. 2002; Bostick et al. 2007). DNMT1 shows a preferential affinity for hemi-methylated CpG sites (Hermann et al. 2004) resulting from newly synthesized DNA (***Figure 12***).

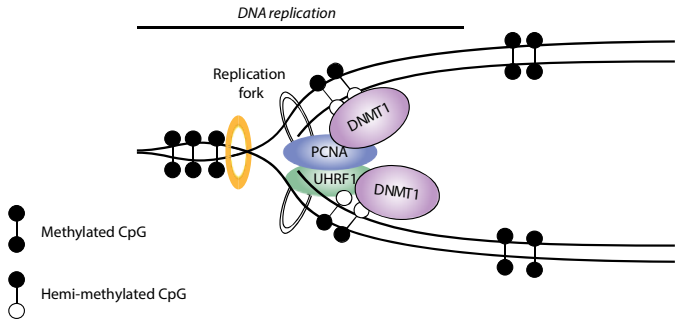


Figure 12: Mechanism of DNA methylation maintenance involving DNMT1 methyltransferase interacting with PCNA and UHRF1 to methylate the newly synthesized hemi-methylated CpGs.

Other members of the DNMTs family proteins

The last members of the DNMTs family, DNMT3L and DNMT2, have no DNA methylation activity. However, DNMT3L knockout mice are not able to establish maternal methylation imprints in oocytes and show male sterility due to spermatogenesis defects (Bourc'his et al. 2001). This observation is linked to the fact that DNMT3L is a co-factor that recruits DNMT3A and DNMT3B on specific genomic regions, such as imprinted genes (Hata et al. 2002). The role of DNMT2 is not essential for the establishment or maintenance of DNA methylation profiles (Okano et al. 1998). It is suggested that DNMT2 has a methylation activity on RNA rather than DNA (Goll et al. 2006).

5.2 Erasing DNA methylation

Although DNA methylation is a modification that is stably maintained through cell divisions, several mechanisms are able to remove this epigenetic mark. Loss of DNA methylation can result from a passive mechanism or from a TET-dependent reaction.

The "simplest" way to lose DNA methylation is to inhibit the maintenance of DNA methylation. In this case, DNA methylation marks are not recopied during cell divisions. Such passive demethylation process occurs for example in early embryos. DNMT1o is a DNMT1 isoform that is specifically expressed at this stage. This protein is sequestered in the cytoplasm during the 1-4 cell stage of embryonic development, and does not enter the nucleus until the 8-cell stage. This results in a dilution of the DNA methylation marks through cell divisions (Monk et al. 1991).

DNA methylation can also occur in a more active way, involving enzymes actively implicated in the process. Such enzymes have only been identified recently and have raised a great deal of interest. They belong to the TET methylcytosine dioxygenase family, which comprises 3 members TET1, TET2 and TET3. These enzymes are Fe^{2+} and 2-oxoglutarate-dependant dioxygenases that catalyze the successive conversion of a 5-methylcytosine (5mC) to a 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5CaC) (**Figure 13 & 14**)(Tahiliani et al. 2009). These oxidized methylcytosines are thought to serve as intermediates in the demethylation process, being further replaced by a cytosine. There are at least three mechanisms by which TET proteins could mediate DNA demethylation.

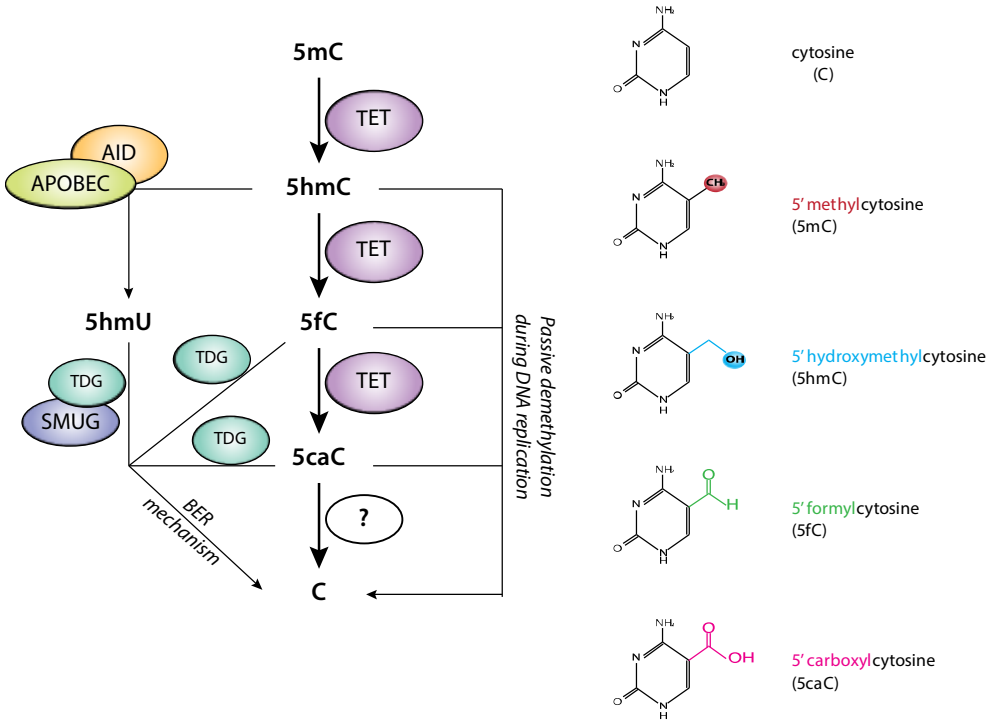


Figure 13: Proposed DNA demethylation pathways involving TET enzymes. Right: Chemical structure of cytosine (C), 5-methylcytosine (5mC), 5hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), 5-carboxylcytosine (5caC) and 5-hydroxymethyluracil (5hmU). Left: TET proteins catalyze 5mC oxidation to 5hmC, which is further converted to 5fC and 5caC. 5hmC can be converted in 5hmU by AID/APOBEC deaminases. 5fC, 5caC and 5hmU can be excised by glycosylases (TDG and SMUG) and be replaced by an unmodified cytosine by base excision repair (BER) mechanism. 5hmC, 5fC and 5caC can be passively converted in unmodified cytosine through DNA replication. caC can also be directly converted in unmodified cytosine through an unknown mechanism.

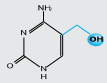
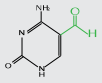
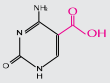
	 5' methylcytosine (5mC)	 5' hydroxymethylcytosine (5hmC)	 5' formylcytosine (5fC)	 5' carboxylcytosine (5caC)
Of all cytosines	4%	0.7% - 0.1%	< 0.002%	< 0.0003%

Figure 14 Relative level of the different forms of modified cytosines in mammals. The percentage is expressed on the total cytosines. 5mC is the major form of all the modified cytosines, and 5mC levels are relatively similar in all tissues. 5hmC levels varied significantly between different tissues (Kriaucionis and Heintz 2009; Tahiliani et al. 2009; Globisch et al. 2010; Szwagierczak et al. 2010). 5fC is nearly undetectable but Ito et al. in 2011 detected 3 to 20 5fC per million cytosines (0,0003% to 0,002%) in different mouse tissues with the maximum detected in mouse ES cells and brain cortex. 5caC is the rarest form, and was only reliably detected in mouse ES cells (Ito et al. 2011; Wu and Zhang 2011a).

Facilitation of passive DNA demethylation: The conversion of 5mC into 5hmC has been shown to reduce the DNA maintenance activity. DNMT1 activity is reduced from 12 to 50 folds at sites of hemi-5hmC compared to sites of hemi-5mC (Valinluck and Sowers 2007; Hashimoto et al. 2012). Furthermore, UHRF1 hemi-5hmC binding is tenfold less efficient than UHRF1 hemi-5mC (Hashimoto H 2012). This implies that TET-mediated oxidation of 5mC can block DNA methylation maintenance and hence facilitate a passive loss of 5mC during cell divisions. It is hence not properly an "active" demethylation.

Active DNA demethylation through DNA repair: Two replication-independent mechanisms have been described, coupling TET activity to DNA repair processes. 5hmC can be further converted to 5fC and to 5CaC, which can then be excised and replaced by an unmodified cytosine by TDG (Thymine DNA glycosylase enzyme) (He et al. 2011; Maiti and Drohat 2011; Zhang et al. 2012). A second mechanism involving AID and APOBEC enzymes has been proposed but is still controversial. In this case, the 5hmC would be deaminated in 5hmU, and further removed by SMUG1 or TDG glycosylases and ultimately replaced by an unmodified cytosine (Guo et al. 2011).

Enzymatic decarboxylation of 5caC: 5-carboxylcytosine may also be directly decarboxylated by a yet unknown enzyme. *In vitro* experiments using embryonic stem (ES) cells lysates have shown a direct conversion of 5CaC into cytosine without BER (Schiesser et al. 2012).

TET enzymes and 5hmC are present at various levels in embryonic tissues. TET3, but not TET1 nor TET2, is particularly abundant in the zygote. It is responsible for the massive oxidation of the male pronucleus that occurs just after fecundation. 5mC and 5hmC are then lost in a replication-dependent fashion, resulting in a drop of modified cytosines by the 16-cell (Gu et al. 2011; Wossidlo et al. 2011). The second wave of global DNA demethylation occurs in PGCs. This time TET1 and TET2, but not TET3, are involved in the massive oxidation of 5mC. As in the early embryo, the resulting 5hmC are not excised by an active mechanism but are rather diluted out through cell divisions (Hackett et al. 2013).

Besides their involvement in embryonic demethylation processes, TET enzymes are considered to be important regulators of DNA methylation fidelity (Williams et al. 2012). Genome wide analyses in ES cells have shown that TET1 binds to the promoter region of a large number of genes (Wu and Zhang 2011b). Most TET1-bound promoters contain a CpG island and are maintained unmethylated, suggesting that TET1 is a major player in maintaining CpG islands free of methylation (Wu and Zhang 2011b; Xu et al. 2011). Consistently, 5hmC also localizes on TSSs (Pastor et al. 2011; Williams et al. 2011) and is enriched in region of higher CpG density than 5mC, indicating that 5mC is converted into 5hmC by TET1 specifically at CpG islands (Ito et al. 2011). This supports the notion that TET1 is involved in protecting CpG islands against aberrant methylation that might occur sporadically. This function is thought to be particularly important in the embryo when *de novo* methylation takes place. At this stage, TET1 and TET2 become highly expressed (Ito et al. 2010; Koh et al. 2011) and are believed to be important to maintain specific regions unmethylated, protecting them from illegitimate methylation that could result from the high activity of the DNMT3A and DNMT3B.

The mechanisms allowing TET enzymes to target specific DNA sequences are not fully understood. Genes and CpG islands targeted by TET1, TET2 and TET3 overlap extensively but not completely. TET1 and TET3 mainly associate with promoters of high CpG content (Williams et al. 2011; Wu et al. 2011; Xu et al. 2011), and this is supported by the fact that both enzymes have a CXXC domain, which typically binds CpG dinucleotides (Zhang et al. 2010a; Xu et al. 2011). TET2 functions more prominently at low CpG density promoters. Furthermore, TET1, TET2 and TET3 are enriched at polycomb-marked H3K27me3 and H3K4me3-rich promoters (Williams et al. 2011; Wu et al. 2011).

Compared to the relatively constant levels of 5mC in somatic tissues (3–4% of total cytosines), 5hmC levels are significantly lower and vary greatly depending on the cell type (0.1%–0.7% of total cytosines) with the highest levels in brain tissues (Globisch et al. 2010; Szwagierczak et al. 2010; Nestor et al. 2012). Interestingly, although 5hmC exists in mouse embryonic stem (ES) cells at high levels, it decreases significantly after ES cell differentiation but rises again in

terminally differentiated cells, such as Purkinje neurons (Kriaucionis and Heintz 2009; Tahiliani et al. 2009; Szwagierczak et al. 2010).

TET1 was found to be activated by neuronal activity, and this leads to demethylation of specific promoters. As neuronal cells are quiescent and do not divide, the TET1-mediated demethylation is here probably more “active”, involving the AID/APOBEC complex (Guo et al. 2011).

ROLE OF DNA METHYLATION IN CANCER

Despite the remarkable stability of DNA methylation profiles, profound alterations occur in tumor cells. Both local gains (hypermethylation) and genome wide losses (hypomethylation) of DNA methylation are observed. These two opposite phenomena co-exist in the same tumor, and are involved in tumor progression. The causes and consequences of the mechanisms leading to these alterations are however still not well understood. Moreover, despite the fact that DNA hypermethylation and hypomethylation co-exist in the same tumors, they have been generally described as the consequence of independent mechanisms (Ehrlich et al. 2002; Ehrlich 2006; Kushwaha et al. 2016).

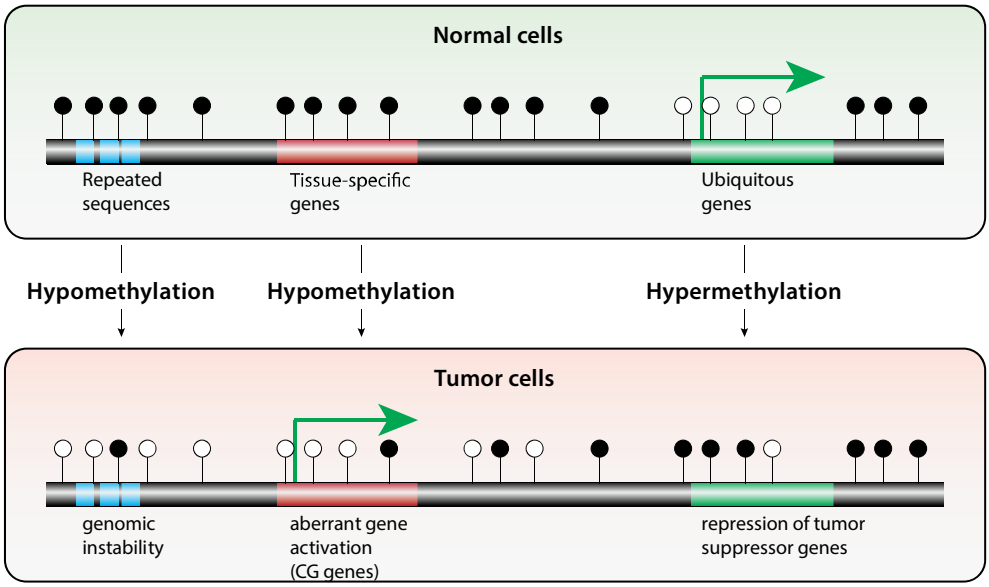


Figure 15: Illustration of aberrant DNA methylation profiles in tumor cells. In normal cells, the genome is globally methylated. Repeated sequences and tissue-specific genes are generally methylated whereas the promoter of ubiquitously expressed genes are unmethylated. In tumor cells, a global loss of DNA methylation (hypomethylation) is observed globally including repeated sequences and the promoter of genes such as cancer-germline (CG) genes. Gain of DNA hypermethylation (hypermethylation) is also observed at specific sites such as tumor suppressor genes promoters inducing the silencing of the corresponding gene. White lollipop: non-methylated CpG, black lollipop: methylated CpG.

6 Local DNA hypermethylation

DNA hypermethylation seems to occur across a large number of malignancies (Jones and Baylin 2007). As the majority of the CpGs of the genome are initially methylated in normal somatic cells, the unique umethylated regions correspond to CGIs. A gain of methylation on these CGIs, frequently localized in gene promoters, is associated with the silencing of neighboring genes. This phenomenon has been reported for several tumor suppressor genes and has been associated with their repression (Baylin 2005; Esteller 2005). It appears that DNA hypermethylation serves as general mechanism of tumor promotion, since it has been linked with loss of expression of a variety of tumor suppressor genes (Jin et al. 2009; Selaru et al. 2009; Baylin and Jones 2011; Agarwal et al. 2012), and affects many cellular pathways such as DNA repair, cell cycle, apoptosis or cell adherence (Choi and Lee 2013).

The mechanisms that direct the DNA methylation machinery to tumor suppressor genes promoters are still unclear. Emerging data suggest that polycomb repressive complex proteins (PRCs) and their associated H3K27me3 histone mark might attract DNMTs and induce local hypermethylation (Rose and Klose 2014). In cancer cell lines there is an extensive overlap between DNA methylation and H3K27me3 methylation (Brinkman et al. 2012). Moreover, promoters that are marked with H3K27me3 in ES cells are more likely to become hypermethylated in cancer, as compared with those lacking H3K27me3 (Schlesinger et al. 2007).

Another mechanism proposed to target DNA hypermethylation to specific regions in tumors is based on the targeting by certain transcription factors. Certain oncogenes have been reported to recruit DNMTs onto the promoter region of target genes (Di Croce et al. 2002). For example, it has been shown that the Myc oncogene induces the recruitment of DNMT3A on the promoter of *CDKN1A*/*p21*, leading to its hypermethylation and subsequent silencing (Brenner et al. 2005).

As already mentioned in the previous chapter, one of the roles of demethylation enzymes TET is to maintain unmethylated regions free of DNA methylation. In cancer, TET enzymes are frequently mutated or downregulated (Delhommeau et al. 2009; Yang et al. 2013). Recent evidence confirmed that loss of TET activity is indeed concomitant with aberrant promoter or enhancer hypermethylation in cancer (Rasmussen et al. 2015; Thomson et al. 2016; Wiehle et al. 2016).

7 Global DNA hypomethylation

Tumor progression is also associated with global DNA hypomethylation. It has been observed that a number of various tumor types can show a reduction of DNA methylation ranging from 10% to 60% relative to the corresponding normal tissues (Esteller et al. 2001; Ehrlich et al. 2002). This alteration affects virtually all tumors even if some tumor types such as leukemia show only rare hypomethylation events (Pfeifer et al. 1988). DNA hypomethylation affects numerous genomic regions including satellite and genomic repeated sequences but also the different parts of a gene (Kushwaha et al. 2016). This alteration seems to be associated with tumoral progression as metastatic tumors show more frequent and deep DNA hypomethylation than primary tumors (Gama-Sosa et al. 1983).

The effect of DNA hypomethylation on tumor biology is still not clear. It was initially proposed that DNA hypomethylation could induce the activation of oncogenes. However, even if it has been reported that well-known oncogenes such as Ras (Feinberg and Vogelstein 1983) are demethylated in cancer, there is no significant correlation with their activation. Conversely, cancer-associated hypomethylation correlates with the activation of genes including cancer-germline (CG) genes, oncogenic role of which has not been clearly demonstrated. We will discuss the potential role of these genes in tumor progression in the next chapter named “potential role of cancer-associated DNA hypomethylation through the activation of cancer-germline genes” (page41).

It has also been suggested that DNA hypomethylation could lead to tumor progression by activating transposable elements. Indeed, DNA hypomethylation is associated with the activation of transposable elements. However, the mutation due to the insertion of such elements has only been rarely observed (Costello and Plass 2001).

A correlation between DNA hypomethylation and genomic instability has been observed in several tumor types (Schulz et al. 2002; Tsuda et al. 2002). This correlation has also been observed in patients with the immunodeficiency centromeric region instability-facial anomalies (ICF) syndrome. These patients harbor a mutation of DNMT3B associated with the hypomethylation of pericentromeric satellite regions and show frequent chromosomal rearrangements (Xu et al. 1999). Additional evidence of a functional role of hypomethylation in tumor progression was provided by a study on mice carrying a hypomorphic DNMT1 allele. These mice developped aggressive lymphoma associated with frequent chromosomal aberrations (Gaudet et al. 2003).

8 Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes (Review article; Cancer letters; 2017)

Aurélie Van Tongelen[1], Axelle Loriot[1], and Charles De Smet[1]

[1] Group of Genetics and Epigenetics, de Duve Institute, Université Catholique de Louvain, Brussels, Belgium

➔ *Review Article : Cancer Lett. 2017 Mar 22.*
pii: S0304-3835(17)30203-3.
doi: 10.1016/j.canlet.2017.03.029.

Abstract

Global loss of DNA methylation is frequently observed in the genome of human tumors. Although this epigenetic alteration is clearly associated with cancer progression, the way it exerts its pro-tumoral effect remains incompletely understood. A remarkable consequence of DNA hypomethylation in tumors is the aberrant activation of "cancer-germline" genes (also known as "cancer-testis" genes), which comprise a diverse group of germline-specific genes that use DNA methylation as a primary mechanism for repression in normal somatic tissues. Here we review the evidence that such cancer-germline genes contribute to key processes of tumor development. Notably, several cancer-germline genes were found to stimulate oncogenic pathways involved in cell proliferation (SSX, DDX43, MAEL, PIWIL1), angiogenesis (DDX53), immortality (BORIS/CTCFL), and metastasis (CT-GABRA3). Others appear to inhibit tumor suppressor pathways, including those controlling growth inhibition signals (MAGEA11, MAGEB2), apoptosis (MAGEA2, MAGEC2), and genome integrity (HORMAD1, NXF2). Cancer-germline genes were also implicated in the regulation of tumor metabolism (MAGEA3/MAGEA6). Together, our survey substantiates the concept that DNA hypomethylation promotes tumorigenesis via transcriptional activation of oncogenes. Importantly, considering their highly restricted pattern of expression, cancer-germline genes may represent valuable targets for the development of anti-cancer therapies with limited side effects.

8.1 Introduction

DNA methylation corresponds to the covalent addition of a methyl group on cytosine to form a 5-methylcytosine (5mC), and represents the predominant DNA modification in vertebrate genomes [1]. DNA methylation occurs mostly in a cytosine-phosphate-guanine (CpG) dinucleotide context. In the human genome, 5mC is present at most CpGs (70-80%), but several regions with densely clustered CpGs, termed CpG islands, remain mostly unmethylated [2]. About half of all human genes comprise a CpG island overlapping their promoter and first exon. DNA methylation profiles are established during development, and are thereafter largely maintained during cell divisions in the adult by a replication-coupled maintenance process involving the DNMT1 DNA methyltransferase [3]. A large number of evidences demonstrated that DNA methylation acts to repress transcription, and it is reasonable to propose that DNA methylation plays critical roles in the regulation of tissue-specific gene expression and human developmental processes [4].

Despite the remarkable stability of DNA methylation profiles, profound alterations occur in tumor cells with both local gains (hypermethylation) and genome-wide losses (hypomethylation) of DNA methylation [5]. Great interest in DNA hypermethylation was stimulated by its direct impact on tumor suppressor genes [6]. DNA hypermethylation has indeed been identified in the CpG island promoter of tumor suppressor genes, and has been linked with irreversible silencing of these genes. DNA hypermethylation contributes thereby to tumor progression by disrupting several key cellular pathways such as DNA repair, cell cycle control, apoptosis or cell adherence [7]. Tumor progression was also associated with DNA hypomethylation, which affects numerous genomic regions and represents a common feature of many tumors [8]. The effect of DNA hypomethylation on tumor biology is however less understood. A first possibility is that DNA hypomethylation would have a direct impact on chromatin integrity, thereby increasing genome instability. Supporting evidence was provided by a study with mice carrying a hypomorphic allele of DNMT1. The study showed that DNA hypomethylation was associated with the development of tumors displaying a higher frequency of chromosomal rearrangements [9]. The association between DNA hypomethylation and genomic rearrangements

has also been observed in human cell lines [10, 11], and in cells from patients with the immunodeficiency centromeric region instability-facial anomalies (ICF) syndrome [12]. In many tumor types, however, the frequency of genomic rearrangements does not correlate with the extent of DNA hypomethylation, suggesting that this epigenetic alteration contributes in other ways to tumor initiation and progression [13, 14]. A prevailing alternative hypothesis to the role of DNA hypomethylation in tumor initiation and development is that it leads to the transcriptional activation of oncogenes [5, 15]. Evidence for this remains however equivocal. On the one hand, well-known oncogenes were shown to be hypomethylated in tumors [16, 17], but this has not been convincingly associated with their transcriptional activation [18, 19]. On the other hand, genes that evidently become activated in tumors through promoter hypomethylation have been identified [20], but information about their oncogenic function remains scarce. Most of the latter genes belong to the so-called cancer-germline (CG) group of genes, as their expression in healthy adults is normally restricted to testicular germline cells [21]. The aim of the present review is to give an overview of the potential pro-tumoral role of such DNA-hypomethylation-responsive genes, and thereby to substantiate the concept that DNA hypomethylation contributes to tumorigenesis through the activation of oncogenes.

8.2 Insight into the definition of cancer-germline genes

Previous work focusing on genes coding for tumor-specific antigens has led to the isolation of a group of genes with a unique expression profile, i.e. they are normally expressed exclusively in germline cells and show aberrant activation in a wide variety of tumors [22]. This particular expression profile earned them the name “cancer-germline” (CG) (also termed “cancer-testis” (CT) genes). Certain types of tumors display frequent co-activation of CG genes, including lung cancer, head and neck cancer, bladder cancer, and melanoma. In other types of tumors, like colon cancer, renal cancer, or leukemia, activation of CG genes is less frequent [23].

More than 270 human CG genes have been identified, and were registered in the CT database (http://www.cta.lncc.br) [24]. These genes are

dispersed on various chromosomes, but with a striking enrichment on the X chromosome. While the function of many CG genes remains unclear, it appears that these genes exert a variety of cellular functions.

As Hofmann pointed out, the definition of CG genes differs vastly according to the literature source [25]. Whereas some authors restrain the definition to genes with an exclusive expression in testis and cancer, others are more permissive accepting genes with some level of expression in a few non-gametogenic tissues [23]. We therefore re-evaluated the tissue-specificity of CG genes currently recorded in the CT database, by exploiting the Genotype-Tissue Expression (GTEx) database [26], which provides RNA-seq data obtained from a wide variety of normal human tissues with, for each of these, a large amount of samples. As a result, 77% of the genes recorded in the CT database displayed the expected expression pattern, as they were expressed exclusively in testicular samples (testis-specific). Another group of genes (12%) had their highest expression in testis, but were nevertheless expressed at some level in somatic tissues (testis-preferential). Finally, several genes (11%) did not exhibit the expected tissue-specificity, as they showed comparable expression levels in testis and in one or several other tissues.

8.3 A significant subset of CG genes are regulated by DNA methylation

It is commonly recognized that many CG genes use DNA methylation as a primary mechanism of transcriptional regulation. This has been demonstrated extensively for MAGEA1, the archetypal member of CG genes. Initial experiments revealed that MAGEA1 can be induced in non-expressing cancer cells following exposure to 5-aza-deoxycytidine (5-azadC), an inhibitor of DNA methyltransferases [27]. It was subsequently shown that the MAGEA1 promoter is methylated in all normal adult tissues, except testicular germline cells, and that its activation in cancer cells correlates strictly with promoter demethylation [28, 29]. Importantly, activation of MAGEA1 in tumor was found to be associated with global genome hypomethylation [29]. Transfection experiments with in vitro methylated MAGEA1 constructs further confirmed the dominant role of DNA methylation on MAGEA1 promoter repression [30]. More recently, experiments using specific antisense oligonucleotides or siRNAs indicated that

transient inhibition of DNMT1 suffices to induce long term activation of MAGEA1, and that this was associated with local changes of the chromatin, which converted from a repressive to a fully active configuration [31, 32]. Altogether, these observations indicate that DNA methylation serves as the primary layer of repression for the MAGEA1 gene, and that DNA hypomethylation is a sufficient trigger to induce activation of this gene in tumor cells.

Besides MAGEA1, other CG genes have been reported to be dependent on DNA methylation for transcriptional regulation, on the basis of both induction upon treatment with a DNA methylation inhibitor, and strict correlation between expression and promoter methylation status. Most of these genes map on the X chromosome, and show specific expression during the pre-meiotic stages of germline development [33]. We have listed all CG genes for which evidence was provided indicating that their regulation relies on DNA methylation (Table 1). For some of these genes, DNA methylation may not be the sole mechanism of transcriptional repression. Some CG genes, indeed, show a more restricted pattern of activation in specific tumor types [34, 35], suggesting that they require the presence of tissue-specific transcription factors in addition to promoter demethylation to permit stable transcriptional activation.

Finally, it appears that a significant proportion of CG genes rely on mechanisms other than DNA methylation for transcriptional regulation. Activation of these genes in tumors is therefore not related to DNA hypomethylation. During both male and female gametogenesis, most CG genes in this group are expressed after the onset of meiosis [33].

Germline-specific (exclusive expression in testis/ovary)	
X -linked	*CT45A1, CT47B1, CTAG2,* ***DDX53****, FTHL17, GAGE1, GAGE2A, LUZP4; MAGEA1,* ***MAGEA2****,* ***MAGEA3****, MAGEA4,* ***MAGEA6****, MAGEA8, MAGEA9, MAGEA10,* ***MAGEA11****, MAGEA12,* ***MAGEB2****, MAGEB6, MAGEC1,* ***MAGEC2****, SPANXA1, SPANXD,* ***SSX1****,* ***SSX2****,* ***SSX5****, SSX6, XAGE, ZNF645*
Autosomal	*ACTLB, BAGE2, BRDT, CAGE1, CRISP2, LDHC,* ***CTCFL****, DMRT1, DPPA2,* ***MAEL****,* ***PIWIL1****, TDRD1, TSPY1, SYCP1*
Germline-preferential (highest expression level in testis/ovary)	
X-linked	***NXF2***
Autosomal	***DDX43****,* ***HORMAD1,*** *LY6K,* ***PIWIL2****, PRAME, SPA17, SPAG6, SYCE1, ZNF165*

Table 1. List of validated germline-specific and methylation-sensitive CG genes. Genes recorded in the CT database were screened through pubmed search for evidences of DNA methylation-dependent transcriptional regulation. The tissue-specificity of selected methylation-sensitive genes was confirmed through analysis of the GTEx database. Validated CG genes were divided into two groups: the "Germline-specific" genes are expressed only in testis or ovary (RPKM > 0.5), and in no other normal tissues (RPKM < 0,5); the "Germline-preferential" genes show some level expression in somatic tissues, but nevertheless have an expression level in testis or ovary at least 5 times higher than in any other tissue. Several CG genes (bold) exhibit an oncogenic potential, and are discussed in this review.

8.4 Methylation-dependent CG genes are involved in multiple cancer pathways

As we proposed earlier in this review, it is possible that the tumor-promoting impact of DNA hypomethylation is mediated to some extent by the activation of CG genes, which represent prime targets of this epigenetic alteration. This implies however that DNA hypomethylation induces activation of CG genes with oncogenic potential. The oncogenic function of CG genes has been discussed in recent reviews [36], but with no particular focus on CG genes that owe their activation in tumors to DNA demethylation. Here, we examined the cellular functions that were so far attributed to DNA methylation-sensitive

CG genes (Table 1), in order to determine which of these genes might contribute to the capabilities that cells must acquire during tumor initiation and development. Oncogenic roles ascribed to CG genes were subdivided according to the hallmarks of cancer previously identified by Hanahan and Weinberg [37, 38].

Self-sufficiency in growth signal

The most evident capability of cancer cells is to sustain chronic proliferation. Contrasting with normal cells, which require the presence of mitogenic signals to proliferate, cancer cells acquire self-sufficiency in growth signal. Central to the acquisition of autonomous stimulation of proliferation in cancer cells are alterations that lead to uncontrolled activation of the RAS-mitogen activated protein kinase (MAPK) signaling pathway [39], or the phosphatidylinositol 3-kinase (PI3K)/AKT pathway [40].

The CG gene Maelstrom (MAEL) has been reported to be upregulated in multiple cancer types, including breast cancer, lung cancer, colon cancer, hepatocellular carcinoma, bladder cancer and prostate cancer [41-43]. In healthy individuals, MAEL mainly contributes with Piwi proteins to the repression of transposable elements during spermatogenesis through a process involving piRNAs, a class of small non-coding RNAs [44]. Recent evidence suggests that MAEL is also involved in activation of the PI3K/AKT pathway [41, 42]. This appears to be mediated by the ability of MAEL to repress the expression of miR-7, a known negative regulator of the AKT pathway [45]. In human hepatocellular and urothelial carcinoma cells, knockdown of MAEL by small hairpin RNA (shRNA) was sufficient to reduce cell proliferation, whereas overexpression of MAEL promoted cell growth in vitro, and increased tumor volume in vivo in xenograft mouse models [41, 42].

Intriguingly, PIWIL1 (Piwi-like RNA-mediated gene silencing 1), which is another CG gene associated with the piRNA process, appears to be involved in the regulation of the PI3K/AKT pathway as well. PIWIL1 activation was observed in a significant proportion of gastric cancer, prostate adenocarcinoma, glioma, and non-small-cell lung carcinomas [46-49]. In gastric cancer cells, PIWIL1 was

shown to support cell proliferation [47]. This appears to be linked with a positive effect of Piwil1 on the PI3K/Akt pathway, which is due to its ability to inhibit expression of PTEN (the phosphatase and tensin homolog), a well-known negative regulator of AKT. Indeed, PIWIL1 was found to promote DNA hypermethylation of the promoter of PTEN, in part by increasing DNMT1 expression [50]. Recently, a proliferation-inducing role was also attributed to PIWIL2, a paralog of PIWIL1 [51].

In several uveal melanoma samples, constitutive hyperactivation of RAS was found to be dependent on upregulation of DDX43 (DEAD-box helicase 43, also termed HAGE), a gene categorized in the "germline-preferential" group of CG genes. Experimental depletion of DDX43 in uveal melanoma cells resulted in inhibition of both MAPK and PI3K/AKT pathways [52, 53]. Moreover, the expression of DDX43 has been associated with poor clinical outcome in breast cancer patients [54].

In cutaneous melanoma cells, members of the SSX family of CG genes were shown to contribute to cellular proliferation. Experimental evidence suggests that SSX proteins contribute to the MAPK signaling pathway [55]. The precise function of SSX proteins is still unclear, but it is proposed that they may act as transcriptional regulators.

Insensitivity to growth inhibition

In normal tissues, cell proliferation is under the control of robust negative regulators to ensure tissue homeostasis. Most such anti-proliferative signals converge onto the family of retinoblastoma proteins (RB1, RBL1/p107, RBL2/p130), which mediate cell cycle arrest. In conditions of growth restriction signaling, these proteins are in a hypo-phosphorylated state, which allows them to bind and inhibit a class of transcription factors (E2F) that control proliferation-associated genes. In many tumors, the RB pathway is altered, leading to loss of sensitivity to anti-proliferative signals and unleashed proliferation [56]. Recent evidence suggests that CG genes may interfere with the RB pathway.

MAGEA11 belongs to the large family of MAGE genes, many of which were identified as CG genes [57]. Proteins encoded by the MAGE gene family members are characterized by the presence of a conserved stretch of about 200 amino acids, named the MAGE homology domain. It is believed that this domain serves for protein-protein interactions. Outside of the MAGE homology domain, MAGE proteins display great sequence variation, and the physiological role of most of them remains unclear. Activation of MAGEA11 is frequently observed in prostate cancer, and has been associated with increased tumor cell growth [58]. It was demonstrated that the MAGEA11 protein interacts with RBL1/p107, which is then converted into an activator, rather than a repressor, of E2F1 transcription activity [59].

MAGEB2, another member of the MAGE family, is activated in a variety of human tumors, including lung carcinoma, and head and neck carcinoma [60]. The growth promoting potential of MAGEB2 was initially demonstrated in a transformed human keratinocyte cell line [61]. More recently, this observation was extended to other cell types, and was found to rely on the ability of the MAGEB2 protein to enhance E2F transcriptional activity. This effect probably involves binding of MAGEB2 to HDAC1, a histone deacetylase that normally contributes to sequester E2F transcription factors together with RBL1/P107 in a repression complex [62].

Resisting cell death

Programmed cell death, or apoptosis, is a natural process that can be induced in cells that sense abnormal conditions such as excessive proliferation, hypoxia, DNA damage, and loss of normal cell-cell or cell-matrix contact. Whether cells live or die is dictated by a delicate balance between pro-apoptotic and anti-apoptotic factors. It is well documented that the development of aggressive tumors requires abrogation of pro-apoptotic pathways [63].

Many cell death promoting signals converge on p53, a transcription factor that directs upregulation of pro-apoptotic genes. Experimental evidence indicates that several members of the MAGE protein family exert a negative impact on the p53 pathway in tumors through different mechanisms. Proteins of

the MAGEA subfamily, including MAGEA2, interact directly with p53, impeding its access onto the promoter of target genes [64]. MAGEA2 was also shown to recruit a histone deacetylase (HDAC3), thereby favoring erasure of acetylation marks on p53, a modification that is essential for p53 pro-apoptotic activity [65, 66]. Other MAGE proteins, including MAGEC2, were found to associate with TRIM28, an E3 ubiquitin ligase that targets p53 for proteolytic degradation [67]. MAGEC2 was shown to increase the ubiquitin ligase activity of TRIM28, thereby accelerating destruction of p53.

The anti-apoptotic role of MAGE proteins was recently supported by studies in mice carrying a deletion encompassing six murine members of the MAGEA gene subfamily [68]. It was observed that these mice had smaller testes, and that this was due to increased apoptosis of germ cells in the first wave of spermatogenesis. The authors also observed increased amounts of p53 and enhanced induction of downstream target genes in germ cells of genetically modified mice.

Not only MAGE proteins, but also other CG proteins appear to protect cancer cells against apoptosis. This was suggested for instance for SPAG6 [69], BORIS/CTCFL [70], and GAGE7 [71]. However, the underlying molecular mechanisms, which do not appear to involve direct interaction with p53, remain less precisely understood.

Immortality

Most normal cells cannot divide indefinitely. This limit in proliferation capacity is largely dictated by the shortening of chromosome ends, called telomeres, which lose a portion of DNA sequence at each replication cycle. Upon excessive shortening, telomeres trigger a signaling cascade leading to a permanent cell cycle arrest, termed senescence [72].

It is believed that senescence serves as a natural barrier to the development of cancer. Progressing tumors, however, avoid this fate by activating a mechanism of elongation of telomeres. This is often achieved through upregulation of hTERT, the gene that encodes the catalytic subunit of

the telomere lengthening complex, termed telomerase [72]. In several tumors, induction of telomerase activity is associated with DNA hypermethylation of the first exon of hTERT, which encompasses a binding site for the CTCF transcriptional repressor. This leads to inhibition of binding of CTCF, and hence de-repression of hTERT [73, 74].

BORIS/CTCFL (brother of regulator of imprinted sites) is a germline-specific homolog of CTCF, which displays a typical cancer-germline pattern of expression, as its activation was detected in various types of cancers [75]. In testicular and ovarian tumors, where hypermethylation within the hTERT gene is not observed, expression of BORIS/CTCFL was found to be essential for hTERT mRNA up-regulation [76]. When present in such tumor cells, BORIS/CTCFL competes with CTCF for binding to the hTERT gene. Unlike CTCF, however, BORIS/CTCFL exerts a positive effect on hTERT expression. Activation of BORIS/CTCFL appears therefore to represent an alternative mechanism of induction of hTERT in cancer cells.

Angiogenesis

Formation of new blood vessels, termed angiogenesis, is a normal physiological process that is particularly active during embryonic development, wound healing, and periodically in female reproductive organs. Angiogenesis is a complex phenomenon resulting from the action of pro- and anti-angiogenic factors, which regulate the growth, migration, and tube formation ability of vascular endothelial cells. Beyond a critical size, tumors must develop a vascular network to provide oxygen and nutrients to the cells located at the center. This is rendered possible in tumors by modifications in the balance of pro- and anti-angiogenic factors [77].

The CG gene DDX53 (DEAD-box helicase 53, also named CAGE) appears to promote angiogenesis. Activation of DDX53 was observed in a variety of tumors, including gastric and endometrial cancers, as well as in hematological cancers [78]. The gene encodes a protein with a DEAD box domain, suggestive of an RNA helicase activity. DDX53 protein is mainly localized in the nucleus, but it is believed that it can also be secreted out of the cell. Experiments showed

that addition of recombinant DDX53 protein to the culture medium of endothelial cells stimulated their tube formation ability, and increased their invasion potential [79]. Consistently, endothelial cells exposed to recombinant DDX53 showed increased expression of PAI-1 (plasminogen activation inhibitor 1), a known stimulator of angiogenesis [80]. Similar results were obtained when endothelial cells were exposed to the supernatant of DDX53-expressing cancer cells. It remains unclear however how DDX53 enters endothelial cells and stimulates their angiogenic potential.

Invasion and metastasis

Metastasis, i.e. the formation of tumor masses in distant tissues, is a critical event in cancer progression and one of the most life-threatening events for patients with cancer. The development of metastases is a highly complex process, as it implies that cancer cells invade the adjacent tissue, enter the bloodstream, and establish in a distant new environment. Recent evidence revealed that activation of a CG gene (CT-GABRA3) in tumors helps weakening the endothelial barrier, an essential step to permit intravasation of cancer cells into the blood vessel [81].

GABRA3 is known as a brain-specific gene that encodes a subunit of a receptor for the Gamma-aminobutyric acid (GABA) neurotransmitter. Recently, an alternative transcript of the gene was identified, which uses a germline-specific promoter located about 250kb upstream of the classical brain-specific promoter. The new transcript, termed CT-GABRA3, displays typical CG features, as it is expressed exclusively in germline cells but shows promoter hypomethylation and activation in several tumor types, including melanoma and lung cancer. CT-GABRA3 harbors two microRNAs, miR-767 and miR-105, which similarly show ectopic expression in tumors [81]. Importantly, in a mouse model of xenografted human breast cancer cells, miR-105 was found to stimulate metastasis [82]. It was demonstrated that miR-105 is secreted out of the tumor cells via exosomes, and reaches nearby endothelial cells. In endothelial cells, miR-105 inhibits the expression of the tight junction protein

TJP1, thereby disrupting the endothelial barrier and facilitating intravasation of cancer cells into the bloodstream [82].

Metastatic adaptation

Cells have evolved molecular mechanisms allowing adaptation of energy usage according to nutrient availability. A central actor in this process is AMPK (AMP activated protein kinase), which serves as a sensor of intracellular ATP concentration [83]. Upon decreased ATP levels, AMP takes its place in the complex formed with AMPK. This results in the activation of AMPK, which inhibits anabolic energy-consuming pathways, and at the same time stimulates catabolic processes. AMPK activation can ultimately lead to autophagy, a process during which cytoplasmic components are enclosed in vesicles before degradation and subsequent recycling of catabolites [84]. One of the targets of AMPK is mTOR (mammalian target of rapamycin), a serine/threonine kinase that regulates several cellular processes such as protein synthesis and autophagy. Upon ATP depletion, activated AMPK inhibits mTOR, thereby inhibiting protein synthesis and promoting autophagy.

In several tumors, the AMPK pathway was found to be suppressed through mutations in upstream regulators [85, 86]. This enables high energy-demanding cancer cells to avoid the growth-restrictive consequences of AMPK activation. Recently, two members of the MAGEA subfamily (MAGEA3 and MAGEA6) were found to be involved in AMPK suppression [87]. Experimental evidence indicates that MAGEA3 and MAGEA6 proteins interact with the TRIM28 ubiquitin ligase (as other MAGE proteins do), and recruit at the same time AMPK, leading to its ubiquitination and degradation. Consistently, MAGEA3/A6-expressing tumor cells exhibited reduced activation of mTOR, inhibition of autophagy, and increased viability.

Of note, MAGEA3 and MAGEA6 show a high frequency of activation in tumors, particularly in melanoma and lung cancer, where their expression is observed in more than 60% of the lesions. In lung cancer, expression of both genes was found to correlate with poor survival [88].

Genome instability

Cells are normally endowed with a capacity to repair DNA damages, which may arise from chemical and physical insults or during replication. This genome maintenance system is often altered in cancer cells, leading to increased mutation rates, and thereby accelerating acquisition of genomic alterations that favor tumor progression [89]. Many breast tumors for instance, show loss of function of BRCA1 and BRCA2 genes, which encode proteins involved in a process of DNA repair that is based on homologous recombination (HR) [90, 91]. HR is an essential component of the genome maintenance machinery, as it contributes to a process of high fidelity repair of DNA double strand breaks by using the sister chromatid as a template for sequence restoration. Impaired BRCA1 function is prevalent in triple negative breast cancer (TNBC), a subtype of breast tumors that is characterized by lack of overexpression of HER2 and loss of expression of estrogen and progesterone receptors [92]. Consistently, TNBC exhibit high levels of genomic instability, including copy number alterations and loss of heterozygosity. In several TNBC samples, however, genomic instability is observed in the absence of loss of BRCA1 function, suggesting that other mechanisms may contribute to HR dysfunction.

Recent evidence indicates that activation of the CG gene HORMAD1 contributes to HR dysfunction in TNBC [93]. The protein encoded by HORMAD1 has important roles in meiosis, especially for the establishment of crossing-overs between homologous chromosomes during prophase 1 [94]. By inhibiting RAD51, a driver of HR-mediated repair between sister chromatids, HORMAD1 favors inter-homolog rather than inter-sister recombinations. This ensures the formation of chiasmata between homologous chromosomes, which are essential for normal segregation during the first meiotic division. It was demonstrated that aberrant expression of HORMAD1 in mitotic cells, such as breast cancer cells, induces HR deficiencies, and thereby increases the rate of copy number alterations [93]. It appears therefore that, besides loss of BRCA1 function, activation of HORMAD1 represents an alternative way to promote genomic instability in TNBC.

Abnormal spindle formation during mitosis is believed to represent another mechanism for the generation of tumor cells with altered genomes, as it leads to chromosome missegregation. This mitotic defect, however, is often associated with high rates of cellular death, and thereby with restricted population growth [95]. Hence, it has been proposed that tumors experiencing spindle dysfunctions require activation of mechanisms that restore mitotic fidelity to permit further expansion [96]. NXF2 (Nuclear RNA Export Factor 2) belongs to the "germline-preferential" group of CG genes, and is activated in several tumor types, including bladder, colorectal and lung carcinomas [35]. NXF2 was also identified in a genetic screen for genes that modulate responsiveness of lung cancer cells to paclitaxel, an inhibitor of microtubule depolymerization that interferes with normal spindle formation during mitosis [97]. Further experimental evidence was provided indicating that NXF2 minimizes spindle dysfunctions in tumor cells. It is therefore proposed that activation of this CG gene in cancer cells contributes to avoid excessive mitotic errors, which would otherwise impair tumor outgrowth [97]. The mechanism by which NXF2 regulates spindle formation is however unknown.

8.5 CONCLUSION

It is commonly asserted that DNA hypomethylation promotes cancer development in part through the activation of genes with oncogenic potential [5, 98]. However, examples of oncogenes that truly rely on DNA hypomethylation for activation in tumors remain scarce. CG genes represent ideal candidates, since they belong to a unique group of genes that primarily use DNA methylation for repression in somatic tissues, and are therefore a prime target of transcriptional activation in tumors with a hypomethylated genome. As reported in this review, a growing body of evidence indicates that CG genes can contribute to several key processes of tumor development, including uncontrolled proliferation, resistance to apoptosis, metastasis, adaptation to cellular energetics constraints, and genome instability (Fig. 1). These observations bring strong support to the notion that DNA hypomethylation induces activation of oncogenes in tumors cells, and it appears that germline-specific genes represent a major component of this process.

The observation that CG genes exert oncogenic functions may provide an explanation to the immunological paradox underlying their activation in tumors. Thus, many CG genes encode potent tumor-specific antigens, and there is strong evidence that their expression in cancer cells triggers attacks by immune cells [99]. This implies that DNA hypomethylation and the consequent activation of CG genes would have a negative impact on tumor development, at least at some stage of tumorigenesis. It is reasonable to propose that the oncogenic potential of CG genes compensates for their negative immunogenic effect.

Importantly, CG genes often show co-activation in tumors, and are thus expected to act concertedly in many cases. It is therefore possible that even if CG genes display only poor oncogenic potential on their own, they may act more potently when jointly expressed in the tumor cell. Interestingly, several CG proteins were reported to interact with other CG proteins [100], or to act as transcriptional regulators for other CG genes [101], thereby suggesting functional cooperativity. Cooperation between CG proteins may also be mediated through their combined effects on common pathways. For instance, multiple members of the MAGE family of CG proteins appear to converge onto the p53 regulator of apoptosis [67], while other CG proteins seem to inhibit apoptosis by targeting other regulators in the apoptotic pathway [70, 71].

There is currently growing interest in determining if CG genes exert crucial oncogenic roles [102]. It is expected indeed that oncogenes with such a restricted pattern of expression will represent ideal targets for the development of anti-cancer therapies with limited side effects. Our analysis revealed, however, that several of the genes that were reported to display a CG pattern of expression, showed in fact some level of expression in somatic tissues. Nevertheless, we identified CG genes with a highly testis-specific expression, and several among these appear to encode oncogenic functions. Further studies on these genes may therefore open the way to the development of highly selective anti-cancer therapies.

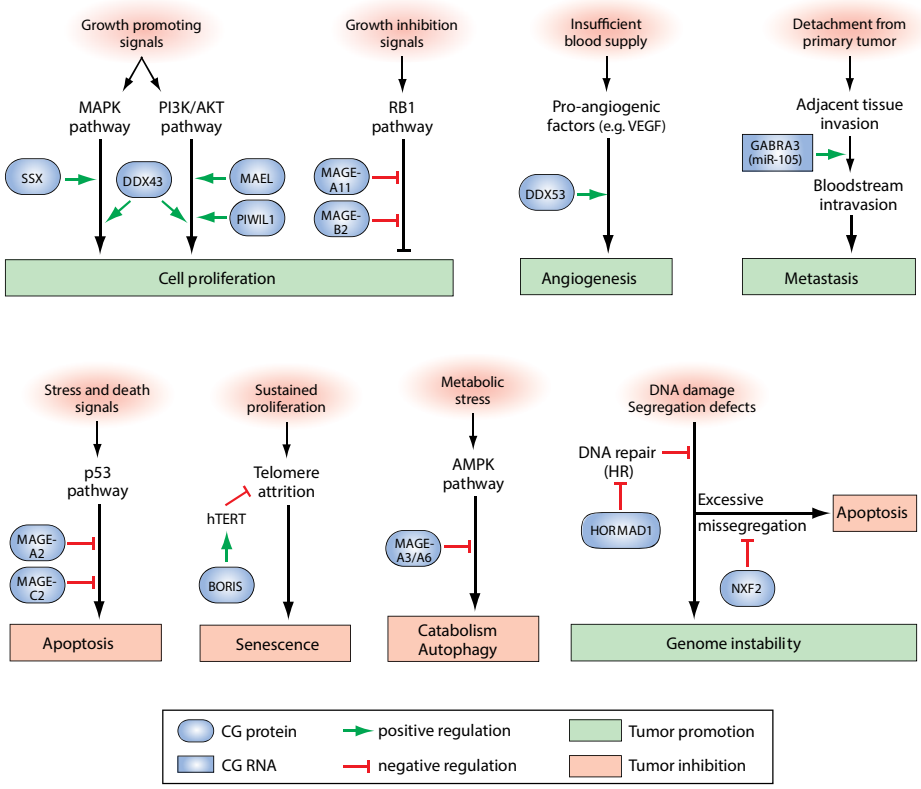


Figure 1. CG proteins and microRNAs contribute to various pro-tumoral processes, either by stimulating oncogenic functions or by inhibiting tumor suppressor pathways. The figure depicts CG proteins and miRNA for which strong evidence was provided that: (i) their expression in tumors is linked to DNA hypomethylation, and (ii) their function favors tumor development. See text for details.

Acknowledgments

This work was supported by grants from the D.G. Higher Education and Scientific Research of the French Community of Belgium (Action de Recherches Concertées) and from the Fonds special de recherche (FSR) of the Université catholique de Louvain, Belgium. A.V.T. was recipient of a Télévie grant from the FRS-FNRS, Belgium [#7.4581.13]. A.L. is supported by the de Duve Institute, Brussels, Belgium.

8.6 *REFERENCES*

[1] A. Bird, The essentials of DNA methylation., Cell, 70 (1992) 5-8.

[2] S.H. Cross, A.P. Bird, CpG islands and genes., Curr. Opin. Genet. Dev., 5 (1995) 309-314.

[3] T.H. Bestor, The DNA methyltransferases of mammals, Hum Mol Genet, 9 (2000) 2395-2402.

[4] A. Bird, DNA methylation patterns and epigenetic memory, Genes Dev, 16 (2002) 6-21.

[5] A.P. Feinberg, B. Tycko, The history of cancer epigenetics, Nat Rev Cancer, 4 (2004) 143-153.

[6] P.A. Jones, S.B. Baylin, The fundamental role of epigenetic events in cancer, Nat Rev Genet, 3 (2002) 415-428.

[7] M. Esteller, Epigenetics in cancer, N Engl J Med, 358 (2008) 1148-1159.

[8] M. Ehrlich, DNA methylation in cancer: too much, but also too little, Oncogene, 21 (2002) 5400-5413.

[9] F. Gaudet, J.G. Hodgson, A. Eden, L. Jackson-Grusby, J. Dausman, J.W. Gray, H. Leonhardt, R. Jaenisch, Induction of tumors in mice by genomic hypomethylation, Science, 300 (2003) 489-492.

[10] A. Vilain, N. Vogt, B. Dutrillaux, B. Malfoy, DNA methylation and chromosome instability in breast cancer cell lines, FEBS Lett, 460 (1999) 231-234.

[11] A. Vilain, J. Bernardino, M. Gerbault-Seureau, N. Vogt, A. Niveleau, D. Lefrancois, B. Malfoy, B. Dutrillaux, DNA methylation and chromosome instability in lymphoblastoid cell lines, Cytogenet Cell Genet, 90 (2000) 93-101.

[12] D. Gisselsson, C. Shao, C.M. Tuck-Muller, S. Sogorovic, E. Palsson, D. Smeets, M. Ehrlich, Interphase chromosomal abnormalities and mitotic missegregation of hypomethylated sequences in ICF syndrome cells, Chromosoma, 114 (2005) 118-126.

[13] M. Ehrlich, N.E. Hopkins, G. Jiang, J.S. Dome, M.C. Yu, C.B. Woods, G.E. Tomlinson, M. Chintagumpala, M. Champagne, L. Dillerg, D.M. Parham, J. Sawyer, Satellite DNA hypomethylation in karyotyped Wilms tumors, Cancer Genet Cytogenet, 141 (2003) 97-105.

[14] M. Ehrlich, DNA hypomethylation in cancer cells, Epigenomics, 1 (2009) 239-259.

[15] A.S. Wilson, B.E. Power, P.L. Molloy, DNA hypomethylation and human diseases, Biochim Biophys Acta, 1775 (2007) 138-162.

[16] M.S. Cheah, C.D. Wallace, R.M. Hoffman, Hypomethylation of DNA in human cancer cells: a site-specific change in the c-myc oncogene, J Natl Cancer Inst, 73 (1984) 1057-1065.

[17] A.P. Feinberg, B. Vogelstein, Hypomethylation of ras oncogenes in primary human cancers, Biochem Biophys Res Commun, 111 (1983) 47-54.

[18] M.G. Borrello, M.A. Pierotti, E. Tamborini, D. Biassoni, M.G. Rizzetti, S. Pilotti, G. Della Porta, DNA methylation of coding and non-coding regions of the human H-RAS gene in normal and tumor tissues, Oncogene, 7 (1992) 269-275.

[19] O.Y. Dokun, A.R. Florl, H.H. Seifert, I. Wolff, W.A. Schulz, Relationship of SNCG, S100A4, S100A9 and LCN2 gene expression and DNA methylation in bladder cancer, Int J Cancer, 123 (2008) 2798-2807.

[20] C. De Smet, A. Loriot, DNA hypomethylation and activation of germline-specific genes in cancer, Adv Exp Med Biol, 754 (2013) 149-166.

[21] A.J. Simpson, O.L. Caballero, A. Jungbluth, Y.T. Chen, L.J. Old, Cancer/testis antigens, gametogenesis and cancer, Nat Rev Cancer, 5 (2005) 615-625.

[22] P. Van Der Bruggen, Y. Zhang, P. Chaux, V. Stroobant, C. Panichelli, E.S. Schultz, J. Chapiro, B.J. Van Den Eynde, F. Brasseur, T. Boon, Tumor-specific shared antigenic peptides recognized by human T cells, Immunol Rev, 188 (2002) 51-64.

[23] M.J. Scanlan, A.J. Simpson, L.J. Old, The cancer/testis genes: review, standardization, and commentary, Cancer Immun, 4 (2004) 1.

[24] L.G. Almeida, N.J. Sakabe, A.R. deOliveira, M.C. Silva, A.S. Mundstein, T. Cohen, Y.T. Chen, R. Chua, S. Gurung, S. Gnjatic, A.A. Jungbluth, O.L. Caballero, A. Bairoch, E. Kiesler, S.L. White, A.J. Simpson, L.J. Old, A.A. Camargo, A.T. Vasconcelos, CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens, Nucleic Acids Res, 37 (2009) D816-819.

[25] O. Hofmann, O.L. Caballero, B.J. Stevenson, Y.T. Chen, T. Cohen, R. Chua, C.A. Maher, S. Panji, U. Schaefer, A. Kruger, M. Lehvaslaiho, P. Carninci, Y. Hayashizaki, C.V. Jongeneel, A.J. Simpson, L.J. Old, W. Hide, Genome-wide analysis of

cancer/testis gene expression, Proc Natl Acad Sci U S A, 105 (2008) 20422-20427.
[26] G.T. Consortium, The Genotype-Tissue Expression (GTEx) project, Nat Genet, 45 (2013) 580-585.
[27] J. Weber, M. Salgaller, D. Samid, B. Johnson, M. Herlyn, N. Lassam, J. Treisman, S.A. Rosenberg, Expression of the MAGE-1 tumor antigen is up-regulated by the demethylating agent 5-aza-2'-deoxycytidine, Cancer Res, 54 (1994) 1766-1771.
[28] C. De Smet, C. Lurquin, B. Lethé, V. Martelange, T. Boon, DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter., Mol. Cell. Biol., 19 (1999) 7327-7335.
[29] C. De Smet, O. De Backer, I. Faraoni, C. Lurquin, F. Brasseur, T. Boon, The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation, Proc Natl Acad Sci U S A, 93 (1996) 7149-7153.
[30] C. De Smet, A. Loriot, T. Boon, Promoter-dependent mechanism leading to selective hypomethylation within the 5' region of gene MAGE-A1 in tumor cells, Mol Cell Biol, 24 (2004) 4781-4790.
[31] J. Cannuyer, A. Loriot, G.K. Parvizi, C. De Smet, Epigenetic hierarchy within the MAGEA1 cancer-germline gene: promoter DNA methylation dictates local histone modification, PLoS ONE, 8 (2013) e58743.
[32] A. Loriot, E. De Plaen, T. Boon, C. De Smet, Transient down-regulation of DNMT1 methyltransferase leads to activation and stable hypomethylation of MAGE-A1 in melanoma cells, J Biol Chem, 281 (2006) 10118-10126.
[33] M. Koslowski, U. Sahin, C. Huber, O. Tureci, The human X chromosome is enriched for germline genes expressed in premeiotic germ cells of both sexes, Hum Mol Genet, 15 (2006) 2392-2399.
[34] C. Grunwald, M. Koslowski, T. Arsiray, K. Dhaene, M. Praet, A. Victor, A. Morresi-Hauf, M. Lindner, B. Passlick, H.A. Lehr, S.C. Schafer, G. Seitz, C. Huber, U. Sahin, O. Tureci, Expression of multiple epigenetically regulated cancer/germline genes in nonsmall cell lung cancer, Int J Cancer, 118 (2006) 2522-2528.
[35] A. Loriot, T. Boon, C. De Smet, Five new human cancer-germline genes identified among 12 genes expressed in spermatogonia, Int J Cancer, 105 (2003) 371-376.
[36] A.W. Whitehurst, Cause and consequence of cancer/testis antigen activation in cancer, Annual review of pharmacology and toxicology, 54 (2014) 251-272.
[37] D. Hanahan, R.A. Weinberg, The hallmarks of cancer, Cell, 100 (2000) 57-70.
[38] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell, 144 (2011) 646-674.
[39] J. Downward, Targeting RAS signalling pathways in cancer therapy, Nat Rev Cancer, 3 (2003) 11-22.
[40] T.L. Yuan, L.C. Cantley, PI3K pathway alterations in cancer: variations on a theme, Oncogene, 27 (2008) 5497-5510.
[41] X.D. Li, J.X. Zhang, L.J. Jiang, F.W. Wang, L.L. Liu, Y.J. Liao, X.H. Jin, W.H. Chen, X. Chen, S.J. Guo, F.J. Zhou, Y.X. Zeng, X.Y. Guan, Z.W. Liu, D. Xie, Overexpression of maelstrom promotes bladder urothelial carcinoma cell aggressiveness by epigenetically downregulating MTSS1 through DNMT3B, Oncogene, (2016).
[42] L. Liu, Y. Dai, J. Chen, T. Zeng, Y. Li, L. Chen, Y.H. Zhu, J. Li, Y. Li, S. Ma, D. Xie, Y.F. Yuan, X.Y. Guan, Maelstrom promotes hepatocellular carcinoma metastasis by inducing epithelial-mesenchymal transition by way of Akt/GSK-3beta/Snail signaling, Hepatology, 59 (2014) 531-543.
[43] L. Xiao, Y. Wang, Y. Zhou, Y. Sun, W. Sun, L. Wang, C. Zhou, J. Zhou, J. Zhang, Identification of a novel human cancer/testis gene MAEL that is regulated by DNA methylation, Mol Biol Rep, 37 (2010) 2355-2360.
[44] S.F. Soper, G.W. van der Heijden, T.C. Hardiman, M. Goodheart, S.L. Martin, P. de Boer, A. Bortvin, Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis, Developmental cell, 15 (2008) 285-297.
[45] J.W. Pek, A.K. Lim, T. Kai, Drosophila maelstrom ensures proper germline stem cell lineage differentiation by repressing microRNA-7, Developmental cell, 17 (2009) 417-424.
[46] L.F. Grochola, T. Greither, H. Taubert, P. Moller, U. Knippschild, A. Udelnow, D. Henne-Bruns, P. Wurl, The stem cell-associated Hiwi gene in human adenocarcinoma of the pancreas: expression and risk of tumour-related death, Br J Cancer, 99 (2008) 1083-1088.
[47] X. Liu, Y. Sun, J. Guo, H. Ma, J. Li, B. Dong, G. Jin, J. Zhang, J. Wu, L. Meng, C. Shou, Expression of hiwi gene in human gastric cancer was associated with proliferation of cancer cells, Int J Cancer, 118 (2006) 1922-1929.
[48] A. Navarro, R. Tejero, N. Vinolas, A. Cordeiro, R.M. Marrades, D. Fuster, O. Caritg, J. Moises, C. Munoz, L. Molins, J. Ramirez, M. Monzo, The significance of PIWI family expression in human lung embryogenesis and non-small cell lung cancer, Oncotarget, 6 (2015) 31544-31556.
[49] G. Sun, Y. Wang, L. Sun, H. Luo, N. Liu, Z. Fu, Y. You, Clinical significance of Hiwi gene expression in gliomas, Brain Res, 1373 (2011) 183-188.

[50] Z. Chen, Q. Che, F.Z. Jiang, H.H. Wang, F.Y. Wang, Y. Liao, X.P. Wan, Piwil1 causes epigenetic alteration of PTEN gene via upregulation of DNA methyltransferase in type I endometrial cancer, Biochem Biophys Res Commun, 463 (2015) 876-880.

[51] X. Qu, J. Liu, X. Zhong, X. Li, Q. Zhang, PIWIL2 promotes progression of non-small cell lung cancer by inducing CDK2 and Cyclin A expression, Journal of translational medicine, 13 (2015) 301.

[52] G. Ambrosini, R. Khanin, R.D. Carvajal, G.K. Schwartz, Overexpression of DDX43 mediates MEK inhibitor resistance through RAS Upregulation in uveal melanoma cells, Mol Cancer Ther, 13 (2014) 2073-2080.

[53] A.J. Linley, M.G. Mathieu, A.K. Miles, R.C. Rees, S.E. McArdle, T. Regad, The helicase HAGE expressed by malignant melanoma-initiating cells is required for tumor cell proliferation in vivo, J Biol Chem, 287 (2012) 13633-13643.

[54] T.M. Abdel-Fatah, S.E. McArdle, C. Johnson, P.M. Moseley, G.R. Ball, A.G. Pockley, I.O. Ellis, R.C. Rees, S.Y. Chan, HAGE (DDX43) is a biomarker for poor prognosis and a predictor of chemotherapy response in breast cancer, Br J Cancer, 110 (2014) 2450-2461.

[55] P. D'Arcy, W. Maruwge, B. Wolahan, L. Ma, B. Brodin, Oncogenic functions of the cancer-testis antigen SSX on the proliferation, survival, and signaling pathways of cancer cells, PLoS One, 9 (2014) e95136.

[56] C.J. Sherr, F. McCormick, The RB and p53 pathways in cancer, Cancer Cell, 2 (2002) 103-112.

[57] P. Chomez, O. De Backer, M. Bertrand, E. De Plaen, T. Boon, S. Lucas, An overview of the MAGE gene family with the identification of all human members of the family, Cancer Res, 61 (2001) 5544-5551.

[58] J.T. Minges, S. Su, G. Grossman, A.J. Blackwelder, E.A. Pop, J.L. Mohler, E.M. Wilson, Melanoma antigen-A11 (MAGE-A11) enhances transcriptional activity by linking androgen receptor dimers, J Biol Chem, 288 (2013) 1939-1952.

[59] S. Su, J.T. Minges, G. Grossman, A.J. Blackwelder, J.L. Mohler, E.M. Wilson, Proto-oncogene activity of melanoma antigen-A11 (MAGE-A11) regulates retinoblastoma-related p107 and E2F1 proteins, J Biol Chem, 288 (2013) 24809-24824.

[60] C. Lurquin, C. De Smet, F. Brasseur, F. Muscatelli, V. Martelange, E. De Plaen, R. Brasseur, A.P. Monaco, T. Boon, Two members of the human MAGEB gene family located in Xp21.3 are expressed in tumors of various histological origins, Genomics, 46 (1997) 397-408.

[61] K.M. Pattani, E. Soudry, C.A. Glazer, M.F. Ochs, H. Wang, J. Schussel, W. Sun, P. Hennessey, W. Mydlarz, M. Loyo, S. Demokan, I.M. Smith, J.A. Califano, MAGEB2 is activated by promoter demethylation in head and neck squamous cell carcinoma, PLoS One, 7 (2012) e45534.

[62] L.Y. Peche, M.F. Ladelfa, M.F. Toledo, M. Mano, J.E. Laiseca, C. Schneider, M. Monte, Human MageB2 Protein Expression Enhances E2F Transcriptional Activity, Cell Proliferation, and Resistance to Ribotoxic Stress, J Biol Chem, 290 (2015) 29652-29662.

[63] S.W. Lowe, A.W. Lin, Apoptosis in cancer, Carcinogenesis, 21 (2000) 485-495.

[64] L. Marcar, N.J. Maclaine, T.R. Hupp, D.W. Meek, Mage-A cancer/testis antigens inhibit p53 function by blocking its interaction with chromatin, Cancer Res, 70 (2010) 10362-10370.

[65] M. Monte, M. Simonatto, L.Y. Peche, D.R. Bublik, S. Gobessi, M.A. Pierotti, M. Rodolfo, C. Schneider, MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents, Proc Natl Acad Sci U S A, 103 (2006) 11160-11165.

[66] L.Y. Peche, M. Scolz, M.F. Ladelfa, M. Monte, C. Schneider, MageA2 restrains cellular senescence by targeting the function of PMLIV/p53 axis at the PML-NBs, Cell Death Differ, 19 (2012) 926-936.

[67] J.M. Doyle, J. Gao, J. Wang, M. Yang, P.R. Potts, MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases, Mol Cell, 39 (2010) 963-974.

[68] S. Hou, L. Xian, P. Shi, C. Li, Z. Lin, X. Gao, The Magea gene cluster regulates male germ cell apoptosis without affecting the fertility in mice, Scientific reports, 6 (2016) 26735.

[69] B. Yang, L. Wang, X. Luo, L. Chen, Z. Yang, L. Liu, SPAG6 silencing inhibits the growth of the malignant myeloid cell lines SKM-1 and K562 via activating p53 and caspase activation-dependent apoptosis, Int J Oncol, 46 (2015) 649-656.

[70] C.J. Dougherty, T.E. Ichim, L. Liu, G. Reznik, W.P. Min, A. Ghochikyan, M.G. Agadjanyan, B.N. Reznik, Selective apoptosis of breast cancer cells by siRNA targeting of BORIS, Biochem Biophys Res Commun, 370 (2008) 109-112.

[71] Z.M. Cilensek, F. Yehiely, R.K. Kular, L.P. Deiss, A member of the GAGE family of tumor antigens is an anti-apoptotic gene that confers resistance to Fas/CD95/APO-1, Interferon-gamma, taxol and gamma-irradiation, Cancer Biol Ther, 1 (2002) 380-387.

[72] J.W. Shay, W.E. Wright, Role of telomeres and telomerase in cancer, Seminars in cancer biology, 21 (2011) 349-353.
[73] I. Guilleret, J. Benhattar, Demethylation of the human telomerase catalytic subunit (hTERT) gene promoter reduced hTERT expression and telomerase activity and shortened telomeres, Exp Cell Res, 289 (2003) 326-334.
[74] S. Renaud, D. Loukinov, F.T. Bosman, V. Lobanenkov, J. Benhattar, CTCF binds the proximal exonic region of hTERT and inhibits its transcription, Nucleic Acids Res, 33 (2005) 6850-6860.
[75] E.M. Klenova, H.C. Morse, 3rd, R. Ohlsson, V.V. Lobanenkov, The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer, Seminars in cancer biology, 12 (2002) 399-414.
[76] S. Renaud, D. Loukinov, L. Alberti, A. Vostrov, Y.W. Kwon, F.T. Bosman, V. Lobanenkov, J. Benhattar, BORIS/CTCFL-mediated transcriptional regulation of the hTERT telomerase gene in testicular and ovarian tumor cells, Nucleic Acids Res, 39 (2011) 862-873.
[77] D. Hanahan, J. Folkman, Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis, Cell, 86 (1996) 353-364.
[78] B. Cho, Y. Lim, D.Y. Lee, S.Y. Park, H. Lee, W.H. Kim, H. Yang, Y.J. Bang, D.I. Jeoung, Identification and characterization of a novel cancer/testis antigen gene CAGE, Biochem Biophys Res Commun, 292 (2002) 715-726.
[79] Y. Kim, D. Park, H. Kim, M. Choi, H. Lee, Y.S. Lee, J. Choe, Y.M. Kim, D. Jeoung, miR-200b and cancer/testis antigen CAGE form a feedback loop to regulate the invasion and tumorigenic and angiogenic responses of a cancer cell line to microtubule-targeting drugs, J Biol Chem, 288 (2013) 36502-36518.
[80] S. Stefansson, G.A. McMahon, E. Petitclerc, D.A. Lawrence, Plasminogen activator inhibitor-1 in tumor growth, angiogenesis and vascular remodeling, Curr Pharm Des, 9 (2003) 1545-1564.
[81] A. Loriot, A. Van Tongelen, J. Blanco, S. Klaessens, J. Cannuyer, N. Van Baren, A. Decottignies, C. De Smet, A novel cancer-germline transcript carrying pro-metastatic miR-105 and TET-targeting miR-767 induced by DNA hypomethylation in tumors, Epigenetics, 9 (2014) 1163-1171.
[82] W. Zhou, M.Y. Fong, Y. Min, G. Somlo, L. Liu, M.R. Palomares, Y. Yu, A. Chow, S.T. O'Connor, A.R. Chin, Y. Yen, Y. Wang, E.G. Marcusson, P. Chu, J. Wu, X. Wu, A.X. Li, Z. Li, H. Gao, X. Ren, M.P. Boldin, P.C. Lin, S.E. Wang, Cancer-Secreted miR-105 Destroys Vascular Endothelial Barriers to Promote Metastasis, Cancer Cell, 25 (2014) 501-515.
[83] D.G. Hardie, F.A. Ross, S.A. Hawley, AMPK: a nutrient and energy sensor that maintains energy homeostasis, Nat Rev Mol Cell Biol, 13 (2012) 251-262.
[84] J. Kim, M. Kundu, B. Viollet, K.L. Guan, AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1, Nat Cell Biol, 13 (2011) 132-141.
[85] E. Conde, A. Suarez-Gauthier, E. Garcia-Garcia, F. Lopez-Rios, A. Lopez-Encuentra, R. Garcia-Lujan, M. Morente, L. Sanchez-Verde, M. Sanchez-Cespedes, Specific pattern of LKB1 and phospho-acetyl-CoA carboxylase protein immunostaining in human normal tissues and lung carcinomas, Hum Pathol, 38 (2007) 1351-1360.
[86] B. Zheng, J.H. Jeong, J.M. Asara, Y.Y. Yuan, S.R. Granter, L. Chin, L.C. Cantley, Oncogenic B-RAF negatively regulates the tumor suppressor LKB1 to promote melanoma cell proliferation, Mol Cell, 33 (2009) 237-247.
[87] C.T. Pineda, S. Ramanathan, K. Fon Tacer, J.L. Weon, M.B. Potts, Y.H. Ou, M.A. White, P.R. Potts, Degradation of AMPK by a cancer-specific ubiquitin ligase, Cell, 160 (2015) 715-728.
[88] C.T. Pineda, P.R. Potts, Oncogenic MAGEA-TRIM28 ubiquitin ligase downregulates autophagy by ubiquitinating and degrading AMPK in cancer, Autophagy, 11 (2015) 844-846.
[89] O.M. Sieber, K. Heinimann, I.P. Tomlinson, Genomic instability--the engine of tumorigenesis?, Nat Rev Cancer, 3 (2003) 701-708.
[90] M.E. Moynahan, J.W. Chiu, B.H. Koller, M. Jasin, Brca1 controls homology-directed DNA repair, Mol Cell, 4 (1999) 511-518.
[91] A. Tutt, D. Bertwistle, J. Valentine, A. Gabriel, S. Swift, G. Ross, C. Griffin, J. Thacker, A. Ashworth, Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences, EMBO J, 20 (2001) 4704-4716.
[92] W.D. Foulkes, I.M. Stefansson, P.O. Chappuis, L.R. Begin, J.R. Goffin, N. Wong, M. Trudel, L.A. Akslen, Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer, J Natl Cancer Inst, 95 (2003) 1482-1485.
[93] J. Watkins, D. Weekes, V. Shah, P. Gazinska, S. Joshi, B. Sidhu, C. Gillett, S. Pinder, F. Vanoli, M. Jasin, M. Mayrhofer, A. Isaksson, M.C. Cheang, H. Mirza, J. Frankum, C.J. Lord, A. Ashworth, S. Vinayak, J.M. Ford, M.L. Telli, A. Grigoriadis, A.N. Tutt, Genomic Complexity Profiling Reveals That HORMAD1 Overexpression Contributes to Homologous Recombination Deficiency in

Triple-Negative Breast Cancers, Cancer discovery, 5 (2015) 488-505.
[94] K. Daniel, J. Lange, K. Hached, J. Fu, K. Anastassiadis, I. Roig, H.J. Cooke, A.F. Stewart, K. Wassmann, M. Jasin, S. Keeney, A. Toth, Meiotic homologue alignment and its quality surveillance are controlled by mouse HORMAD1, Nat Cell Biol, 13 (2011) 599-610.
[95] B.A. Weaver, A.D. Silk, C. Montagna, P. Verdier-Pinard, D.W. Cleveland, Aneuploidy acts both oncogenically and as a tumor suppressor, Cancer Cell, 11 (2007) 25-36.
[96] M. Kwon, S.A. Godinho, N.S. Chandhok, N.J. Ganem, A. Azioune, M. Thery, D. Pellman, Mechanisms to suppress multipolar divisions in cancer cells with extra centrosomes, Genes Dev, 22 (2008) 2189-2203.
[97] K.M. Cappell, R. Sinnott, P. Taus, K. Maxfield, M. Scarbrough, A.W. Whitehurst, Multiple cancer testis antigens function to support tumor cell mitotic fidelity, Mol Cell Biol, 32 (2012) 4131-4140.
[98] M. Szyf, P. Pakneshan, S.A. Rabbani, DNA methylation and breast cancer, Biochem Pharmacol, 68 (2004) 1187-1197.
[99] H.T. Khong, Q.J. Wang, S.A. Rosenberg, Identification of multiple antigens recognized by tumor-infiltrating lymphocytes from a single patient: tumor escape by antigen loss and loss of MHC expression, J Immunother, 27 (2004) 184-190.
[100] H.J. Cho, O.L. Caballero, S. Gnjatic, V.C. Andrade, G.W. Colleoni, A.L. Vettore, H.H. Outtz, S. Fortunato, N. Altorki, C.A. Ferrera, R. Chua, A.A. Jungbluth, Y.T. Chen, L.J. Old, A.J. Simpson, Physical interaction of two cancer-testis antigens, MAGE-C1 (CT7) and NY-ESO-1 (CT6), Cancer Immun, 6 (2006) 12.
[101] Y. Kang, J.A. Hong, G.A. Chen, D.M. Nguyen, D.S. Schrump, Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells, Oncogene, 26 (2007) 4394-4403.
[102] M. Scudellari, A ballsy search for cancer targets, Nat Med, 17 (2011) 916-918.

OBJECTIVES

My thesis work is part of the overall objective to better understand the role of DNA hypomethylation in tumor development, in particular via the transcripitonal activation of genes with oncogenic potential. Up to now, only the transcription of genes coding for proteins, and more particularly the Cancer-Germline (CG) genes, have been described as a consequence of DNA hypomethylation in tumors. The aim of my thesis was to investigate whether DNA hypomethylation drives the aberrant activation of oncogenic non-coding transcripts such as CG miRNAs. There is indeed mounting evidence that miRNAs, which exert important regulatory functions through their ability to induce post-transcriptional inhibition of target miRNAs play important roles in cancer development.

1. Do CG-type transcripts produce miRNAs?

As first part of my thesis, the objectives were to identify new CG-type transcripts carrying miRNAs. In order to do so, we have chosen to base our analysis on two known characteristics of the CG genes: testis-specific expression, and preferential location on the X chromosome. By this way we identified the transcript *CT-GABRA3* as a novel CG transcript which produces the microRNAs miR-105 and miR-767.

2. Do CG miRNAs exert oncogenic functions?

Following identification of two CG-type microRNAs, our second goal was to identify their oncogenic potential. miR-105 has been described by Zhou et al. (Zhou, 2014) as a miRNA with metastatic potential. In our lab, we investigated the role of miR-767 and more specifically its role on its predicted targets: TET1 and TET3 genes, two members of the *t*en-*e*leven-*t*ranslocation family of tumor suppressor genes.

3. What epigenetic mechanisms are involved in the regulation of the transcripts comprised within *GABRA3* locus?

Our experiments revealed a striking inverse correlation between the methylation status of the promoter of the two transcripts, *CT-GABRA3* and *BT-GABRA3*, located in the *GABRA3* locus. Our third objective was to investigate the

mechanisms underlying this process of interdependent epigenetic modifications in cancer cells.

CHAPTER I

A novel cancer-germline transcript carrying pro-metastatic miR-105 and TET-targeting miR-767 induced by DNA hypomethylation in tumors

Axelle Loriot,[1,†] Aurélie Van Tongelen,[1,†] Jordi Blanco,[1,‡] Simon Klaessens,[1] Julie Cannuyer,[1] Nicolas van Baren,[2] Anabelle Decottignies,[1] and Charles De Smet[1,*]

[1]Group of Genetics and Epigenetics; de Duve Institute; Université Catholique de Louvain; Brussels, Belgium

[2]Ludwig Institute for Cancer Research Ltd; Centre du Cancer des Cliniques; Universitaires Saint-Luc; Brussels, Belgium

[†]These authors contributed equally to this work.

[‡]Current affiliation: Physiology Unit; School of Medicine; Universitat Rovira i Virgili; Reus, Spain

➔ *Epigenetics*. 2014;9(8):1163-1171. doi:10.4161/epi.29628.

Abstract

Genome hypomethylation is a common epigenetic alteration in human tumors, where it often leads to aberrant activation of a group of germline-specific genes, commonly referred to as "cancer-germline" genes. The cellular functions and tumor promoting potential of these genes remain, however, largely uncertain. Here, we report identification of a novel cancer-germline transcript (*CT-GABRA3*) displaying DNA hypomethylation-dependent activation in various tumors, including melanoma and lung carcinoma. Importantly, *CT-GABRA3* harbors a microRNA (miR-105), which has recently been identified as a promoter of cancer metastasis by its ability to weaken vascular endothelial barriers following exosomal secretion. *CT-GABRA3* also carries a microRNA (miR-767) with predicted target sites in *TET1* and *TET3*, two members of the *t*en-*e*leven-*t*ranslocation family of tumor suppressor genes, which are involved in the conversion of 5-methylcytosines to 5-hydroxymethylcytosines (5hmC) in DNA. Decreased TET activity is a hallmark of cancer; here, we provide evidence that aberrant activation of miR-767 contributes to this phenomenon. We demonstrate that miR-767 represses *TET1/3* mRNA and protein expression and regulates genomic 5hmC levels. Additionally, we show that high *CT-GABRA3* transcription correlates with reduced *TET1* mRNA levels *in vivo* in lung tumors. Together, our study identified a cancer-germline gene that produces microRNAs with oncogenic potential. Moreover, our data indicate that DNA hypomethylation in tumors can contribute to reduced 5hmC levels via activation of a *TET*-targeting microRNA.

1 Introduction

DNA methylation in mammalian genomes, which occurs mostly at cytosines within CpG dinucleotides, is a potent mechanism of gene repression, and contributes thereby to the establishment and maintenance of cell-type specific gene expression programs.[1] Genome methylation patterns often undergo profound alterations in human tumors.[2] Both gains (hypermethylation) and losses (hypomethylation) of CpG methylation are observed. DNA hypermethylation often affects CpG-rich promoters of tumor suppressor genes, leading to their irreversible silencing. DNA hypomethylation, on the other hand, has been associated with aberrant activation of a limited group of protein-coding genes, most of which have their transcription normally restricted to the germ line.[3,4] Genes in this group, commonly referred to as 'cancer-germline' (CG) genes, are indeed characterized by their strict reliance on DNA methylation for repression in normal somatic tissues.[5,6] Intriguingly, most CG genes map on the X chromosome.[4] It is still unclear, however, if activation of CG genes, which appear to exert a variety of cellular functions, plays a major oncogenic role in hypomethylated tumor cells.[3]

In the present study, we searched to determine if DNA hypomethylation in tumors also induces aberrant expression of miRNA-producing cancer-germline transcripts. There is indeed mounting evidence that dysregulated expression of miRNAs, which exert important regulatory functions through their ability to induce post-transcriptional inhibition of target mRNAs, contributes to cancer development.[7,8]

2 Results

2.1 Aberrant activation of GABRA3 and hosted miR-105 and miR-767 in tumors

As an initial step in our search for CG-type miRNAs, we performed an *in silico* screening in miRNA databases (microRNA.org and miRBase.org), using as filtering criteria two characteristics of CG genes: predominant expression in testis and localization on the X chromosome. This led to the selection of 21 X-linked miRNAs with predicted expression in testis and in no more than one normal somatic tissue. Among these, we noticed a pair of miRNAs (miR-105 and miR-767), deriving from the first intron of *GABRA3*, a gene encoding a receptor subunit for the γ-aminobutyric-acid neurotransmitter. Our interest for these miRNAs was prompted by the fact that, although *GABRA3* expression is normally restricted to brain and testis, aberrant transcription of the gene was reported in several tumor types, and was identified as a significant predictor of poor survival in lung cancer patients.[9-12] Moreover, *GABRA3* is located within a region of the X chromosome (Xq28) that harbors many known CG genes. RT-qPCR experiments with primers located in exons 5 and 6 of *GABRA3* confirmed specific expression of this gene in brain and testis, and revealed its activation in melanoma cell lines and tissues (Fig. 1A,B). In parallel, RT-qPCR directed toward miR-105 and miR-767 indicated that expression of these miRNAs strictly mirrors that of their host gene (Fig. 1B). Additional analyses in larger sets of tumor samples detected *GABRA3* transcripts in 65% of melanoma tissues and in 40% of lung tumors (Fig. 1C).

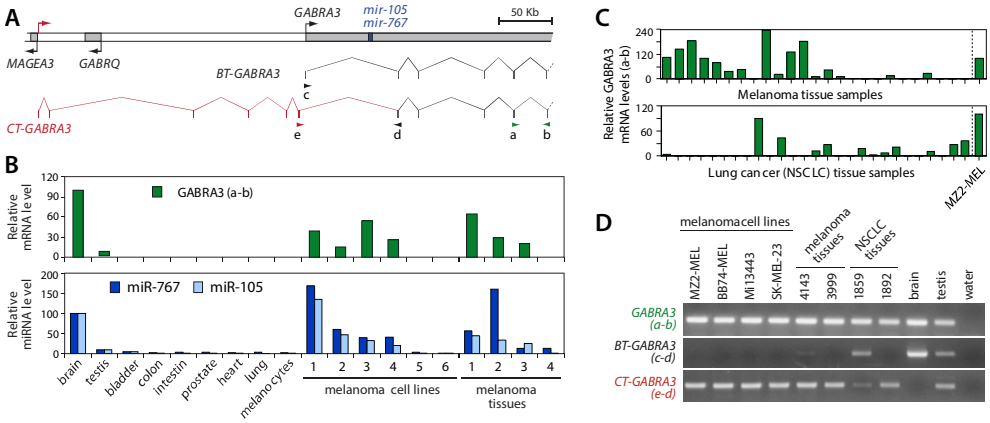


Figure 1. Tumors show aberrant expression of a testis-specific transcript variant of *GABRA3*, and of the miRNAs it harbors. (A) Schematic representation of the human *GABRA3* locus, with broken arrows indicating transcription start sites. The exon/intron structure of the referenced *GABRA3* transcript (re-named *BT-GABRA3*) and of the newly characterized transcript variant (*CT-GABRA3*) is shown below, with arrowheads indicating the orientation and location of PCR primers used in subsequent expression analyses. (B) Expression analysis of *GABRA3*, miR-105 and miR-767 in normal human tissues and melanocytes, as well as in melanoma cell lines (1, MI13443-MEL; 2, SK-MEL-23; 3, BB74-MEL; 4, MZ2-MEL; 5, LB2667-MEL; 6, Mi665/2-MEL) and tissue samples (Table S1). Primers *a* and *b* amplify both *BT-GABRA3* and *CT-GABRA3* transcripts. Normalized mRNA (ratio to *ACTB*) and miRNA (ratio to *SNORD44*) levels are expressed relative to the brain sample taken as 100% reference. (C) RT-qPCR analysis of *GABRA3* with primers *a* and *b*, in a larger series of melanoma samples (n = 25), and in non-small-cell lung carcinoma (NSCLC) tissue samples (n = 27). See table S1 for sample descriptions. *GABRA3* mRNA levels are expressed relative to the MZ2-MEL melanoma cell line taken as 100% reference. (D) Gel analysis of RT-PCR experiments with primers recognizing either both *GABRA3* transcript variants (primers *a* and *b*), only *BT-GABRA3* transcripts (primers *c* and *d*), or only *CT-GABRA3* transcripts (primers *e* and *d*).

2.2 Tumors express a cancer-testis variant of GABRA3*:* CT-GABRA3

To further characterize the composition of *GABRA3* transcripts in tumor cells, RT-PCR experiments with primers located in different exons were performed. Surprisingly, RT-PCR with primers located in exon 1 and 2 of *GABRA3* amplified the transcript in brain and testis, but failed to detect it in most tumor cells (Fig. 1A,D). This suggested the existence of an alternative form of *GABRA3* transcript in tumors. In order to identify this transcript variant, we performed 5′ RACE experiments in *GABRA3*-expressing melanoma cell lines. This led to the identification of an alternative transcription start site located 247-kb upstream of the reference *GABRA3* start site. We isolated several novel transcript variants originating from this start site, which contained alternatively spliced exons in the 5′ part followed by all exons but exon 1 of *GABRA3* (Fig. 1A, and Fig. S1). *GABRA3* transcripts originating from this alternative start site were named *CT-GABRA3* (*C*ancer-*T*estis), as opposed to the reference *GABRA3* transcript, which, for sake of clarity, we re-named *BT-GABRA3* (*B*rain-*T*estis). Unlike *BT-GABRA3*, *CT-GABRA3* displayed a typical cancer-germline pattern of expression, as it was expressed in testis but not in brain, and was commonly activated in tumor cells (Fig. 1D). *CT-GABRA3* transcripts comprise several short upstream open reading frames, which were found to inhibit translation of the GABRA3 protein (Fig. S2). Interestingly, the transcription start site of *CT-GABRA3* is located nearby that of a known CG gene, *MAGEA3*, oriented in the opposite direction (Fig. 1A). Both genes appear therefore to share a bidirectional promoter, as supported by their frequent co-activation in melanoma and lung tumor samples (Fig. S3).

2.3 **CT-GABRA3** ***activation in tumors is dependent on DNA demethylation***

We next investigated whether activation of *CT-GABRA3* and its hosted miRNAs in tumors is linked to DNA hypomethylation. Sensitivity of *CT-GABRA3* expression to DNA demethylation-dependent activation was demonstrated in an experiment showing induction of this transcript, but not of *BT-GABRA3*, following treatment of non-expressing cells with the DNA methylation inhibitor, 5-aza-2′-deoxycytidine (5-azadC; Figure 2A). Not surprisingly, the DNA methylation inhibitor also induced expression of miR-105 and miR-767 (Fig. 2B). Moreover, sodium bisulfite sequencing revealed that *CT-GABRA3* expression in testis and tumor cells is associated with extensive promoter demethylation (Fig. 2C). Consistent with a primary role of genome demethylation in the activation of *CT-GABRA3* in tumor cells, we observed a significant trend of co-activation of this gene with other DNA methylation-sensitive CG genes in melanoma cell cultures (Fig. S4). Together, our results indicate that miR-105 and miR-767 are carried by two transcript variants of *GABRA3*: *BT-GABRA3*, which is transcribed in brain and testis; and *CT-GABRA3*, which normally displays specific transcription in testis, but also becomes activated as a result of DNA demethylation in tumor cells. We conclude that DNA hypomethylation in tumors is associated with aberrant activation of miR-105 and miR-767 expression.

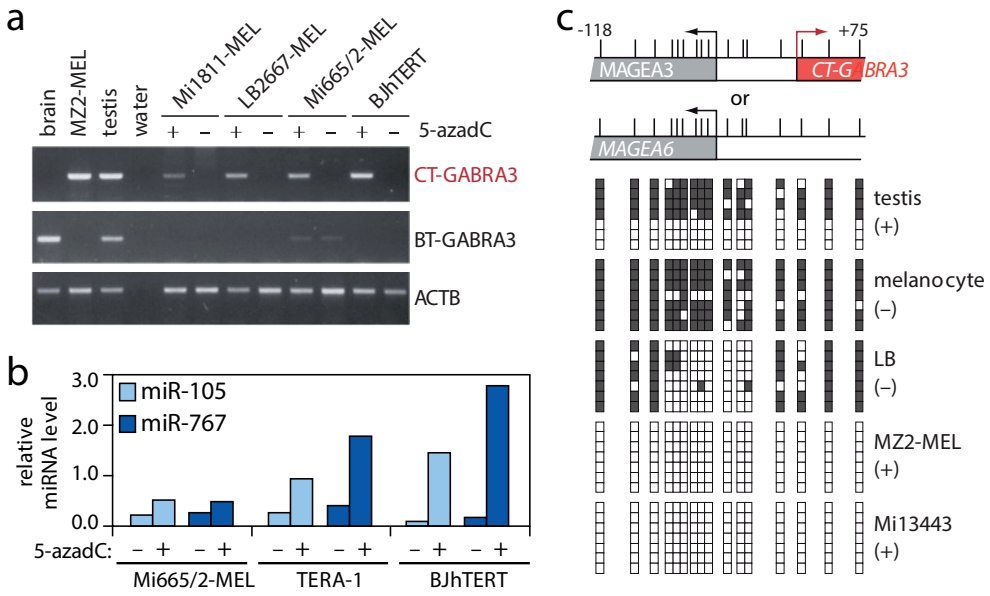


Figure 2. Expression of *CT-GABRA3*, miR-105 and miR-767 is induced by DNA demethylation. (A) Three *GABRA3*-negative melanoma cell lines (-MEL), and an immortalized fibroblast cell line (BJhTERT) were cultured in the presence (+) or in the absence (-) of 5-azadC. Expression of *CT-GABRA3, BT-GABRA3* and *ACTB* (control) was analyzed by RT-PCR. (B) Expression of miR-105 and miR-767 was analyzed by RT-qPCR in similarly treated cell lines, including the TERA-1 embryonal carcinoma cell line. Relative miRNA levels are expressed as ratio to *SNORD44* (x 10^4). (C) Bisulfite sequencing of the *MAGEA3/CT-GABRA3* promoter region. Sequences could not be distinguished from those deriving from the *MAGEA6* promoter region, as both loci show 100% sequence identity. Vertical bars indicate location of CpG sites with positions relative to the *CT-GABRA3* start site. Open and filled squares represent unmethylated and methylated CpG sites, respectively, and each row represents a single clone. *CT-GABRA3* expression status (+) or (-) in samples is indicated (positive samples also express *MAGEA3* and *MAGEA6*). Highly methylated sequences in testis likely derive from somatic cells in the tissue sample.

2.4 TET1 *and* TET3 *are targets of miR-767*

During the course of our research project, a study was published showing that miR-105 is expressed in metastatic breast cancer cells, and acts as a crucial promoter of cancer metastasis.[13] The study revealed indeed that miR-105 undergoes exosome-mediated secretion, and destroys vascular barriers by inhibiting expression of the tight junction protein ZO-1 in endothelial cells. Our data indicate that miR-105 is expressed in other tumor types as well, and reveal that DNA hypomethylation accounts for its tumor-specific activation. In order to get more insight into the function of miR-767 and its potential contribution to tumor development, we decided to search relevant mRNA targets of this miRNA

with the help of prediction algorithms. We focused our attention on two potential targets of miR-767, *TET1* and *TET3*, because *TET* genes were recently shown to exert tumor-suppressive functions.[14] The *TET* family of genes (*TET1*, *TET2*, and *TET3*) encode dioxygenases that are recruited to specific regions of the genome, where they contribute to processes of localized DNA demethylation by converting 5-methylcytosines (5mC) to 5-hydroxymethylcytosines (5hmC).[15,16] TET activities are often downregulated in a wide variety of cancers, leading to marked reduction in genomic 5hmC levels.[17-19] In a significant proportion of hematopoietic malignancies, this has been associated with mutations in *TET2*.[17] In most solid tumors, however, the origin of reduced TET activities has remained unclear.

Effective targeting of *TET1* and *TET3* genes by miR-767 was confirmed by transfection experiments showing that synthetic miR-767 molecules, but not control miRNA molecules, induce downregulation of luciferase reporter genes linked to the 3′-UTR of either *TET1* or *TET3* (Fig. 3A,B). Importantly, we also demonstrated that expressing tumor cells contain sufficient amounts of endogenous miR-767 to inhibit expression of the *TET1* 3′-UTR luciferase reporter gene, and we confirmed that this inhibition involves a direct interaction, by showing impaired inhibition of a reporter that carries a mutant version of the *TET1* 3′-UTR lacking miR-767 target sequences (Fig. 3C, and Fig. S5). We then assessed whether miR-767 can regulate endogenous *TET1* and *TET3* mRNA levels. In HEK293T human embryonic kidney cells and in TERA-1 human embryonal carcinoma cells, which both lack constitutive expression of miR-767, transfection of synthetic miR-767 molecules resulted in reduced *TET1* and *TET3* mRNA levels, although this reduction was not significant for *TET3* in TERA-1 cells (Fig. 3D). Conversely, inhibition of miR-767 by antisense oligonucleotides in expressing tumor cell lines resulted in a significant elevation in *TET1* mRNA levels (Fig. 3E). For *TET3*, we observed a less constant effect of the inhibitor, as only one out of the four treated cell lines showed an increased level of *TET3* mRNA (Fig. 3E). We also confirmed the ability of miR-767 to inhibit TET1 and TET3 expression at the protein level, by showing decreased amounts of these proteins in HEK293T cells upon transfection with synthetic miR-767 molecules (Fig. 3F). Finally, we assessed whether miR-767 can regulate 5hmC levels. Slot

blot assays with anti-5hmC antibodies showed that the transfection of synthetic miR-767 molecules in both TERA-1 and HEK293T cells indeed resulted in a significant reduction of global 5hmC levels (Fig. 3G). Taken together, these data indicate that miR-767 can function as a regulator of cellular 5hmC levels via targeting of *TET* genes. In our experiments, miR-767 exhibited a preferential effect on *TET1*.

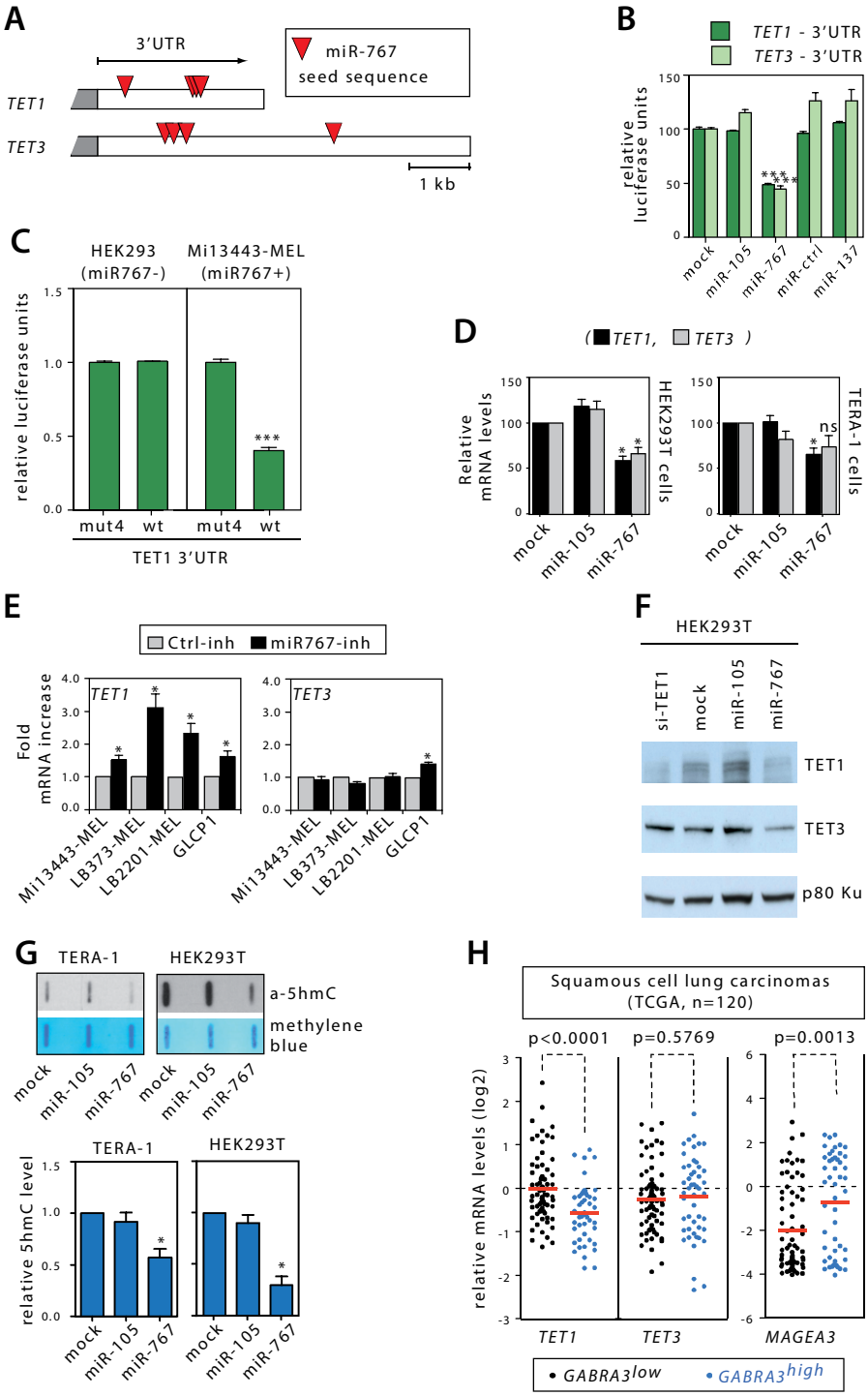

A
3'UTR
miR-767
seed sequence
TET1
TET3
1 kb
B
TET1 - 3'UTR
TET3 - 3'UTR
relative luciferase units
mock
miR-105
miR-767
miR-ctrl
miR-137
C
HEK293 (miR767-)
Mi13443-MEL (miR767+)
relative luciferase units
mut4
wt
TET1 3'UTR
D
(TET1, TET3)
Relative mRNA levels
HEK293T cells
TERA-1 cells
ns
E
Ctrl-inh
miR767-inh
Fold mRNA increase
TET1
TET3
Mi13443-MEL
LB373-MEL
LB2201-MEL
GLCP1
F
HEK293T
si-TET1
mock
miR-105
miR-767
TET1
TET3
p80 Ku
G
TERA-1
HEK293T
a-5hmC
methylene blue
relative 5hmC level
H
Squamous cell lung carcinomas (TCGA, n=120)
p<0.0001
p=0.5769
p=0.0013
relative mRNA levels (log2)
TET1
TET3
MAGEA3
GABRA3low
GABRA3high

Figure 3. miR-767 controls *TET1* and *TET3* expression levels and regulates cellular 5hmC levels. (A) Red triangles indicate the location of miR-767 seed sequences (7-mer and 8-mer) in *TET1* and *TET3* 3′-UTRs. (B) HEK293T cells were co-transfected with a luciferase reporter linked to the 3′-UTR of either *TET1* or *TET3*, and with the indicated miRNAs (miR-Ctrl and miR-137 served as a negative control). Normalized luciferase activities, which were measured 24h after transfection, are expressed relative to mock cells (no miRNA transfected), and represent means ± SEM (n = 3). *** $P < 0.001$ for comparison to mock cells (Bonferroni's multiple comparison test). (C) HEK293 cells (miR-767 negative) and Mi13443-MEL cells (miR-767 positive) were stably transfected with a luciferase reporter vector carrying either the wild-type (wt) 3′-UTR of *TET1* or a mutant version of the 3′-UTR (mut4), lacking the four miR-767 target sequences. Relative luciferase units represent mean ± SEM (n = 3). *** $P < 0.001$. (D)*TET1* and *TET3* mRNA levels were evaluated by RT-qPCR three days after transfection of miR-105 (serving as negative control) or miR-767. Normalized mRNA levels (ratio to *ACTB*) are expressed relative to mock cells (no miRNA transfected). Values represent mean ± SEM ($n \geq 3$). * $P < 0.05$, for comparison to mock cells (Friedman test with Dunn's multiple comparison test). (E) miR-767-expressing melanoma cell lines (-MEL) and non-small cell lung carcinoma cell line (GLCP1) were transfected with a LNA inhibitor of miR-767 (miR767-inh) or an irrelevant control LNA inhibitor (Ctrl-inh). *TET1* and *TET3* mRNA expression levels were assessed by RT-qPCR. Values represent mean ± SEM (n = 3). * $P < 0.05$ (Wilcoxon signed rank tests). (F) HEK293T cells transfected with the indicated miRNA or siRNA were subjected to western blot analysis for TET1 and TET3 proteins (p80-Ku was assessed to verify equal loading). (G) Genomic 5hmC levels were evaluated by slot blot analysis at day 3 after transfection with the indicated miRNA. Blots were stained with methylene blue to control for loading. Representative blots are shown out of three to five repeats. Quantification of slot blot data was performed. Normalized 5hmC levels are expressed relative to mock-transfected cells, and represent mean ± SEM * $P < 0.05$ (Friedman test with Dunn's multiple comparison test). (H) Analysis of microarray data derived from the TCGA collection of lung squamous cell carcinomas (n = 120) revealed significant downregulation of *TET1* (but not *TET3*) in tumor cells that show upregulation of miR-767-harboring *GABRA3* transcripts. Positive correlation between *GABRA3* and *MAGEA3* activation confirmed the validity of the test. Red bars indicate mean mRNA levels. *P* values were determined by unpaired *t* tests.

We next searched to determine if miR-767 expression is indeed correlated with reduced expression of *TET* genes *in vivo* in tumor tissues. To this end, we analyzed publicly available microarray data, and searched if we could find a negative correlation between the expression of *TET1/3* genes and the presence of miR-767-harboring *GABRA3* transcripts (probes on microarrays detect both *CT*- and *BT-GABRA3* transcripts). miRNAs usually induce only limited decrease in the level of their target mRNAs. We nevertheless observed a significant correlation between increased *GABRA3* expression and reduced *TET1* mRNA levels in lung carcinoma tissues (Fig. 3H), a tumor type where *GABRA3* upregulation was identified as a predictor of poor survival. *TET3* did not seem to be affected, suggesting that *TET1* is a preferred target of miR-767 in lung carcinoma.

3 Discussion

One common target of DNA hypomethylation in tumors is the group of CG genes, which normally displays specific expression in germline cells. As a consequence, aberrant transcription of these genes is observed in a large variety of tumor types. Evidence that CG genes exert tumor-promoting roles has however remained scarce. Our identification of *CT-GABRA3* provides the first example, to the best of our knowledge, of a CG gene that harbors miRNAs with oncogenic potential. One of these miRNAs, miR-105, was shown very recently to be expressed in breast cancer cells, and to be released in the extracellular environment via exosome secretion.[13] Secreted miR-105 targets tight junction protein ZO-1 in endothelial cells, thereby facilitating cancer cell migration to distant locations. Our study establishes therefore a connection between DNA hypomethylation, CG gene activation, and cancer metastasis.

CT-GABRA3 also produces miR-767, and we demonstrate that this miRNA targets genes of the *TET* family. *TET* genes encode epigenetic regulators, which were found to exert tumor suppressive functions, notably through their impact on cell differentiation and invasion.[18,20,21] A recent study revealed the existence of an extensive network of *TET*-targeting miRNAs.[22] Among these, several miRNAs, including miR-29b and miR-22, were shown to contribute to the development of hematopoietic and breast cancers, in a manner that depended on their effect on *TET* gene expression.[22,23] This demonstrates that *TET*-targeting miRNAs can have a critical impact on tumor cell functions, even though their effects on *TET* expression levels are generally very subtle.[22,24] In this regard, it is worth noting that monoallelic mutation of *Tet2* was sufficient to promote myeloproliferative disorders in mice models, suggesting physiologically relevant gene dosage effects.[25,26] The specific contribution of miR-767 to the regulation of TET activities in tumor cells will likely vary according to the level of expression of other *TET*-targeting miRNAs. Compared with these miRNAs, miR-767 displays a much more contrasted pattern of expression, as it is completely silenced in all tissues, except brain and testis, and becomes activated in a wide variety of tumors. We anticipate therefore that miR-767 can exert a critical tumor-promoting function in several tumor types. Interestingly, a recent study, which

conducted a systematic screening for miRNAs involved in cell migration, identified miR-767 as a potential candidate.[27] This finding supports a possible pro-metastatic role of miR-767, when expressed in tumor cells.

Intriguingly, our results concerning miR-767 reveal the existence of an unexpected link between DNA methylation and *TET* genes, whereby genome demethylation can lead to inhibition of TET activities via activation of a *TET*-targeting miRNA. DNA-hypomethylation-mediated activation of miR-767 in tumors is therefore expected to cause subsequent epigenetic remodeling events. For instance, because TET proteins are required to maintain select genomic sequences in a DNA methylation-free status,[28] their downregulation by miR-767 may facilitate DNA hypermethylation at specific gene promoters. This raises the interesting possibility of a link between DNA hypomethylation and DNA hypermethylation in tumor cells.

In a more physiological context, we hypothesize that miR-767 may contribute to an epigenetic regulatory circuit in developing germ line cells, where transient processes of DNA demethylation involving TET enzymes are taking place.[29,30] One possibility is that miR-767 becomes activated upon completion of these DNA demethylation processes, and then inhibits TET activities in order to allow subsequent re-methylation of specific DNA sites (Fig. 4). The timeframe of activation of CG genes in developing germline cells is compatible with this proposed role of miR-767.[31-33] Experimental validation of this hypothesis can be envisaged in the mouse, where miR-767 location in the X-linked *Gabra3* gene, and target sites in the 3′-UTR of *Tet* genes, are conserved.

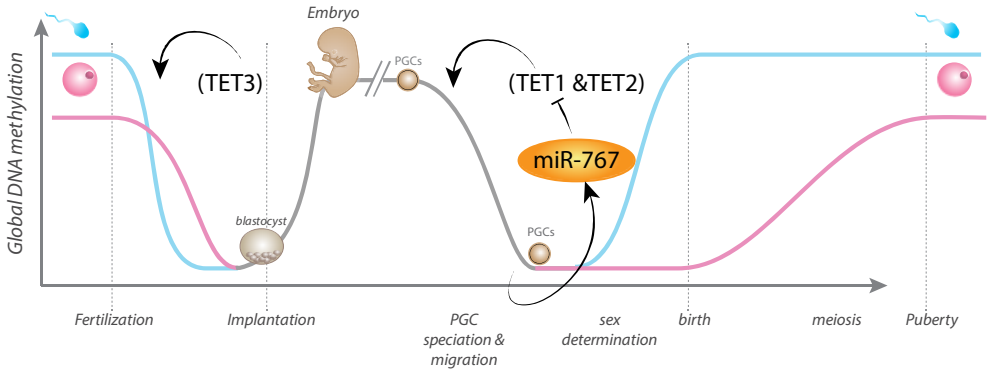


Figure 4. Illustration of a possible physiological mechanism where miR-767 becomes activated upon completion of DNA demethylation processes, and then inhibits TET activities in order to allow subsequent re-methylation of specific DNA sites.

The other site of constitutive miR-767 expression is brain, where 5hmC levels are generally high.[34] Interestingly, TET1 was found to be significantly downregulated by neuronal activity,[35] suggesting the existence of intricate regulatory mechanisms, in which miR-767 might be implicated. This may have important implications in memory formation and extinction.[35,36]

4 Material and methods

Cell lines and tumor tissue samples

All human melanoma cell lines, which derive from cutaneous melanoma metastases, and the GLCP1 cell line, which derives from a human NSCLC, were obtained from the Brussels Branch of the Ludwig Institute for Cancer Research, and were cultured as previously described.[5] TERA-1 human embryonal carcinoma cells were kindly provided by W. Schultz (Heinrich Heine University, Germany), and their culture conditions are described elsewhere.[37] HEK293T cells, which were purchased from Thermo Fisher, HEK293 cells, which were kindly provided by B. Lauwerys (Université catholique de Louvain, Belgium), and BJhTERT cells, which were a gift from F. d'Adda di Fagagna (IFOM foundation, Italy), were maintained in high glucose DMEM (Life Technologies), supplemented with GlutaMAX™ (Life Technologies), 1 x non-essential amino acids (Life Technologies), 1 x Penicillin/Streptavidin (Life Technologies), and 10% fetal bovine serum (Hyclone). Early passage human normal epidermal melanocytes (HNEM) were received from E. De Plaen (Ludwig Institute for Cancer Research, Belgium), and were cultured in Ham's F10 medium (Life Sciences) supplemented with 6 mM Hepes, 1 x MelanoMax supplement (Gentaur), and 10% fetal bovine serum. A description of human melanoma and NSCLC tissue samples is provided in the table S1. They were obtained from the Brussels Branch of the Ludwig Institute for Cancer Research

RT-PCR and RT-qPCR analyses

Total RNA samples were purchased from Ambion Life Technologies, or prepared from cell lines and surgical specimens (obtained from the Ludwig Institute for Cancer Research, Belgium) using either TriPure Isolation Reagent (Roche Diagnostics GmbH) or the guanidinium-isothiocyanate/cesium chloride procedure.[38] Reverse transcription was performed on 2 µg of total RNA using either PrimeScript Reverse transcriptase (Takara) and random hexamers primers, or M-MLV Reverse transcriptase (Invitrogen) and dT18 primers. The transcripts were amplified from 1/40 of the reverse transcription reaction. Conventional PCRs were performed using the DreamTaq polymerase (Thermo Fisher

Scientific), in a final reaction volume of 25 µl. Quantitative RT-PCR amplifications were performed using the qPCR Core kit (Eurogentec). All qPCRs where SybrGreen assays, except for the *GABRA3* qPCR, which was a Taqman assay. Primers and probe sequences are listed in the table S2.

For miRNA RT-qPCR analyses, 20ng of total RNA was used for the RT with the Universal cDNA Synthesis Kit II (Exiqon). miRNAs were amplified from 1/320 of the reverse transcription reaction, using LNA primers specific for hsa-mir-767–5p (#204238, Exiqon) and hsa-mir-105–5p (#204389, Exiqon). LNA primers specific for SNORD44 (#203902, Exiqon) were used for normalization. qPCR was performed using the qPCR Core Kit (Eurogentec).

Rapid amplification of 5' cDNA ends (5'-RACE)

Two protocols of 5'-RACE were applied, which were based on the use of either ThermoScript Reverse Transcriptase (Fig. S1B) or PrimeScript Reverse Transcriptase (Fig. S1C).

For Thermoscript 5'-RACE: reverse transcription was performed on 3 µg of total RNA using ThermoScript Reverse Transcriptase (Invitrogen), according to manufacturer's instructions. Ten pmoles of *GABRA3*-specific primer (GABRA3b) and 10 pmoles of SMART IV oligo from the SMART™ cDNA Library Construction Kit (Clontech) were used for the reaction. The reaction was incubated during 1 h at 58 °C, and stopped by heating 5 min at 80 °C. The reaction was subsequently incubated for 20 min at 37 °C in the presence of 1 unit of RNaseH (Invitrogen). The 5'-RACE products were amplified from 1/20 of the reverse transcription reaction, with 0.625 units of TaKaRa Taq DNA polymerase (Takara). Three rounds of nested PCR, with primers indicated in Figure S1 (see sequence in accompanying table), were applied. PCR products were run on an agarose gel, and purified with the QIAquick Gel Extraction Kit (Qiagen) before sequencing.

For PrimeScript 5'-RACE: Reverse transcription was performed on 1 µg of total RNA using PrimeScript Reverse Transcriptase (Takara), according to manufacturer's instructions. Ten pmoles of random hexamers and 10 pmoles of SMART IV oligo were used in the reaction. The reaction was incubated for 1 h at

42 °C, and stopped by heating 7 min at 72 °C. Subsequent steps were as described here above.

DNA fragments generated by 5′-RACE and RT-PCR analyses were sequenced with the BigDye® Terminotor V3.1 Cycle Sequencing kit (Life technologies), according to the manufacturer's recommendation. Sequencing reactions were purified with the BigDye® XTerminator kit (Life technologies) and were run on an ABI 3130xL Genetic Analyzer (Life technologies). Sequence of the full-length *CT-GABRA3* transcript has been submitted to GenBank under the accession number KJ620007.

5-Aza-2′-deoxycytidine treatment

Cells grown to 60–70% confluence were exposed to a single dose of 2 μM 5-aza-2′-deoxycytidine (Sigma-Aldrich), and maintained in culture during 6 d (a period of time corresponding to at least 3 population doublings) before RNA extraction.

Sodium bisulfite genomic sequencing

Sodium bisulfite genomic sequencing of the *MAGEA3/CT-GABRA3* promoter region was performed as described previously.[39] Primer used for nested PCR amplification of bisulfite treated DNA were respectively TYGATTTTTATTTAGGTAGAATTT and TAAAATAATAACRACCCAACCTAA (1st PCR) and ATTTAGGTAGAATTTAGTTTTAT and CCCTACRAAATAACCCAAA (2nd PCR). These primers sets also amplify the *MAGEA6* promoter region, which shows 100% sequence identity.

Construction of luciferase reporter vectors and luciferase assays

TET1 and *TET3* 3′-UTRs were amplified by PCR using the high fidelity PrimeStar HS DNA polymerase (Takara), with primers carrying a 5′ overhang containing a restriction site for either *Xho*I (sense primer) or *Not*1 (antisense primer) (see table S2). PCR fragments were cloned between the corresponding restriction sites into the psiCHECK™-2 vector (Promega). Vector inserts were sequenced to verify the presence of error-free miR-767 target sequences.

The luciferase reporter vector containing the mutant version of *TET1* 3′-UTR was generated by site directed mutagenesis of the wild-type vector, using the QuickChange multi site-directed mutagenesis kit (Stratagene), and according to the manufacturer's instructions. Mutagenic primers (TET1-mut1 to -mut4, see Table S1) were designed using the QuickChange primer design program (www.agilent.com/genomics/qcpd).

In co-transfection of miRNA and luciferase reporter experiments, $2.x10^4$ HEK293T cells were seeded in each well of a 96-well plate. After 24h, cells in each well were transfected using the Lipofectamine 2000 reagent (Invitrogen, Life Sciences), with 0.05 ng of the luciferase reporter vector and mirVana miRNA Mimics (Ambion, ThermoScientific) at a final concentration of 10 nM (Fig. 3B or 3 nM (Fig. S3B). Luciferase activities were measured 24h after transfection by using the Dual-Glo® Luciferase Assay System (Promega) and a Glomax® 96 Microplate luminometer (Promega).

In the experiments of stable transfection of luciferase reporters in HEK293 cells, cells were seeded at ~40% confluency in 75 cm^2 flasks, and were transfected 24h later, using Lipofectamine 2000, with 20 µg of the reporter vector, 10 µg of genomic DNA, and 1 µg of pCDNA3 (carrying a neomycin resistance gene; Invitrogen). Transfectants were selected in 0.8 mg/ml G418 Geneticin® (Invitrogen, Life Sciences), and were harvested after 19 d of selection for analysis of luciferase activities. For Mi13443-MEL cells, cells were seeded at ~50% confluency in 75 cm^2 flasks, and were transfected 24h later, using the Genius DNA transfection reagent (Westburg) according the manufacturer's instructions, with 10 µg of the reporter vector, and 1 µg of pCDNA3. Transfectants were selected in 0.8 mg/ml G418 Geneticin®, and were harvested at different time points (day 14, 17 and 22) for analysis of luciferase activities.

Transfection of synthetic miRNAs and miRNA inhibitors

For analysis of the effect of miRNA transfection on endogenous *TET* mRNA and protein levels, and on 5hmC levels, synthetic mirVana miRNA Mimics were transfected at a final concentration of 50 nM using the Lipofectamine 2000 reagent (Invitrogen, Life Sciences) in HEK293T cells; and at a final concentration

of 50 nM using the Lipofectin reagent (Invitrogen, Life Sciences) in TERA-1 cells. In both cases, the manufacturer's instructions were applied.

For analysis of miRNA inhibitors on *TET* expression levels in tumor cells, 5'-fluorescein-labeled miRCURY LNA™ Power microRNA inhibitors (Exiqon) were transfected at a final concentration of 100 nM, using the ExGen 500 transfection reagent (Thermo Scientific) for Mi13443-MEL cells, the Genius DNA transfection reagent for LB373-MEL cells, and the Lipofectamine 2000 transfection reagent for LB2201-MEL and GLCP1 cells. In all cases, the manufacturer's instructions were applied, and visual analysis under a fluorescent microscope revealed nearly 100% transfection efficiencies. Cells were harvested at day one after transfection for *TET1/3* gene expression analysis.

Western blotting

Whole cell lysates were obtained by harvesting cells in 1 x Laemmli buffer complemented with the cOmplete Mini protease inhibitor cocktail (Roche), PhosphoStop phosphatase inhibitor cocktail (Roche), and 1 mM of phenylmethanesulfonyl fluoride (Sigma-Aldrich). Whole cell lysates were denatured for 10 min at 99 °C, sonicated with a Bioruptor sonicator (Diagenode), and reheated for 5 min at 99 °C before loading and electrophoresis in a 4–15% acrylamide Mini-Protean® TGX gel (Biorad). Proteins were thereafter submitted to an overnight electrotransfer on a polyvinydilene difluoride Immobilon®-P transfer membrane (Millipore) at 4 °C. The membrane was thereafter saturated in a PBS solution containing 4% non-fat milk and 0.05% Tween 20 during 1h at room temperature. Incubation with the primary antibodies was performed in the same solution either overnight at 4 °C for TET1 and TET3, or during 1h at room temperature for p80-Ku. Primary antibodies were: anti-TET1 rabbit polyclonal antibody (1:5000, GT1462, Genetex), anti-TET3 rabbit polyclonal antibody (1:5000, C3, Genetex), and anti-p80-Ku mouse monoclonal antibody (1/500, GE2.9.5, Millipore). Following incubation with the primary antibody, the membrane was washed 3 times in PBS-Tween 0.05% and then incubated at room temperature for 45 min in the presence of either HRP-conjugated goat anti-rabbit IgG antibody (1:10000, Enzo, Life Sciences) or HRP-conjugated goat

anti-mouse IgG antibody (1:2000, Santa Cruz). Signals on the membrane were revealed using the SuperSignal West Pico Chemiluminescent Substrate (Pierce, Thermo Scientific), and after exposure to Fuji Medical X-RAY films (Fujifilm). Before applying a new antibody, the membrane was subjected to a 10 min incubation in 0.4 M NaOH at room temperature, three washes in PBS-Tween 0.05%, and incubation in a PBS solution containing 4% non-fat milk and 0.05% Tween 20 during 1h at room temperature.

Slot Blot analysis of 5hmC

Genomic DNA was extracted using the SDS/Proteinase K lysis method as described previously.[40] It was then treated with RNase A (1µg/µl final concentration) for 30 min, and purified by phenol/chloroform extractions and ethanol precipitation. The Nanodrop ND-1000 spectrophotometer (Isogen, Life Sciences) was used for quantification. DNA samples were denatured by incubation in a solution of 0.4 mM NaOH and 10 mM EDTA at 99 °C for 10 min, and then chilled on ice. Aliquots (500 ng for HEK293T and 2 µg for TERA-1) were slotted on positively charged nylon Hybond™ membranes (Amesham, GE Healthcare) using a Hybri.Slot 24 blotting apparatus (Core, Life Sciences). Membranes were thereafter washed quickly in 2 x SSC, dried, and cross-linked by UV exposure. Blocking of the membranes was performed by incubation in PBS containing 5% dry milk and 0.05% Tween, during 1h at room temperature. They were then probed with an anti-5hmC rabbit polyconal antibody (1:10000, #39769, Active Motif), during an overnight incubation at 4 °C. After three washes in PBS-Tween 0.05%, the membranes were incubated at room temperature for 45 min in the presence of HRP-conjugated goat anti-rabbit IgG antibody (1:10000, Enzo, Life Sciences). Signals on the membranes were revealed using the SuperSignal West Pico Chemiluminescent Substrate (Pierce, Thermo Scientific), and after exposure to Fuji Medical X-RAY films (Fujifilm). Membranes were thereafter washed quickly in PBS-Tween 0.05%, before staining in a methylene blue solution (0.02% methylene blue, 0.3M NaOH, pH 5.2). Films and stained membranes were scanned, and slot intensities were quantified using ImageJ.

Analysis of lung squamous cell carcinoma mRNA data sets

We used data sets from the TCGA (Nature 2012), which were obtained on Agilent microarrays.[41] The gene expression value in each sample was reported to the mean value in all samples. Relative gene expression values were then $\log_2$ transformed for subsequent analyses. Samples were sorted according to their relative *GABRA3* expression value. *GABRA3*^{low} and *GABRA3*^{high} subgroups were defined so as to match the proportion of *GABRA3* negative/positive (60% neg. and 40% pos.) samples determined by RT-qPCR experiments in NSCLC lung carcinomas, see Figure 1C). Two-tailed unpaired *t* test was used for statistical analyses.

Acknowledgments

The authors wish to acknowledge the excellent technical assistance of Marjorie Mercier. This work was supported by the Fonds special de recherche (FSR), Université catholique de Louvain, Belgium. A.L. was supported by a special grant from the FSR, and by the de Duve Institute, Brussels, Belgium. A.V.T. and J.C. are the recipients of a Télévie grant from the FRS-FNRS, Belgium [#7.4581.13, and #7.4517.13]

5 References

1. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16:6–21. doi: 10.1101/gad.947102. [PubMed] [Cross Ref]
2. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet. 2002;3:415–28. [PubMed]
3. De Smet C, Loriot A. DNA hypomethylation and activation of germline-specific genes in cancer. Adv Exp Med Biol. 2013;754:149–66. doi: 10.1007/978-1-4419-9967-2_7. [PubMed] [Cross Ref]
4. Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ. Cancer/testis antigens, gametogenesis and cancer. Nat Rev Cancer. 2005;5:615–25. doi: 10.1038/nrc1669. [PubMed] [Cross Ref]
5. Cannuyer J, Loriot A, Parvizi GK, De Smet C. Epigenetic hierarchy within the MAGEA1 cancer-germline gene: promoter DNA methylation dictates local histone modifications. PLoS One. 2013;8:e58743. doi: 10.1371/journal.pone.0058743. [PMC free article] [PubMed] [Cross Ref]
6. De Smet C, Lurquin C, Lethé B, Martelange V, Boon T. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. Mol Cell Biol. 1999;19:7327–35. [PMC free article] [PubMed]
7. Croce CM. Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet. 2009;10:704–14. doi: 10.1038/nrg2634. [PMC free article] [PubMed] [Cross Ref]
8. Melo SA, Esteller M. Dysregulation of microRNAs in cancer: playing with fire. FEBS Lett. 2011;585:2087–99. doi: 10.1016/j.febslet.2010.08.009. [PubMed] [Cross Ref]
9. Liu Y, Guo F, Dai M, Wang D, Tong Y, Huang J, Hu J, Li G. Gammaaminobutyric acid A receptor alpha 3 subunit is overexpressed in lung cancer. Pathol Oncol Res. 2009;15:351–8. doi: 10.1007/s12253-008-9128-7. [PubMed] [Cross Ref]
10. Liu Y, Li YH, Guo FJ, Wang JJ, Sun RL, Hu JY, Li GC. Gamma-aminobutyric acid promotes human hepatocellular carcinoma growth through overexpressed gamma-aminobutyric acid A receptor alpha 3 subunit. World J Gastroenterol. 2008;14:7175–82. doi: 10.3748/wjg.14.7175. [PMC free article] [PubMed] [Cross Ref]
11. Lu Y, Lemon W, Liu PY, Yi Y, Morrison C, Yang P, Sun Z, Szoke J, Gerald WL, Watson M, et al. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. PLoS Med. 2006;3:e467. doi: 10.1371/journal.pmed.0030467. [PMC free article] [PubMed] [Cross Ref]
12. Zhang X, Zhang R, Zheng Y, Shen J, Xiao D, Li J, Shi X, Huang L, Tang H, Liu J, et al. Expression of gamma-aminobutyric acid receptors on neoplastic growth and prediction of prognosis in non-small cell lung cancer. J Transl Med. 2013;11:102. doi: 10.1186/1479-5876-11-102. [PMC free article] [PubMed] [Cross Ref]
13. Zhou W, Fong MY, Min Y, Somlo G, Liu L, Palomares MR, Yu Y, Chow A, O'Connor ST, Chin AR, et al. Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis. Cancer Cell. 2014;25:501–15. doi: 10.1016/j.ccr.2014.03.007. [PMC free article] [PubMed] [Cross Ref]
14. Pfeifer GP, Kadam S, Jin SG. 5-hydroxymethylcytosine and its potential roles in development and cancer. Epigenetics Chromatin. 2013;6:10. doi: 10.1186/1756-8935-6-10. [PMC free article] [PubMed] [Cross Ref]
15. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature. 2010;466:1129–33. doi: 10.1038/nature09303. [PMC free article] [PubMed] [Cross Ref]
16. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324:930–5. doi: 10.1126/science.1170116. [PMC free article] [PubMed] [Cross Ref]
17. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, An J, Lamperti ED, Koh KP, Ganetzky R, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature. 2010;468:839–43. doi: 10.1038/nature09586. [PMC free article] [PubMed] [Cross Ref]
18. Lian CG, Xu Y, Ceol C, Wu F, Larson A, Dresser K, Xu W, Tan L, Hu Y, Zhan Q, et al. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. Cell. 2012;150:1135–46. doi: 10.1016/j.cell.2012.07.033. [PMC free article] [PubMed] [Cross Ref]
19. Yang H, Liu Y, Bai F, Zhang JY, Ma SH, Liu J, Xu ZD, Zhu HG, Ling ZQ, Ye D, et al. Tumor development is associated with decrease of

TET gene expression and 5-methylcytosine hydroxylation. Oncogene. 2013;32:663–9. doi: 10.1038/onc.2012.67. [PMC free article] [PubMed] [Cross Ref]

20. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell. 2010;18:553–67. doi: 10.1016/j.ccr.2010.11.015. [PMC free article] [PubMed] [Cross Ref]
21. Hsu CH, Peng KL, Kang ML, Chen YR, Yang YC, Tsai CH, Chu CS, Jeng YM, Chen YT, Lin FM, et al. TET1 suppresses cancer invasion by activating the tissue inhibitors of metalloproteinases. Cell Rep. 2012;2:568–79. doi: 10.1016/j.celrep.2012.08.030. [PubMed] [Cross Ref]
22. Cheng J, Guo S, Chen S, Mastriano SJ, Liu C, D'Alessio AC, Hysolli E, Guo Y, Yao H, Megyola CM, et al. An extensive network of TET2-targeting MicroRNAs regulates malignant hematopoiesis. Cell Rep. 2013;5:471–81. doi: 10.1016/j.celrep.2013.08.050. [PMC free article] [PubMed] [Cross Ref]
23. Song SJ, Ito K, Ala U, Kats L, Webster K, Sun SM, Jongen-Lavrencic M, Manova-Todorova K, Teruya-Feldstein J, Avigan DE, et al. The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. Cell Stem Cell. 2013;13:87–101. doi: 10.1016/j.stem.2013.06.003. [PMC free article] [PubMed] [Cross Ref]
24. Song SJ, Poliseno L, Song MS, Ala U, Webster K, Ng C, Beringer G, Brikbak NJ, Yuan X, Cantley LC, et al. MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. Cell. 2013;154:311–24. doi: 10.1016/j.cell.2013.06.026. [PMC free article] [PubMed] [Cross Ref]
25. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, Figueroa ME, Vasanthakumar A, Patel J, Zhao X, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell. 2011;20:11–24. doi: 10.1016/j.ccr.2011.06.001. [PMC free article] [PubMed] [Cross Ref]
26. Quivoron C, Couronné L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O, Do Cruzeiro M, Delhommeau F, Arnulf B, Stern MH, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell. 2011;20:25–38. doi: 10.1016/j.ccr.2011.06.003. [PubMed] [Cross Ref]
27. Zha R, Guo W, Zhang Z, Qiu Z, Wang Q, Ding J, Huang S, Chen T, Gu J, Yao M, et al. Genome-wide screening identified that miR-134 acts as a metastasis suppressor by targeting integrin β1 in hepatocellular carcinoma. PLoS One. 2014;9:e87665. doi: 10.1371/journal.pone.0087665. [PMC free article] [PubMed] [Cross Ref]
28. Williams K, Christensen J, Helin K. DNA methylation: TET proteins-guardians of CpG islands? EMBO Rep. 2012;13:28–35. doi: 10.1038/embor.2011.233. [PMC free article] [PubMed] [Cross Ref]
29. Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, Surani MA. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science. 2013;339:448–52. doi: 10.1126/science.1229277. [PMC free article] [PubMed] [Cross Ref]
30. Yamaguchi S, Shen L, Liu Y, Sendler D, Zhang Y. Role of Tet1 in erasure of genomic imprinting. Nature. 2013;504:460–4. doi: 10.1038/nature12805. [PMC free article] [PubMed] [Cross Ref]
31. Gjerstorff MF, Kock K, Nielsen O, Ditzel HJ. MAGE-A1, GAGE and NY-ESO-1 cancer/testis antigen expression during human gonadal development. Hum Reprod. 2007;22:953–60. doi: 10.1093/humrep/del494. [PubMed] [Cross Ref]
32. Koslowski M, Bell C, Seitz G, Lehr HA, Roemer K, Müntefering H, Huber C, Sahin U, Türeci O. Frequent nonrandom activation of germ-line genes in human cancer. Cancer Res. 2004;64:5988–93. doi: 10.1158/0008-5472.CAN-04-1187. [PubMed] [Cross Ref]
33. Maatouk DM, Kellam LD, Mann MR, Lei H, Li E, Bartolomei MS, Resnick JL. DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. Development. 2006;133:3411–8. doi: 10.1242/dev.02500. [PubMed] [Cross Ref]
34. Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic Acids Res. 2011;39:5015–24. doi: 10.1093/nar/gkr120. [PMC free article] [PubMed] [Cross Ref]
35. Kaas GA, Zhong C, Eason DE, Ross DL, Vachhani RV, Ming GL, King JR, Song H, Sweatt JD. TET1 controls CNS 5-methylcytosine hydroxylation, active DNA demethylation,

gene transcription, and memory formation. Neuron. 2013;79:1086–93. doi: 10.1016/j.neuron.2013.08.032. [PMC free article] [PubMed] [Cross Ref]
36. Rudenko A, Dawlaty MM, Seo J, Cheng AW, Meng J, Le T, Faull KF, Jaenisch R, Tsai LH. Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. Neuron. 2013;79:1109–22. doi: 10.1016/j.neuron.2013.08.003. [PMC free article] [PubMed] [Cross Ref]
37. Loriot A, Parvizi GK, Reister S, De Smet C. Silencing of cancer-germline genes in human preimplantation embryos: evidence for active de novo DNA methylation in stem cells. Biochem Biophys Res Commun. 2012;417:187–91. doi: 10.1016/j.bbrc.2011.11.120. [PubMed] [Cross Ref]
38. Davis LG, Dibner MD, Battey JF. Basic methods in molecular biology. New York: Elsevier, 1986.
39. De Smet C, Loriot A, Boon T. Promoter-dependent mechanism leading to selective hypomethylation within the 5′ region of gene MAGE-A1 in tumor cells. Mol Cell Biol. 2004;24:4781–90. doi: 10.1128/MCB.24.11.4781-4790.2004. [PMC free article] [PubMed] [Cross Ref]
40. De Smet C, De Backer O, Faraoni I, Lurquin C, Brasseur F, Boon T. The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. Proc Natl Acad Sci U S A. 1996;93:7149–53. doi: 10.1073/pnas.93.14.7149. [PMC free article] [PubMed] [Cross Ref]
41. Cancer Genome Atlas Research Network Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489:519–25. doi: 10.1038/nature11404. [PMC free article] [PubMed] [Cross Ref]

6 Supplemental data

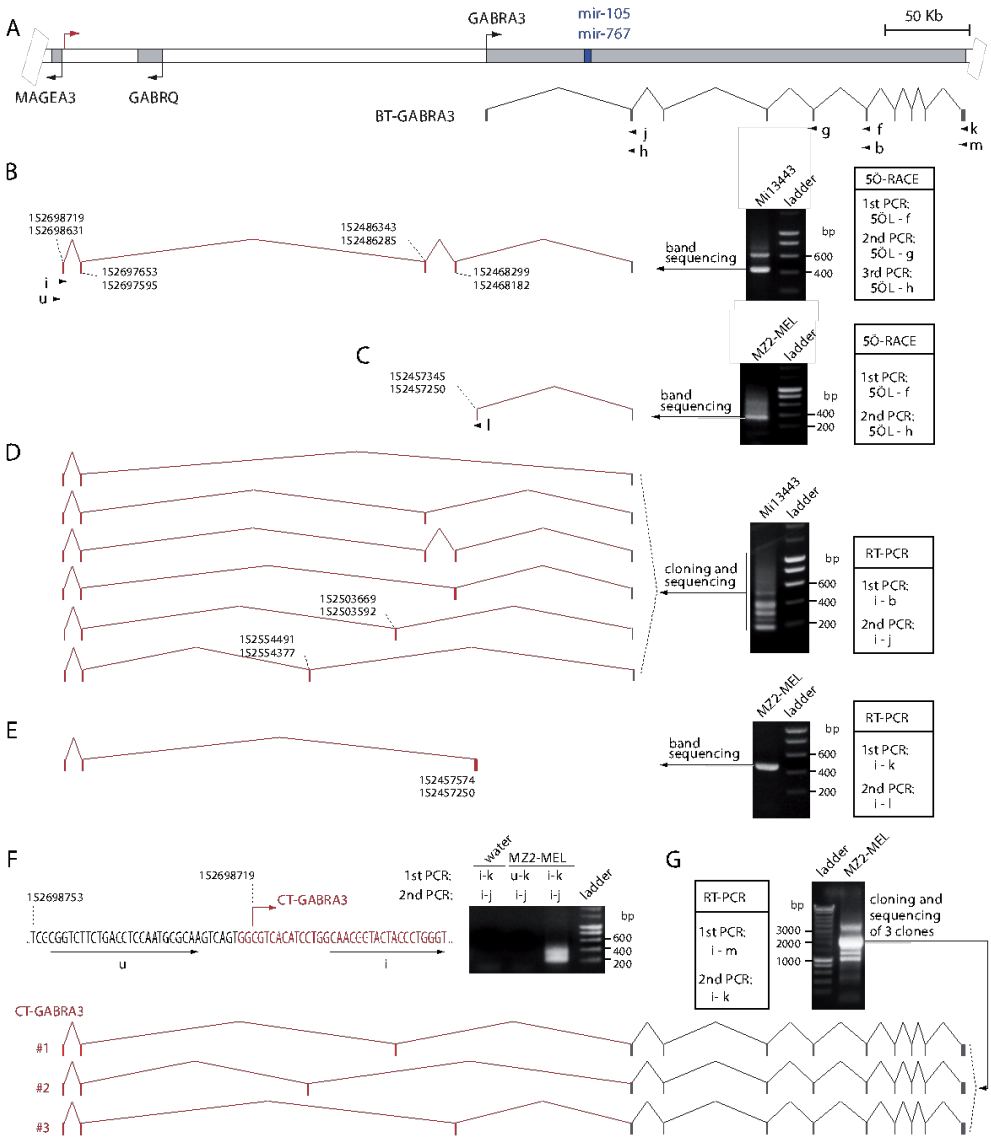


Figure S1. Characterization of the novel *CT-GABRA3* transcript by 5′-RACE and RT-PCR experiments. (A) Schematic representation of the GABRA3 locus, with the exon/intron structure of the reference GABRA3 gene (XM_005274659.1, NCBI RefSeq), which we re-named *BT-GABRA3*. Arrowheads indicate the approximate position and orientation of PCR primers used in subsequent experiments. (B)A first 5′-RACE experiment, using the SMART technology, was performed on total RNA extracted from the Mi13443-MEL melanoma cell line. Three rounds of nested PCR were applied, using the 5′L adapter sense primer and different antisense primers, as indicated. PCR products resulting from the third amplification round were submitted to gel electrophoresis, and the indicated band was extracted and sequenced. Exon/intron structure of the corresponding transcript is shown, and positions of exons relative to the reference GRCh38 human genome assembly are given. Of note, the upper band was found to result from non-specific amplification. (C)A second 5′-RACE experiment, with two rounds of nested PCR, was applied to melanoma cell line MZ2-MEL, and led to the identification of an alternative, likely truncated (see panel E), transcript variant. (D)Total RNA from Mi13443-MEL was reverse-transcribed, and submitted to two nested PCR with the indicated primers. Gel analysis revealed amplification of several bands. The bulk PCR product was cloned into a vector, and individual clones were sequenced. This led to the characterization of several splicing variants. (E) Inclusion of the exon identified by 5′-RACE in MZ2-MEL (see panel C) in *CT-GABRA3* transcripts also containing upstream exons was verified by an RT-PCR experiment using corresponding primers, as indicated. (F)Initiation of *CT-GABRA3* transcription at the position defined by 5′-RACE experiments was corroborated by the lack of RT-PCR signal when a PCR primer (u) located just upstream of that position was used. (G)A final experiment, combining RT-PCR, bulk cloning, and individual clone sequencing, was performed to verify that *CT-GABRA3* transcripts comprise newly indentified upstream exons linked to all exons but exon 1 of the referenced *BT-GABRA3* gene.

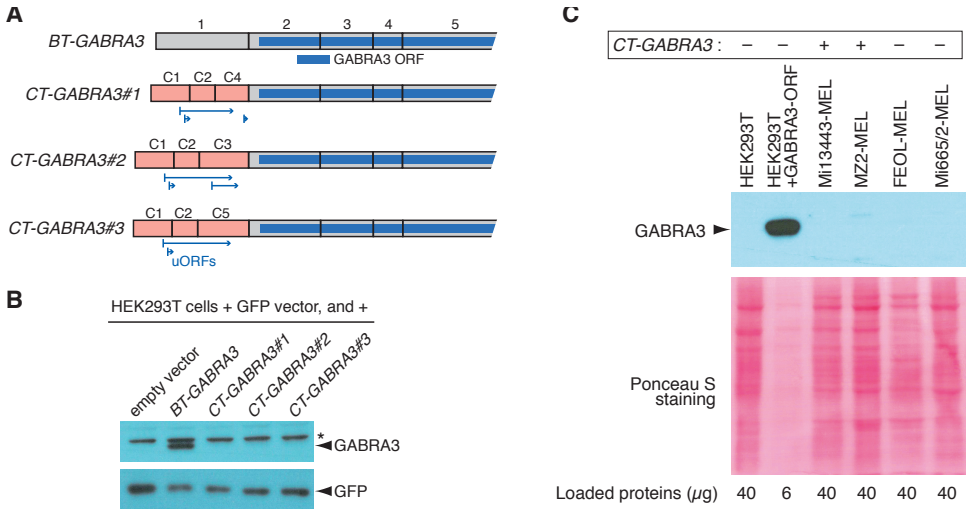


Figure S2. Short upstream open reading frames (uORFs) in *CT-GABRA3* mRNAs inhibit translation of the GABRA3 protein.(A)Expression vectors were constructed, which contained the full length ORF of GABRA3 preceded by the 5′ UTR of either *BT-GABRA3* or three splice variants of *CT-GABRA3* (#1 to #3; see Fig. E2G). Blue arrows below *CT-GABRA3* transcripts delineate uORFs. Of note, two of these originate in the most 5′ exon (C1), which is generally present in all *CT-GABRA3* splice variants. (B)These expression vectors, as well as an empty control vector, were transfected together with a GFP-expressing vector into HEK293T cells. Proteins were extracted from transfected cells and analyzed by Western blotting for the presence of the GABRA3 protein (anti-GABRA3; #ab23334, Abcam). Immunodetection of GFP (anti-GFP; #ab290, Abcam) served to verify similar transfection efficiencies. The results revealed marked inhibition of GABRA3 translation in *CT-GABRA3* transcripts. * non-specific band. (C)The presence of the GABRA3 protein was examined in melanoma cell lines that do (+) or do not (-) express *CT-GABRA3*. None of these cell lines expressed *BT-GABRA3*. HEK293T cells that had been transfected with an expression vector carrying the GABRA3 ORF were used as positive control. Loaded protein amounts are indicated, and were

verified after Ponceau S staining. The results show that the GABRA3 protein remains undected in *CT-GABRA3*-expressing cell lines.

melanoma tissues (n=24)		
	GABRA3 +	GABRA3 Ð
MAGEA3 +	17	1
MAGEA3 Ð	0	7

$p = 1.66 \times 10^{-5}$ (Fisher exact test)

NSCLC tissues (n=24)		
	GABRA3 +	GABRA3 Ð
MAGEA3 +	11	8
MAGEA3 Ð	0	8

$p = 5.80 \times 10^{-3}$ (Fisher exact test)

Figure S3. GABRA3 and MAGEA3 show frequent co-activation in melanoma and non-small cell lung carcinoma (NSCLC) tissue samples.Tumor samples were analyzed by RT-PCR for the expression of GABRA3 and MAGEA3. For each tumor type, the number of samples that scored either positive or negative for one or both genes was reported in a 2X2 contingency table. Significant tendency of co-expression of the two genes was validated by a Fisher exact test.

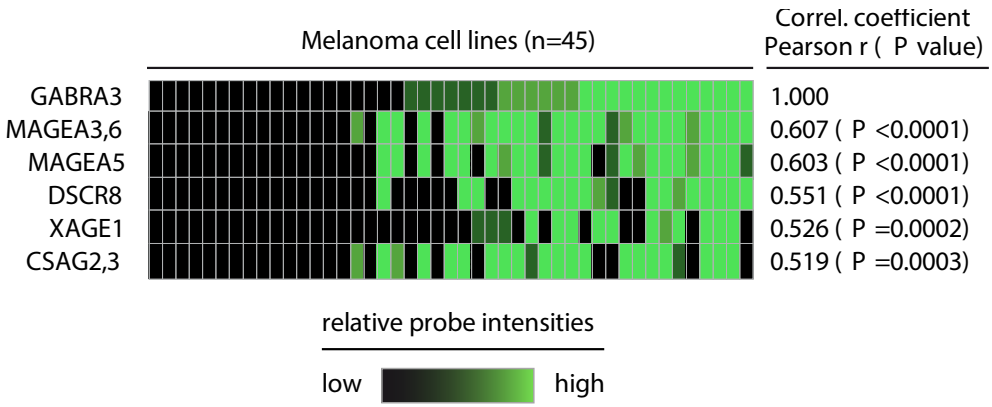


Figure S4. Activation of GABRA3 transcripts in melanoma cell cultures correlates with that of CG genes. Publicly available gene expression microarray data obtained from 45 melanoma cell cultures (dataset GSE4843) were downloaded. Relative probe intensities are reported by the indicated color code in the grid, with each column corresponding to a melanoma cell culture, and each row to a gene. Melanoma cell cultures were ordered according to GABRA3 relative probe intensity values. Probe intensities for several known CG genes are also depicted (MAGEA3 and MAGEA6, as well as CSAG2 and CSAG3 cannot be distinguished because of high sequence identity). Correlation coefficients between relative probe intensity values of CG genes and GABRA3 were calculated. These revealed a significant trend of co-activation of GABRA3 with CG genes.

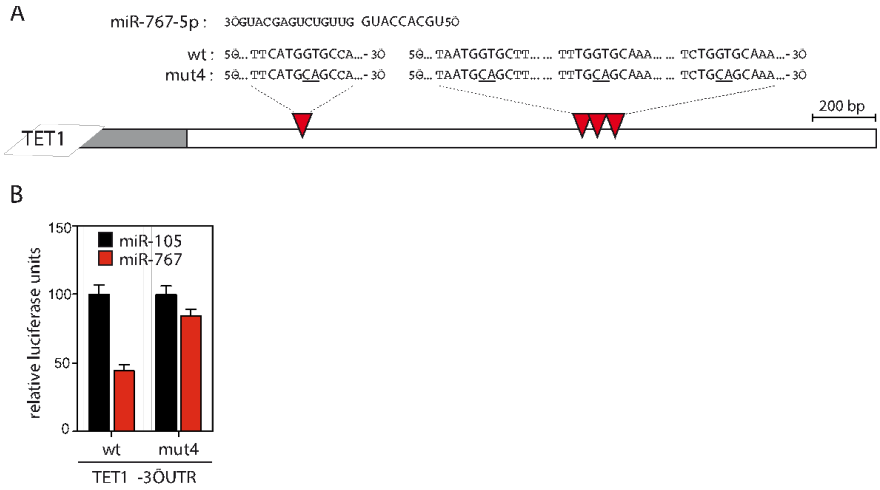


Figure S5. TET1 regulation by miR-767 is impaired by mutations in 3′-UTR target sequences.(A)Schematic representation of the 3′-UTR of TET1 (empty bar) with the locationof miR-767 seed sequences indicated by red triangles (only 7-mer and 8-mer seeds are depicted). The sequence of miR-767(-5p) is given, as well as the seed match sequences in the 3′-UTR of TET1 (wt). A mutant form of the 3′UTR of TET1 was generated (mut4), which contains the indicated nucleotide changes (underligned) in each of the four seed sequences. (B) HEK293T cells were co-transfected with a luciferase reporter gene linked to either the wt or the mut4 TET1 3′-UTR, and with the indicated miRNAs. Renilla luciferase were measured 24h after transfection, and were normalized with respect to the Firefly luciferase reporter carried on the same vector. Luciferase units are expressed relative to miR-105-transfected control cells, and represent means ±SD (n=4).

Table S1. Description of human tumor tissue samples

Melanoma tissue samples				
Ref	Patient code	Sex*	Age**	type of lesion
#1	LB1616	F	77	cutaneous melanoma, breast metastasis
#2	LG37	F	36	cutaneous melanoma , iliac node metastasis
#3	EN7	F	35	cutaneous melanoma , leg metastasis
#4	CP30	F	n/a	cutaneous melanoma, lymph node metastasis
#5	LB2077	F	28	metastatic cutaneous melanoma
#6	DDHK0002	M	56	metastatic cutaneous melanoma
#7	DDHK0002	M	56	cutaneous melanoma , leg in transi t metastasis
#8	KUL73	F	71	cutaneous melanoma , satellite nodule
#9	LB2201	F	45	cutaneous melanoma, epithrochlear lymph node metastasis
#10	CP64	M	82	cutaneous melanoma , clavicular su bcutaneous nodule
#11	KUL73	F	71	metastatic cutaneous melanoma
#12	VUB39	F	73	cutaneous melanoma , axillary node metastasis
#13	BB132	M	70	cutaneous melanoma , thoracic subcutaneous nodule
#14	LB168	M	53	cutaneous melanoma , mesenteric lymph node metastasis
#15	KUL73	F	71	cutaneous melanoma , inflammatory cutaneous nodule
#16	LB207 7	F	28	cutaneous melanoma , intracardiac metastasis
#17	EB81	F	70	cutaneous melanoma , lymph node metastasis
#18	LB2269	M	62	cutaneous melanoma, axillary lymph node metastasis
#19	LB2259	F	74	cutaneous melanoma , thoracic subcutaneous nodule
#20	LB2370	M	39	cutaneous melanoma , inguinal lymph node metastasis
#21	LB2293	F	53	cutaneous melanoma, lymph node metastasis
#22	LB2357	M	62	cutaneous melanoma , supraclavicular lymph node metastasis
#23	CP67	F	44	cutaneous melanoma , left axillary lymphadenopathy
#24	LB2174	F	49	cutaneous melanoma , subcutaneous nodule
#25	LB2439	F	33	cutaneous melanoma , mediastinal lymph node metastasis
#26	DDHK0062	M	42	cutaneous melanoma , leg in transit metastasis
#27	LB2174	F	49	cutaneous melanoma , subcuta neous nodule
#28	LB2652	M	n/a	cutaneous melanoma, temporal lobe metastasis
#29	LB1572	M	36	benign nevus

Non -small cell lung carcinoma tissus samples				
Ref	Patient code	Sex	Age	type of lesion
#1	LB1214	M	65	epidermoid carcinoma, primary tumor
#2	LB498	M	55	epidermoid carcinoma, calf muscle metastasis
#3	LB973	M	62	epidermoid carcinoma, primary tumor
#4	LB1005	M	68	epidermoid carcinoma, primary tumor
#5	LB1006	F	67	epidermoid carcinoma, primary tumor
#6	LB1007	n/a	60	epidermoid carcinoma, primary tumor
#7	LB498	M	55	epidermoid carcinoma, recurrent calf muscle metastasis
#8	LB1061	M	62	epidermoid carcinoma, primary tumor
#9	LB1080	M	68	epidermoid carcinoma, primary tumor
#10	LB1102	M	49	epidermoid carcinoma, primary tumor
#11	LB1104	M	77	epidermoid carcinoma, primary tumor
#12	LB1123	M	n/a	epidermoid carcinoma, primary tumor
#13	LB1124	M	n/a	epidermoid carcinoma, primary tumor
#14	LB1125	M	66	epidermoid carcinoma, primary tumor
#15	LB1135	M	74	epidermoid carcinoma, primary tumor
#16	LB1136	F	68	epidermoid carcinoma, primary tumor
#17	LB1152	M	64	epidermoid carcinoma, primary tumor
#18	LB1156	M	54	epidermoid carcinoma, primary tumor
#19	LB911	M	73	epidermoid carcinoma, primary tumor
#20	LB1201	M	72	epidermoid carcinoma, primary tumor
#21	LB1211	M	72	epidermoid carcinoma, primary tumor
#22	LB1210	M	63	epidermoid carcinoma, primary tumor
#23	LB1222	M	69	epidermoid carcinoma, primary tumor
#24	LB1229	M	73	epidermoid carcinoma, primary tumor
#25	LB1239	M	64	epidermoid carcinoma, primary tumor
#26	LB1261	M	74	epidermoid carcinoma, primary tumor
#27	LB1303	M	n/a	epidermoid carcinoma, primary tumor

* M = male; F = female
** Age at tissue sample resection

Table S2. Oligonucleotidic primers and probe sequences

Primer/probe	Sequence	Ref.
GABRA3 a	GGCAGACATGGCATGATGAA	
GABRA3 b	CTTCAGCTGTTGTATAGGCATAGCTT	
GABRA3 c	CGAGGGCTCAACCTCCAACTT	
GABRA3 d	CATCATGCCATGTCTGCCGAAA	
GABRA3 e	GGAGGCGGAGATTGCACA	
GABRA3 f	GGCATAGCTTCCAAACTTCAGTGGGCAGGCA	
GABRA3 g	GGGGCCATCAAATTTCAGTCTTTCATCATGCCAT	
GABRA3 h	CCTTGACCAGTGGTTCCAGGGAGA	
GABRA3 i	CAACGGTACTACCCTGGGT	
GABRA3 j	CCAAGGCTGGTCATGTAACAGT	
GABRA3 k	ATGAATTCTACTGTTTGCGGATCATGCCCT	
GABRA3 l	GGAGAGACCTGTGACCTTTCT	
GABRA3 m	TGCTGCACTGCCACCACTAT	
GABRA3 u	CGGTCTTCTGACCTCCAATGCGCAA	
GABRA3 -probe	FAM -CTGGACACCGGACACCTTCTTCCACAATGGCA -TAMRA	
TET1_F1	CCCGAATCAAGCGGAAGAATA	1
TET1_R1	TACTTCAGGTTGCACGGT	1
ACTINF	CCCTGGACTTCGAGCAAGAGAT	1
ACTINR	AAGGTAGTTTCGTGGATGCCACA	1
TET3_F	GTTCCTGGAGCATGTACTTC	2
TET3_R	CTTCCTCTTTGGGATTGTCC	2
TET1_Xho	ATGACTCGAGAGGCTTTTCTCCCCCTCT	
TET1_NotI	TATGCGGCCGCCCTTGCTTCATGAGAAAGAGCA	
TET3_NotI	TATGCGGCCGCGGGAGGGTAAGGAGGGGTA	
TET3_XhoI	ATGACTCGAGGTGCCAGGGAGCCAGCGT	
TET1_mut1	AAAATAAGCTGAATTATTATTTCATGCAGCCATTGTTCCAACATCTTCCAATC	
TET1_mut2	GAGTATGGAAAACCTAATGCAGCTTCTCCCTTGGAAATGC	
TET1_mut3	CCGCTAACACTTACAATTTTGCAGCAAAAGCA AACAGTTCCAGC	
TET1_mut4	AAACTCATTGTAACTTATTAAAATAATATCTGCAGCAAAGTATCTGTTTTGAGCTTTTGAC	

1. Hsu et al., 2012 Cell Rep , 2:568 -579
2. Lian et al., 2012 Cell , 150:1135 -1146
!

CHAPTER II

Potential mechanistic link between DNA hypomethylation and hypermethylation of alternative *GABRA3* promoters in tumors

Van Tongelen A., Fain J., Loriot A.,and De Smet C.

1 Introduction

Through our previous research on the *GABRA3* locus, we discovered a new cancer-germline (CG) transcript called *CT-GABRA3* (Loriot et al. 2014). This transcript starts its transcription from an alternative promoter located 247-kb upstream of the canonical promoter of *GABRA3* which we renamed *BT-GABRA3* (**Fig. 1A**). We observed that in normal tissues, *BT-GABRA3* is only expressed in brain and testis. The *CT-GABRA3* transcript is normally expressed exclusively in testis and becomes frequently active in melanoma cell lines (**Fig. 1B**). We also showed that the tissue-specific expression of *CT-GABRA3* is regulated by DNA methylation. Indeed, in normal somatic tissues, *CT-GABRA3* is methylated while its transcriptional activation is associated with its promoter hypomethylation in melanoma. Moreover, the expression of *CT-GABRA3* can be induced by a DNA methylation inhibitor, 5-azadC (Loriot et al. 2014).

An intriguing observation was that in tumor cell lines where *CT-GABRA3* is demethylated and expressed, the promoter of *BT-GABRA3*, located downstream, is hypermethylated. This raised the possible existence of a crosstalk between DNA hypomethylation and DNA hypermethylation in the *GABRA3* locus. In order to verify this possibility, we initiated a research project aiming at generalizing our observation of opposing DNA methylation alterations in *CT-* and *BT-GABRA3* promoters in a larger set of samples, and at deciphering the mechanisms underlying this process of interdependent epigenetic modification. Preliminary results are presented in the following pages.

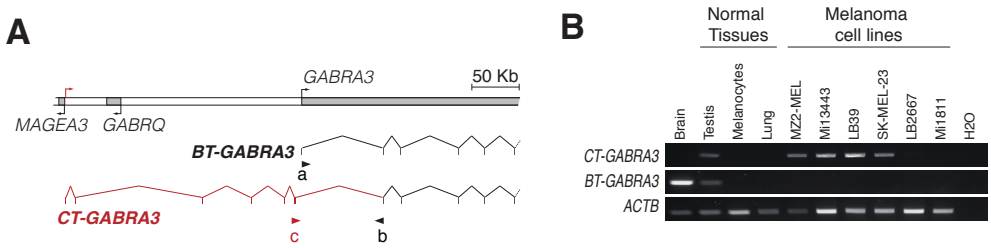


Figure 1: (A) Schematic representation of the GABRA3 locus, and localization of primers used for specific amplification of *BT-GABRA3* and *CT-GABRA3* transcript variants. (B) Gel analysis of RT-PCR with primers recognizing only *BT-GABRA3* transcript (primers a & b) or only *CT-GABRA3* transcripts (primers c & b) in normal tissues and melanoma cell lines.

2 Results

2.1 CT-GABRA3 is activated in tumors

In our previous research, we analyzed *CT-GABRA3* and *BT-GABRA3* expression in only a limited number of samples of two melanoma and non-small cell lung cancer (NSCLC) samples. To confirm that *CT-GABRA3* is activated more frequently than *BT-GABRA3* in tumors, we performed an RT-PCR analysis to examine the expression profile of each of these transcripts in a larger set of melanoma and NSCLC tissue samples (**Fig. 2**). We observed that *CT-GABRA3* is expressed in 64% (16/25) of melanoma tissue samples and 30% (8/27) of NSCLC tissue samples. On the other hand, *BT-GABRA3* is generally silent with only sporadic activation in 8% (2/25) melanoma and 7% (2/27) NSCLC tissue samples. These observations confirm our previous data showing that *CT-GABRA3* is a CG-gene frequently activated in melanoma and lung cancer, whereas *BT-GABRA3* is generally silent.

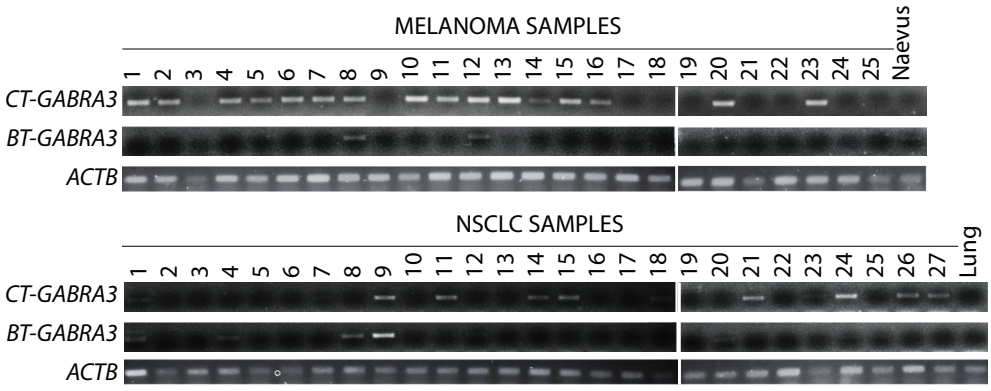


Figure 2: Gel analysis of RT-PCR of *BT-GABRA3* and *CT-GABRA3* transcripts in melanoma and lung carcinoma tissue samples.

2.2 *CT-GABRA3 promoter hypomethylation correlates with BT-GABRA3 promoter hypermethylation in melanoma*

Both promoters, CT-GABRA3 and BT-GABRA3, contain CpG islands (**Fig. 3**). Bisulfite sequencing analyses were performed in order to determine DNA methylation status of *CT-GABRA3* and *BT-GABRA3* promoters in normal tissues and melanoma cell lines (**Fig. 4**). In normal somatic cells (melanocytes and lung), and in two melanoma cell lines (Mi1811 and LB2667) where *CT-GABRA3* is not expressed, the promoter of *CT-GABRA3* is methylated and that of *BT-GABRA3* is unmethylated. Conversely, in melanoma cell lines MZ2-MEL, Mi13443, LB29 and SK23, where the promoter *CT-GABRA3* is hypomethylated and activated, the *BT-GABRA3* promoter is hypermethylated (**Fig. 4**).

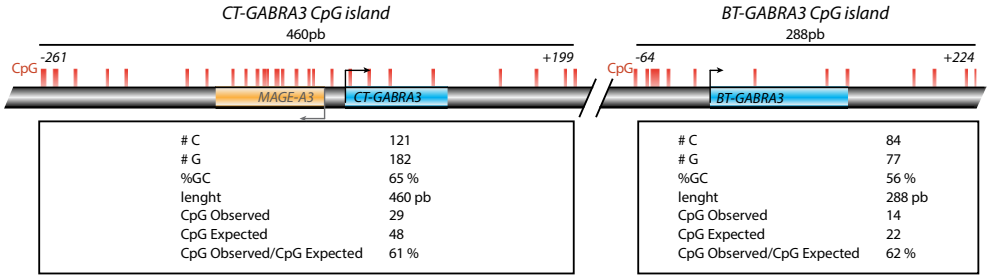


Figure 3: Illustration of CpG islands of the *CT-GABRA3* and *BT-GABRA3* promoters according the criteria : length > 200pb ; GC% > 0,5 and CpG expected on CpG expected > 0,6 (Gardiner-Garden and Frommer 1987). Grey box represents genomic DNA, colored boxes represent first exon of the indicated gene, broken arrows represent transcription initiation and red lines represent CpG sites.

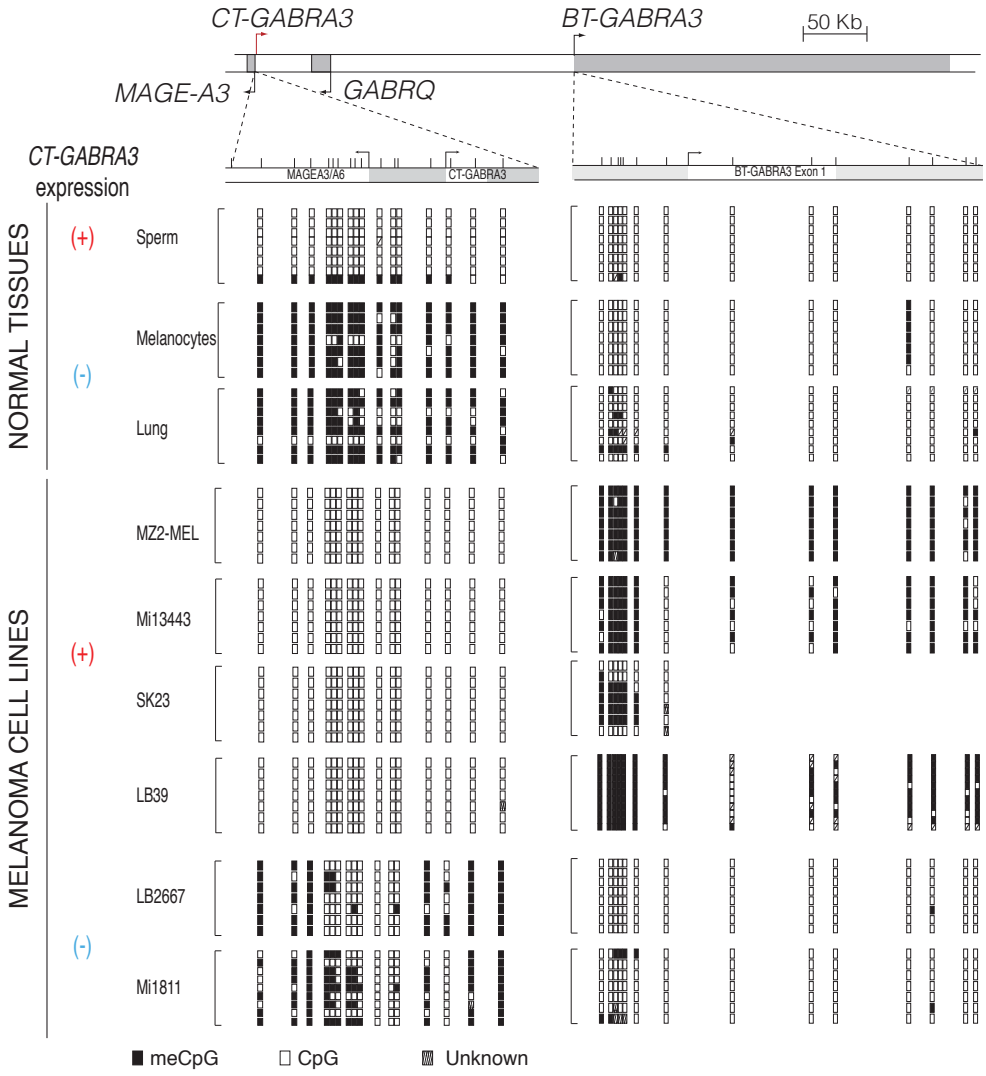


Figure 4: Bisulfite sequencing of the MAGEA3/*CT-GABRA3* promoter region and *BT-GABRA3* promoter in normal tissues and melanoma cell lines. Vertical bars indicate location of CpG sites with positions relative to the *CT-GABRA3* start site. Open and filled squares represent unmethylated and methylated CpG sites respectively, and each row represents a single clone. *BT-GABRA3* and *CT-GABRA3* expression status (+) or (-) in samples is indicated (positive samples also express MAGEA3).

It is known that cancer cell lines display DNA methylation changes that differ from the corresponding primary malignancies, and that these changes are likely due to long-term culture. Therefore, we searched to verify if the inverse

correlation of DNA methylation profiles we observed in *CT-GABRA3* and *BT-GABRA3* promoters in melanoma cell lines, also existed in melanoma samples *in vivo*. Because tumor samples are heterogeneous, we resorted to a more sensitive method to evaluate a relative methylation level of both promoters in melanoma samples by using quantitative methyl-specific PCR (qMSP) analysis.

A relative hypomethylation level was calculated for the *CT-GABRA3* promoter and a relative hypermethylation level for the *BT-GABRA3* promoter. Similarly to the observations in melanoma cell lines, there was a trend towards an inverse correlation can be observed between the methylation level of *CT-GABRA3* and *BT-GABRA3* promoters *in vivo* in melanoma samples (**Fig. 5**).

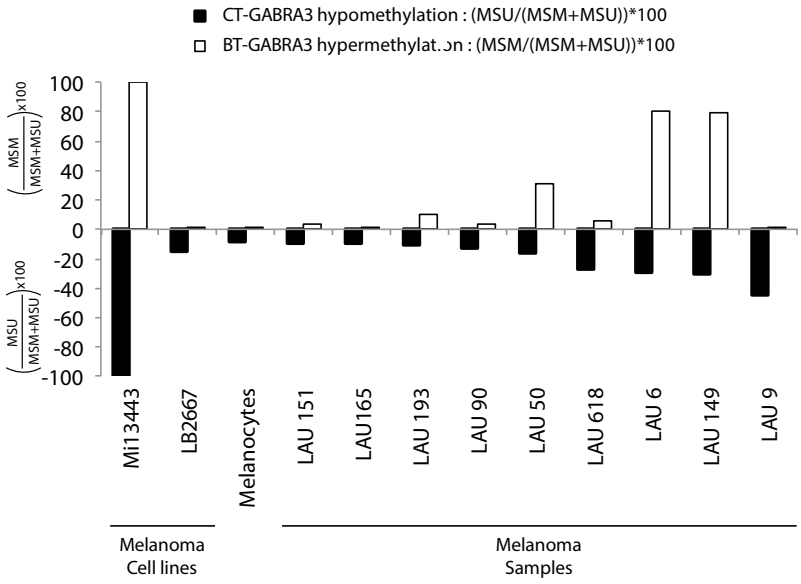


Figure 5: Methylation profile of *BT-GABRA3* and *CT-GABRA3* by quantitative methyl-specific PCR (qMSP) analysis on melanoma samples. The relative methylation level corresponds to the number of amplification by primers specific for unmethylated (Methyl-Specific Unmethylated - MSU) or methylated (Methyl-Specific Methylated - MSM) regions on the number of amplification by both primer pairs, MSM and MSU, multiplied by 100. The upper part of the graph shows a gain of methylation of the *BT-GABRA3* promoter region whereas the lower part of the graph indicates a loss of methylation on the *CT-GABRA3* promoter region. Melanoma samples data are sorted by the level of *CT-GABRA3* hypomethylation. As a control, we took the cell lines Mi13443 that showed in previous bisulfite sequencing analysis a complete hypomethylation of *CT-GABRA3* and extremely high methylation of *BT-GABRA3*, as well as the cell line LB2667 that showed a high methylation level of *CT-GABRA3* and an absence of methylation in *BT-GABRA3* promoter.

2.3 Activation/hypomethylation promoter CT-GABRA3 correlates with BT-GABRA3 promoter hypermethylation in lung adenocarcinoma

CT-GABRA3 is expressed not only in melanoma, but also in lung cancer (**Fig. 6**). Therefore, in order to verify the association between *CT-GABRA3* demethylation and *BT-GABRA3* hypermethylation in lung cancer, we performed an *in silico* analysis. The Database Transcription Start Site (DBTSS; *www. http://dbtss.hgc.jp*) contains the precise position of transcriptional start sites (TSS) in the genome, various multiomics data such as the DNA methylation profile of gene promoter regions, the enrichment of histone marks across the genome and also gene expression data. All these data are available for 26 lung adenocarcinoma cell lines. As *GABRA3* locus is located on the X chromosome, we studied 15 lung adenocarcinoma cell lines that derived from men. Through this selection, we avoided the bias of allele differences, as men harbor only one copy of the X chromosome. Of note, RNA-seq RNA-seq data of DBTSS does not differentiate the two variants of *GABRA3*.

We classified the 15 cell lines in two groups according to the expression profile of *GABRA3* (**Fig. 6A**). The first group contains nine cell lines showing an expression of *GABRA3* (>1RPKM) and six cell lines showing an absence of *GABRA3* expression (<1RPKM). It is reasonable to consider that positive RNA-seq results for *GABRA3* expression in these cell lines correspond to the presence of *CT-GABRA3* transcripts, as this variant is much more frequently activated in tumors than *BT-GABRA3*. Moreover, in these cell lines, positive RNA-seq results for *GABRA3* always correlated with expression of gene *MAGEA3*, which shares the same promoter with *CT-GABRA3* (Loriot et al. 2014). The frequent activation of *CT-GABRA3* in lung cancer cell lines is consistent with previous observations made in lung cancer samples.

Furthermore, bisulfite sequencing (BS-Seq) data from DBTSS show that the expression of *CT-GABRA3* in lung adenocarcinoma cell lines is correlated with *CT-GABRA3* promoter hypomethylation (**Fig. 6B**). Indeed, in *CT-GABRA3* non-expressing cell lines, *CT-GABRA3* promoter is methylated in 100% (6/6) of the cell lines, while in the group of cell lines expressing *CT-GABRA3*, nearly 80 % (7/9) of the cell lines show hypomethylation of *CT-GABRA3* promoter. Remarkably, 100%

(9/9) of the *CT-GABRA3* expressing cell lines showed hypermethylation within the 5' region of *BT-GABRA3* whereas this was observed for only 50% (3/6) of the *CT-GABRA3* non-expressing cell lines (**Table 1**).

Overall, the methylation profile analysis of *BT-GABRA3* and *CT-GABRA3* promoters in lung cancer shows an inverse correlation between the methylation profile of the two promoters *BT-GABRA3* and *CT-GABRA3,* as we also observed in melanoma.

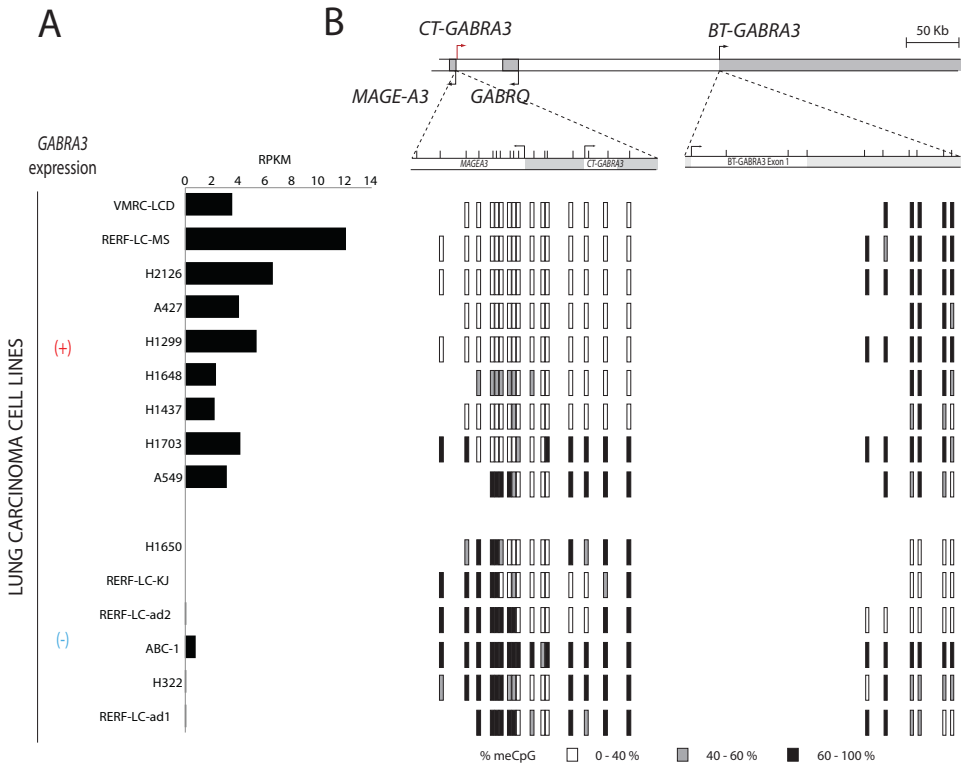


Figure 6 in silico (http://dbtss.hgc.jp) analysis of GABRA3 expression and methylation profile in lung carcinoma cell lines. (A) Relative expression of GABRA3 (RPKM). (B) Bisulfite-sequencing (BS-Seq) data showing methylation state of the *BT-GABRA3* and *CT-GABRA3* promoter region.

Table 1: Number of lung adenocarcinoma cell lines sorted according GABRA3 expression and methylation profile from in silico (http://dbtss.hgc.jp)

		GABRA3 expression +	*GABRA3* expression -
Methylation of *CT-GABRA3* promoter	+	2 (22%)	6 (100%)
	-	7 (78%)	0

		GABRA3 expression +	*GABRA3* expression -
Methylation of *BT-GABRA3* promoter	+	9 (100%)	3 (50%)
	-	0 (0%)	3 (50%)

2.4 CT-GABRA3 expression and hypermethylation of BT-GABRA3 promoter is associated with H3K36me3 enrichment in BT-GABRA3 promoter region

In order to investigate potential epigenetic mechanisms associated with *BT-GABRA3* hypermethylation, we performed an *in silico* analysis on the 15 male lung carcinoma cell lines through the DBTSS platform where ChIP seq data are available for several histone marks. We classified the 15 lung adenocarcinoma cell lines in two groups according to the expression of *CT-GABRA3*. These analysis allowed us to identify an enrichment of H3K36me3 histone mark in association with *BT-GABRA3* hypermethylation (**Fig. 7**). Moreover, we observed that the enrichment of H3K36me3 mark is associated with the transcription of *CT-GABRA3*, as it extends from the TSS of *CT-GABRA3* up to the end of the transcript, located far downstream of the *BT-GABRA3* promoter. Additionally, only H3K36me3 histone mark was associated with *BT-GABRA3* hypermethylation while no specific enrichment or loss of other histone marks such as H3K27me3, H3K4me2/3, H3ac, H3K9me3 or H3K27ac had been observed in the *BT-GABRA3* promoter region (data not shown).

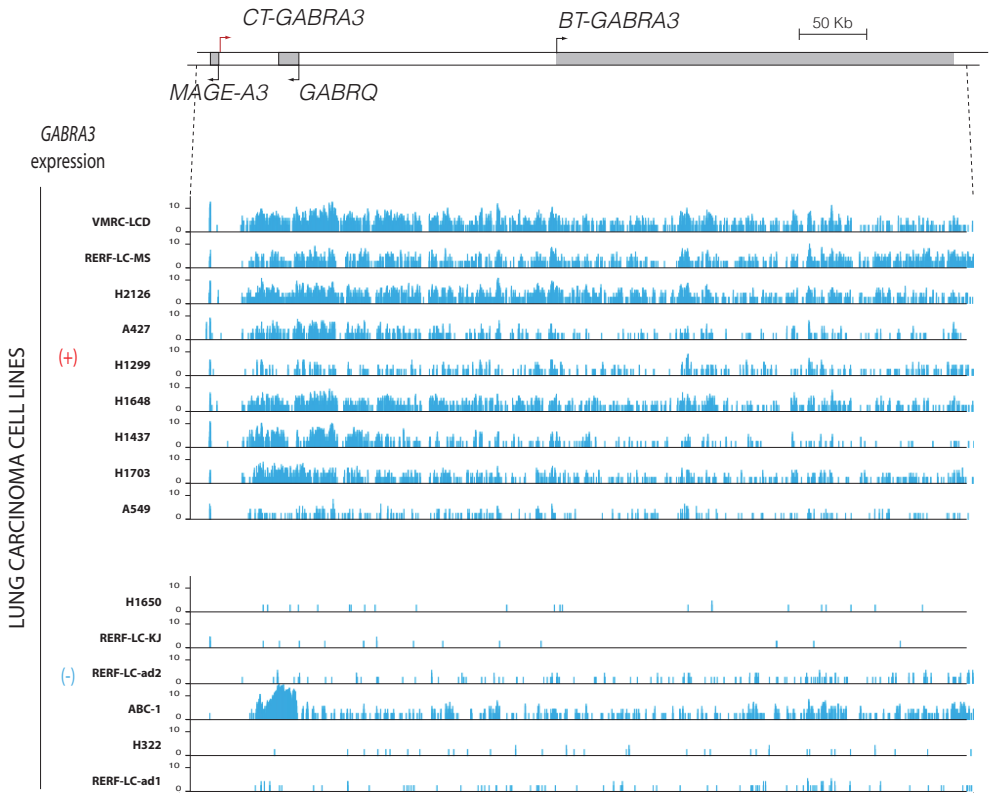


Figure 7: in silico data of ChIP-Seq analysis targeting H3K36me3 mark in 15 lung carcinoma cell lines showing the presence of H3K36me3 mark all along *CT-GABRA3* transcript when it's expressed.

2.5 Methylated BT-GABRA3 promoter but not unmethylated BT-GABRA3 promoter is sensitive to DNA demethylation treatment

Since *BT-GABRA3* is silent in most cell types regardless of its promoter methylation states, hypermethylation of its promoter is expected to have little impact. An experiment using a DNA demethylating agent produced however unexpected results (**Fig. 8**). Thus, four melanoma cell lines displaying *CT-GABRA3* expression and *BT-GABRA3* promoter hypermethylation (MZ2-MEL, Mi13443, LB39 and SK23), and two melanoma cell lines lacking *CT-GABRA3* expression and *BT-GABRA3* promoter hypermethylation (LB2667 and Mi1811) were treated with

the DNA methylation inhibitor, 5-azadC. The treatment showed no effect on *BT-GABRA3* expression in melanoma cell lines LB2667 and Mi1811, but surprisingly, induced activation of this gene where its promoter was initially hypermethylated (**Fig. 8**).

These results indicate that in is default state, *BT-GABRA3* is insensitive to DNA methylation. However, following promoter hypermethylation, *BT-GABRA3* becomes dependent on DNA methylation and can be induced by a DNA demethylation agent.

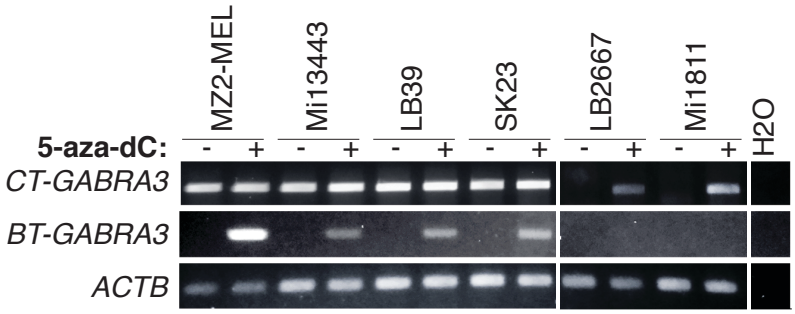


Figure 8: Six melanoma cell lines cultured in the presence (+) or in the absence (-) of a DNA methylation inhibitor, 5-aza-deoxycytidine (5-azadC). Expression of *CT-GABRA3*, *BT-GABRA3* and *ACTB* (control) were analyzed by RT- PCR.

3 Discussion & Perspectives

Today, three main mechanisms have been proposed to direct hypermethylation of gene promoters in tumors. These mechanisms are based on the activation of transcription factors (Di Croce et al. 2002; Thillainadesan et al. 2012), the co-existence of H3K27me3 and H3K4me2 histone marks at gene promoter (Varambally et al. 2002; Schlesinger et al. 2007), and the downregulation of TET demethylase activity (Thienpont et al. 2016). However, these mechanisms have been demonstrated for a restricted proportion of hypermethylation events suggesting that DNA hypermethylation in tumors can result from multiple mechanisms.

Our preliminary results on the *GABRA3* locus uncover a potential new molecular process leading to DNA hypermethylation in tumors. Our analysis revealed that the hypomethylation and transcription of the upstream promoter of the *GABRA3* locus, *CT-GABRA3,* are associated with the hypermethylation of the downstream promoter of the *GABRA3* locus, *BT-GABRA3*, in cancer. This observation reveals an unexpected connection between DNA hypomethylation and hypermethylation in tumors.

Moreover, we showed that the transcription of *CT-GABRA3* was associated with the enrichment of H3K36me3 histone mark through its transcribed region. Based on our preliminary results, we propose a model whereby DNA hypermethylation and hypomethylation are connected by a mechanism of transcriptional overlap (**Figure 9**). The demethylation of the *CT-GABRA3* induces its transcriptional activation. Then, the transcriptional machinery including RNA PolII, and the H3K36 methyltransferase SETD2 (Edmunds et al. 2008), catalyzes H3K36me3 marks throughout the transcribed region. These marks subsequently recruit the *de novo* DNA methyltransferases (DNMT3s) to methylate the region including the *BT-GABRA3* promoter (Dhayalan et al. 2010; Baubec et al. 2015).

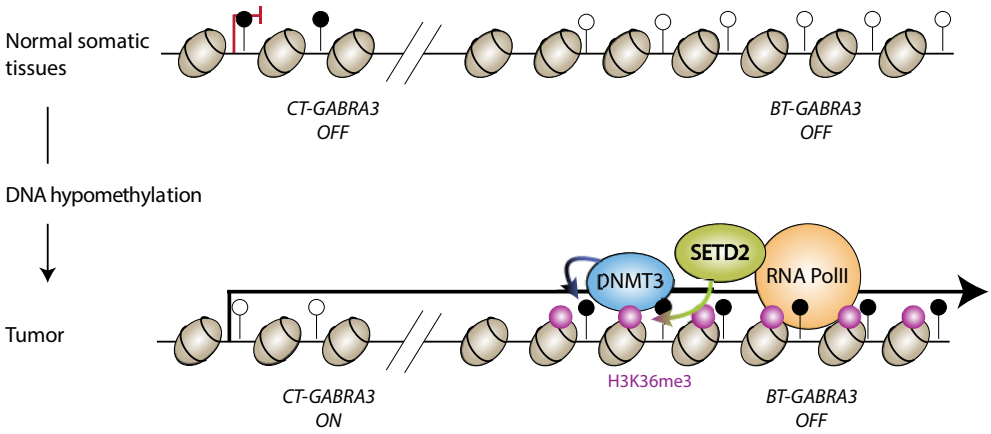


Figure 9: Proposed mechanism of activation of *CT-GABRA3* by hypomethylation creates a methylation-dependent chromatin environment on the downstream *BT-GABRA3* promoter. Top panel: in normal somatic cells, *CT-GABRA3* and *BT-GABRA3* transcripts are both silenced. *CT-GABRA3* displays a hypermethylated promoter and *BT-GABRA3* displays a non-methylated promoter. Lower panel: in cancer, DNA hypomethylation induces *CT-GABRA3* transcription through the promoter of *BT-GABRA3* resulting in *BT-GABRA3* promoter hypermethylation. Probably, during the transcription of *CT-GABRA3*, the interaction between RNA PolII and SETD2 methyltransferase induces tri-methylation of H3K36. Then, H3K36me3 would attract DNA methyltransferases (DNMT3), which would subsequently induce DNA hypermethylation.

Further experimentation is required to confirm our model. First, we would like to demonstrate in an *in vitro* model that experimental activation of *CT-GABRA3* leads to subsequent DNA hypermethylation of *BT-GABRA3*. Such type of experiment is being conducted in a melanoma cell line (LB2667), where we observed that experimental activation of *CT-GABRA3*, following transient exposure to 5-azadC, leads after a few days to hypermethylation of the downstream *BT-GABRA3* promoter. The results are however difficult to reproduce, suggesting a very weak effect. Challenges could come from several problems such as the use of an inconvenient cell line model, which lacks necessary factors that we still have to identify.

ChIP experiments against H3K36me3 mark should also be performed to confirm that the transcriptional activation of *CT-GABRA3* directs the deposition of H3K36me3 mark along its transcribed region. This analysis could be performed in different melanoma or lung cancer cell lines that express or not *CT-GABRA3*. An enrichment of H3K36me3 mark on the *BT-GABRA3* promoter in cell lines expressing *CT-GABRA3* would strengthen the *in silico* analysis. Moreover, an

enrichment of H3K36me3 on the *BT-GABRA3* promoter in melanoma cell lines displaying a similar methylation profile to that of normal somatic cells after 5-azadC treatmentswould provide evidence of a relation between *CT-GABRA3* transcription and H3K36 trimethylation on the *BT-GABRA3* promoter.

The fact that *BT-GABRA3* becomes sensitive to DNA demethylation after it becomes hypermethylated suggests a switch of repression mechanisms of the promoter. Hypermethylation could induce this switch by inhibiting the binding of the repressors initially involved in *BT-GABRA3* silencing. These repressors remain however to be characterized.

We anticipate that the model we propose here, linking *CT-GABRA3* hypomethylation and *BT-GABRA3* hypermethylation via transcriptional overlap, applies to other parts of the genome and contributes more largely to redefine the DNA methylation landscape in tumor cells. Identifying other genomic loci that undergo a similar process of epigenetic alteration in tumors also represents an important perspective.

4 Material and methods

Cell lines and tumor tissue samples

All human melanoma cell lines MZ2-MEL, Mi13443, SK-MEL-23, LB39, Mi1811 and LB2667 which derive from cutaneous melanoma metastases were obtained from the Brussels Branch of the Ludwig Institute for Cancer Research, and were cultured as previously described. Early passage human normal epidermal melanocytes (HNEM) were received from E. De Plaen (Ludwig Institute for Cancer Research, Belgium), and were cultured in Ham's F10 medium (Life Sciences) supplemented with 6 mM Hepes, 1 x MelanoMax supplement (Gentaur), and 10% fetal bovine serum. RNA from human melanoma and NSCLC tissue samples used were obtained from the Brussels Branch of the Ludwig Institute for Cancer Research. Genomic DNAs used for bisulfite analysis from melanoma samples were obtained from Donata Rimoldi from the Lausanne Branch of the Ludwig Institute for Cancer Research.

Isolation of clones

To isolate clones showing homogenous genetic and epigenetic background, the cells were plated at limiting dilutions for an average density of 1,5, 0,5 or 0,15 cells/well into 96-well plates. After four weeks each well showing one colony likely to derive from a single clone was chosen to be amplified for further analysis.

5-Aza-2′-deoxycytidine treatment

Cells grown to 60–70% confluence were exposed to a single dose of 2 μM 5-aza-2′-deoxycytidine (Sigma-Aldrich), and maintained in culture during 6 d (a period of time corresponding to at least 3 population doublings) before RNA extraction. After treatment, LB2667.3 cells were cultured a subsequent ten days without treatment before cloning or 20-45 days before bisulfite analysis.

DNA isolation and bisulfite modification

DNA was isolated by phenol/chloroform extraction method and 1µg of genomic DNA was treated with sodium bisulfite.

Sodium bisulfite genomic sequencing

Sodium bisulfite genomic sequencing of the *MAGEA3*/*CT-GABRA3* and *BT-GABRA3* promoter regions were performed as described previously. Primer used for *MAGEA3*/*CT-GABRA3* nested PCR amplification of bisulfite treated DNA were respectively TYGATTTTTATTTAGGTAGAATTT and TAAAATAATAACRACCCAACCTAA (1st PCR) and ATTTAGGTAGAATTTAGTTTTAT and CCCTACRAAATAACCCAAA (2nd PCR). These primer sets also amplify the *MAGEA6* promoter region, which shows 100% sequence identity. Primer used for *BT-GABRA3* nested PCR amplification of bisulfite treated DNA were respectively AGTGGTGAATTTAAAGTTAGTAAAGG and CCCCAATATCTCCCTACTCAAAT (1st PCR) and GATAGAGAGGGAGGGAGGTAGA and CTAAACTTCCACCAACCCCACT (2nd PCR).

RT-PCR and RT-qPCR analyses

Total RNA samples were purchased from Ambion Life Technologies, or prepared from cell lines and surgical specimens (obtained from the Ludwig Institute for Cancer Research, Belgium) using either TriPure Isolation Reagent (Roche Diagnostics GmbH) or the guanidinium-isothiocyanate/cesium chloride procedure. Reverse transcription was performed on 2 µg of total RNA using M-MLV Reverse transcriptase (Invitrogen) and random hexamers primers or dT18 primers. Conventional PCRs were performed using the DreamTaq polymerase (Thermo Fisher Scientific), in a final reaction volume of 20 µl. Quantitative RT-PCR amplifications were performed using the qPCR Core kit (Eurogentec). All qPCRs where SybrGreen assays. Primers used for the expression of *CT-GABRA3* were CATCATGCCATGTCTGCCGAAA and GGAGGCGGAGATTGCACA, of *BT-GABRA3* were CATCATGCCATGTCTGCCGAAA and GGGCTCAACCTCCAACTT and of *ACTIN-beta* were GGCATCGTGATGGACTCCG and GCTGGAAGGTGGACAGCGA.

Quantitative methylation-specific PCR (qMSP)

Bisulfite-treated DNA was amplified using *CT-GABRA3* or *BT-GABRA3* promoter specific primers. A non-methylation-specific PCR was first performed and a quantitative methylation-specific PCR was subsequently performed on it. The primers used for the primary amplification of *CT-GABRA3* are TYGATTTTTATTTAGGTAGAATTT and TAAAATAATAACRACCCAACCTAA (55°C of annealing during 36 cycles) and of *BT-GABRA3* are CCCCAATATCTCCCTACTCAAAT and AGTGGTGAATTTAAAGTTAGTAAAGG (60°C of annealing during 36 cycles). Primers GTTGAGGGAGGATTGAGGC with CTACGAAATAACCCAAACCCG and GTTGAGGGAGGATTGAGGT with CTACAAAATAACCCAAACCCA are used to quantitatively amplify the methylated and unmethylated forms of *CT-GABRA3* promoter respectively. Similarly, primers AGCGAGAGAGCGTGAGCGC with TCTCTCCTAATCCCTCTCTCG and AGTGAGAGAGTGTGAGTGT with TCTCTCCTAATCCCTCTCTCA are used to quantitatively amplify the methylated and unmethylated forms of *BT-GABRA3* respectively.

Relative methylation level was calculated by the ratio of copies of methylated (MSM) or unmethylated (MSU) on the total of the copies of methylated and unmethylated region (MSM+MSU) multiplied by 100.

5 References

Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schubeler D. 2015. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**: 243-247.

Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A. 2010. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem* **285**: 26114-26120.

Di Croce L, Raker VA, Corsaro M, Fazi F, Fanelli M, Faretta M, Fuks F, Lo Coco F, Kouzarides T, Nervi C et al. 2002. Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* **295**: 1079-1082.

Edmunds JW, Mahadevan LC, Clayton AL. 2008. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J* **27**: 406-420.

Gardiner-Garden M, Frommer M. 1987. CpG islands in vertebrate genomes. J Mol Biol 196: 261-282.

Loriot A, Van Tongelen A, Blanco J, Klaessens S, Cannuyer J, van Baren N, Decottignies A, De Smet C. 2014. A novel cancer-germline transcript carrying pro-metastatic miR-105 and TET-targeting miR-767 induced by DNA hypomethylation in tumors. *Epigenetics* **9**: 1163-1171.

Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE 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.

Thienpont B, Steinbacher J, Zhao H, D'Anna F, Kuchnio A, Ploumakis A, Ghesquiere B, Van Dyck L, Boeckx B, Schoonjans L et al. 2016. Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature* **537**: 63-68.

Thillainadesan G, Chitilian JM, Isovic M, Ablack JN, Mymryk JS, Tini M, Torchia J. 2012. TGF-beta-dependent active demethylation and expression of the p15ink4b tumor suppressor are impaired by the ZNF217/CoREST complex. *Mol Cell* **46**: 636-649.

Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RG, Otte AP et al. 2002. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**: 624-629.

DISCUSSION

As already mentioned, cancer cells display profound epigenetic alterations with local hypermethylation and global hypomethylation of DNA methylation. Through the study of the *GABRA3* locus we tried to extend our knowledge about the causes and consequences of these two types of epigenetic alterations.

One effect of DNA hypomethylation could be the activation of genes with oncogenic function. In cancer cells, loss of DNA methylation leads to the activation of a small group of genes called "cancer-germline" (CG). As we observed for *CT-GABRA3*, some of these CG genes appear to have oncogenic functions in cancer. In the case of *CT-GABRA3*, the oncogenic function is mediated through the synthesis of two microRNAs: one (miR-105) that participates to metastasis, and the other (miR-767) that impacts on epigenetic processes.

During the study of the *GABRA3* locus, we noted a striking correlation between DNA hypomethylation of the *CT-GABRA3* promoter and DNA hypermethylation of the *BT-BABRA3* promoter. This observation suggests an epigenetic crosstalk between these two promoters.

The following discussion summarizes the results obtained during the thesis and develops the possible additional functions of miR-105 and miR-767. Moreover, mechanisms leading to the hypermethylation of the *BT-GABRA3* promoter and the potential role of *CT-GABRA3* transcription in this mechanism will also be discussed.

A) Roles of CG miRNAs in tumor

1 Functions of miR-105 and miR-767 in cancer

One role of miR-105 in tumorigenesis has already been identified, as Zhou *et al*. (Zhou et al. 2014) reported that miR-105 exerts an oncogenic function by promoting metastasis. The authors showed that overexpression of miR-105 in non-metastatic breast cancer cells induces metastasis in distant organs and vascular permeability *in vivo*. MiR-105 seems to exert its function by repressing the tight-junction protein 1 (TJP1). Exosome-mediated transfer of cancer-secreted miR-105 in endothelial cells disrupts the endothelial barrier. Besides its role in metastasis, miR-105 has been described as a tumor suppressor microRNA through inhibition of cell proliferation in ovarian, prostate and liver cancer (Sirotkin et al. 2010; Honeywell et al. 2013; Shen et al. 2014). In hepatocellular carcinoma (HCC) cells, miR-105 overexpression experiments in a xenograft mouse model showed a decrease of tumor size whereas an increase of tumor size could be observed when miR-105 is inhibited (Shen et al. 2014). The authors suggest that the effect of miR-105 on cell proliferation would be mediated by an inhibition of the PI3K/Akt signaling pathway. Careful examination of transcriptiomic data indicate, however, that *GABRA3* transcripts (and therefore miR-105) are not present in normal liver tissues. This is of course incompatible with a role of miR-105 as a tumor suppressor. Nevertheless, it cannot be excluded that miR-105 can be bivalent, and that its function might depend on tumor-type.

Whereas the putative function of miR-105 is known, the function of miR-767 has never been studied except in our lab (Loriot et al. 2014). Mir-767 overexpression and inhibition studies revealed that miR-767 induces repression of the TET DNA demethylase. Overexpression of miR-767 also led to decreased levels of 5hmC, the oxidative product of TET enzymes. In this manner, DNA demethylation could lead to downregulation of TETs and induce subsequent epigenetic remodeling events. Downregulation of TET by miR-767 may facilitate DNA hypermethylation at specific gene promoters. This effect could also affect

the *BT-GABRA3* promoter and could potentially explain the hypermethytion of this region when *CT-GABRA3* is expressed.

TET enzymes are described as tumor suppressor genes through their effect on cell differentiation and invasion. The synergic expression of miR-105 and miR-767 could participate to tumor progression by enhancing metastasis, but further investigations are needed. It is known that miRNAs target multiple mRNAs, and that they might have more than one function. Thus, it is likely that miR-105 and miR-767 exert additional functions.

2 Research of additional functions?

Until now, miR-105 and miR-767 functions were studied independently from each other, although they are always co-expressed. To investigate a possible synergic function of miR-105 and miR-767, we developed a conditional knockout (cKO) model for these two miRNAs. This model consists in the addition of loxP sites on each side of the miR-105/767 locus in the HT1080 fibrosarcoma cell line. Transfection of these cells with a Cre recombinase combined with a fluorescent protein, and subsequent fluorescent cell sorting induces loss of miR-105 and miR-767 with 94% efficiency. The construction of this model by the CRISPR/Cas9 technique paired with homologous recombination resulted in the publication of a technical article (Van Tongelen et al. 2015) & (Annex 1). With this model we investigated the role of miR-105 and miR-767 on cell proliferation, migration and anchorage-independent growth. These functional tests showed that inhibition of these two miRNAs in HT1080 cells does not result in detectable phenotype modifications. Besides, we did not detect an effect on the expression of the respective targets TJP1 and TET1 either. The weak initial expression level of miR-105 and miR-767 in HT1080 could explain the results we obtained. Moreover, this cell line derives from another tumor type than those where the functions of these miRNAs were investigated (melanoma, breast, and lung cancer). The impact of miR-105 and miR-767 could vary according to tumor-type.

To better investigate the function of miR-105 and miR-767, a *miR-105/767* cKO model should be tested in other melanoma or lung cancer cell

lines. We tried to carry out conditional KO (cKO) in lung carcinoma and melanoma cell lines. However, we did not succeed to create a cKO in these cell lines.

3 Possible effect on tumor microenvironment?

Studies on miR-105 and other miRNAs, such as miR-210, revealed that their function involves a crosstalk between cancer cells and stromal cells (Kosaka et al. 2013; Zhou et al. 2014). We can imagine that a similar action could also be extended to miR-767. The intercellular communication processes between stromal cells and cancer cells can be mediated by exosomes, which are small vesicles secreted by eukaryotic cells into the extracellular environment. Some experiments in the lab allowed us to confirm that not only miR-105, but also miR-767, is present in exosomes emanating from tumor cells that express both miRNAs. Via exosomes, miR-105 and miR-767 could deregulate the function of stromal cells, such as endothelial cells, fibroblasts or lymphocytes. Cancer-associated fibroblasts (CAF) are a subpopulation of fibroblasts that reside within the tumor microenvironment, and act as critical actors to support tumor growth (Kalluri and Zeisberg 2006; Franco et al. 2010; Pietras and Ostman 2010). DNA methylation alterations were reported in these cells, including widespread DNA hypomethylation (Jiang et al. 2008) with concomitant focal gains of DNA methylation compared to normal fibroblasts. The hypermethylation of certain genes in CAF has been associated with gene repression and the pro-invasive phenotype of CAF (Albrengues et al. 2015; Vizoso et al. 2015). If downregulation of TET activity is responsible for DNA hypermethylation of certain genes (Thienpont et al. 2016), the presence of miR-767 and the subsequent downregulation of TET could participate in the DNA hypermethylation process in CAF. According to this hypothesis, co-culture experiments of fibroblasts and tumor cells expressing miR-767 might show an increase of DNA methylation of specific genes in fibroblasts, and promote their transformation into CAFs.

4 Additional miR-105 and miR-767 targets?

The physiological role of miRNAs can be investigated through the cellular pathway in which they are involved. They can also be determined by their specific targeted genes. Indeed, one of the ongoing projects in the lab aims at investigating other genes targeted by miR-105 and miR-767. In order to do this, a multistep procedure has been carried out. First, a list of potential target genes was generated using prediction algorithms. Secondly, RNA sequencing data of melanoma cell lines was used to identify among the predicted target genes those whose expression is inversely correlated with the expression of *GABRA3.* Validation and functional characterization of selected candidate genes are underway.

5 Additional cancer-germline miRNAs?

The initial objective of the thesis was to search for miRNAs with cancer-germline (CG) characteristics, and to test whether they have oncogenic functions. In the first article of this thesis (Loriot et al. 2014), we identified miR-105 and miR-767 as CG miRNAs with potential oncogenic functions. These miRNAs are located in the newly identified CG gene *CT-GABRA3*. Other intragenic miRNAs could be located in known CG genes. However, through an analysis of RNA-seq databases, we did not identify any other miRNA in a known GC gene. The study of the *GABRA3* locus revealed that some loci not referenced as CG genes could nevertheless harbor alternative transcripts with a CG pattern of expression. By crossing data from RNA-seq and transcription start site (TSS) sequencing, new CG genes from alternative transcripts could be identified, and we could detect new CG miRNAs in these CG genes.

CG-miRNAs could also exist outside of gene units. Such type of intergenic CG miRNA was recently identified (miR-888), and its expression was shown to correlate with aggressiveness of endometrial cancers (Hovey et al. 2015). The authors identified the gene encoding the progesterone receptor (PR) as a direct target of miR-888. Consistently, PR has been described as a potent suppressor of endometrial tumor.

6 What about the GABRA3 protein?

The presence of *GABRA3* transcripts have been previously associated with poor survival in breast and lung cancer patients (Lu et al. 2006; Liu et al. 2009; Zhang et al. 2013; Gumireddy et al. 2016; Liu et al. 2016). The studies, however, did not permit to distinguish between *CT-GABRA3* and *BT-GABRA3* transcripts. It is therefore unclear whether this correlation should be ascribed to the presence of the GABRA3 protein, which is solely produced from *BT-GABRA3* mRNAs, or to that of miRNAs, which are generated from both *CT-* and *BT-GABRA3* transcripts.

Very recent studies demonstrated that the GABRA3 protein promotes metastasis by increasing the expression of matrix metalloproteinases (MMPs) via the AKT pathway in lung and breast cancer (Gumireddy et al. 2016; Liu et al. 2016). MMPs are secreted proteins that are able to degrade extracellular matrix (ECM) proteins such as collagen, but also proteins involved in cell-cell junctions. Their upregulation in cancer is associated with pro-metastatic functions (Stamenkovic 2000; Deryugina and Quigley 2006). It is therefore possible that the oncogenic function of the GABRA3 locus results from a synergic action of miR-105, miR-767 and the GABRA3 protein.

Expression analysis of the different variants of *GABRA3* indicates that in melanoma and lung cancer, it is predominantly the *CT-GABRA3* transcript that is activated rather than *BT-GABRA3* (Loriot et al. 2014). Therefore we suggest that in melanoma and lung cancer, the oncogenic function of the *GABRA3* locus is primarily mediated by the two miRNAs.

B) An inverse correlation between methylation of *CT-GABRA3* and *BT-GABRA3* promoters in tumors.

In normal somatic cells, the *CT-GABRA3* (upstream) promoter is methylated while the *BT-GABRA3* (downstream) promoter is unmethylated. Intriguingly, we observed that in cancer cells, hypomethylation of *CT-GABRA3* is often associated with hypermethylation of *BT-GABRA3*. This observation reveals a new potential link between the process of DNA hypomethylation and DNA hypermethylation in tumors, which have until now been described as the result of independent mechanisms.

We showed that hypomethylation of the *CT-GABRA3* promoter is associated with its transcription in tumors. Therefore, we can suppose that *BT-GABRA3* hypermethylation is linked with *CT-GABRA3* transcription. By our observations, we might have discovered a new mechanism that induces DNA hypermethylation in cancer and potentially in normal tissues.

In silico analyses indicated that *CT-GABRA3* activation is associated with deposition of the histone mark H3K36me3 all over the transcribed region, which includes the *BT-GABRA3* promoter. As several studies demonstrated a link between H3K36me3 and *de novo* DNA methylation (Dhayalan et al. 2010; Hahn et al. 2011; Baubec et al. 2015), we suggest that the mechanism directing DNA methylation in the *BT-GABRA3* promoter involves a process of transcriptional overlap.

Overlapping transcription can induce SETD2-mediated establishment of H3K36me3 histone modification on the overlapped promoter and reduce the recruitment of transcription factors (van Werven et al. 2012). This epigenetic mark is also known to recruit DNMT3A/B which causes *de novo* DNA methylation in the bodies of transcribed genes (Dhayalan et al. 2010; Baubec et al. 2015). According to this observation the deposition of the H3K36me3 histone mark on the entire *CT-GABRA3* transcribed region, which includes *BT-GABRA3* promoter, might recruit DNMT3A/B along the transcribed region and their recruitment would promote the hypermethylation of the region. To summarize, we suggest

that the hypermethylation of *BT-GABRA3* could be orchestrated by the interaction between RNA Pol II, the histone methyltransferase SETD2 and DNMT3s (Figure 16).

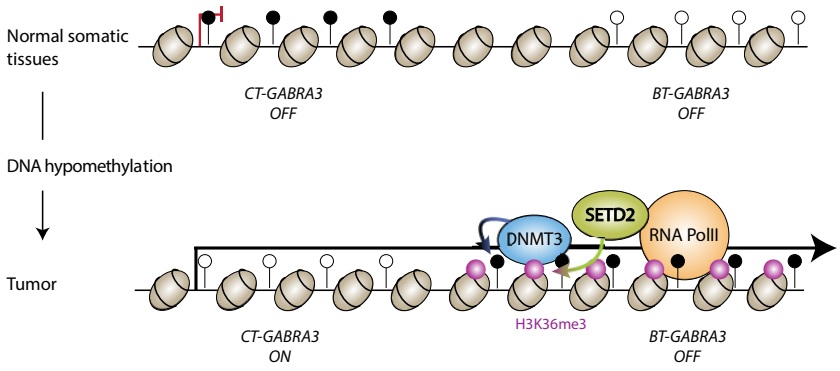


Figure 16: Proposed mechanism of activation of *CT-GABRA3* by hypomethylation which creates a methylation-dependent chromatin environment on the downstream *BT-GABRA3* promoter. Top panel: In normal somatic cells, *CT-GABRA3* and *BT-GABRA3* transcripts are both silenced. *CT-GABRA3* displays hypermethylated promoter and *BT-GABRA3* displays a non methylated promoter. Lower panel: In cancer, DNA hypomethylation induces *CT-GABRA3* transcription through the promoter of *BT-GABRA3* resulting in *BT-GABRA3* promoter hypermethylation. Probably, during the transcription of *CT-GABRA3*, the interaction between RNA PolII and SETD2 methyltransferase induces tri-methylation of H3K36. Then, H3K36me3 would attract DNA methyltransferases (DNMT3), which would induce DNA hypermethylation.

We could however imagine another hypothesis based on a possible recruitment of H3K36 histone-methyltransferase by the *CT-GABRA3* transcript. It is well known that lncRNA can induce hypermethylation *in cis.* The best example is the *Xist* lncRNA which promotes the silencing of the whole X chromosome. This RNA is able to interact with polycomb protein complexes which leads to chromatin rearrangement, DNA hypermethylation and exclusion of the transcription machinery (Wutz 2011).

These correlation-based evidences have however to be validated by further *in vitro* experiments. For instance, in cell lines treated with 5-azadC, we would expect to see *CT-GABRA3* activation together with H3K36me3 deposition and DNA hypermethylation within the *BT-GABRA3* promoter.

7 Overlapping transcription and promoter methylation

Several examples of promoter hypermethylation subsequent to transcriptional overlap have already been described in the context of human diseases. First, in α-thalassemia patients, the *LUC7L/HBA2* locus undergoes chromosomal deletion (Figure 17A). In healthy cells, these two genes are expressed in an opposed direction, but do not overlap thanks to the presence of a transcriptional stop (polyadenylation) signal. In α-thalassemia patients, a chromosomal deletion of *LUC7L* transcriptional stop signal causes transcription of *LUC7L* gene to proceed further through the promoter of *HBA2*. This causes DNA hypermethylation and silencing of HBA2 promoter. The mechanism leading to promoter hypermethylation remains unknown (Tufarelli et al. 2003; Kornienko et al. 2013). A second example is illustrated by a chromosomal deletion in a subset of Lynch syndrome patients (Figure 17B). In these patients, DNA methylation and inactivation of the *MSH2* (mutS homolog 2) gene correlate with transcriptional expression from the neighboring *EPCAM* gene when the transcription stop signal is deleted (Ligtenberg et al. 2009). The deletion that targets only one of the two alleles, leads to DNA hypermethylation of the corresponding allele. Therefore, even if the mechanism is unknown, it seems that it depends on the process of transcription in *cis* rather than the transcriptional product. These two examples illustrate how transcriptional overlap can contribute to aberrant promoter methylation in disease.

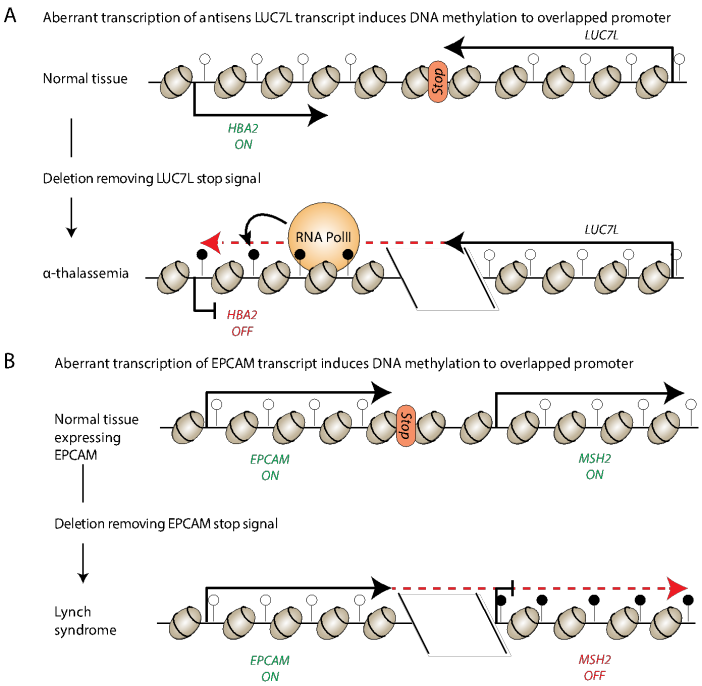


Figure 17: Mechanisms described in the literature showing a hypermethylation event associated with aberrant transcriptional overlap. (A) In normal cells, LUC7L and HBA2 are both expressed, do not overlap and display hypomethylated promoters. In α-thalassemia patients, chromosomal deletion of the LUC7L transcriptional stop signal causes transcription through the promoter of HBA2. This aberrant transcription causes DNA methylation and silencing of HBA2 promoter. (B) In healthy human, EPCAM and MSH2 are both expressed, do not overlap and display hypomethylated promoters. In Lynch syndrome, chromosomal deletion of the EPCAM stop signal induces transcription through the promoter of MSH2. This aberrant transcription causes DNA hypermethylation and silencing of MSH2 promoter.

Transcriptional overlap is also an essential component for the establishment of DNA methylation in imprinted regions during germline development (Smallwood and Kelsey 2012). Supporting evidence for this concept was provided by the observation that truncating transcription from an upstream oocyte-specific promoter of the *Gnas* locus causes loss of germline methylation at the downstream maternal germline gDMR (Chotalia et al. 2009).

8 Is transcriptional overlap a general cause of DNA hypermethylation in cancer?

Among the epigenetic alteration associated with tumor development, DNA hypermethylation is the most studied. It was shown indeed that this epigenetic alteration affects on average a few hundred genes in a tumor (Rauch et al. 2008), and that several among these genes have a tumor suppressor function (Esteller 2008). DNA hypermethylation has therefore a well-established pro-tumoral role. Based on this finding, epigenetic cancer therapies aimed at re-activating hypermethylated tumor suppressor genes are being developed. These are essentially based on inhibitors of DNA methyltransferases (mainly DNMT1). Nevertheless, points of concern remain regarding the efficacy and especially the specificity of this type of treatment (Juo et al. 2015). It is obvious that improvements in this therapeutic approach will necessitate a more precise knowledge of the mechanisms that induce DNA hypermethylation in tumors.

Two mechanisms have classically been proposed to explain the process of DNA hypermethylation in tumors. The first involves abnormal activation of transcriptional repressors, which induce hypermethylation of target promoters by favoring local recruitment of DNMTs (Di Croce et al. 2002). The second mechanism suggests that promoters that are initially tagged with the H3K27me3 histone modification, and therefore targeted by the polycomb repressor complex, constitute privileged chromatin regions for recruitment of DNMTs (Schlesinger et al. 2007). However, it can be estimated that these two mechanisms explain no more than half of the DNA hypermethylation events in tumors. It is therefore likely that other mechanisms of DNA hypermethylation exist.

Our observations on the *GABRA3* locus reveal a novel mechanism of epigenetic alteration in tumors, whereby DNA hypomethylation and hypermethylation are linked via a process of transcriptional overlap. It seems very likely that this novel mechanism takes place not only in the *GABRA3* locus, but also in other parts of the cancer genome, and can contribute to epigenetic repression of genes with tumor suppressor function. Bioinformatic approaches

could be used to seek for other genes for which promoter hypermethylation in tumors results from an overlapping transcription induced by hypomethylation and activation of a neighboring gene. Computational investigation of cancer-derived genomic datasets could take into account criteria such as: neighboring CpG islands showing opposite methylation changes; promoters with H3K36me3 enrichment; CpG-rich promoter-derived transcripts.

9 Epigenetic switch in the hypermethylated BT-GABRA3 promoter

DNA hypermethylation of promoters in cancer cells is commonly considered to cause transcriptional silencing of corresponding genes. Recent evidence revealed however that many of the genes that become hypermethylated in tumors are already repressed in the original tissue (Schlesinger et al. 2007). This is the case for *BT-GABRA3*, which is repressed in most normal tissues, except brain and testis. It is therefore unlikely that hypermethylation of *BT-GABRA3* serves to inhibit tumor suppressor functions.

Interestingly, we observed that, in melanoma cell lines where the *BT-GABRA3* promoter is hypermethylated, *BT-GABRA3* transcription is induced upon treatment with a DNA methylation inhibitor. In cells lacking *BT-GABRA3* hypermethylation, the gene appeared to be insensitive to DNA methylation inhibitors. This suggests that when *CT-GABRA3* activation leads to the *BT-GABRA3* promoter hypermethylation, *BT-GABRA3* undergoes an epigenetic switch whereby it becomes sensitive to DNA methylation (**Figure 18**).

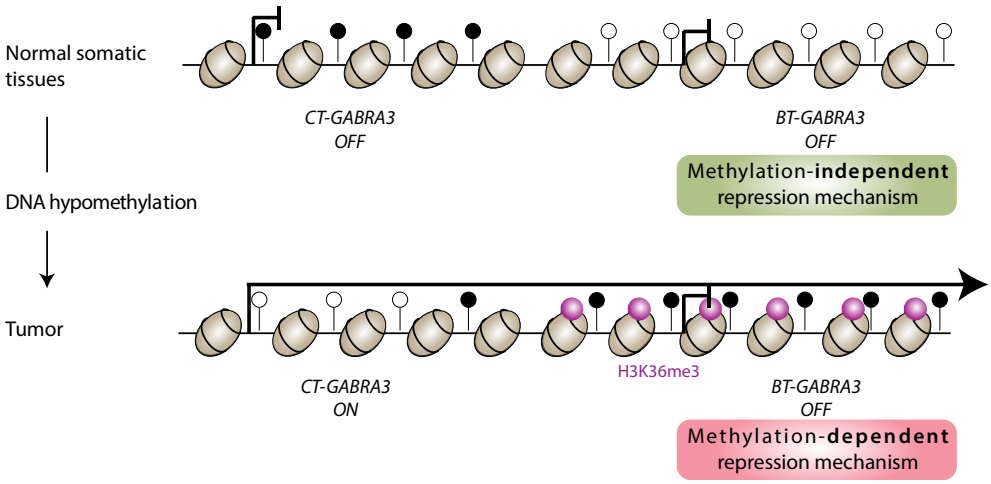


Figure 18: Illustration of the epigenetic switch that *BT-GABRA3* undergoes due to its hypermethylation induced by *CT-GABRA3* transcription: conversion of the *BT-GABRA3* promoter from a methylation-independent repression mechanism to a methylation-dependent one.

In a few melanoma and lung tumor samples, expression of both *CT-GABRA3* and *BT-GABRA3* transcripts is observed. We believe that this may be the result of a two-step process of DNA demethylation. The first demethylation phase would have induced *CT-GABRA3* transcription, and consequently *BT-GABRA3* hypermethylation. Under this configuration, *BT-GABRA3* would have undergone the epigenetic switch that renders it sensitive to DNA methylation. A second phase of demethylation would therefore have led to *BT-GABRA3* activation.

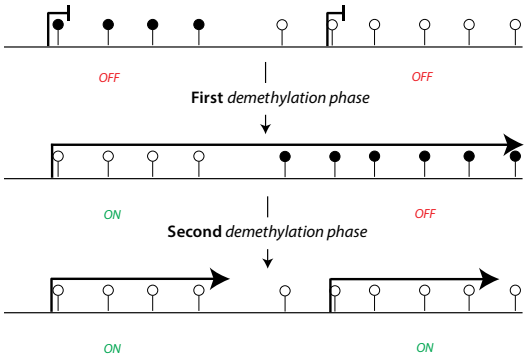


Figure 19: Epigenetic switch and consequences of successive DNA demethylation episodes.

10 What are the potential mechanisms responsible for the epigenetic switch?

We propose that *BT-GABRA3* promoter could be initially repressed by transcriptional repressors, and that binding of these repressors would be lost when the promoter becomes hypermethylated. The hypermethylated *BT-GABRA3* would therefore switch to a purely DNA methylation-dependent mechanism of regulation.

Since *BT-GABRA3* is specifically expressed in brain, repression in the other normal somatic tissues could rely on the transcriptional repressor complex REST/coREST, which is involved in the repression of neuronal genes in non-neuronal cells (Coulson 2005). As REST establishes transcriptional repression by the deposition of histone repressive marks such as H3K9me3 (Gal-Yam et al. 2008), it should be interesting to check the epigenetic landscape of the *BT-GABRA3* promoter in normal brain tissues.

When *BT-GABRA3* is transcribed, as in testis, *CT-GABRA3* transcription does not induce *BT-GABRA3* hypermethylation and silencing. This suggests that the actively transcribed *BT-GABRA3* promoter is protected from DNA hypermethylation (**Figure 20**). It has been reported that distribution of the histone mark H3K4me3 on the genome is inversely correlated with that of DNA methylation (Meissner et al. 2008), H3K4me3 inhibits DNMT activity by impairing the binding of DNMT3L (Ooi et al. 2007; Illingworth and Bird 2009). Protection of the active *BT-GABRA3* promoter could also rely on the presence of transcription–initiation machinery including Pol II, and transcription factors (Gebhard et al. 2010). Besides passive protection, the unmethylated state of *BT-GABRA3* could also be maintained by the presence of demethylating enzymes (TET), which would remove unwanted DNA methylation (Takeshima et al. 2009).

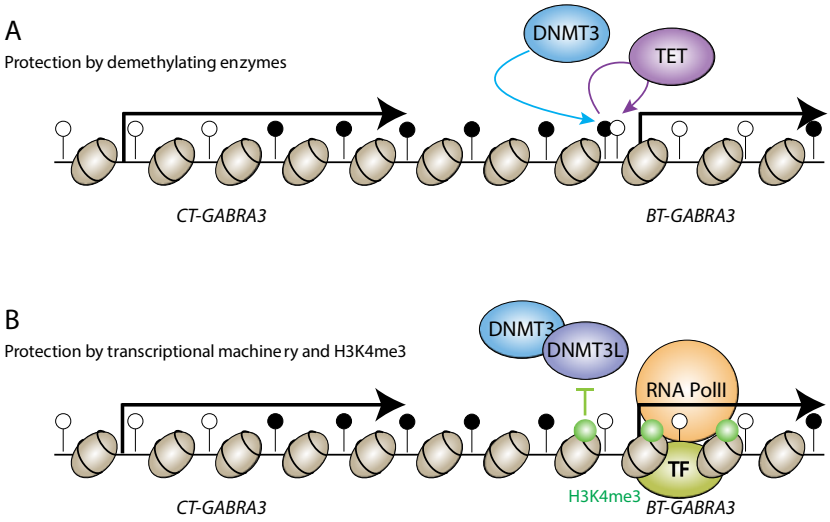


Figure 20: *Potential hypermethylation-protection mechanisms of the transcriptionally active BT-GABRA3 promoter. (A) BT-GABRA3 can be methylated by DNMT3 but the aberrant methylation is oxidized and replaced by unmethylated cytosine due to the presence of TET demethylating enzymes on the BT-GABRA3 promoter. (B) The basal transcriptional machinery (RNA PolII and TF) and histone H3K4me3 exclude the DNMTs from the BT-GABRA3 transcription start site.*

REFERENCES

Agarwal A, Polineni R, Hussein Z, Vigoda I, Bhagat TD, Bhattacharyya S, Maitra A, Verma A. 2012. Role of epigenetic alterations in the pathogenesis of Barrett's esophagus and esophageal adenocarcinoma. *Int J Clin Exp Pathol* **5**: 382-396.

Albrengues J, Bertero T, Grasset E, Bonan S, Maiel M, Bourget I, Philippe C, Herraiz Serrano C, Benamar S, Croce O et al. 2015. Epigenetic switch drives the conversion of fibroblasts into proinvasive cancer-associated fibroblasts. *Nat Commun* **6**: 10204.

Avner P, Heard E. 2001. X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet* **2**: 59-67.

Bachman KE, Rountree MR, Baylin SB. 2001. Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *J Biol Chem* **276**: 32282-32287.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823-837.

Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, Akalin A, Schubeler D. 2015. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**: 243-247.

Baylin SB. 2005. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol* **2 Suppl 1**: S4-11.

Baylin SB, Jones PA. 2011. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer* **11**: 726-734.

Bird AP. 1980. DNA methylation and the frequency of CpG in animal DNA. *Nucleic acids research* **8**: 1499-1504.

Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE. 2007. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**: 1760-1764.

Bourc'his D, Bestor TH. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**: 96-99.

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. 2001. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**: 2536-2539.

Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. 2006. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* **20**: 1123-1136.

Brandeis M, Kafri T, Ariel M, Chaillet JR, McCarrey J, Razin A, Cedar H. 1993. The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. *EMBO J* **12**: 3669-3677.

Brenner C, Deplus R, Didelot C, Loriot A, Vire E, De Smet C, Gutierrez A, Danovi D, Bernard D, Boon T et al. 2005. Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO J* **24**: 336-346.

Brinkman AB, Gu H, Bartels SJ, Zhang Y, Matarese F, Simmer F, Marks H, Bock C, Gnirke A, Meissner A et al. 2012. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res* **22**: 1128-1138.

Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, McCabe VM, Norris DP, Penny GD, Patel D, Rastan S. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351**: 329-331.

Brown SJ, Stoilov P, Xing Y. 2012. Chromatin and epigenetic regulation of pre-mRNA processing. *Hum Mol Genet* **21**: R90-96.

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581-592.

Chen S, Ma J, Wu F, Xiong LJ, Ma H, Xu W, Lv R, Li X, Villen J, Gygi SP et al. 2012. The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. *Genes Dev* **26**: 1364-1375.

Choi JD, Lee JS. 2013. Interplay between Epigenetics and Genetics in Cancer. *Genomics Inform* **11**: 164-173.

Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, James K, Dean W, Kelsey G. 2009. Transcription is required for

establishment of germline methylation marks at imprinted genes. *Genes Dev* **23**: 105-117.

Costello JF, Plass C. 2001. Methylation matters. *J Med Genet* **38**: 285-303.

Coulson JM. 2005. Transcriptional regulation: cancer, neurons and the REST. *Curr Biol* **15**: R665-668.

D'Urso A, Brickner JH. 2014. Mechanisms of epigenetic memory. *Trends Genet* **30**: 230-236.

De Smet C, Lurquin C, Lethe B, Martelange V, Boon T. 1999. DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol Cell Biol* **19**: 7327-7335.

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

Deryugina EI, Quigley JP. 2006. Matrix metalloproteinases and tumor metastasis. *Cancer Metastasis Rev* **25**: 9-34.

Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, Jeltsch A. 2010. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem* **285**: 26114-26120.

Di Croce L, Raker VA, Corsaro M, Fazi F, Fanelli M, Faretta M, Fuks F, Lo Coco F, Kouzarides T, Nervi C et al. 2002. Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* **295**: 1079-1082.

Edmunds JW, Mahadevan LC, Clayton AL. 2008. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J* **27**: 406-420.

Ehrlich M. 2002. DNA methylation in cancer: too much, but also too little. *Oncogene* **21**: 5400-5413.

-. 2006. Cancer-linked DNA hypomethylation and its relationship to hypermethylation. *Curr Top Microbiol Immunol* **310**: 251-274.

Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C. 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* **10**: 2709-2721.

Ehrlich M, Jiang G, Fiala E, Dome JS, Yu MC, Long TI, Youn B, Sohn OS, Widschwendter M, Tomlinson GE et al. 2002. Hypomethylation and hypermethylation of DNA in Wilms tumors. *Oncogene* **21**: 6694-6702.

Esteller M. 2005. Aberrant DNA methylation as a cancer-inducing mechanism. *Annu Rev Pharmacol Toxicol* **45**: 629-656.

-. 2008. Epigenetics in cancer. *N Engl J Med* **358**: 1148-1159.

Esteller M, Fraga MF, Guo M, Garcia-Foncillas J, Hedenfalk I, Godwin AK, Trojan J, Vaurs-Barriere C, Bignon YJ, Ramus S et al. 2001. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum Mol Genet* **10**: 3001-3007.

Feinberg AP, Vogelstein B. 1983. Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun* **111**: 47-54.

Franco OE, Shaw AK, Strand DW, Hayward SW. 2010. Cancer associated fibroblasts in cancer pathogenesis. *Semin Cell Dev Biol* **21**: 33-39.

Fuks F, Burgers WA, Godin N, Kasai M, Kouzarides T. 2001. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J* **20**: 2536-2544.

Gama-Sosa MA, Slagel VA, Trewyn RW, Oxenhandler R, Kuo KC, Gehrke CW, Ehrlich M. 1983. The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res* **11**: 6883-6894.

Gardiner-Garden M, Frommer M. 1987. CpG islands in vertebrate genomes. *J Mol Biol* **196**: 261-282.

Gaudet F, Hodgson JG, Eden A, Jackson-Grusby L, Dausman J, Gray JW, Leonhardt H, Jaenisch R. 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.

Globisch D, Munzel M, Muller M, Michalakis S, Wagner M, Koch S, Bruckl T, Biel M, Carell T. 2010. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One* **5**: e15367.

Goll MG, Bestor TH. 2005. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74**: 481-514.

Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, Golic KG, Jacobsen SE, Bestor TH. 2006. Methylation of tRNAAsp by the DNA

methyltransferase homolog Dnmt2. *Science* **311**: 395-398.

Gruenbaum Y, Stein R, Cedar H, Razin A. 1981. Methylation of CpG sequences in eukaryotic DNA. *FEBS letters* **124**: 67-71.

Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG et al. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**: 606-610.

Gumireddy K, Li A, Kossenkov AV, Sakurai M, Yan J, Li Y, Xu H, Wang J, Zhang PJ, Zhang L et al. 2016. The mRNA-edited form of GABRA3 suppresses GABRA3-mediated Akt activation and breast cancer metastasis. *Nat Commun* **7**: 10715.

Guo JU, Su Y, Zhong C, Ming GL, Song H. 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**: 423-434.

Hackett JA, Reddington JP, Nestor CE, Dunican DS, Branco MR, Reichmann J, Reik W, Surani MA, Adams IR, Meehan RR. 2012. Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. *Development* **139**: 3623-3632.

Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, Surani MA. 2013. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**: 448-452.

Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. 2011. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. *PLoS One* **6**: e18844.

Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X, 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, 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 YF, Li BZ, 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.

Heard E, Disteche CM. 2006. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* **20**: 1848-1867.

Herman JG, Baylin SB. 2003. Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med* **349**: 2042-2054.

Hermann A, Goyal R, Jeltsch A. 2004. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* **279**: 48350-48359.

Holliday R, Grigg GW. 1993. DNA methylation and mutation. *Mutation research* **285**: 61-67.

Honeywell DR, Cabrita MA, Zhao H, Dimitroulakos J, Addison CL. 2013. miR-105 inhibits prostate tumour growth by suppressing CDK6 levels. *PLoS One* **8**: e70515.

Hovey AM, Devor EJ, Breheny PJ, Mott SL, Dai D, Thiel KW, Leslie KK. 2015. miR-888: A Novel Cancer-Testis Antigen that Targets the Progesterone Receptor in Endometrial Cancer. *Transl Oncol* **8**: 85-96.

Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM, Chaillet JR. 2001. Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104**: 829-838.

Huff JT, Plocik AM, Guthrie C, Yamamoto KR. 2010. Reciprocal intronic and exonic histone modification regions in humans. *Nat Struct Mol Biol* **17**: 1495-1499.

Iida T, Suetake I, Tajima S, Morioka H, Ohta S, Obuse C, Tsurimoto T. 2002. PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. *Genes Cells* **7**: 997-1007.

Illingworth RS, Bird AP. 2009. CpG islands--'a rough guide'. *FEBS Lett* **583**: 1713-1720.

Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr AR, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R, Bird AP. 2010. Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet* **6**: e1001134.

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**: 1129-1133.

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**: 1300-1303.

Jiang L, Gonda TA, Gamble MV, Salas M, Seshan V, Tu S, Twaddell WS, Hegyi P, Lazar G, Steele I et al. 2008. Global hypomethylation of genomic DNA in cancer-associated myofibroblasts. *Cancer Res* **68**: 9900-9908.

Jiang N, Wang L, Chen J, Wang L, Leach L, Luo Z. 2014. Conserved and divergent patterns of DNA methylation in higher vertebrates. *Genome Biol Evol* **6**: 2998-3014.

Jin Z, Cheng Y, Gu W, Zheng Y, Sato F, Mori Y, Olaru AV, Paun BC, Yang J, Kan T et al. 2009. A multicenter, double-blinded validation study of methylation biomarkers for progression prediction in Barrett's esophagus. *Cancer Res* **69**: 4112-4115.

Jones PA. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**: 484-492.

Jones PA, Baylin SB. 2007. The epigenomics of cancer. *Cell* **128**: 683-692.

Jones PA, Takai D. 2001. The role of DNA methylation in mammalian epigenetics. *Science* **293**: 1068-1070.

Juo YY, Gong XJ, Mishra A, Cui X, Baylin SB, Azad NS, Ahuja N. 2015. Epigenetic therapy for solid tumors: from bench science to clinical trials. *Epigenomics* **7**: 215-235.

Kafri T, Ariel M, Brandeis M, Shemer R, Urven L, McCarrey J, Cedar H, Razin A. 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev* **6**: 705-714.

Kalluri R, Zeisberg M. 2006. Fibroblasts in cancer. *Nat Rev Cancer* **6**: 392-401.

Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G et al. 2011. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8**: 200-213.

Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. 2009. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* **41**: 376-381.

Kornienko AE, Guenzl PM, Barlow DP, Pauler FM. 2013. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol* **11**: 59.

Kosaka N, Iguchi H, Hagiwara K, Yoshioka Y, Takeshita F, Ochiya T. 2013. Neutral sphingomyelinase 2 (nSMase2)-dependent exosomal transfer of angiogenic microRNAs regulate cancer cell metastasis. *J Biol Chem* **288**: 10849-10859.

Kriaucionis S, Heintz N. 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**: 929-930.

Kushwaha G, Dozmorov M, Wren JD, Qiu J, Shi H, Xu D. 2016. Hypomethylation coordinates antagonistically with hypermethylation in cancer development: a case study of leukemia. *Hum Genomics* **10 Suppl 2**: 18.

Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* **16**: 2893-2905.

Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**: 1021-1036.

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**: 1192-1200.

Li E, Beard C, Jaenisch R. 1993. Role for DNA methylation in genomic imprinting. *Nature* **366**: 362-365.

Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**: 915-926.

Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, Lee TY, Bodmer D, Hoenselaar E, Hendriks-Cornelissen SJ et al. 2009. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat Genet* **41**: 112-117.

Liu L, Yang C, Shen J, Huang L, Lin W, Tang H, Liang W, Shao W, Zhang H, He J. 2016. GABRA3 promotes lymphatic metastasis in lung adenocarcinoma by mediating upregulation of matrix metalloproteinases. *Oncotarget* **7**: 32341-32350.

Liu Y, Guo F, Dai M, Wang D, Tong Y, Huang J, Hu J, Li G. 2009. Gammaaminobutyric acid A receptor alpha 3 subunit is overexpressed in lung cancer. *Pathol Oncol Res* **15**: 351-358.

Lorincz MC, Dickerson DR, Schmitt M, Groudine M. 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol* **11**: 1068-1075.

Loriot A, Van Tongelen A, Blanco J, Klaessens S, Cannuyer J, van Baren N, Decottignies A, De Smet C. 2014. A novel cancer-germline transcript carrying pro-metastatic miR-105 and TET-targeting miR-767 induced by DNA hypomethylation in tumors. *Epigenetics* **9**: 1163-1171.

Lu Y, Lemon W, Liu PY, Yi Y, Morrison C, Yang P, Sun Z, Szoke J, Gerald WL, Watson M et al. 2006. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med* **3**: e467.

Lyon MF. 1999. X-chromosome inactivation. *Curr Biol* **9**: R235-237.

Maiti A, Drohat AC. 2011. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* **286**: 35334-35338.

Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y et al. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**: 253-257.

Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature* **403**: 501-502.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**: 766-770.

Michalowsky LA, Jones PA. 1989. DNA methylation and differentiation. *Environ Health Perspect* **80**: 189-197.

Monk M, Adams RL, Rinaldi A. 1991. Decrease in DNA methylase activity during preimplantation development in the mouse. *Development* **112**: 189-192.

Monk M, Boubelik M, 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.

Nestor CE, Ottaviano R, Reddington J, Sproul D, Reinhardt D, Dunican D, Katz E, Dixon JM, Harrison DJ, Meehan RR. 2012. Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res* **22**: 467-477.

Nishioka K, Chuikov S, Sarma K, Erdjument-Bromage H, Allis CD, Tempst P, Reinberg D. 2002. Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes Dev* **16**: 479-489.

Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**: 247-257.

Okano M, Xie S, Li E. 1998. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* **26**: 2536-2540.

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD et al. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**: 714-717.

Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W, Walter J. 2000. Active demethylation of the paternal genome in the mouse zygote. *Curr Biol* **10**: 475-478.

Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P et al. 2011. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**: 394-397.

Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. 1996. Requirement for Xist in X chromosome inactivation. *Nature* **379**: 131-137.

Pfeifer GP, Steigerwald S, Boehm TL, Drahovsky D. 1988. DNA methylation levels in acute human leukemia. *Cancer Lett* **39**: 185-192.

Pietras K, Ostman A. 2010. Hallmarks of cancer: interactions with the tumor stroma. *Exp Cell Res* **316**: 1324-1331.

Plass C, Soloway PD. 2002. DNA methylation, imprinting and cancer. *Eur J Hum Genet* **10**: 6-16.

Ptashne M. 2007. On the use of the word 'epigenetic'. *Curr Biol* **17**: R233-236.

Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, Porse BT, Bernard OA, Christensen J, Helin K. 2015. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev* **29**: 910-922.

Rauch TA, Zhong X, Wu X, Wang M, Kernstine KH, Wang Z, Riggs AD, Pfeifer GP. 2008. High-resolution mapping of DNA hypermethylation and hypomethylation in lung cancer. *Proc Natl Acad Sci U S A* **105**: 252-257.

Razin A, Riggs AD. 1980. DNA methylation and gene function. *Science* **210**: 604-610.

Reik W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**: 425-432.

Reik W, Dean W, Walter J. 2001. Epigenetic reprogramming in mammalian development. *Science* **293**: 1089-1093.

Reik W, Walter J. 2001. Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2**: 21-32.

Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD. 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell* **12**: 1591-1598.

Riggs AD. 1975. X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet* **14**: 9-25.

Robertson KD. 2005. DNA methylation and human disease. *Nat Rev Genet* **6**: 597-610.

Rollins RA, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, Bestor TH. 2006. Large-scale structure of genomic methylation patterns. *Genome Res* **16**: 157-163.

Rose NR, Klose RJ. 2014. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim Biophys Acta* **1839**: 1362-1372.

Roth SY, Denu JM, Allis CD. 2001. Histone acetyltransferases. *Annu Rev Biochem* **70**: 81-120.

Rougeulle C, Chaumeil J, Sarma K, Allis CD, Reinberg D, Avner P, Heard E. 2004. Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* **24**: 5475-5484.

Saha RN, Pahan K. 2006. HATs and HDACs in neurodegeneration: a tale of disconcerted acetylation homeostasis. *Cell Death Differ* **13**: 539-550.

Saxonov S, Berg P, Brutlag DL. 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.

Schiesser S, Hackner B, Pfaffeneder T, Muller M, Hagemeier C, Truss M, Carell T. 2012. Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. *Angew Chem Int Ed Engl* **51**: 6516-6520.

Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE 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.

Schubeler D. 2015. Function and information content of DNA methylation. *Nature* **517**: 321-326.

Schulz WA, Elo JP, Florl AR, Pennanen S, Santourlidis S, Engers R, Buchardt M, Seifert HH, Visakorpi T. 2002. Genomewide DNA hypomethylation is associated with alterations on chromosome 8 in prostate carcinoma. *Genes Chromosomes Cancer* **35**: 58-65.

Selaru FM, David S, Meltzer SJ, Hamilton JP. 2009. Epigenetic events in gastrointestinal cancer. *Am J Gastroenterol* **104**: 1910-1912.

Sharp AJ, Stathaki E, Migliavacca E, Brahmachary M, Montgomery SB, Dupre Y, Antonarakis SE. 2011. DNA methylation profiles of human active and inactive X chromosomes. *Genome Res* **21**: 1592-1600.

Shen G, Rong X, Zhao J, Yang X, Li H, Jiang H, Zhou Q, Ji T, Huang S, Zhang J et al. 2014. MicroRNA-105 suppresses cell proliferation and inhibits PI3K/AKT signaling in human hepatocellular carcinoma. *Carcinogenesis* **35**: 2748-2755.

Sirotkin AV, Laukova M, Ovcharenko D, Brenaut P, Mlyncek M. 2010. Identification of microRNAs controlling human ovarian cell proliferation and apoptosis. *J Cell Physiol* **223**: 49-56.

Smallwood SA, Kelsey G. 2012. De novo DNA methylation: a germ cell perspective. *Trends Genet* **28**: 33-42.

Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Segonds-Pichon A, Sato S, Hata K, Andrews SR, Kelsey G. 2011. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat Genet* **43**: 811-814.

Spies N, Nielsen CB, Padgett RA, Burge CB. 2009. Biased chromatin signatures around polyadenylation sites and exons. *Mol Cell* **36**: 245-254.

Stamenkovic I. 2000. Matrix metalloproteinases in tumor invasion and metastasis. *Semin Cancer Biol* **10**: 415-433.

Sun XJ, Wei J, Wu XY, Hu M, Wang L, Wang HH, Zhang QH, Chen SJ, Huang QH, Chen Z. 2005. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J Biol Chem* **280**: 35261-35271.

Surani MA, Barton SC, Norris ML. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308**: 548-550.

Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. 2010. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* **38**: e181.

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al. 2009. Conversion of 5-

methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**: 930-935.

Takeshima H, Yamashita S, Shimazu T, Niwa T, Ushijima T. 2009. The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands. *Genome Res* **19**: 1974-1982.

Thienpont B, Steinbacher J, Zhao H, D'Anna F, Kuchnio A, Ploumakis A, Ghesquiere B, Van Dyck L, Boeckx B, Schoonjans L et al. 2016. Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature* **537**: 63-68.

Thomson JP, Ottaviano R, Unterberger EB, Lempiainen H, Muller A, Terranova R, Illingworth RS, Webb S, Kerr AR, Lyall MJ et al. 2016. Loss of Tet1-Associated 5-Hydroxymethylcytosine Is Concomitant with Aberrant Promoter Hypermethylation in Liver Cancer. *Cancer Res* **76**: 3097-3108.

Tsuda H, Takarabe T, Kanai Y, Fukutomi T, Hirohashi S. 2002. Correlation of DNA hypomethylation at pericentromeric heterochromatin regions of chromosomes 16 and 1 with histological features and chromosomal abnormalities of human breast carcinomas. *Am J Pathol* **161**: 859-866.

Tufarelli C, Stanley JA, Garrick D, Sharpe JA, Ayyub H, Wood WG, Higgs DR. 2003. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet* **34**: 157-165.

Valinluck V, Sowers LC. 2007. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* **67**: 946-950.

Van der Meulen J, Speleman F, Van Vlierberghe P. 2014. The H3K27me3 demethylase UTX in normal development and disease. *Epigenetics* **9**: 658-668.

Van Tongelen A, Loriot A, De Backer O, De Smet C. 2015. Application of CRISPR/cas9-Directed Homologous Recombination to the Generation of Human Tumor Cells with Conditional Knockout of an X-Linked MicroRNA Locus. in *Gene Technology*.

van Werven FJ, Neuert G, Hendrick N, Lardenois A, Buratowski S, van Oudenaarden A, Primig M, Amon A. 2012. Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. *Cell* **150**: 1170-1181.

Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM et al. 2006. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**: 871-874.

Vischer E, Chargaff E. 1947. The separation and characterization of purines in minute amounts of nucleic acid hydrolysates. *The Journal of biological chemistry* **168**: 781.

Vizoso M, Puig M, Carmona FJ, Maqueda M, Velasquez A, Gomez A, Labernadie A, Lugo R, Gabasa M, Rigat-Brugarolas LG et al. 2015. Aberrant DNA methylation in non-small cell lung cancer-associated fibroblasts. *Carcinogenesis* **36**: 1453-1463.

Wade PA. 2001. Methyl CpG-binding proteins and transcriptional repression. *Bioessays* **23**: 1131-1137.

Walsh CP, Bestor TH. 1999. Cytosine methylation and mammalian development. *Genes Dev* **13**: 26-34.

Walsh CP, Chaillet JR, Bestor TH. 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* **20**: 116-117.

Warnecke PM, Clark SJ. 1999. DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation. *Mol Cell Biol* **19**: 164-172.

Wiehle L, Raddatz G, Musch T, Dawlaty MM, Jaenisch R, Lyko F, Breiling A. 2016. Tet1 and Tet2 Protect DNA Methylation Canyons against Hypermethylation. *Mol Cell Biol* **36**: 452-461.

Wilkinson LS, Davies W, Isles AR. 2007. Genomic imprinting effects on brain development and function. *Nat Rev Neurosci* **8**: 832-843.

Williams K, Christensen J, Helin K. 2012. DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep* **13**: 28-35.

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K. 2011. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**: 343-348.

Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, Walter J. 2011. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun* **2**: 241.

Wu H, Coskun V, Tao J, Xie W, Ge W, Yoshikawa K, Li E, Zhang Y, Sun YE. 2010. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* **329**: 444-448.

Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y. 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**: 389-393.

Wu H, Zhang Y. 2011a. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev* **25**: 2436-2452.

-. 2011b. Tet1 and 5-hydroxymethylation: a genome-wide view in mouse embryonic stem cells. *Cell Cycle* **10**: 2428-2436.

Wutz A. 2011. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet* **12**: 542-553.

Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, Viegas-Pequignot E. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**: 187-191.

Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, 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.

Yang H, Liu Y, Bai F, Zhang JY, Ma SH, Liu J, Xu ZD, Zhu HG, Ling ZQ, Ye D et al. 2013. Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. *Oncogene* **32**: 663-669.

Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. 2014. Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell* **26**: 577-590.

Yang XJ, Seto E. 2003. Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. *Curr Opin Genet Dev* **13**: 143-153.

Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. 2010a. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res* **20**: 1390-1393.

Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu GL, Luo C, Jiang H, He C. 2012. Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat Chem Biol* **8**: 328-330.

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE et al. 2006. Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* **126**: 1189-1201.

Zhang X, Zhang R, Zheng Y, Shen J, Xiao D, Li J, Shi X, Huang L, Tang H, Liu J et al. 2013. Expression of gamma-aminobutyric acid receptors on neoplastic growth and prediction of prognosis in non-small cell lung cancer. *J Transl Med* **11**: 102.

Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, Rathert P, Brandt O, Reinhardt R, Fischle W et al. 2010b. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res* **38**: 4246-4253.

Zhou W, Fong MY, Min Y, Somlo G, Liu L, Palomares MR, Yu Y, Chow A, O'Connor ST, Chin AR et al. 2014. Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis. *Cancer Cell* **25**: 501-515.

ANNEX

Application of CRISPR/cas9-Directed Homologous Recombination to the Generation of Human Tumor Cells with Conditional Knockout of an X-Linked MicroRNA Locus

Aurélie Van Tongelen[1], Axelle Loriot[1], Olivier De Backer[2] and Charles De Smet[1]

[1]Department of Genetics and Epigenetics, de Duve Institute, Université catholique de Louvain, Brussels, Belgium

[2]URPhyM NARILIS, University of Namur-FUNDP, Namur, Belgium

Received July 01, 2015; **Accepted** July 17, 2015; **Published** July 20, 2015

Gene Technol 4:124. doi: 10.4172/2329-6682.1000124

Abstract

Studying the cellular function of microRNAs requires genetic strategies to generate their loss-of-function. Recently, a novel approach of targeted genomic deletion was proposed, which is based on induction of site-specific DNA cuts with Cas9/gRNA ribonucleoprotein complexes, combined with homologous recombination-dependent insertion of cassettes that contain sequences for Cre-lox or FLP-FRT systems. Here, we provide a technical report describing application of this CRISPR/Cas9-directed homologous recombination procedure to the generation of human tumor cells in which conditional knockout of an X-linked cluster of microRNAs (miR-105/miR-767) can be induced. We describe the successive steps of genetic engineering and cell clone selection that allowed us to generate cells with the expected genome editing.

Introduction

MicroRNAs (miRNAs) are small noncoding RNAs that exert important cellular functions by repressing post-transcriptional gene expression through binding to target mRNAs [1]. More than 60% of all human genes are predicted to be regulated by a total of over 2.000 mature miRNAs. Some miRNAs are expressed in virtually all cell types, whereas others are highly tissue-specific with a distinct function depending on cell type or organ. miRNAs play key roles in most biological processes, including cell division and death, cellular metabolism, intracellular signaling, immunity and cell movement. Not surprisingly, disregulation of miRNA expression has been functionally linked to various pathological and occasionally malignant outcomes [2]. Specific miRNA expression patterns, which have been associated with particular diseases, hold great prognostic value. In cancer, oncogenic miRNAs as well as tumor suppressor miRNAs have been identified, and therapeutic strategies directed against such cancer-related miRNA are being considered [3,4].

The physiological roles of most miRNAs still need to be deciphered. Therefore, tools allowing manipulation of miRNA activity are required. Gain-of-function by over-expression of the miRNA is relatively easy, and can be achieved

by transfection of synthetic microRNAs or by enforcing expression of primary miRNA transcripts [5,6]. Loss of- function analyses are less obvious. miRNA inhibitors based on synthetic antisense molecules or miRNA sponges, which act by binding and sequestering miRNAs away from their natural targets, are available [7]. However, such inhibitors have some disadvantages, including incomplete masking of miRNA function, and uncontrolled off-target effects.

Genome engineering is a powerful tool for dissecting biological mechanisms. The Clustered, Regularly Interspaced, Short Palindromic Repeats (CRISPR) / CRISPR-associated protein 9 (Cas9) system provides a rapid and efficient technology for targeted genome editing. In association with a specifically designed guide RNA (gRNA), Cas9 can achieve site-specific DNA recognition and cleavage [8]. A constraint in this genome editing process is the compulsory presence of a Protospacer Adjacent Motif (PAM) sequence near the Cas9 cleavage site. In most cases, the site-specific DNA double-strand breaks (DSB) induced by Cas9 triggers a non-homologous end-joining (NHEJ) process of DNA repair, leading to small insertions or deletion in the nucleotidic sequence. This procedure is therefore exploited to generate loss-of-function of protein coding genes, via alteration of the open reading frame. Short miRNAs are less amenable to this editing process, as many of them do not carry the required PAM motif within their sequence.

An alternative use of the CRISPR/Cas9 technology, which would be better adapted to manipulate miRNAs, has been proposed. In this setting, Cas9 cleavages are used to direct homologous recombinationdependent insertion of cassettes that contain LoxP or FRT sequences up- and downstream of the miRNA locus, thereby allowing subsequent deletion of the embedded sequence upon expression of Cre or FLP recombinases. This procedure requires that template DNA sequences with homologous arms are provided to the cells, in order to stimulate homology-directed repair (HDR) of the Cas9-iduced DSB [9]. Examples where this technology was used to delete a miRNA locus remain however scarce [10,11]. Here, we report successful application of this procedure to the generation of human tumor cells in which a miRNA locus can be conditionally deleted. The locus targeted in our study corresponds to an X-linked cluster of

two miRNAs (miR-105 and miR-767), which were recently shown to display aberrant activation in wide variety of tumors, and to exert oncogenic potential [12,13].

Results and discussion

General strategy for the generation of cells with conditional knockout of MIR105/767

The general strategy to obtain cells where the *X-linked MIR105/767* locus can be conditionally deleted is described in Figure 1 and each step is described in details below. In brief, the procedure involved successive steps of CRISPR/cas9-directed homologous recombination to permit integration of neomycin-*loxP/FRT* cassettes 3' and 5' to the *MIR105/767* genomic locus, as well as exposure of cells to Cre and FLP (Flippase) recombinases to leave only single loxP sites on both sides. At each step of the procedure, a cell clone with the appropriate integration/ recombination was selected, and used for the subsequent step.

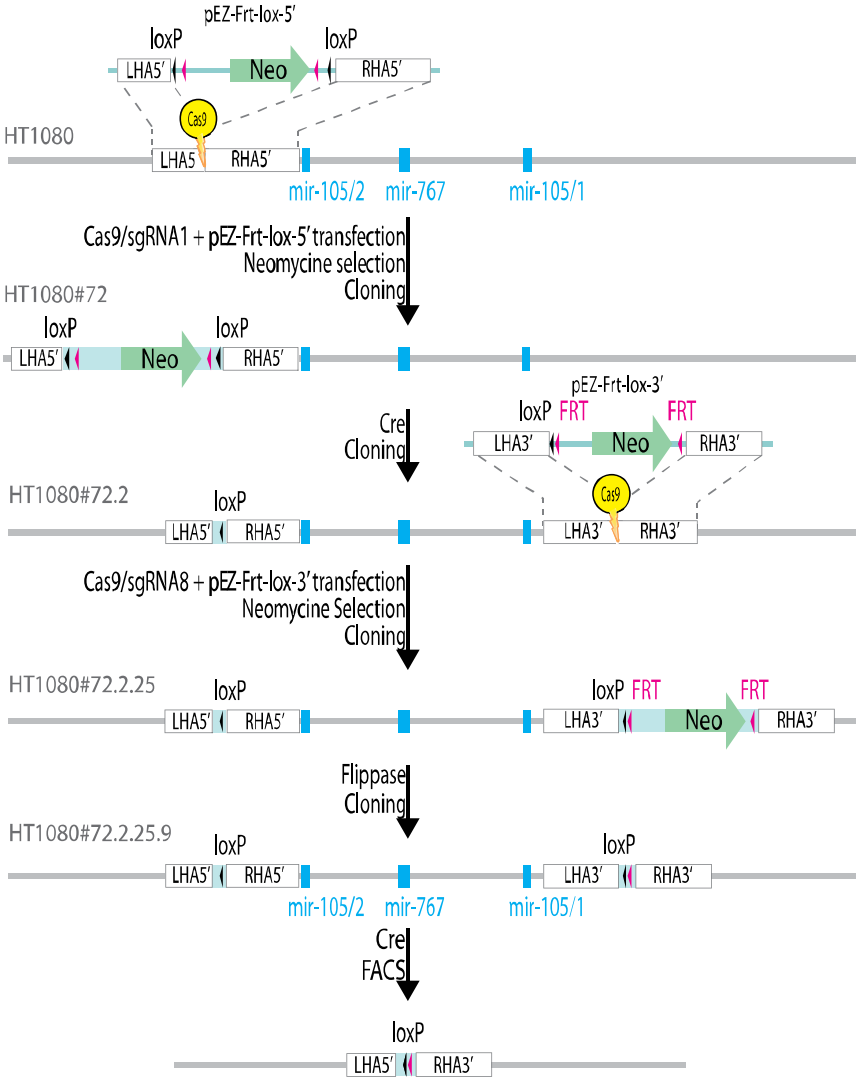


Figure 1: Schematic draw depicting the strategy used for generation of a cell line with a "floxed" MIR105/767 locus. The MIR105/767 locus comprises two pre-microRNA sequences (MIR105-1 and MIR105-2) for miR-105 and one (MIR767) for miR-767, represented by blue boxes. Refer to the text for detailed explanation on the targeting vectors, and the successive steps of cell transfection and clone selection.

Selection of efficient single guide RNAs (sgRNAs)

Targeting of the Cas9 nuclease to a specific site on the genome is commonly achieved through its association with a chimeric RNA construct, termed single guide RNA (sgRNA). The sgRNA includes the tracrRNA sequence, which allows association of the RNA molecule with the Cas9 enzyme, and the spacer sequence, which binds the genomic target by nucleotide complementarity. Not all sgRNAs are equally efficient. The first step to achieve our construction was therefore to test different sgRNAs, and to select those that stimulate site-specific cleavage by Cas9 with the highest efficiency (Figure 2). Using a spacer RNA design algorithm (http://crispr.mit.edu), we designed four sgRNAs for each side of the *MIR105/767* locus. The corresponding sequences were inserted in the pX330-U6-Chimeric_BB-CBh-hSpCas9 vector, which also carries a human codon-optimized Cas9 coding sequence. These different Cas9/sgRNA vectors were then transfected into HT1080 cells, generating DSBs at precise sgRNA targeted sites. The error-prone NHEJ process results in the formation of mutation or small insertions or deletion. Genomic DNA was collected from transfectants after three days. DNA cleavage efficiency was evaluated by mutation detection assay using the Surveyor® Mutation Detection kit. Thus, DNA fragments carrying the expected cleavage site were amplified by PCR. Amplicons were denatured by heat, and then allowed the re-anneal slowly. They were then exposed to the Surveyor nuclease, which cuts heteroduplexes but not homoduplexes of DNA. Nuclease-treated DNA fragments were electrophoresed in an agarose gel, and the cleavage efficiency was evaluated by calculating the ratio between the cleaved and uncleaved band (percentage of indel). Using this method, we identified the sgRNAs that directed the most efficient Cas9-mediated cleavage either 5' (sgRNA1, 9.3% indel) or 3' (sgRNA8, 12.9% indel) to the *MIR105/767* locus (Figure 2).

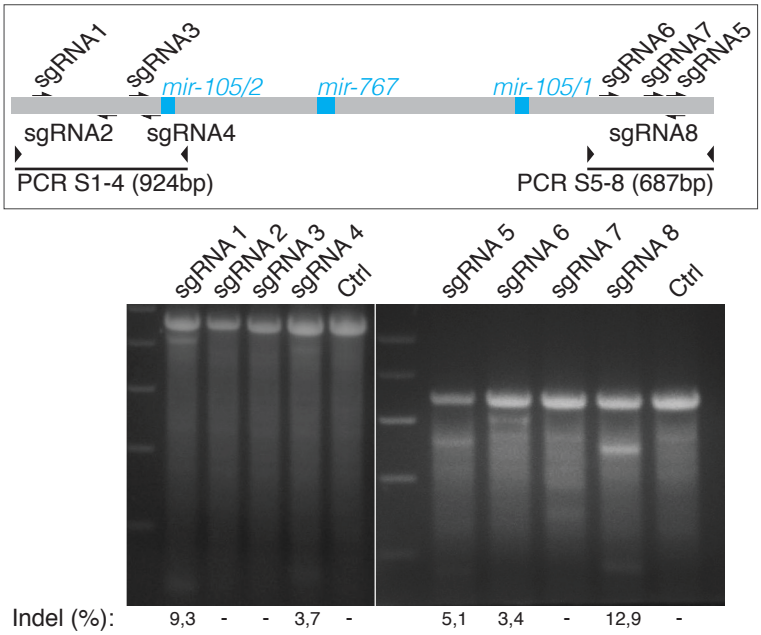


Figure 2: Surveyor assay comparing the cleavage efficiency induced by eight different sgRNAs at corresponding DNA regions. On the top panel, the schematic draw depicts localisation of each sgRNA on the MIR105/767locus, as well as the position of primers used for PCR amplification prior to the Surveyor nuclease treatment. Sizes of initial (uncleaved) PCR products are also given. Transfections with each of the Cas9/sgRNA vectors (sgRNA1 to -8) were performed in HT1080 cells, and DNA was extracted from transfectants 3 days later. DNA samples were submitted to PCR amplification with indicated primers, and then exposed to Surveyor nuclease. Resulting DNA products were analyzed by gel electrophoresis. Presence of DNA fragments with a size lower than the original PCR product indicates efficient sgRNA-directed cleavage. The ratio between the intensity of the lower band and that of the starting PCR band provides a measurement of cleavage efficiency (% indel).

CRISPR/Cas9 mediated integration of a *Neo-2loxP/2FRT* cassette 5' to the *MIR105/767* locus

In order to direct homologous recombination towards the sgRNA1-induced Cas9 cleavage site, a vector (pEZ-FRT-lox-5') was constructed in which left and right homology arms were inserted on each side of a neomycin resistance cassette (*Neo*) flanked by two loxP and FRT sites (*Neo-2loxP/2FRT;* Figure 3). The left homology arm (LHA5', 306bp) and the right homology arm

(RHA5', 692bp) corresponded to genomic sequences located upstream and downstream the Cas9/sgRNA1 cutting site, respectively.

We transfected Cas9/sgRNA1 and pEZ-FRT-lox-5' vectors in six different human cell lines: embryonic kidney cells (HEK293), two melanoma cell lines (Mi13443 and LB2201-MEL), two lung carcinoma cell lines (GLCP1 and SKMES1), and a fibrosarcoma cell line (HT1080). We selected cells of male origin, as these contain only one X chromosome. Following transfection, cells were selected for two weeks in neomycin-containing medium, and part of the resistant populations was harvested for DNA extraction. Integration of the Neo-2loxP/2FRT cassette was assessed by PCR amplification with one primer matching a genomic sequence upstream of the LHA5' and another primer corresponding to a sequence within the cassette. Of the six transfected cell lines, only three (HT1080, GLCP1 and SK-MES1) showed evidence of integration of the *Neo-2loxP/2FRT* cassette (Figure 3).

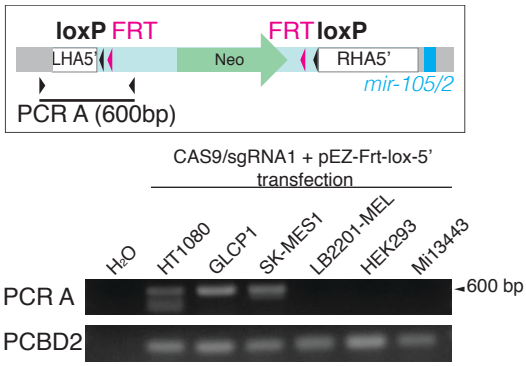


Figure 3: PCR screening for insertion of Neo-2loxP/2FRTcassette in six cell lines transfected with pX330-sgRNA1 and pEZ-FRT-lox-5' vectors. Following transfection, indicated cell lines were submitted to 2 weeks neomycin selection, and their genomic DNA was harvested for PCR analyses. The upper panel depicts the annealing sites of the PCR primers used to test integration of the Neo-2loxP/2FRTcassette (PCR A). One primer is located in the genomic region outside of the left homology arm (LHA5') and the other primer is located in the inserted sequence. Gel electrophoresis shows PCR amplification products obtained with primers for PCR A, and for the PCBD2, a gene located on chromosome 5 used as positive control.

We decided to derive clones from the transfected HT1080 cell population, in order to isolate clones harboring the appropriate integration event. Out of 113 clones, 4 showed integration of the *Neo-2loxP/2FRT*, as evidenced by positive signals following PCR amplification with appropriate primers (Figure 4). Following verification of PCR products by sequencing, only clones #72, #76 and #83 showed perfect recombination. The remaining clone (#29) showed deletions in the intersection between homology arms and the cassette sequence. Overall, we observed a 2.5% (3/113) efficiency of homologous recombination directed by CRISPR/Cas9 towards the 5' side of the MIR105/767 locus in HT1080 cells. Clone HT1080#72 was selected for the subsequent steps.

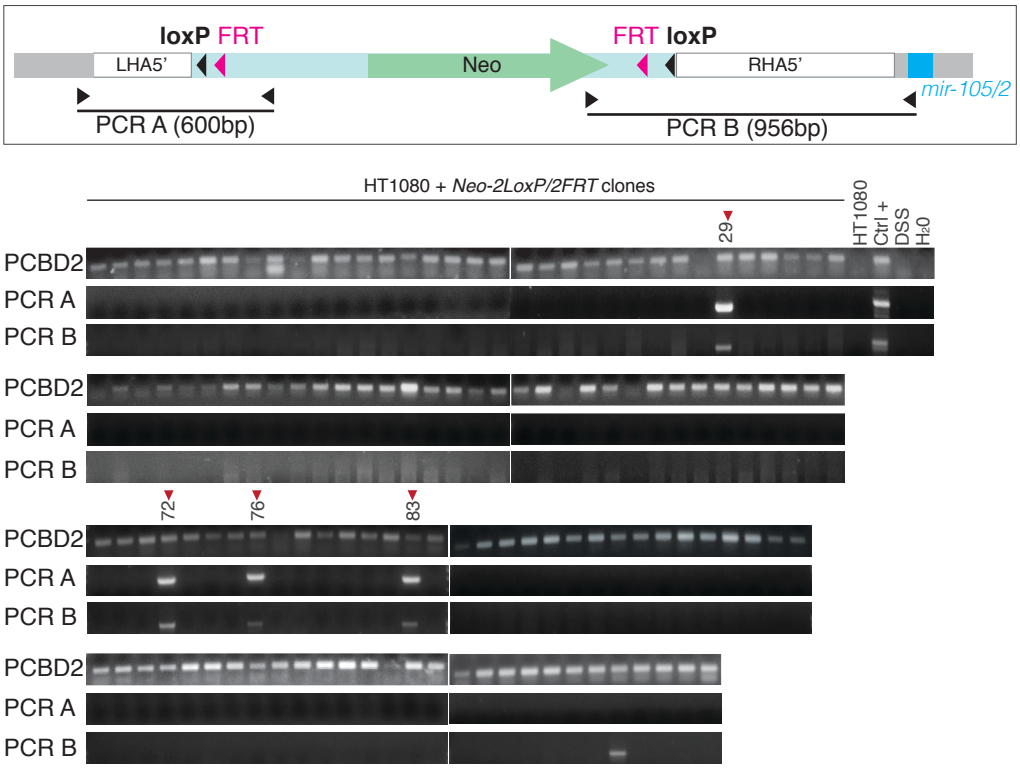


Figure 4: PCR screening for Neo-2loxP/2FRTcassette insertion in HT1080 cell clones. Following transfection with pX330-sgRNA1 and pEZ-FRT-lox-5′ vectors, HT1080 cells were selected in neomycin, and 140 cell clones were derived. Their DNA was submitted to PCR analyses. The picture above depicts the annealing sites of the PCR primers used for screening clones with effective homologous integration of the Neo-2loxP/2FRT cassette. PCR A (600bp) and PCR B (956bp) have both one primer located in genomic DNA outside the homology arms, and another primer located in the donor vector sequence. Recombinant clones are expected to produce a signal with both PCRs. PCR amplification of PCBD2 served as control. According to PCBD2 amplification, only 113 of these were analyzable. Control samples included the total population of HT1080 cells transfected withpX330-sgRNA1/pEZ-FRT-lox-5′ (HT1080 sgRNA1), untransfected HT1080 cells, salmon sperm DNA (DSS), and water.

Cre recombinase-mediated removal of the 5′ cassette

In order to obtain cells where the Neo-2loxP/2FRT cassette was removed and only one loxP sequence remained 5′ to the MIR105/767 locus, HT1080#72 cells were transiently transfected with a vectors encoding Cre recombinase (CMV-Cre). Sub-clones were isolated, and screened for effective loxP recombination by PCR amplification with primers matching sequences in LHA5′ and RHA5′ homology arms (Figure 5). Out of five tested clones, two clones (#72.2 and #72.3) showed effective recombination. We chose clone HT1080 #72.2 for the subsequent steps aiming at introducing a second loxP site 3′ to the *MIR105/767* locus.

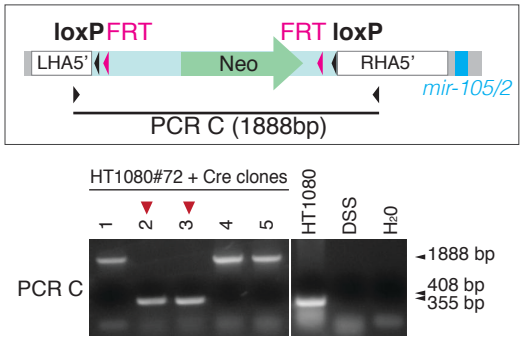


Figure 5: PCR screening for Cre-induced loxP recombination at the 5′ cassette. The schematic draw depicts the annealing sites of the PCR C primers used to detect "floxing" of the 5′ cassette. The PCR product is expected at a size of 408bp in case of appropriate loxP recombination (1888bp in non-recombinant cells). PCR in naïve cells is expected to yield a 355bp product. Five clones isolated from the CMV-Cre-transfected HT1080#72 cell population were tested. Control samples included untransfected HT1080 cells, salmon sperm DNA (DSS), and water.

CRISPR/Cas9-directed integration of a Neo-1LoxP/2FRT cassette 3' to the MIR105/767 locus

In HT1080#72.2 cells, we next performed directed integration of a Neo-1loxP/2FRT cassette towards the sgRNA8-induced Cas9 cleavage site, located on the 3' side of the *MIR105/767* locus. To this end, we constructed a vector (pEZ-FRT-lox-3') in which left and right homology arms were inserted on each side of the Neo-1loxP/2FRT cassette (Figure 6). The left homology arm (LHA3', 774bp) and the right homology arm (RHA3', 692bp) corresponded to genomic sequences located upstream and downstream of the Cas9/sgRNA8 cutting site, respectively. Restriction sites used for insertion of the RHA3' into the vector were chosen in order to delete the neighboring loxP site.

HT1080#72.2 cells were transfected with Cas9/sgRNA8 pEZFRT- lox-3' vectors, and following neomycin selection, resistant clones were derived. These clones were screened for integration of the *Neo-1loxP/2FRT* cassette by PCR amplification on genomic DNA, using primer pairs matching either sequences within the cassette or genomic sequences outside of the homology regions (Figure 6). PCR signals of the expected size were detected in 4 out 33 tested clones, indicating that Cas9/sgRNA8-directed homologous recombination in HT1080#72.2 cells was obtained with an efficiency of 12,1%. Clone HT1080#72.2.25 was selected for the subsequent steps.

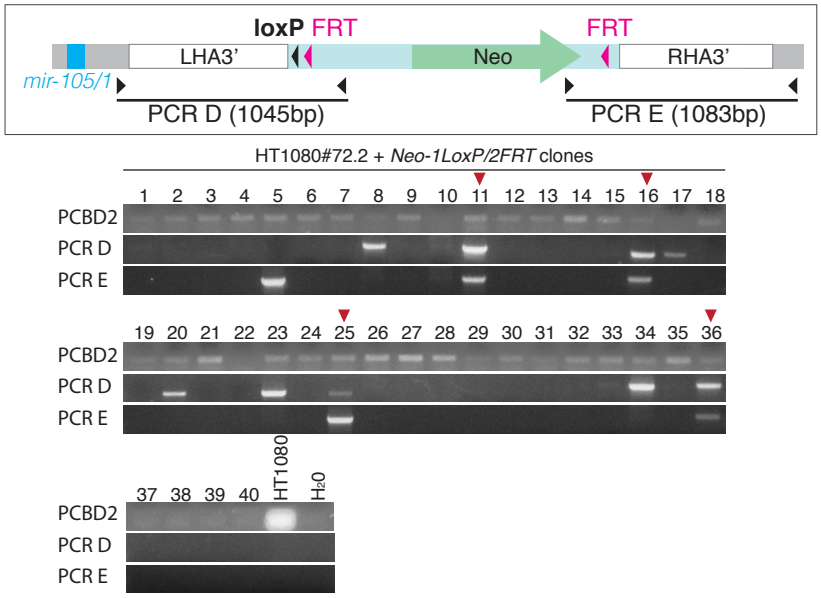


Figure 6: PCR screening for Neo-1loxP/2FRTcassette insertion 3' to the MIR105/767locus in HT1080 cell clones. HT1080#72.2 cells, which had been transfected with pX330-sgRNA8 and pEZ-FRT-lox-3' vectors, were selected in neomycin and cloned. DNA from 40 clones was tested for insertion of the Neo-1loxP/2FRTcassette insertion 3' to the MIR105/767by PCR. The upper panel indicates the position of the primers and expected product size of the two PCR (D and E). PCR results for the PCBD2control sequence shows that only 33/44 clones were analyzable. Control samples included untransfected HT1080 cells, and water.

Flippase mediated deletion of the 3' cassette

We next aimed to delete the cassette integrated 3' to the MIR105/767 locus, but without removing the single loxP site. To this end, HT1080#72.2.25 cells were transiently transfected with a vector encoding Flippase, which induces recombination between FRT sites. Following transfection, cellular clones were isolated, and screened for effective FRT recombination by PCR amplification with primers matching sequences in LHA3' and RHA3' homology arms (Figure 7). Out of 21 tested clones, 8 clones (38%) showed effective FRT recombination. Among these, clone HT1080 #72.2.25.9, in which the MIR105/767 locus is now surrounded by two loxP sites, was chosen for subsequent analyses.

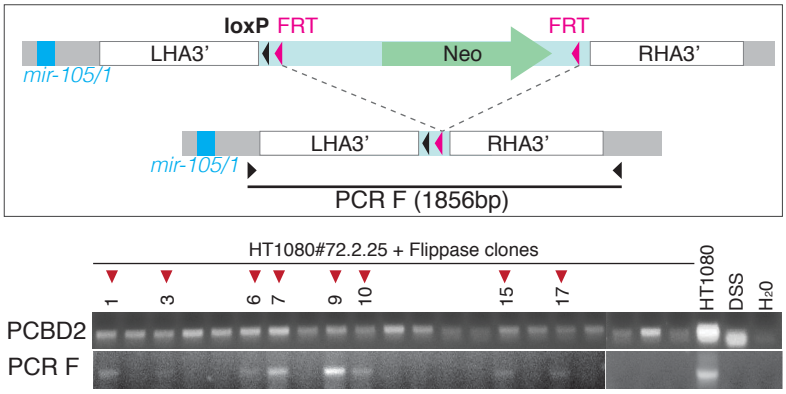


Figure 7: PCR screening for Flippase-induced FRTrecombination at the 3' cassette.Clones were derived from HT1080#72.2.25 that had been transiently transfected with a Flippase-encoding vector. DNA was extracted from 21 clones, and analyzed by PCR to detect FRT recombination. The schematic draw depicts the annealing sites of the PCR primers used for the screening (PCR F), and the expected PCR product size in case of effective FRT recombination. Amplification of PCBD2served as positive control. Control samples included untransfected HT1080 cells, and water.

Cre recombinase-mediated knockout of the *MIR105/767* locus

Since HT1080#72.2.25.9 cells contain loxP sequences on both sides of the *MIR105/767* locus, their exposure to Cre recombinase is expected to induce deletion of the corresponding genomic region. To test this, HT1080#72.2.25.9 cells were transiently transfected with the CMV-Cre vector. Previous experiments established a 40% efficiency of transfection in HT1080 cells, implying that only part of the transfected HT1080#72.2.25.9 will effectively express Cre recombinase. We therefore decided to verify successful deletion of the MIR105/767 locus in clones isolated from the CMV-Cre transfected HT1080#72.2.25.9 cell population. Following derivation of the clones, their DNA was extracted and submitted to PCR screening with three different pairs of primers allowing unambiguous determination of the presence or absence of the *MIR105/767* locus (Figure 8). Among 24 tested clones, 7 clones (29%) showed effective deletion of the MIR105/767 locus (Figure 8).

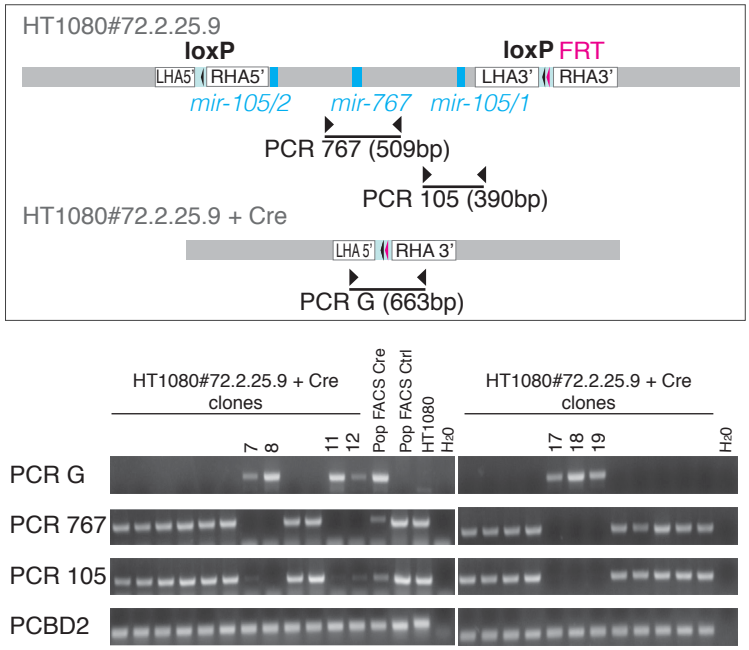


Figure 8: PCR analysis of Cre-induced deletion of the MIR105/767locus. HT1080#72.2.25.9 were transiently transfected with either CMV-Cre or pTM945-Cre-IRES-mCherry. CMV-Cre transfected cells were cloned, whereas pTM945-Cre-IRES-mCherry transfectants were sorted by FACS to isolate mCherry-expressing cells (Pop FACS Cre). A control population of cells was obtained after sorting cells that had been transfected with pTM945-mCherry (Pop FACS Ctrl). PCR screening was applied to cell clones (n=24) or sorted populations to detect deletion of the MIR105/767locus. Three PCR amplifications were performed (PCR 767, PCR 105, PCR G). The upper panel indicates the position of PCR primers, and expected sizes of PCR products (following deletion for PCR G). Control samples included untransfected HT1080 cells, and water.

We next tried to facilitate the knockout procedure by avoiding the process of cellular cloning that followed CMV-Cre transfection. To this end, we replaced CMV-Cre with the pTM945-Cre-IRES-mCherry vector, which expresses both Cre recombinase and the mCherry fluorescent protein. Thus, HT1080#72.2.25.9 cells were transfected with pTM945-Cre-IRES-mCherry, and two days later cells with the highest level of fluorescence were selected by fluorescence-activated cell sorting (FACS) and reseeded. A control group, transfected with a similar vector lacking Cre recombinase (pTM945-mCherry), was treated in parallel. Genomic DNA was extracted from FACS sorted cells, and submitted to PCR screening with the above-mentioned primer pairs (Figure 8).

The results showed efficient enrichment of cells with *MIR105/767* deletion, although a small proportion of non-recombined cells remained (Figure 8). We also analyzed the expression of miR-105 and miR-767 by RT-qPCR in FACS-sorted cells. The results confirmed marked loss of expression of these two miRNAs in Cre-exposed cells, as only 5,6% and 6% residual expression level was observed for miR-105 and miR-767, respectively (Figure 9).

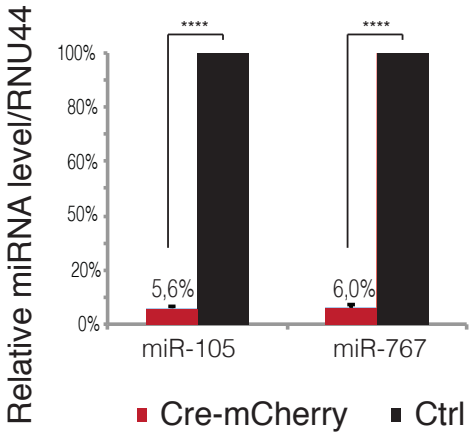


Figure 9: RT-qPCR expression analysis of miR-105 and miR-767 in FACS-sorted Cre-induced HT1080#72.2.25.9 cells.RNA was extracted from Cherry-expressing (FACS-sorted) HT1080#72.2.25.9 cells that had been previously transfected with either pTM945-Cre-IRES-mCherry (Cre-mCherry) or pTM945-mCherry (Ctrl). RT-qPCR analyses were performed at 6 different time points (days 6 to 34) post-transfection. Expression levels are expressed relative to the level in Ctrl cells. 2 way ANOVA was used for statistical analysis (***, p<0,0001).

Altogether, we described a genetic engineering strategy based on CRISPR/Cas9-directed homologous recombination, which allowed us to create a human tumor cell line in which the *MIR105/767* locus can be conditionally deleted. Our data therefore confirm the feasibility of this strategy, and provide several guidelines concerning the experimental procedures. CRISPR/Cas9-directed homologous recombination was found to occur with an expected efficiency (2.5% and 12.1% efficiency for the 5' and 3' position, respectively). One of the pitfalls of the CRISPR/ Cas9 technology is the presence of many off-target mutations induced by Cas9, which renders comparisons between genetically modified cells uncertain. This is however not a problem in our cellular model, as

comparisons will be made within the final clone (HT1080#72.2.25.9), between two cell groups that differ only for the presence or absence of the *MIR105/767* locus.

Material and methods

Cell culture and transfection

HT1080, LB2201-MEL, Mi13443, SK-MES1, GLCP1 cell lines were cultured in Iscove's modified Dulbecco's medium (IMDM, Life Technologies) and HEK293 cell line in 4,5 g/l D-glucose Dulbecco's modified Eagle's medium (DMEM, Life Technologies). All media were supplemented with 10% fetal bovine serum (FBS, Life Technologies) and contained 1x non-essential amino acids (Life Technologies) and 1x Penicillin/Streptomycin (Life Technologies).

DNA transfection was performed in a T75 flasks, with 2 to 6 million cells, using Lipofectamine 2000® Reagent (Invitrogen). Five hours after transfection, the medium was replaced. The amount of transfected DNA was 5 µg of Cas9/sgRNAs vectors and 5 µg of donor pEZ-Frt-lox-5' or pEZ-Frt-lox-3' vectors. Prior to transfection, pEZFrt- lox-5' vector was linearized 5' to the left homology arm by digestion with the EcoRV restriction enzyme, and the pEZ-Frt-lox-3' vector was digested by EcoRV and XhoI restriction enzymes. Transfectants were selected for 15 days in medium containing 0.8 to 2 mg/ml of geneticin, depending on the type of cells. Two weeks after transfection, integration of the donor DNA was tested by PCR amplification with primers encompassing the recombinant fragment on genomic DNA. Cell populations showing effective recombination were cloned by limiting dilutions.

For Cre or FLP recombination, cells were transiently transfected using Lipofectamine 2000® Reagent (Invitrogen) in 6-well plates with 2 µg of CMV-Cre or pCAGGS-FLPe vectors, and cloned by limiting dilution 72h after transfection.

Plasmid design and construction

The pX330-U6-Chimeric_BB-CHb-hSpCas9 vector (available at Addgene, and kindly provided by P. Coulie) was digested using Fast- Digest BbsI (Thermo Scientific) and double-stranded oligonucleotides corresponding to different spacer sequences were ligated with T4 DNA ligase (Thermo Scientific) generating different sgRNAs. Double-stranded oligonucleotides for spacer sequences were obtained by annealing pairs of complementary oligos (Eurogentec) composed of the spacer sequence flanked by adapter sequences (Forward: 5′-CAACG…-3′; Complement: 5′-AAAC…C-3′). Design of spacer sequences was based on CRISPR Design Tool of ZhangLab (http://crispr. mit.edu) generating: sgRNA1 5′ -TTTGTGTAATGGGATCGTTTGG; sgRNA2, 5′-GGGGATAGATATGGTATCCCAGG; sgRNA3 5′-CAATCATTGTCTACTTGTACTGG; sgRNA4 5′-CTGCAACATGAGCTGATACCAGG; sgRNA5 5′-TGGTAGACTGAGCACGCGTAAGG; sgRNA6 5′- GGTCGTCTGATGTTATACCTTGG; sgRNA7 5′- AGGAATTCTTTCCCGTATTGTGG; sgRNA8 5′-TACGCGTGCTCAGTCTACCATGG.

Donor vector constructions were performed on the basis of the pEZ-Frt-lox-DT vector (Addgene). Left (LHA) and right (RHA) homology arms (~300 to ~800 pb) were amplified by PrimeSTAR HS DNA Polymerase (Takara) from HT1080 genomic DNA using complementary primers flanked by restriction site sequence at the 5′ end. For the pEZ-Frt-lox-5′ vector, primers to amplify LHA were flanked with EcoRV and NotI sites (sense: 5-attaGATATCtaggcatgtgttcatccgagga; reverse: 5′-attaGCGGCCGCgatcccattacacaaactaacagc), and primers to amplify RHA were flanked with HindIII and XhoI sequences (sense: 5′-attaAAGCTTgtttggatttatatcatgcccaaac; reverse: 5′-attaCTCGAGgataccaggttatagggagtaga). For the pEZ-Frt-lox-3′ vector, primers to amplify LHA were flanked with EcoRV and NotI sites (sense: 5′-atta-GATATCgtgtatgtgcacaatgctcact; reverse: 5′-attaGCGGCCGCccatggagaatgagccaatact), and primers to amplify RHA were flanked with SalI and XhoI sequences (sense: 5′-attaGTCGACtagactgagcacgcgtaaggaa; reverse: 5′-

attaCTCGAGccatagactggtacttaccgaga). PCR products were digested with the appropriate restriction enzymes, and purified (QIAquick Gel Extraction Kit, Qiagen), prior to ligation with T4 DNA ligase (Thermo Scientific) into the pre-digested vector. pEZ-Frt-lox-DT vector was pre-digested with EcoRV-NotI for the insertion of both LHA, and with HindIII-XhoI or SalI-XhoI for introduction of RHA to pEZFrt- lox-5′ or pEZ-Frt-lox-3′ vectors, respectively. All PCR products and final vectors were verified by DNA sequencing.

DNA preparation and PCR

Two weeks after cloning, HT1080 clones were removed from 96-well plates with 2 mM EDTA/PBS. Half of the cells were reseeded in 24-well plate and the other half was used for DNA extraction. Cells were placed in a lysis buffer composed by 10mM Tris pH8, 10 mM EDTA, 1% SDS and 0.1 mg/ml of proteinase K (New England BioLabs). After overnight incubation at 45°C, DNA was extracted by organic extraction with phenol:chloroform and 1 µg of salmon sperm DNA was added to help isopropanol precipitation. PCR S1-4 (Forward: 5′-gtcactggctcctgcagaat; reverse: 5′-gaagcagatggcaatgttgcaa), PCR S5-8 (Forward: 5′-gagaatgtctggttgtgtacctt; Reverse: 5′-caggcaaacacgatactgtcaa), PCR A (Forward: 5′-cactgtcatcttatccttaactga; Reverse: 5′-ctcgaccatatgggagagct), PCR B (Forward: 5′-gacgagttcttctgagggga; Reverse: 5′-atttgaccacgatgcacacaca), PCR C (Forward: 5′-ggctcctgcagaatgggtat; Reverse: 5′-acacctacaggacataggctat), PCR D (Forward: 5′-atatgaggaagcatggcaccat; Reverse: 5′-ctcgaccatatgggagagct), PCR E (Forward: 5′-gacgagttcttctgagggga; Reverse: 5′-gataactatggtgatgcacataca), PCR F (Forward: 5′-atatgaggaagcatggcaccat; Reverse: 5′-gataactatggtgatgcacataca), PCR G (Forward: 5′-ggctcctgcagaatgggtat; Reverse: 5′-tgctacctctaatggatatggtaa), PCR 105 (Forward: attctcgagcccttttagctatggtcttctgct; Reverse: gatacgcgttgatggtgccatgcttcctca) and PCR 767 (Forward: 5′-attctcgagtttaacagtcaaatattagtgttagttgct; Reverse: 5′-gatacgcgtcagccatcacatcataggca) were performed using Dream Taq polymerase (Thermo Scientific) on 1/3 of the total extracted DNA.

DNA extracted from total cell populations was performed similarly, but without addition of salmon sperm DNA. DNA concentrations were quantified with a

NanoDrop ND-1000 (NanoDrop Technologies), and PCR were performed on 100 ng of genomic DNA per reaction.

RNA preparation and RT-qPCR evaluation of miRNA expression levels

We isolated total RNA using TriPure Isolation Reagent (Roche) according to the manufacturer's protocol, and used 100ng of this RNA for reverse transcription with the miRCURY LNATM Universal RT microRNA PCR (Exiqon). The resultant cDNA was diluted 20 times, and a 2.5 µl aliquot was used in a SYBRGREEN qPCR with LNA specific primers for miR-105 and miR-767 (Exiqon). RNU44 was used as internal control.

Flow cytometry cell sorting

HT1080 #72.2.25.9 clone was transiently transfected by pTM945- Cre-IRES-mCherry vector or control vector pTM945-mCherry, coding both for mCherry fluorescent protein. Two days later, cells were sorted by fluorescence-activated cell sorting (FACS, FACSAriaTM III, BD Biosciences) based on mCherry signal. The sorted populations were harvested for RT-qPCR analysis 6 to 34 days after reseeding.

Detection of gRNA efficiency

HT1080 cells were transfected with pX330-U6-Chimeric_BB-CBhhSpCas9 vector carrying 8 different sgRNAs, named Cas9/sgRNA1-8. Three days after transfection, DNA extracted from those cells was treated with RNAse A and submitted to SURVEYOR® Mutation Detection Kit, according to the manufacturer's recommendations (Transgenomic). In brief, the genomic region flanking the CRISPR/Cas9 target site was first amplified by PCR using PrimeSTAR HS DNA polymerase (Takara). After migration through agarose gel, PCR products were purified using QIAquick Gel Extraction Kit (Qiagen). 2 µg of the purified PCR products were mixed with 10x Surveyor Buffer (100 mM Tris, 500 mM KCl, 15mM MgCl2) and water to a final volume of 20 µl, and subjected to a re-annealing process to enable heteroduplex formation: 95°C for 10min, 95°C to 85°C ramping at − 2°C/s, 85°C to 25°C at − 0.25°C/s, and 25°C hold for 1 minute. After re-annealing, PCR products were treated with 1 µl of SURVEYOR nuclease and

SURVEYOR enhancer S (Transgenomics) during 1 hour at 42°C. DNA samples were run in a 2% agarose gel, in TAE 1x buffer supplemented with ethidium bromide. Quantification was performed by calculating relative band intensities using ImageJ software.

Acknowledgments

This work was supported by grants from the D.G. Higher Education and Scientific Research of the French Community of Belgium (Action de Recherches Concertées) and from the Fonds special de recherche (FSR) of the Université catholique de Louvain, Belgium. A.V.T. is recipient of a Télévie grant from the FRSFNRS, Belgium [#7.4581.13]. A.L. is supported by the de Duve Institute, Brussels, Belgium.

References

1. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136: 215-233.
2. Esteller M (2011) Non-coding RNAs in human disease. Nat Rev Genet 12: 861-874.
3. Croce CM (2009) Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet 10: 704-714.
4. Hayes J, Peruzzi PP, Lawler S (2014) MicroRNAs in cancer: biomarkers, functions and therapy. Trends Mol Med 20: 460-469.
5. Chen CZ, Li L, Lodish HF, Bartel DP (2004) MicroRNAs modulate hematopoietic lineage differentiation. Science 303: 83-86.
6. Wang Z (2011) The guideline of the design and validation of MiRNA mimics. Methods MolBiol 676: 211-223.
7. Brown BD, Naldini L (2009) Exploiting and antagonizing microRNA regulation for therapeutic and experimental applications. Nat Rev Genet 10: 578-585.
8. Mali P, Yang L, Esvelt KM, Aach J, Guell M, et al. (2013) RNA-guided human genome engineering via Cas9. Science 339: 823-826.
9. Cong L, Ran FA, Cox D, Lin S, Barretto R, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339: 819-823.
10. Ho TT, Zhou N, Huang J, Koirala P, Xu M, et al. (2015) Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. Nucleic Acids Res 43: e17.
11. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, et al. (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res 41: e141.
12. Loriot A, Van Tongelen A, Blanco J, Klaessens S, Cannuyer J, et al. (2014) A novel cancer-germline transcript carrying pro-metastatic miR-105 and TET-targeting miR-767 induced by DNA hypomethylation in tumors. Epigenetics 9: 1163-1171.
13. Zhou W, Fong MY, Min Y, Somlo G, Liu L, et al. (2014) Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis. Cancer Cell 25: 501-515.