Epigenetic regulation of 5-methylcytosine RNA modification in human cancer

Vanessa Ortiz Barahona

UNIVERSITAT DE BARCELONA

Epigenetic regulation of 5-methylcytosine RNA modification in human cancer

PhD Thesis

Vanessa Ortiz Barahona

Barcelona, 2021

Epigenetic regulation of 5-methylcytosine RNA modification in human cancer

Memoria presentada por Vanessa Ortiz Barahona para optar al grado de Doctor por la Universidad de Barcelona

Universidad de Barcelona – Facultad de Medicina
Programa de Doctorado en Biomedicina

Este trabajo ha sido realizado en el Grupo de Epigenética del Cáncer, primero localizado en el Programa de Epigenética y Biología del Cáncer (PEBC) del Instituto de Investigación Biomédica de Bellvitge (IDIBELL), y posteriormente en el Instituto de Investigación contra la Leucemia Josep Carreras (IJC)

Dr. Manel Esteller Badosa
Director y Tutor

Vanessa Ortiz Barahona
Doctoranda

A vosotros, que no me veréis acabar esta etapa, va dedicada esta tesis

Os quiero

CONTENTS

CONTENTS

Index of contexts

Index of figures

Index of tables

Abbreviations

#	
5-FU	5-fluorouracil
5caC	DNA 5-Carboxylcytosine
5fC	DNA 5-Formylcytosine
5hmC	DNA 5-hydroxymethylcytosine
5mC	DNA 5-methylcytosine
6mA	DNA N6-methyladenosine

A	
AaRS	Aminoacyl-tRNA synthetase
ADC	Adenocarcinoma
ADP	Adenosine diphosphate
ALKBH1	AlkB Homolog 1, Histone H2A dioxygenase
ALKBH5	AlkB homolog 5, RNA demethylase
ALL	Acute lymphoblastic leukemia
ALYREF	Aly/REF export factor
ARE	AU-rich element
ATCC	American Type Culture Collection
ATP2B4	ATPase Plasma Membrane Ca2+ Transporting 4
ATX	Autotaxin
AZA	5-aza-2′-deoxycytidine or Decitabine
AZA-IP	5-azacytidine immunoprecipitation

B	
BER	Base excision repair
BRCA1	Breast cancer 1
BRG1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4
BSP	Bisulfite sequencing PCR

C	
CATSPERG	Cation Channel Sperm Associated Auxiliary Subunit Gamma
CBP	CREB binding protein
CC	Corpus callosum
CCDC9B	Coiled-Coil Domain Containing 9B
CDK1	Cyclin dependent kinase 1
CDKN1B	Cyclin dependent kinase inhibitor 1B
CDS	Coding sequence
CI	Confidence interval
circRNA	Circular RNA
CMCT	N-cyclohexyl-N′-(2-morpholinoethyl)-carbodiimide metho-p-toluenesulfonate
CNS	Central nervous system

CNV	Copy number variation
CRC	Colorectal cancer
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
CTNNB1	Catenin Beta 1

D

D-2-HG	D-2-hydroxyglutarate
ddNTP	Dideoxynucleotide triphosphate
DKC1	Dyskerin pseudouridine synthase 1
DNA	Deoxyribonucleic acid
DNMT1	DNA methyltransferase 1
DNMT2	TRNA Aspartic acid methyltransferase 1
DNMT3A	DNA methyltransferase 3A
DNMT3B	DNA methyltransferase 3B
DNQ	Deoxynyboquinone
dNTP	Deoxyribonucleotide triphosphate

E

eEF	Eukaryotic translation elongation factor
eIF3	Eukaryotic translation initiation factor 3
EMT	Epithelial-to-mesenchymal transition
eRF1	Eukaryotic release factor 1
EZH2	Enhancer Of Zeste homolog 2

F

FBL	Fibrillarin
FBS	Fetal bovine serum
FC	Fold-change
FDA	US Food and Drug Administration
FTO	Fat mass and obesity-associated protein
FTSJ1	FtsJ RNA methyltransferase 1

G

GAR1	H/ACA ribonucleoprotein complex subunit 1
GBM	Glioblastoma multiforme
gCIMP	Glioma CpG island methylator phenotype
GO	Gene ontology
GSEA	Gene set enrichment analysis

H

HAT	Histone acetyl transferase
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDAC	Histone deacetylase
HDGF	Heparin-binding growth factor

HDMT	Histone demethylase
HIF1α	Hypoxia-induced factor 1α
hm5C	RNA 5-hydroxymethylcytidine
HMT	Histone methyltransferase
HNRNP	Heterogeneous nuclear ribonucleoprotein
HR	Hazard ratio

I

IB-DNQ	Isobutyl-deoxynyboquinone
IC50	Half-maximal inhibitory concentration
IDH1	Isocitrate Dehydrogenase 1
IGF2	Insulin-like growth factor 2
IGF2BP	Insulin-like growth factor RNA binding protein
IRE	Internal ribosomal entry site

L

LC	Liquid chromatography
LGG	Low-grade glioma
LINE	Long interspersed nuclear element
lncRNA	Long non-coding RNA
LSD1	Lysine demethylase 1A
LSU	60S large ribosomal subunit

M

m^1A	RNA 1-methyladenosine
m^5C	RNA 5-methylcytosine
m^6A	RNA N6-methyladenosine
m^6Am	RNA N6-2′-O-dimethyladenosine
MBD	Methyl-CpG-binding Domain
MD	Molecular dynamics
MEM	Modified Eagle's Medium
meRIP	Methylated-RNA-immunoprecipitation
METTL14	Methyltransferase 14, N6-adenosine-methyltransferase subunit
METTL16	Methyltransferase 16, N6-methyladenosine
METTL3	Methyltransferase 3, N6-adenosine-methyltransferase complex catalytic subunit
MGMT	O-6-Methylguanine-DNA Methyltransferase
miCLIP	Methylation individual-nucleotide resolution cross-linking and immunoprecipitation
miRNA	MicroRNA
MLH1	MutL homolog 1
MLL	Lysine methyltransferase 2A
mRNA	Messenger RNA
MS	Mass spectrometry
MTERF4	Mitochondrial transcription termination factor 4
MYC	V-Myc avian myelocytomatosis viral oncogene homolog

N

NASH	Non-alcoholic steatohepatitis
ncRNA	Non-coding RNA
NHP2	H/ACA ribonucleoprotein complex subunit 2
Nm	Ribose 2'-O-methylation
NOP10	Nucleolar protein 10
NOP56	Nucleolar protein 5A (56kDa With KKE/D Repeat)
NOP58	Nucleolar protein 58
NQO1	NAD(P)H Quinone Dehydrogenase 1
NSD1	Nuclear receptor binding SET domain protein 1
NSUN	NOL1/NOP2/SUN domain

O

OP-Puro	O-propargyl-puromycin
OPC	Oligodendrocyte precursor cells
ORF	Open reading frame
OS	Overall survival
OXPHOS	Oxidative Phosphorylation System

P

p21	Cyclin dependent kinase inhibitor 1A
PD-1	Programmed cell death
PD-L1	Programmed death-ligand 1
PFS	Progression-free survival
PGC-1α	PPARG Coactivator 1 Alpha
PHD	Plant homeodomain
PI	Propidium iodide
piRNA	PIWI-interacting RNA
PKM2	Pyruvate Kinase Isozymes M1/M2
POU2F3	POU domain class 2 homeobox
PRRC2A	Proline rich coiled-coil 2A
PTC	Peptidyl-transferase center
PUS	Pseudouridine synthase

R

RB1	Retinoblastoma 1
RBP	RNA binding protein
RFM	Rossman-fold domain
RIP-Seq	RNA immunoprecipitation sequencing
RIZ1	Retinoblastoma protein-interacting zinc finger protein
RNA	Ribonucleic acid
RNY1	Y1 Small Cytoplasmic RNA (Associated with Ro protein)
ROS	Reactive oxygen species
RP	Ribosomal protein
RPL30	Ribosomal protein large subunit 30

RPL41	Ribosomal protein large subunit 41
RPMI	Roswell Park Memorial Institute
RPPH1	Ribonuclease P RNA component H1
RRM	RNA-recognition motif
rRNA	Ribosomal RNAs
RT-qPCR	Real-time quantitative PCR

S

SA-β-gal	Senescence-associated β-galactosidase
SAM	S-adenosyl methionine
scaRNA	Small Cajal Body-specific RNA
SCC	Squamous cell carcinoma
SCLC	Small cell lung cancer
SCR	Scramble
shRNAs	Short hairpin RNAs
SILAC	Stable isotopic labeling of amino acids in cell culture
SMRT	Single-molecule real-time sequencing
snRNA	Small nuclear RNA
SNU13	Small nuclear ribonucleoprotein 13
SSU	40S small ribosomal subunit

T

TCGA	The Cancer Genome Atlas
TDG	Thymine DNA glycosylase
TERC	Telomerase RNA component
TERT	Telomerase reverse transcriptase
TET	Ten-Eleven translocation methylcytosine dioxygenase
THOC1	THO Complex 1
THOC7	THO Complex 7
TNM	Tumour-node-metastasis
TP53	Tumor pprotein
TREX	Transcription-and-export complex
TRF	tRNA-derived small RNA fragment
tRNA	Transfer RNA
TSS	Transcription start site

U

UPR	Unfolded protein response
UTR	Unstranslated region

V

VTRNA	Vault RNA

W

WES	Whole exome sequencing
WHO	World Health Organisation

X

X-DC	X-linked dyskeratosis congenita
XCI	X-chromosome inactivation

Y

YBX1	Y-Box binding protein 1
YTH	YT521-B homology domain

Greek alphabet

α-KG	α-ketoglutarate
Ψ	Pseudouridine

INTRODUCTION

INTRODUCTION

Cancer

The term ‘cancer’ designates a set of diseases characterized by two properties: an aberrant tissue growth due to an uncontrolled cell division, and its ability to spread and colonize other tissues. In multicellular organisms, all cells are required to behave in a tightly coordinated manner, including dividing, differentiating or dying, in order to sustain organism survival. Molecular alterations that bypass this cellular homeostasis may give rise to cancer, in a multistep process that has been termed tumorigenesis. Molecular changes driving tumorigenesis, either genetic or epigenetic, confer cells selective advantages by allowing them to divide faster and survive more readily through adapting better to the environment, which progressively turn them carcinogenic and invasive [1].

Cancer is a complex disease, comprising more than one hundred types of malignancies, every of them with a different pattern of molecular alterations. In an attempt to rationalize tumoral diversity, Hanahan and Weinberg proposed, in 2000, that all tumors display six common capabilities, regardless their specific mutational backgrounds: maintaining cell growth through sustaining proliferative signaling and evading growth suppressors, avoiding cell death, enabling replicative immortality, promoting angiogenesis, and activating invasion and metastasis programs [2] (**Figure 1**).

As the cancer research field evolved so did our knowledge, and these authors upgraded the hallmarks list in 2011 with two additional capabilities: reprogramming of the energy metabolism to sustain cell proliferation and evading immune destruction [3] (**Figure 1**).

Moreover, they proposed that two additional tumoral characteristics enable the acquisition of all cancer hallmarks. One is genomic instability, which favors mutability, providing cells with genetic and epigenetic alterations potentially beneficial for tumor progression. Molecular alterations promoting malignancy affect mainly proto-oncogenes and tumor suppressor genes. Proto-oncogenes are activated into oncogenes through a dominant gain-of-function alteration in one allele, occurring normally through genetic mutations, gene amplification and chromosomal rearrangements. However, inactivation of most tumor suppressor genes needs a recessive loss-of-function alteration, achieved

through loss of heterozygosity, genetic mutations or epigenetic silencing through promoter DNA methylation [3].

The second enabling characteristic is the inflammatory state driven by immune cells in order to eradicate the tumor, which several evidence propose that could paradoxically exert a tumor-promoting role. Inflammation contributes to the acquisition of multiple cancer hallmarks by being a source of biomolecules such as growth and survival factors, promoting cancer cell proliferation and cell death avoidance, or proangiogenic factors and matrix-remodeling enzymes, facilitating angiogenesis and promoting invasion and metastasis. In addition, inflammatory cells can contribute to the genomic instability of cancer cells through release of mutagenic substances such as reactive oxygen species (ROS) [3] (**Figure 1**).

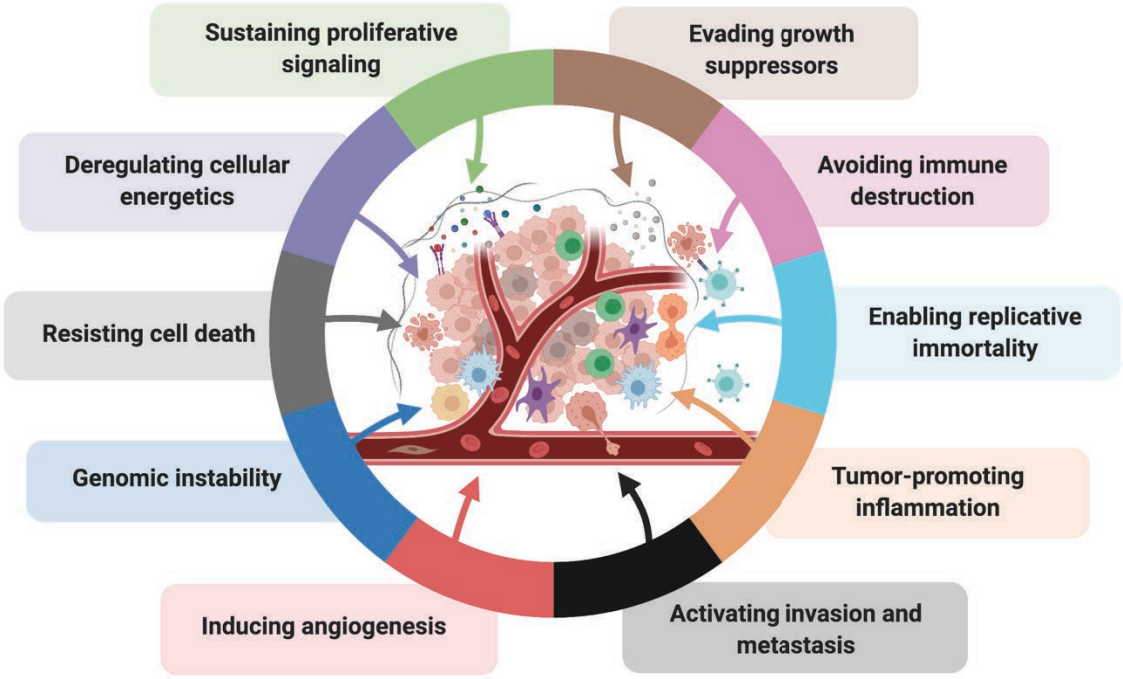


Figure 1. Hallmarks of cancer. Common capabilities that normal cells acquire during the tumorigenesis process, including eight core hallmarks and two enabling characteristics. Based on Hanahan and Weinberg, 2011 [3].

Epigenetics

Development of multicellular organisms is a well-orchestrated process that begins with a totipotent zygote and leads to over 200 different cell types. Despite sharing the same DNA sequence, each cell type expresses specific sets of genes that define its unique functions. This different way of interpreting the genetic information between cell types is accomplished by epigenetic mechanisms. The term 'epigenetics' was first coined by Conrad Waddington in 1939, who defined it as 'the casual interactions between genes and their products, which bring the phenotype into being' [4], definition that was further reformulated to all heritable changes in the phenotype occurring without alterations on the DNA sequence [5].

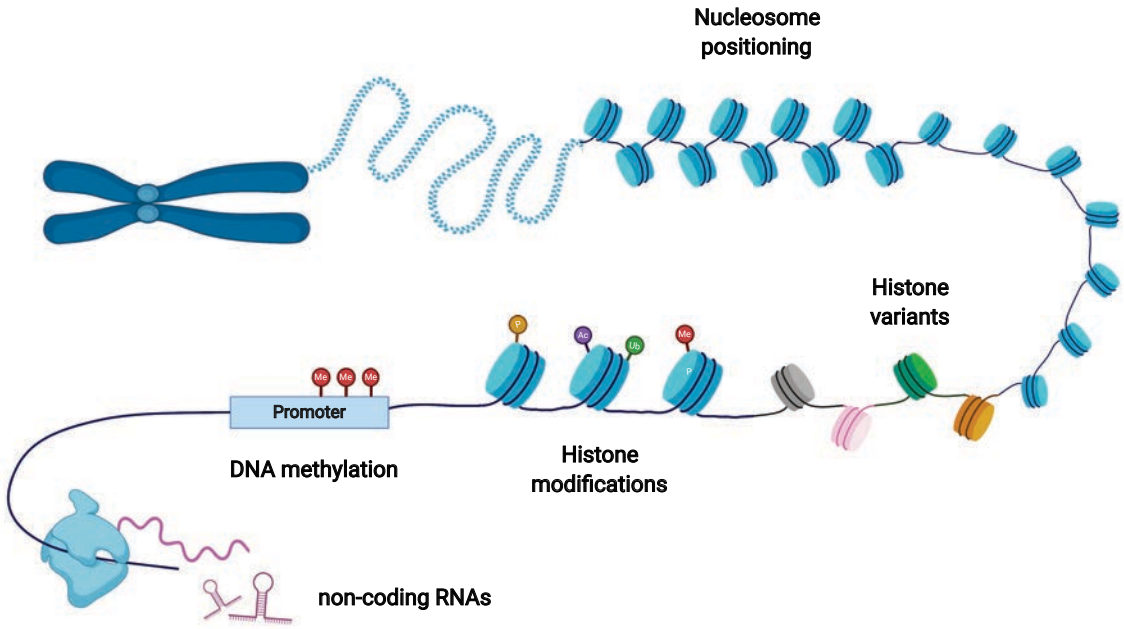


Figure 2. Overview of epigenetic mechanisms. The three classic epigenetic mechanisms consist of DNA methylation, which refers to the addition of a methyl group to the 5' position of a cytosine in the context of a CpG dinucleotide; post-translational modifications of the amino-terminal tails of histone proteins; and the role exerted by non-coding RNAs (ncRNAs) in gene expression. However, the classic picture of the epigenetic landscape can be further extended by the addition of two extra mechanisms regulating chromatin compaction: nucleosome positioning and the substitution of core histones with histone variants, both dictating DNA accessibility.

The classic epigenetic mechanisms are DNA methylation and covalent histone modifications, further discussed, and non-coding RNAs (ncRNAs). Overall, epigenetics affects chromatin compaction degree, regulating the access of different molecular machineries to genes, dictating gene expression or repression. Taking this notion into account, other epigenetic layers affecting chromatin compaction could be added to the classic picture: nucleosome positioning and histone replacement [6] (**Figure 2**), although they will not be further discussed.

DNA methylation

DNA methylation generally consists of the covalent attachment of a methyl group ($-CH_3$) to the fifth carbon of the pyrimidine ring of a cytosine, resulting in 5-methylcytosine (5mC). It was the first covalent modification discovered on the DNA [7], and most likely the most extensively characterized epigenetic mark.

DNA methylation is carried out by DNA methyltransferases (DNMTs), using the cofactor S-adenosyl-L-methionine (SAM) as the methyl group donor. There are three DNMTs in humans. DNMT3A and DNMT3B are responsible for *de novo* methylation, namely, methylation of previously unmethylated regions [8]. Both enzymes are essential during early development, with their inactivation leading to early embryonic lethality [9]. Moreover, DNMT3A/B methyltransferase activity is enhanced by the catalytically inactive partner DNMT3L [10]. On the other hand, DNMT1 is in charge of the maintenance of DNA methylation patterns across cell divisions. DNMT1 binds hemimethylated DNA double strands after replication, guided by the UHRF1 protein, and 'copies' the methylation pattern from the parental strand to the daughter strand [11] (**Figure 3**). UHRF1 expression is regulated in a cell cycle-dependent manner, to ensure the correct timing of maintenance DNA methylation [12].

DNA demethylation can occur either by passive or active mechanisms. Passive demethylation takes place upon DNMT1 inhibition, which leads to loss of DNA methylation after successive cell divisions. Active demethylation is an energy-demanding process, as it involves the catalytic removal of the methyl group from the modified cytosine, and it is carried out by the ten-eleven translocation (TET) family of proteins (TET1-3) [13]. TET-mediated DNA demethylation is a multistep process that sequentially oxidizes 5mC, first to 5-hydroxymethylcytosine (5hmC), which is in turn transformed to 5-formylcytosine (5fC) and finally to 5-carboxylcytosine (5caC).

These three intermediates can be eliminated through replication-dependent dilution, or, additionally, 5fC and 5caC can be excised by the thymine DNA glycosylase (TDG) enzyme, which leaves an abasic site further repaired by the base excision repair (BER) complex. Both ways result in restoration of the cytosine unmethylated status [13] (**Figure 3**).

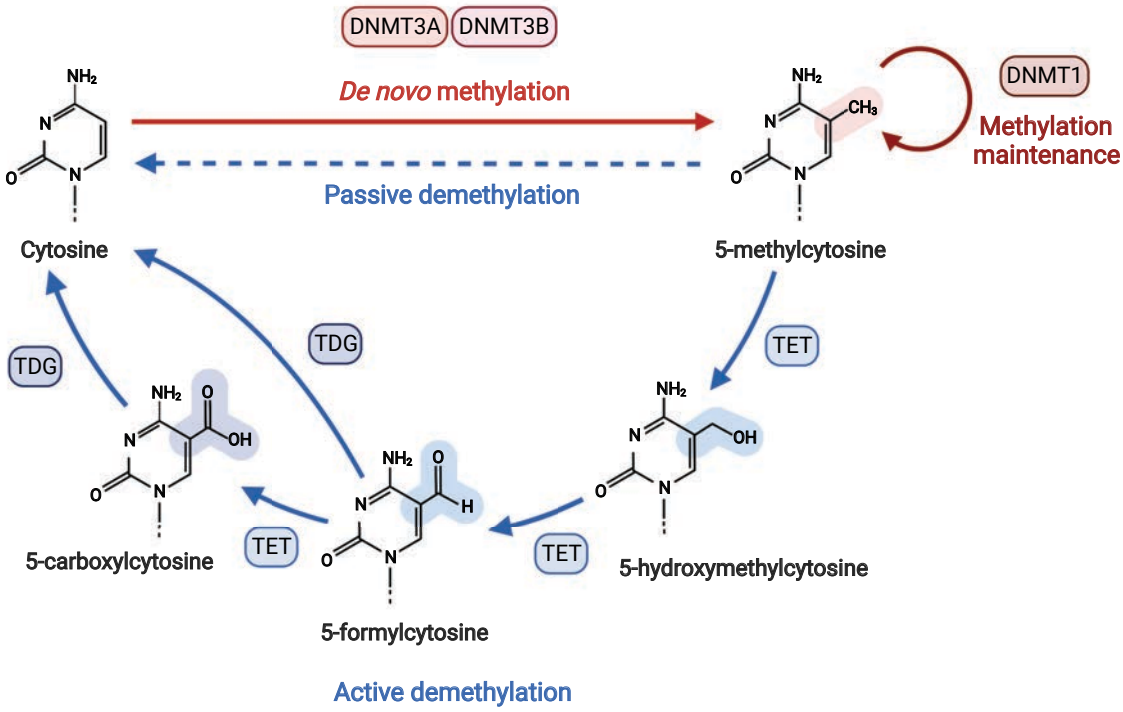


Figure 3. DNA methylation establishment, maintenance and demethylation. In humans, *de novo* DNA methylation is carried out by DNMT3A and DNMT3B enzymes, while maintenance of methylation patterns across cell divisions is performed by DNMT1. Active DNA demethylation from 5mC to 5fC or 5caC is mediated by TET enzymes, and final restoration of the cytosine unmethylated status is achieved by the TDG enzyme.

Distribution of DNA methylation along the genome is not random, but mainly restricted to the context of CpG dinucleotides. Such dinucleotides are especially enriched at gene promoters, where they cluster forming 'CpG islands'. These islands are defined as genomic regions longer than 200 bp, with a GC content higher than 50%, and an observed versus expected CpG ratio greater than 0.6. Importantly, CpG islands are found in approximately 70% of annotated gene promoter regions [14], with their hypermethylation resulting in the transcriptional silencing of gene expression. Functionally, DNA methylation at promoter regions impacts chromatin structure inducing a repressive state, characterized by refractoriness to nuclease digestion and enrichment in hypoacetylated core histones, by different mechanisms. One is the steric impediment exerted by methylation, that directly inhibits gene expression by interfering with

transcription factors binding. Second, by recruiting Methyl-CpG-binding Domain (MBD) proteins, which indirectly repress transcription via interaction with co-repressors that induce a condensed chromatin state (such as histone modifiers or chromatin remodeling complexes) [15].

DNA methylation plays crucial roles during development, differentiation and tissue homeostasis. Some functions for *de novo* DNA methylation include: genomic imprinting, for ensuring monoallelic gene expression in a parent-of-origin-specific manner [16]; X-chromosome inactivation (XCI), to compensate X-linked gene dosage between sexes [17]; germline-specific gene silencing, for the establishment of proper tissue-specific patterns in differentiated cell types [18]; and inactivation of repetitive genomic regions, such as pericentromeric, centromeric and telomeric areas and transposable elements, contributing to chromosomic stability [19].

Finally, it is worth mentioning that cytosines are not the only modified nucleotide on the DNA. With the development of deep sequencing, another nucleotide, adenine, has recently been found to be methylated in a number of eukaryotic genomes [20–22]. However, N6-methyladenosine (6mA) mark was already known, as it is a prevalent DNA modification in prokaryotes, being involved in the host's defense system [23]. Whether 6mA is present in the human genome is still under debate.

Histone modifications

Histones are low molecular weight and highly basic proteins that help organize DNA into nucleosomes, the basic units of chromatin conformation. Nucleosomes are composed of two subunits, every of them containing four canonical histones (H2A, H2B, H3, H4), wrapped around approximately 147 DNA bp. Histones are globular proteins but possess intrinsically unfolded N-terminal tails susceptible to being post-translationally modified. Histone modifications impact their interaction with DNA, affecting chromatin compaction level and, therefore, transcriptional activity of genes. Modifications disrupting histone-DNA interaction cause unwinding of nucleosomes, giving rise to an 'open' chromatin state (euchromatin) that allows binding of the transcriptional machinery for gene expression. Conversely, modifications strengthening histone-DNA interactions give rise to condensed chromatin (heterochromatin), transcriptionally silent [24]. Additionally, some modifications exert their role over the chromatin state by recruitment of chromatin factors [25].

There are several histone modifications, including acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation; being the most studied acetylation and methylation. Such modifications are not static, but dynamically regulated depending on the cellular context by a plethora of histone modifiers. Enzymes catalyzing the deposition of the mark onto histone tails are commonly referred to as 'writers', enzymes removing modifications, 'erasers', and proteins recognizing specific histone modifications and recruiting other molecular machineries to those sites, 'readers' [26].

Writers of histone acetylation are called histone acetyl transferases (HATs), enzymes that employ Acetyl-CoA as the acetyl group donor, and catalyze its transfer to lysines (K) on histone tails. This modification neutralizes lysine's positive charge, weakening the interaction between histones and DNA. Therefore, euchromatin is usually associated with high levels of histone acetylation. On the other hand, erasers of acetylation are called histone deacetylases (HDACs) and exert the opposite role over chromatin compaction [27].

Regarding histone methylation, it mainly takes place on lysine and arginine amino acids, catalyzed by histone methyltransferases (HMTs), which employ SAM as cofactor, and it is removed by histone demethylases (HDMTs). Adding complexity, lysines can be mono-, di- or trimethylated, and arginines mono-, and dimethylated, being HMTs and HDMTs highly specific in terms of substrate specificity and even its degree of methylation. While histone acetylation is normally associated with transcriptional activation, histone methylation does not alter histone charge and can play a dual role (activating or repressing gene expression), depending on the residue and its position, mainly mediated by reader proteins. Methylation readers usually display Chromodomains, Tudor domains, plant homeodomains (PHD) and PWWP domains [27]. As an example, H3K4, H3K36 and H3K79 trimethylations are associated with transcriptional activation, while H3K9, H3K27 and H4K20 monomethylations are responsible for transcriptional repression [28].

Epigenetics and cancer

As previously mentioned, cancer cells acquire several molecular alterations along their malignant transformation, including those affecting epigenetic mechanisms and players. Cancer-associated alterations in the epigenetic landscape include a global DNA hypomethylation together with specific hypermethylation events at promoters of tumor suppressor genes, genomic mutations on genes coding for histone modifiers and chromatin remodeling complexes, and genetic and epigenetic lesions affecting ncRNAs expression.

The lower level of global DNA methylation in tumoral genomes, compared to those of their normal-tissue counterparts, was in fact, one of the first epigenetic alterations found in cancer [29]. The molecular consequences are chromosomal instability, loss of imprinting, reactivation of transposable elements and of silenced oncogenes. As already mentioned, demethylation of repetitive sequences of satellite DNA can favor mitotic recombination, giving rise to chromosomal rearrangements, amplifications or deletions, further promoting tumor progression [30]. Moreover, loss of genomic imprinting increases susceptibility to cancer. Patients of the hereditary Beckwith-Wiedemann syndrome, characterized by a loss of imprinting at the insulin-like growth factor 2 (*IGF2*) gene, suffer an increased risk of cancer [31] . Likewise, loss of *IGF2* imprinting is a described risk factor for colorectal cancer (CRC) [32,33], and contributes to the development of Wilm's tumor, a pediatric nephroblastoma [33]. Loss of global DNA methylation in cancer cells can, besides, reactivate transposable elements such as LINEs and Alu sequences, whose retrotransposition leads to further genomic lesions potentially advantageous for the tumor [34]. And finally, hypermethylation-mediated silenced oncogenes can get active upon global demethylation of the cancer genome [35,36].

Strikingly, despite the global hypomethylation observed in tumoral genomes, the most common epigenetic alteration observed in cancer-related genes is their promoter hypermethylation. In fact, cancer-associated promoter hypermethylation is believed to affect between 5–10 % of CpG island-containing promoters [37,38]. Examples of genes silenced through promoter hypermethylation are found in all regulatory pathways governing cell proliferation and homeostasis, in order for malignant cells to sustain their abnormal growth [38]. Interestingly, many genes that are normally found mutated in specific tumor types, are also found hypermethylated in those types, highlighting the selective advantage that their inactivation involves during clonal selection. Some examples are

BRCA1 hypermethylation in breast and ovarian tumors, *RB1* in retinoblastoma, and *MLH1* in colon cancer [38].

Mutation, translocation, amplification or deletion events on histone modifier genes are known to profoundly impact the epigenetic landscape during carcinogenesis. Regarding histone acetylation, key examples are exerted by p300 and CBP, two HATs carrying out acetylation at multiple positions in all the four core histones, as well as at other non-histone targets, being responsible for both, activating and repressing gene translation. Several mutations and translocations on these two genes have been reported in different cancer types. In leukemia, p300 is fused to MLL through a genomic translocation, which restricts its activity to MLL target genes, leading to a genome-wide histone hypoacetylation. Other tumorigenic effects derived from genetic lesions on p300 and CBP are caused by loss of their acetylation of p53 and BCL6, which abrogates their repressive roles on cell-cycle arrest and apoptosis [39]. Deregulated global histone acetylation patterns may also be achieved by impaired HDACs expression, observed in multiple tumor types [39]. HMTs and HDMTs are also deregulated in cancer, affecting global histone methylation levels. For instance, overexpression of LSD1 and EZH2 in prostate cancer [39]. Finally, not only genetic, but an epigenetic deregulation has been reported to affect histone modifiers expression in cancer. Some instances are DNA methylation-based inactivation of the HMTs RIZ1, in several cancer types, and NSD1 in neuroblastoma and glioma [38].

Chromatin remodeling complexes are also found deregulated in human cancer. For instance, genes coding for components of the SWI/SNF complex show loss-of-function mutations in more than 20% of all tumor types, being BRG1 (*SMARCA4*) the most frequently mutated subunit. Their recurrent loss in cancer evidences their tumor suppressive roles [40].

Non-coding RNAs are also targets of cancer-associated deregulation, both by genetic and epigenetic mechanisms. An important number of micro-RNA (miRNA) genes undergo copy number variations (CNV) in different cancer types, either amplification or deletion. Some examples comprise miR-21, miR-17 and miR-155 genes, frequently amplified in lung cancer and B-cell lymphomas, or miR-15a and miR-16a, usually deleted in patients with chronic lymphocytic leukemia [40]. Epigenetic mechanisms also serve for miRNA repression in cancer. Examples constitute promoter hypermethylation of miR-200 and miR-34 families of miRNAs in several tumor types, negative regulators of the epithelial-to-mesenchymal transition (EMT) [40].

Epitranscriptomics

Technological advances, such as mass spectrometry and high-throughput sequencing, have allowed researchers to perform large-scale analysis of both transcriptomes and proteomes from different organisms. Surprisingly, comparisons between mRNA and protein concentrations within organisms have brought to light too low correlation coefficients [41]. Therefore, protein abundances can only be partially explained by transcript concentrations, implying the existence of further regulatory mechanisms for protein expression. Some of these mechanisms are well established (stability and degradation of mRNAs, posttranslational modifications of proteins and programmed degradation of the proteins themselves), and some have just recently gained attention: RNA modifications.

Although several research efforts during the last decade have confirmed that RNA molecules are decorated with over 170 different chemical modifications [42], these marks have been known to exist in RNA since nearly 70 years ago, with the discovery of pseudouridine (Ψ) in tRNA and rRNA [43,44]. In fact, pseudouridine (Ψ), which is the most abundant RNA modification and has been termed the fifth nucleotide, was discovered only 7 years after the first described DNA modification (5mC) [7]. However, for many decades RNA modifications were just considered to be static features of 'infrastructural' ncRNAs, affecting their biogenesis, stability and function, and therefore, with no direct role in gene expression. Curiously, not long after discovery of pseudouridine (Ψ) the first modifications on mRNAs were also described, including the 5'cap and the polyA tail [45,46]; and in 1975, the first internal mRNA modification, N^6-methyladenosine (m^6A), was discovered [47].

The field remained mainly unattended until 2011, when the discovery that FTO could demethylate m^6A-modified RNAs [48] dramatically changed the static vision of RNA modifications. Ever since, the discovery of several machineries dedicated to dynamically regulate RNA modifications led to the coining of the term 'RNA epigenetics', and further on, 'epitranscriptomics' [49]. Similar to DNA epigenetics, proteins installing RNA modifications are called writers, proteins removing these marks are called erasers, and proteins decoding the RNA chemical repertoire are called readers. Recent technological and methodological advances, including the development of transcriptome-wide sequencing techniques, have allowed to widely increase the catalog of modifications on mRNAs, highlighting epitranscriptomics direct implication in gene expression [49,50].

However, the epitranscriptomics field still faces important challenges regarding reliability of every new discovery, due to the low abundance of some nucleotide modifications, its dynamic nature, and the technical limitations of detection techniques. In fact, there are several examples of overestimated abundances of some chemical marks, due to high false-positive rates in initially reported maps [51,52]. As the only way to study epitranscriptomic mechanisms and their phenotypic consequences is to carefully determine single epitranscriptomes in a reproducible manner, it is important to consider the inherent limitations of every detection technique.

Best characterized RNA modifications

Pseudouridine

Pseudouridine (Ψ), the 5′-ribosyluracil isomer of uridine (**Figure 4**), was the first modification of RNA unraveled [43,44], and is found in all three domains of life [53]. It was first discovered in tRNA and rRNA, and the development of transcriptome-wide Ψ-detection methods by coupling CMCT [*N*-cyclohexyl-*N*′-(2-morpholinoethyl)-carbodiimide metho-*p*-toluenesulfonate] treatment-mediated generation of reverse transcription stops with primer-extension and deep sequencing, we currently know it is present in small nuclear RNAs (snRNAs), small Cajal Body-specific RNAs (scaRNAs), miRNAs, long non-coding (lncRNAs) and mRNAs [53].

Pseudouridine writers can be classified into two distinct groups, based on if they carry out RNA-guided or RNA-independent pseudouridine isomerization. RNA-guided pseudouridylation is carried out by box H/ACA small ribonucleoproteins (snoRNPs), complexes formed by a unique box H/ACA small nucleolar RNA (snoRNA), and a core of four proteins (in humans, comprised by DKC1, NHP2, NOP10 and GAR1). The box H/ACA snoRNA, upon base-pairing interactions with the RNA substrate, demarcates the target uridine (U), which is further isomerized to Ψ by the catalytic activity of DKC1 [53]. Contrarily, RNA-independent pseudouridylation is carried out by single writers, named pseudouridine synthases (PUS), responsible for both, substrate recognition and U-to-Ψ isomerization [53]. In eukaryotes, over fourteen PUS enzymes have been described so far [54]. Conversely, no erasers exist for this modification in mammals, where Ψ constitutes an irreversible mark, and no readers have been described yet.

Depending on the RNA specie, pseudouridine plays different biological roles. In rRNAs, Ψs are found in all functionally important ribosomal domains [mRNA decoding site, peptidyltransferase center (PTC), tRNA binding site], playing important roles in correct ribosome assembly and function [55]. In tRNAs, Ψs can be found at the anticodon stem and loop, and at the D stem, involved in stabilizing tRNA tertiary structure and codon–anticodon base-pairing [56]. In mRNAs, some of the roles described for Ψs are increasing transcript stability under stress conditions [57], promoting stop codon read-through [53] and mediating nonsense-to-sense codon conversion [58].

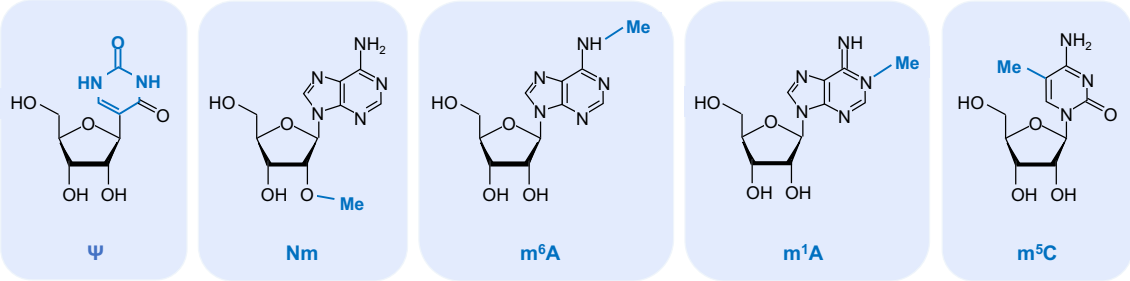


Figure 4. Chemical structure of some of the best characterized RNA modifications. Ψ, pseudouridine; Nm, ribose 2'-O-methylation; m^6A, N6-methyladenosine; m^1A, N1-methyladenosine; m^5C, 5-methylcytosine.

Ribose 2'-O-methylation

Ribose 2'-O-methylation (N_m) is the most widespread RNA modification, being found in virtually all RNA species. It consists of the addition of a methyl group to the ribose 2'-hydroxyl moiety (-OH) of a nucleotide (**Figure 4**), occurring thus, in all canonical and non-canonical nucleotides (A_m, G_m, U_m, C_m, $Ψ_m$) [59]. Despite ribose methylations were first described back in the 60's [60], their study remained hampered for several decades due to the challenges in mapping this mark [61].

Similar to pseudouridylation, writers of 2'- O-methylations fall into two groups based on their dependency on RNA guidance, although all of them make use of SAM as the methyl-donor. RNA-guided 2'-O-methylation is carried out by box C/D small ribonucleoproteins (snoRNPs). The RNA responsible for substrate recognition consist of either a box C/D snoRNA, for rRNA and tRNA methylation, or a scaRNA for methylation of snRNAs in Cajal bodies [62]. The protein core of snoRNPs for 2'-O-methylation is comprised by FBL, the catalytic unit; SNU13, for RNA binding, and NOP56 and NOP58

as scaffold proteins [62]. Most mRNAs, miRNAs, piwi RNAs (piRNAs), some tRNAs and all bacterial rRNA and tRNAs, are in turn, 2'-O-methylated by single writers [63,64].

2′-*O*-methylation, is, together with pseudouridine, the most abundant modification in rRNA. The number of N_m residues increases with evolutionary complexity, from around 10 N_m in *E. coli*, to more than 100 in human rRNA. They are located in functionally important domains of the ribosome, such as the PTC and the mRNA decoding center, which underscore their importance for correct translation [65]. Other functions attributed to this RNA modification are regulation of RNA stability and mRNA splicing, and mediating innate immune response [66].

N^6-methyladenosine

N^6-methyladenosine (m^6A) is the best understood and most abundant internal modification in mRNAs, present from viruses to higher mammals (**Figure 4**). Quantitative approaches suggest that around 0.2–0.6% of all mRNA adenines are N^6-methylated, which results in approximately 1–5 m^6As per mRNA [67]. However, high-throughput sequencing methods have proven its presence in ncRNAs such as rRNAs, miRNAs, lncRNAs, and circular RNAs (circRNAs) [68]. m^6A distribution, rather than random, is enriched near stop codons and at 3′ untranslated regions (3′UTRs), but also found within long exons, introns, and at 5′ untranslated regions (5'UTRs) [69–71].

This RNA modification is dynamically and reversibly regulated by an expanding plethora of writers, readers and erasers. m^6A deposition is carried out by a multicomponent writer complex, composed of METTL3, the SAM-dependent catalytic unit; METTL14, responsible for substrate recognition, and a number of adapter proteins (RBM15, WTAP, KIAA1429, ZC3H13), responsible for complex recruitment and formation at the target site [72]. And more recently, a single enzyme, METTL16, has been described as a m^6A methyltransferase in human cells [73]. Regarding erasers, so far two m^6A demethylases have been identified: FTO and ALKBH5, both members of the nonheme Fe (II)/α-ketoglutarate-dependent (AlkB) family of dioxygenases [72]. Finally, several binding proteins have been proposed to decode m^6A RNA modifications. Some m^6A readers contain the YTH RNA binding domain, including YTHDC1 and YTHDC2, and YTHDF1, YTHDF2 and YTHDF3. Other families of m^6A readers are comprised by insulin-like growth factor RNA binding proteins (IGF2BP1, IGF2BP2, and IGF2BP3), heterogeneous nuclear ribonucleoproteins (HNRNPC and HNRNPA2/B1), and single readers like eIF3 or PRRC2A [72].

m^6A biological role depends on the reader decoding it, through a m^6A-interaction dependent manner. For example, YTHDC proteins, mainly found in the nucleus, mediate pre-mRNA alternative splicing and mRNA export to the cytoplasm [74,75]. The YTHDF family, however, is predominant in the cytoplasm and differently regulates the fate of m^6A-modified mRNAs. While YTHDF1 and YTHDF3 increase mRNA translational efficiency [76], YTHDF2 promotes mRNA decay through the CCR4-NOT deadenylase complex [77]. Another m^6A reader involved in translation efficiency is eIF3, which increases translation of 5'UTR m^6A-modified mRNAs through a cap-independent manner [71]. Similar to mRNAs, m^6A readers affect several aspects of ncRNA life cycles such as processing and maturation, location, stability or degradation. Just an example is the HNRNP family involvement in processing of precursor miRNAs [78].

N^1-methyladenosine

N^1-methyladenosine (m^1A) RNA modification (**Figure 4**) was originally detected in 1961 [79], and so far, has been found in cytoplasmic and mitochondrial tRNAs (mt-tRNAs), rRNAs and nuclear- and mitochondrial-encoded mRNAs [80]. m^1A is quite abundant in tRNAs from all three domains of life, where can appear at positions 9, 14, 22, 57 and 58 of cytoplasmic molecules, and at 9 and 58 in mt-tRNAs [81]. These tRNA methylations are described to increase tRNA structural stability and induce tRNA correct folding. Contrary to the high abundance of m^1A modifications in tRNAs, m^1A is only present on mammalian mRNAs at a limited number of sites, and at many of them, at a low stoichiometry [82]. The presence of this methyl group on mRNAs prevents the effective translation of modified codons as it disrupts normal Watson-Crick base pairing [83].

m^1A writers belong to either the Rossman-fold (RFM) or the SPOUT superfamily of methyltransferases, and they all use SAM as the methyl donor. In human, two members of the RFM family, TRMT61A and TRMT6, are responsible for establishing m^1A at the 58 position of cytoplasmic and mitochondrial tRNAs [84,85], and one member of the SPOUT family, TRMT10C, catalyzes m^1A at mt-mRNAs, and at position 9 of mt-tRNAs [83]. Regarding erasers, similar to m^6A, two members of the nonheme Fe (II)/α-ketoglutarate-dependent (AlkB) family of dioxygenases have been shown to demethylate m^1A: ALKBH1 and ALKBH3 [86,87]. No m^1A readers have been described so far.

5-methylcytosine

5-methylcytosine (m^5C) is one of the common modifications between DNA and RNA (**Figure 4**), whose presence in RNA was postulated only 8 years after its discovery in DNA [7,88]. Although m^5C presence in mRNA was found relatively soon after [89], its low prevalence in this RNA specie is the reason why most of our current knowledge about m^5C RNA methylations derives from studies on tRNA and rRNA. While the emergence of transcriptome-wide m^5C-mapping techniques was thought to come fill the research gap, the reality is that many independent studies have proposed substantially different results regarding abundance and presence of m^5C in mammalian mRNAs [90–92]. Current data suggest the presence of m^5C in several hundreds of mRNAs per tissue, the majority at low stoichiometries [93].

Obtaining precise m^5C maps, thus, has shown to be challenging. Error sources can be found at many levels, ranging from the biology of the RNA molecule itself, to the experimental approach and the data analysis method.

Methods for detecting 5-methylcytosine RNA methylation

Current methods for detecting m^5C RNA methylation in a transcriptome-wide manner can be first classified into two major groups: techniques not allowing to inform about the location of the methyl group within the molecule sequence, and those allowing site-specific determination. The first group mainly consist of chromatography-based methods, especially liquid chromatography (LC), alone or in combination with mass spectrometry (MS). LC-MS methods are highly accurate, quantitative and sensitive, allowing for detection of RNA modifications in the femto- to attomol range [51,94]. However, the drawback of these approaches is that they lose sequence information, as they require enzymatic digestion of the RNA to single nucleotides. On the other hand, techniques providing sequence context of the m^5C mark can be further classified into antibody-based, and antibody-free methods (**Figure 5**) and depend on high-throughput sequencing.

Antibody-based methods for m^5C detection

meRIP-seq

Methylated-RNA-immunoprecipitation (meRIP-seq) relies on RNA immunoprecipitation (RIP) using an antibody that recognizes the modified nucleotide and is followed by massive sequencing of the retrieved RNA (**Figure 5**). It was firstly developed in 2012 to identify the human and mouse m^6A methylomes using a specific anti-m^6A antibody [68], and the use of an antibody against m^5C has already allowed to detect m^5C-modified transcripts across the three domains of life [95]. The major caveat of this technique is its dependency on antibody specificity, as antibody cross-reactivity accounts for the main reason of false-positive results.

AZA-IP and miCLIP

While meRIP-seq detect m^5C-modified transcripts without the need to know the responsible methyltransferase enzyme, miCLIP and AZA-IP methods are mechanism-based techniques. They take advantage of the catalytic mechanism of RNA methyltransferases (RNA-MTs), which includes a transient covalent linkage between the methyltransferase and the target cytosine. AZA-IP makes use of an RNA-MT suicide inhibitor, 5-azacytidine, which locks the enzyme into an enzyme-substrate covalent adduct during the catalysis, preventing its release. Then, an antibody against the RNA-MT is used to immunoprecipitate the protein covalently bound to the target RNA, subsequently sequenced (**Figure 5**) [96]. This technique determines the exact target site of the m^5C RNA-MT as AZA-IP leaves a C>G transversion at the target cytosine. However, it has some limitations: detection of C>G transversion is not quantitative, and its sensitivity is limited, as 5-azacytidine is toxic and cells should not be treated for long periods, reducing its probability to be incorporated, especially for low-abundant RNAs.

Methylation individual-nucleotide-resolution cross-linking and immmunoprecipitation (miCLIP) makes use, instead of 5-azacytidine, of a point-mutated RNA-MT at the cysteine responsible for releasing the enzyme from the target RNA, resulting in the same covalently linked RNA-enzyme adduct. Then, similarly to AZA-IP, the RNA-MT is immunoprecipitated and the bound RNA is subjected to deep sequencing (**Figure 5**) [97]. As it depends on the overexpression of a mutant m^5C RNA-MT, drawbacks of this technique include artifacts in the methylation pattern assessed.

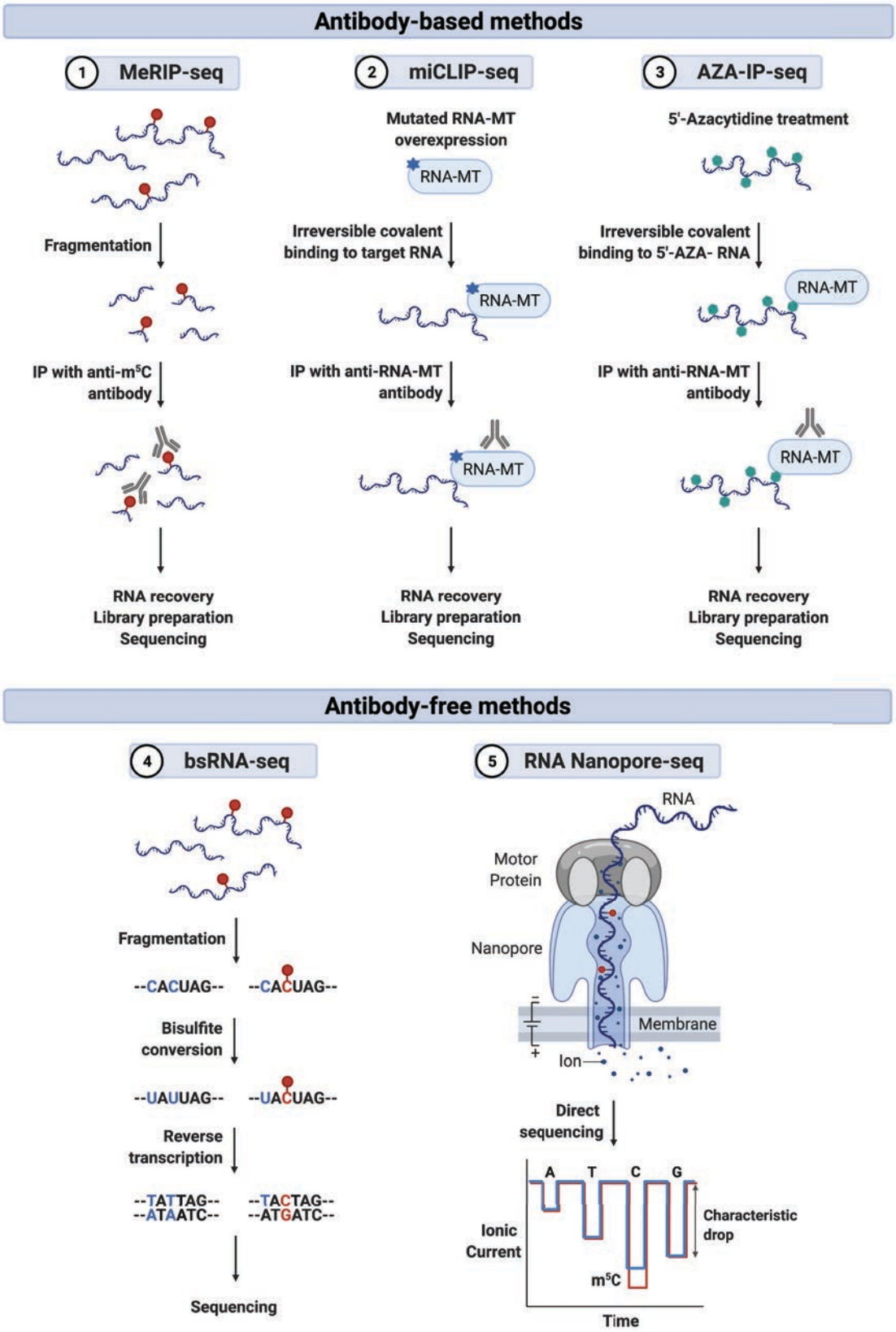


Figure 5. Schematic representation of techniques used for transcriptome-wide detection of m^{5}C modifications in RNA. Top, methods relying on antibody-based enrichment for m^{5}C-methylated RNAs. Bottom, methods not requiring immunoprecipitation of m^{5}C-methylated RNAs.

Antibody-free methods for m^5C detection

bsRNA-seq

Bisulfite sequencing was first developed in 1994 to study cytosine methylation in DNA [98]. The principle behind this technique is the fact that cytosine and 5-methylcytosine nucleotides react differently to an acidic sodium bisulfite solution. Cytosines react with bisulfite ions (HSO_3^-) in acidic pH conditions through deamination and formation of uracil-sulfonates, which under basic pH conditions turn into uracils (U). However, m^5Cs are refractory to this deamination and remain as Cs. Sanger sequencing then allow discriminating between methylated and non-methylated cytosines, by detecting C-to-U conversions of non-methylated Cs. However, due to the harsh reaction conditions (efficient denaturation of DNA is performed at 95°C, in alkaline pH conditions), this method was not applied to RNA until 2009, when Schäfer et al. obtained an efficient RNA bisulfite conversion without RNA degradation by lowering denaturing temperatures and increasing incubation times (**Figure 5**) [99]. This method allows a quantitative determination of the degree of methylation of the cytosine in RNA molecules, unlike all previously described methods with sequence conservation. However, its bigger drawback is bisulfite ions refractoriness to react with double-stranded molecules. Therefore, complete denaturation of complex RNA secondary structures is crucial in order to achieve complete conversion of all non-methylated cytosines, as incomplete conversion results in false positives.

Nanopore sequencing

High-throughput sequencing-based approaches are indirect methods to detect RNA modifications, due to requirement for cDNA conversion and subsequently amplification of the material, which makes them sensitive to PCR artifacts and biases. Development of third generation sequencing technologies, for direct RNA sequencing will revolutionize the epitranscriptomics field.

One example is the platform developed by Oxford Nanopore Technologies (ONT) for direct RNA sequencing. It consists of a membrane with thousands of embedded protein nanopores immersed in an ionic solution, and upon voltage application, an ionic current is created that passes through the nanopore and is measured by a sensor (**Figure 5**). When an RNA molecule traverses the pore, it causes specific disruptions in the current that make possible to establish its nucleotide sequence, and even if those nucleotides

are modified. Therefore, ONT enables RNA modifications detection with minimal sample manipulation, at single nucleotide resolution and in a quantitative manner. Since its release in 2014, ONT has proven able to undoubtedly detect some RNA modifications, including m^5C [100,101]. However, this technology still holds some challenges in its way to virtually detect all RNA modifications, related to the interpretation of raw signals for proper base calling. Accurate base-calling algorithms require training with highly modified datasets, or the use of *in vitro* synthesized RNA templates [102].

m^5C writers

m^5C modifications in RNA are carried out by members of the NOL1/NOP2/SUN domain (NSUN) family of proteins, which in humans is comprised by seven enzymes (NSUN1-7) and, out of this family, by the DNA methyltransferase homologue DNMT2 (also called TRMT1). NSUN enzymes are SAM-dependent methyltransferases characterized by the presence of an RNA-recognition motif (RRM), the catalytic domain and a RFM core that binds the SAM cofactor. Mechanistically, NSUN proteins make use of two cysteines at the active site, one for covalently binding the target cytosine, enabling the transfer of the methyl group from the SAM cofactor, and the other to facilitate the release of the covalently-bound methylated RNA from the protein. However, DNMT2 resembles more DNA methyltransferases and uses a single catalytic cysteine [103].

In the last decades, target RNAs for each of these methyltransferases have been identified, and the list is endlessly growing (**Figure 6**). Molecular and functional characterization of their activities and those of their methylated targets have proven important roles at all levels of gene expression. And importantly, it has enabled a better understanding of several diseases bearing mutations or expression changes of m^5C RNA methyltransferases.

NSUN1

This m^5C RNA methyltransferase, as well as NSUN2 and NSUN5 (further discussed) are highly conserved throughout unicellular and multicellular eukaryotes. Human NSUN1 (primarily called P120 and later NOL1) localizes to nucleolus (**Figure 6**), and its expression depends on the cell cycle stage, being highly expressed during late G1 and S phases [104]. NSUN1 expression was described to be associated with active cell

proliferation, reason why it has been considered a prognostic marker for cancer development since the 90's [105]. Moreover, NSUN1 has been shown to interact with the RNA component of telomerase to promote transcription of cyclin D1 [106]. However, it was not until last decade that P120 was proposed to be an RNA methyltransferase, and in fact, a member of a bigger family. Nop2 (*S. cerevisiae* orthologue) and nsun-1 (*C. elegans* orthologue) have been shown to catalyze the transfer of a methyl group onto C2870 and C2982 of 25S and 26S rRNAs, respectively [107,108]. Although the human target of NSUN1 has not been properly validated yet, m^5C2870 and m^5C2982 correspond to C4447 of the human 28S rRNA (**Table 1**). Moreover, NSUN1 homologues exert further functions demonstrated to be methyltransferase-independent, such as yeast Nop2, required for correct 27S pre-rRNA processing into 25S and 5.8S mature rRNAs, and therefore, assembly of the 60S large ribosomal subunit (LSU) [109,110].

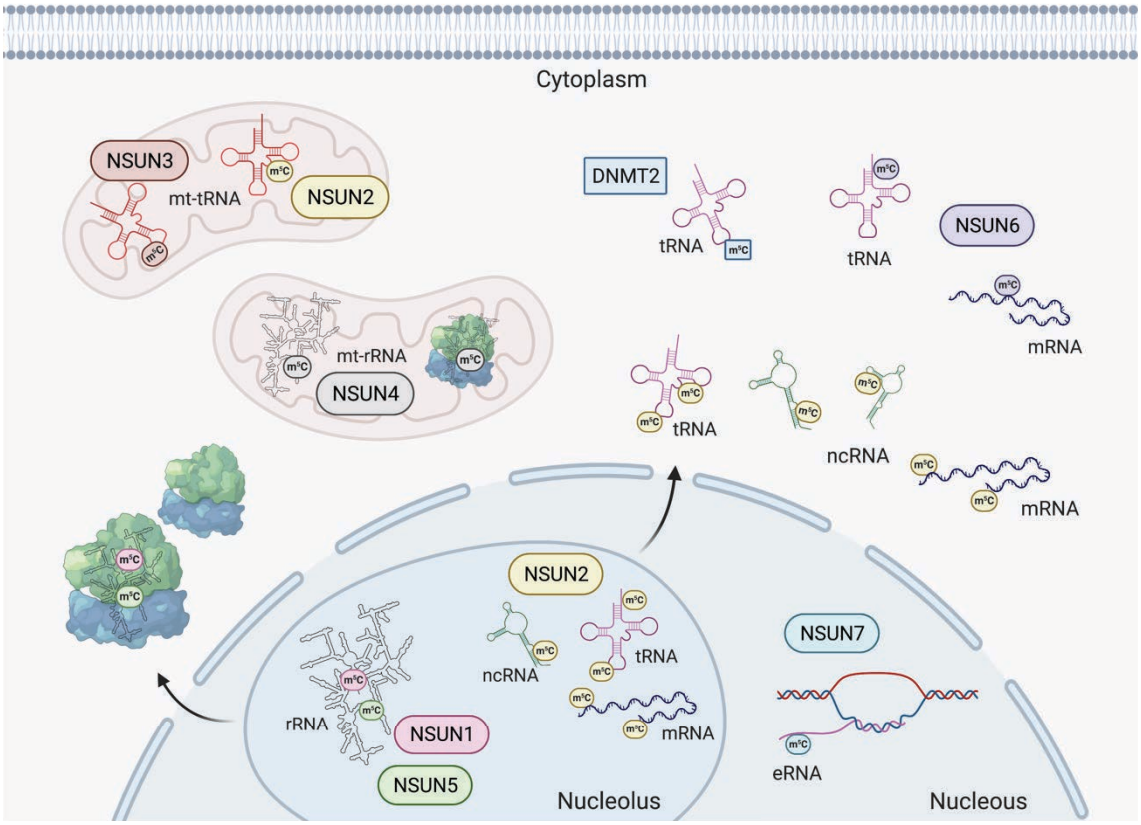


Figure 6. Overview of eukaryotic m^5C RNA methyltransferases and their RNA targets. rRNA, ribosomal RNA; tRNA, transfer RNA; mt-tRNA, mitochondrial tRNA; eRNA, enhancer RNA; ncRNA, non-coding RNA; mRNA, messenger RNA.

NSUN2

Unlike other members of the NSUN family, displaying limited and mainly evolutionary conserved substrate specificity, NSUN2 shows a marked substrate promiscuity, and even within same substrates, target promiscuity. NSUN2 was primarily described to methylate tRNAs, but further research has demonstrated its role in m^5C methylating mRNAs, rRNA, ncRNAs, and mt-tRNAs (**Table 1**). NSUN2 subcellular location varies along cell cycle: at G1 is found at the nucleolus, at the S phase at the nucleoplasm, and later, during mitosis is found in the cytoplasm, at the centrioles, associated to the mitotic spindle [111]. Recently, NSUN2 was shown to be imported into the mitochondrial matrix, for deposition of m^5C onto mitochondrial-encoded tRNAs [112,113] (**Figure 6**). tRNA methylation by NSUN2 occurs mainly at two tRNA regions: the anticodon loop (C34) and the variable loop (C47, C48, C49, C50) (**Figure 7**) (**Table 1**). m^5C at cytosines on the variable loop has proven to play a role in angiogenin-mediated response to cellular stress. Under stress, the ribonuclease angiogenin carries out endonucleolytic cleavage of tRNAs giving rise to 5′ tRNA-derived small RNA fragments (TRFs), whose accumulation represses protein translation and activates stress-resistance pathways. NSUN2-mediated deposition of m^5C prevents angiogenin cleavage of tRNAs [114]. Regarding NSUN2-mediated methylation of mRNAs, it can take place at any location along the transcript (5'UTR, CDS, 3'UTR) but within a clear consensus sequence (CNGGG), and it is involved in transcript stability and translational efficiency [92,115]. ncRNA targets of NSUN2 exert widely different biological functions, including processing of other ncRNAs [*RPPH1* and *RNY1*, components of ribonucleases; *SCARNA2*, for processing of scaRNA, needed for 2'-O-methylation of RNAs in Cajal bodies (see 2'-O-methylation), translation (5S rRNA), or epidermal differentiation (*VTRNAs*, components of the Vault complex [116] (**Table 1**). NSUN2 broad target spectrum can explain the diversity of phenotypes observed upon mutations or *NSUN2* gene expression changes, ranging from cancer (further discussed) to neurodevelopmental, vascular and autoimmune diseases [117–120].

NSUN3

NSUN3 is a nuclear-encoded mitochondrial tRNA methyltransferase (**Figure 6**). Although many components of the oxidative phosphorylation (OXPHOS) system are encoded and transcribed from the mitochondrial genome, their translation requires various nuclear-encoded proteins, including NSUN3. NSUN3 has been described to deposit m^5C at position 34 (the wobble position of the anticodon) (**Figure 7**) of the mt-

tRNA Met [121,122] (**Table 1**). m^5C34 undergoes further sequential oxidations to 5-formylcytosine (f^5C), catalyzed by ALKBH1 [123], which is necessary for the mt-tRNA Met to recognize AUA and AUU codons, in addition to the AUG codon. Expanding codon recognition is needed for proper mitochondrial translation, and therefore, for proper oxidative phosphorylation, as the mitochondrial genome counts with as little as 22 tRNA genes [124]. Accordingly, loss-of-function mutations of NSUN3 are observed in patients with mitochondrial disorders like OXPHOS deficiency or early-onset mitochondrial encephalomyopathy and seizures [122,125].

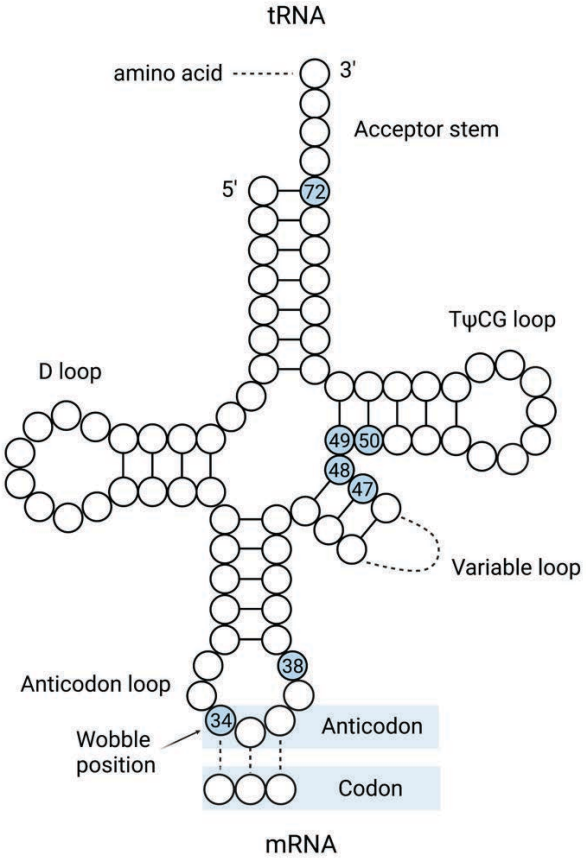


Figure 7. Simplified tRNA structure, including all known m^5C sites in human. Nucleotides are depicted as black circles, the main features of the tRNA structure are stated, and the position of cytosines described to undergo m^5C modification is displayed in light blue.

NSUN4

Another nuclear-encoded protein required for proper mitochondrial translation is the mitochondrial rRNA methyltransferase NSUN4 (**Figure 6**). NSUN4 is a dual protein, with both, a methyltransferase-dependent activity and a catalytic-independent function. Nsun4 is responsible for methylating the C911 position of the mouse 12S rRNA (corresponding to C841 of 12S rRNA in humans), located within the decoding center of

the small mitochondrial ribosomal subunit (mt-SSU) [126] (**Table 1**). Moreover, NSUN4 forms a heterodimeric complex with the mitochondrial transcription terminator factor MTERF4 and is recruited to the large mitochondrial ribosomal subunit (mt-LSU), where it participates in correct ribosomal biogenesis through assembly of both the mt-SSU and the mt-LSU into a monosome [126,127].

NSUN5

NSUN5 is, as NSUN1, a nucleolus-located rRNA methyltransferase highly conserved across eukaryotes [103] (**Figure 6**). Although its human target cytosines have not been formally described yet, Rcm1 (*S. cerevisiae* orthologue) and *nsun-5* (*C. elegans* orthologue) are responsible for m^5C methylating the C2278 and the C2381 of 25S and 26S rRNAs, respectively [107,128]. 25S rRNA m^5C2278 has shown to play a role in stress resistance, with its loss, due to deletion of Rcm1, leading to an increased translation of stress-response mRNAs, such as those coding for cell cycle regulators or DNA repair components [128]. Moreover, NSUN5-orthologues depletion in yeasts, worms, and flies increases their lifespan under dietary restriction, seemingly due to a higher oxidative stress resistance [128]. In humans, the *NSUN5* gene is located in a large chromosomal region deleted in patients with Williams-Beuren syndrome, a rare multisystemic disorder characterized by mental retardation [129], suggesting NSUN5 contribution to the disease phenotype. Moreover, mice lacking Nsun5 have been described to show agenesis of the corpus callosum (CC), the structure that connects the left and right hemispheres of the brain; hypomyelination of CC axons, disrupted development of oligodendrocyte precursor cells (OPC) leading to cognitive defects, and disrupted development of cerebral cortex [130–132].

NSUN6

Although modification of tRNAs takes place mainly at early stages of their biogenesis, while they still remain in the nucleus, NSUN6 is a tRNA methyltransferase located in the cytoplasm, carrying out tRNA modification at late biogenesis stages (**Figure 6**). It was described to catalyze m^5C at the position 72 (at the acceptor stem) (**Figure 7**) of tRNA Thr and tRNA Cys [133], requiring their tRNA substrates full-lengthen and well-folded for proper recognition [134]. Surprisingly, in 2020 two independent research groups reported NSUN6-mediated m^5C methylation of mRNAs, in a sequence motif dependent-manner (CTCCA) [135,136] (**Table 1**). m^5Cs deposited by NSUN6 are proposed to be enriched at 3'UTR regions, and mediate transcript stability and stop codon read-through [136].

Table 1. Overview of described targets for the human m^5C RNA methyltransferases.

Writer	Subcellular localization	RNA specie	Target RNA	Target site	Ref
NSUN1	Nucleolus	rRNA	28S	m^5C4447*	107,110
NSUN2	Variable	tRNA	pre-tRNA Leu (CAA)	m^5C34 m^5C48	137,138
		tRNA	tRNA Leu (CAA)	m^5C34 m^5C48	90,138
		tRNA	tRNA Asp (GUC)	m^5C47 m^5C48	90
		mRNA	CINP	m^5C748	90
		mRNA	NAPRT1	m^5C779	90
		ncRNA	RPPH1	m^5C174	90,97
		tRNA	tRNA Gly (GCC)	m^5C48 m^5C49 m^5C50	138
		rRNA	5S	m^5C92	97,139
		ncRNA	VTRNA1-1	m^5C69	97,139
		ncRNA	VTRNA1-3	m^5C27 m5C59	97,139
		ncRNA	VTRNA2-1	m^5C37	139
		ncRNA	SCARNA2	m^5C316	139
		ncRNA	RNY1	m^5C64	139
		ncRNA	7SL RNA	m^5C257	139
		mRNA	CDK1	m^5C1733	140
		mRNA	p27	m^5C64	141
		mRNA	ICAM-1	various	119
		mRNA	SHC	various	142
		mRNA	p21	m^5C2079	143
		mRNA	IL-17A	m^5C466	144
		mRNA	HDGF	-	145
		mt-tRNA	mt-tRNA Tyr	m^5C48	112,113
		mt-tRNA	mt-tRNA His	m^5C48	112,113
		mt-tRNA	mt-tRNA Leu (UUR)	m^5C48	112,113
		mt-tRNA	mt-tRNA Phe	m^5C48	112,113

		mt-tRNA	mt-tRNA Glu	m^5C49	[112,113]
		mt-tRNA	mt-tRNA Ser(AGY)	m^5C48 m^5C49 m^5C50	[112,113]
		mRNA	ATX	m^5C2756	[146]
NSUN3	Mitochondria	mt-tRNA	mt-tRNA Met	m^5C34	[121–123]
NSUN4	Mitochondria	mt-rRNA	mt-12S	m^5C841**	[103,126]
NSUN6	Cytoplasm	tRNA	tRNA Thr	m^5C72	[133]
		tRNA	tRNA Cys	m^5C72	[133]
		mRNA	various	various	[135,136]
DNMT2	Cytoplasm	tRNA	tRNA Asp	m^5C38	[147]

Footnote: rRNA, ribosomal RNA; tRNA, transfer RNA; mt-tRNA, mitochondrial tRNA; eRNA; ncRNA, non-coding RNA; mRNA, messenger RNA. *by analogy to yeast; **by analogy to mouse.

NSUN7

NSUN7 is the least studied member of the NSUN family. NSUN7 expression has been shown to be specially enriched in spermatocytes and spermatids [148], where it is thought to play a role in sperm motility, as point mutations in its gene lead to reduced sperm motility and consequent infertility in mice, and are found in asthenospermic men [149–152]. Only recently a catalytic activity for this methyltransferase has been deciphered in mouse cells, to methylate enhancer RNAs (eRNA) [153], short ncRNAs transcribed from enhancers and involved in transcriptional regulation (**Figure 6**). Nsun7 was described to interact with the transcriptional coactivator PGC-1α and colocalize at its target genes, where Nsun7 promotes target gene expression through m^5C methylation of their cognate eRNAs [153]. However, no catalytic activity for this m^5C RNA methyltransferase has been described in humans yet.

DNMT2

DNMT2 was initially proposed to be a DNA methyltransferase, as it shows sequence and structural similarities to those enzymes. However, depletion of Dnmt2 in different model organisms showed unaltered DNA methylomes, and besides, it mainly localizes to the cytoplasm [154]. Soon after, DNMT2 was described to catalyze the transfer of a methyl

group onto the C38 position (at the anticodon loop) (**Figure 7**) of tRNA Asp [147]. This methyl group is proposed to play two roles. First, it prevents the endonucleolytic cleavage of angiogenin to TRFs, similar to what occurs to NSUN2-methylated tRNAs, suggesting a role for DNMT2 in stress resistance [155]. And secondly, it increases tRNA affinity towards the aspartyl-tRNA synthetase, enhancing aminoacylation efficiency, which in turns enhances translation of poly-Asp-containing transcripts. In human, many poly-Asp-containing proteins are described to localize to the nucleus and participate in gene expression regulation [156].

m^5C readers

To date, only two proteins have been described to bind m^5C-modified RNAs: the nuclear export factor ALYREF, and the RNA-binding protein YBX1. ALYREF has been proposed to preferentially export m^5C-modified mature mRNAs to the cytoplasm [157], and YBX1 to increase the stability of its bound m^5C-modified mRNAs by recruiting the RNA-binding protein ELAVL1 [145]. ELAVL1 is a member of the ELAVL family of mRNA-binding proteins that selectively binds AU-rich elements (AREs) in 3'UTR regions. These elements usually target mRNAs for degradation, but binding of ELAVL proteins stabilizes the mRNAs, avoiding their degradation [158,159].

m^5C erasers

While m^5C writers have been profoundly characterized in recent decades, the existence of m^5C erasers is still under debate, as 5-hydroxymethylcytosine (hm^5C) has been reported in RNA by MS bulk analyses [160], but no precise sites have been mapped yet in the transcriptome. So far, some researchers have observed that m^5C RNA modifications can be oxidized to hm^5C by members of the TET family *in vitro* [161], but if they catalyze hm^5C formation *in vivo*, and if this modification and its subsequent derivatives exert any biological function is still unknown. The only enzyme demonstrated to carry out a m^5C erasing role is ALKBH1, that apart from demethylating m^1A, is described to sequentially oxidize m^5C34 on the mt-tRNA Met (catalyzed by NSUN3) to 5-formylcytidine (f^5C) [123], whose role has been explained above (see NSUN3).

Role of m^{5}C RNA modification in cancer

Although the precise molecular and biological functions of the majority of m^{5}C-related modifiers are still poorly understood, protein expression alteration of these epitranscriptomic regulators are linked to a wide range of human diseases, many of them already depicted in the previous section. Regarding cancer, there is growing evidence of the contribution of an aberrant m^{5}C epitranscriptome to the tumorigenesis process.

NSUN1 overexpression was first shown to increase proliferation of mouse fibroblasts [162], being soon afterwards described to be a valuable proliferation biomarker [163]. Since then, it has been found upregulated in lung, breast, and prostate cancer, correlating with decreased patient survival [164–166]. A proposed mechanism for NSUN1 role in cell proliferation involves its contribution to cyclin D1 transcription. NSUN1 interacts with the T-cell factor (TCF)-binding element of the cyclin D1 promoter, and recruits the telomerase by interaction with the telomerase RNA component (TERC), activating cyclin D1 transcription [106]. However, if NSUN1 m^{5}C-catalytic activity of rRNA is also involved in its tumorigenic role remains to be established.

The *NSUN2* gene has been shown to undergo different genomic mutations leading to expression alterations linked to several cancer types, including breast, skin, colon, head and neck, gastric and bladder cancer [167–169]. Moreover, NSUN2 has been described as a downstream target of MYC oncogene [111]. As already depicted, initial research into NSUN2 role showed that it was required for the stability of the mitotic spindle, with NSUN2 overexpression in cancer leading to aberrant mitosis [111], although in a methylation-independent manner. However, recent research works have proven that NSUN2 methyltransferase activity promotes tumorigenesis, in a tissue and context-dependent manner. For instance, NSUN2-mediated mRNA methylation enhances the translation of CDK1 and p21 proteins, promoting cell proliferation in different tumors [140,143]. Also, NSUN2 methylation of the autotaxin (*ATX*) mRNA promotes glioma cell migration [146], and its 3'UTR methylation of the heparin-binding growth factor (*HDGF*) mRNA favors bladder cancer progression [145]. However, while these studies focus on the carcinogenic role of tissue-specific methylation of single transcripts, they do not consider the conserved role of NSUN2 in methylating tRNAs. Hence, it is not clear if NSUN2 role in cancer is actually exclusively mediated by m^{5}C modification on mRNA. In fact, NSUN2-mediated tRNA methylation has been proposed to exert a role in tumorigenesis. It has been shown that loss of NSUN2 in cancer cells leads to hypomethylated tRNAs that are more sensitive to the catalytic cleavage of the ribonuclease angiogenin, resulting in

accumulation of TRFs [114]. Growing evidence suggest that biogenesis of TRFs is enhanced under stress conditions, and is commonly found in cancer [170]. The accumulation of TRFs can reprogram translation to promote the expression of stress response pathways and decrease those of differentiation pathways [171]. Indeed, Nsun2 deficiency in mouse cells demonstrated to block cells in a proliferative and undifferentiated state necessary for self-renewal of tissue or tumor stem cells [172]. However, this state made cells more sensitive to chemotherapeutics such as cisplatin and 5-FU, which could be rescued by treatment with angiogenin inhibitors [172]. Further data supporting tRNA methylation involvement in cell responses is the finding that Dnmt2 deletion in mice and Drosophila leads to defects in stress responses and differentiation [155,173]. These studies, therefore, show how cancer cells require controlling tRNA methylation, tRNA biogenesis and protein synthesis for stress responses and survival.

Regarding DNMT2, it shows altered expression in different cancer types [174]. Moreover, more than 60 *DNMT2* somatic mutations have been identified in tumors, some of them strongly affecting DNMT2 methyltransferase activity, which supports the functional role for a deregulated m^5C tRNA methylation in tumorigenesis [175].

NSUN3 has been proposed to help modulate chromatin organization and 5-azacytidine response in leukemia. Mechanistically, NSUN3 binds the RNA-binding protein hnRNPK, which interacts with lineage-determining TFs that recruit RNA polymerase II into nascent RNA, forming 5-azacytidine-sensitive chromatin structures [176].

Regarding NSUN4, a genome-wide meta-analysis reported its gene to be a risk locus for the development of breast, ovarian and prostate cancer [177]. In addition, NSUN4 constitutes a prognostic biomarker in hepatocellular carcinoma (HCC), being its expression negatively correlated with patient survival [178].

NSUN6 has been described as a prognostic biomarker of different cancer types as well. It was found downregulated in different tumors, where NSUN6 higher expression was associated with better clinical outcome [136]. Moreover, another study supported NSUN6 tumor suppressive role by demonstrating that NSUN6 inhibits development of pancreatic cancer (PC) by regulating cell proliferation, potentially through CDK10 expression modulation [179]. However, if CDK10 modulation is achieved by an RNA methylation-mediated mechanism remains unexplored.

The m^5C reader protein ALYREF shows altered expression in various tumor types, with prognostic value [180]. In bladder cancer, ALYREF promotes tumor growth by contributing to the Warburg effect through enhancing PKM2-mediated glycolysis. ALYREF binds to m^5C-modified *PKM2* mRNAs, increasing their stability and, thus, upregulating PKM2 protein expression [181].

The second described m^5C reader, YBX1, is also reported to exert an oncogenic role in bladder cancer, in a m^5C-dependent manner as well. Mechanistically, YBX1 increases the stability of the NSUN2-methylated *HDGF* mRNA through recruitment of the ARE element-binding protein ELAVL1 (see m^5C readers section) [145].

Finally, genomic mutations at the m^5C eraser ALKBH1 have been proposed to increase susceptibility to childhood acute lymphoblastic leukemia (ALL) [182].

HYPHOTESIS AND OBJECTIVES

HYPHOTESIS AND OBJECTIVES

Hypothesis

Epitranscriptomics, the emerging field studying biochemical modifications of RNA, has experienced an upgrowth boost in the last decade as a result of modern technological and methodological approaches, especially the development of transcriptome-wide sequencing techniques. This has led to an enormous map of post-transcriptional RNA marks that have been proven essential for normal cell fate and development. Importantly, this scientific interest has also discovered aberrant epitranscriptomes associated with several human diseases, including cancer. Growing scientific evidence exposes epitranscriptomics dysregulation in cancer, mainly by genetic mutations and aberrant expression of epitranscriptomic enzymes [183,184]. However, no epigenetic mechanisms affecting the expression of epitranscriptomic players had been described in cancer by the time this thesis started. As epigenetics represents an essential facet for achieving malignant transformation [38,185,186], and several recent data point out that an altered epitranscriptome contribute to tumorigenesis, we hypothesized that "there could be epigenetic mechanisms affecting the expression of epitranscriptomic enzymes in cancer".

Since the RNA modification repertoire comprises more than 170 biochemical marks [42], we decided to focus our first dive into the epitranscriptomic field on a mark already familiar to the laboratory, 5-methylcytosine. As already depicted, (see Introduction chapter) mounting evidence begins to show the contribution of aberrantly deposited and decoded m^5C marks to the tumorigenesis process. We, thus, sought to find tumor-associated epigenetic alterations giving rise to m^5C RNA modification defects. We found interesting to explore how these defects could participate in tumor biology, how they could correlate with the outcome of the disease, and to evaluate their suitability to be targeted by novel therapeutic approaches.

Objectives

The general aim of this thesis is to identify and characterize if there are DNA methylation patterns giving rise to cancer-related m^5C RNA modification defects, potentially promoting malignant transformation and tumor progression. For this aim, the following specific objectives are proposed:

1. To determine if there is any m^5C RNA-related enzyme that is epigenetically lost in cancer by DNA promoter hypermethylation.

2. To select a candidate gene of interest and experimentally validate its epigenetic silencing in cancer.

3. To characterize the cellular implications of the observed epigenetic lesion *in vitro*, in cancer cell line models resulting from the depletion or the recovery of the enzyme of interest, and *in vivo,* in murine models.

4. To identify and validate target RNAs whose m^5C methylation pattern is altered as a consequence of the gene of interest silencing in cancer.

5. To test whether the methylation status of the selected gene may be valuable as a biomarker in clinical management, as well as to assess its potential as putative druggable target for cancer treatment.

MATERIALS AND METHODS

MATERIALS AND METHODS

All the below explained methods describe the standard procedures carried out in both studies unless otherwise specified.

Human cell lines and tissues

Study I

Human glioblastoma cell lines were purchased from the American Type Culture Collection (ATCC) (DBTRG-05MG, MO59J, A172 and LN229); the Biological Bank of the IRCCS Azienda Ospedaliera Universitaria San Martino (CAS-1) and the Japanese Collection of Research Bioresources (JCRB) Cell Bank (KS-1). Cell lines were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS) (Innovative), and 1% v/v penicillin/streptomycin (Gibco). Human grade III glioma cell lines were purchased from the ATCC (SW1088 and BT142 mut/-) and from Sigma-Aldrich (MOG-G-CCM). SW1088 cell line was cultured in Leibovitz's L-15 medium (LGC Standards), BT142 mut/- was cultured in NeuroCult TM NS-A proliferation kit (Stem Cell Technologies), and MOG-G-CCM in Ham's F-10 Nutrient mix (Life Technologies). All the aforementioned cell lines were maintained at 37ºC in an atmosphere of 5% (v/v) carbon dioxide, and regularly tested for the absence of mycoplasma. Human astrocytes were purchased from Innoprot and maintained in culture for a limited number of passages in poly-L-lysine-coated plates. Brain white matter samples were kindly provided by Raul Delgado (IDIBELL, Barcelona, Spain).

Study II

Human HCC cell lines were purchased from the ATCC (HEP3B2-1-7, SNU-475, SNU-387 and SNU-423); the JCRB Cell Bank (HUH-7) and the SNU-398 cell line was kindly provided by Dr.Josep M.Llovet (Icahn School of Medicine at Mount Sinai, New York, USA, and IDIBAPS-Hospital Clínic, Barcelona, Spain). Cell lines were cultured in Roswell Park Memorial Institute (RPMI) medium supplemented with 10% (v/v) fetal bovine serum (FBS) (Innovative), and 1% v/v penicillin/streptomycin (Gibco). Human liver samples were kindly provided by Dr. Alberto Villanueva (IDIBELL, Barcelona, Spain).

DNA Methylation Analyses

The methylation status of 5'-end promoter-associated CpG islands of candidate genes was determined by DNA Methylation microarrays and validated through bisulfite sequencing PCR (BSP) and/or bisulfite pyrosequencing.

DNA extraction

DNA was extracted from cell pellets firstly by overnight incubation with lysis buffer (10 mM Tris pH 8, 5 mM EDTA, 100 mM NaCl, 1% SDS, 10 mg/ml proteinase K) at 37°C. Then, 5M NaCl was added to cell lysates, and samples were centrifuged 15 minutes at 4,000 RPM. Upper phases were collected, and DNA was precipitated by isopropanol addition. DNA was washed with 70% ethanol, air-dried and resuspended in water. DNA concentration was measured by Nanodrop.

DNA Bisulfite conversion

Genomic DNA was bisulfite converted using the EZ DNA Methylation-Gold kit (Zymo Research) following manufacturer's instructions. Bisulfite treatment converts unmethylated cytosines into uracil, whereas methylated cytosines remain unchanged.

DNA methylation arrays

Genome-wide promoter methylation status was determined either using the HumanMethylation450 Beadchip (450K) or Infinium MethylationEPIC BeadChip (EPIC) (Illumina). In both cases, bisulfite converted DNA was used for hybridization onto the beadchip. Briefly, samples were whole-genome amplified, enzymatically fragmented, and hybridized onto the beadchip for 16 hours at 48ºC. Single nucleotide extension with differentially labelled dideoxynucleotides (ddNTPs) was performed, and ddNTPs were next differentiated by incubation with fluorescence labelled antibodies. Raw fluorescence intensity values were normalized with Illumina Genome Studio software, and used to calculate DNA methylation levels (β-values). β-values should range from 0 to 1. By consensus, a promoter region is considered to be hypermethylated when the average of the β-values of the interrogated CpGs is, in cell lines, higher than 0.65, and in human tissues, higher than 0.33. The reason to lower down this threshold is the potential sample contamination with normal adjacent tissues.

Bisulfite sequencing PCR (BSP)

DNA methylation status of 5'-end promoter-associated CpG island of candidate genes was analyzed by bisulfite sequencing. Promoter regions were obtained from the UCSC Genome Browser, and CpG islands were *in silico* identified using Methyl Primer Express v1.0 (Applied Biosystems). Primers encompassing the transcription start site (TSS)-associated CpGs were designed by using the same program. Primer sequences can be found in Table 2. A specific region of every promoter of interest was then PCR amplified, cloned into pGEM®-T Easy vector (Promega), and transformed into competent bacteria. Bacterial plasmid DNA was extracted and purified. A minimum of eight single clones were submitted to sequencing PCR using BigDye Terminator V3.1 cycle sequencing kit (Applied Biosystems) and sequenced in a 3730 DNA Analyzer (Applied Biosystems). Reads were aligned using the BioEdit software, and the methylation status of each promoter-associated island region was represented using the BSMap software**.** Initially methylated cytosines, which remained unchanged after bisulfite treatment may be identified, whereas non-methylated cytosines, uracil-converted, become thymines upon PCR amplification.

Bisulfite Pyrosequencing

Bisulfite-treated DNA was PCR amplified using primers (see Table 2) designed with PyroMark Assay Design 2.0 software (Qiagen). PCR amplicons were then pyrosequenced with the PyroMark Q24 system and analysed with PyroMark CpG software (Qiagen).

Table 2. List of the oligonucleotides employed for bisulfite sequencing

DNA bisulfite sequencing		
Name	**Study**	**Sequence**
NSUN5 BSP Fw	I	TAAAAATTTAAGAATTAGAGAAGAAGTG
NSUN5 BSP Rv	I	AAAACAAACTAAATAAAAACTCCCC
NSUN7 BSP Fw	II	TTTTAGGAATTTGTGAGTTAAAATT
NSUN7 BSP Rv	II	TCAAATACAACCACTCCCAC
DNA bisulfite pyrosequencing		
Name	**Study**	**Sequence**
NSUN5 PCR Pyroseq Fd	I	Biot-ATGGGGTTGTATGTTGTAGTTG
NSUN5 PCR Pyroseq Rv	I	CCCACCCCRAATTCCTCACT
NSUN5 Seq1 Pyroseq Rv	I	TCCCCACCCCTACTA
NSUN5 Seq2 Pyroseq Rv	I	CCCTTAATAAAACCC
RNA bisulfite sequencing		
Name	**Study**	**Sequence**
NSUN5 C3782 28S rRNA Fw	I	GTGATTTTTGTTTAGTGTTTTGAATGTTAA
NSUN5 C3782 28S rRNA Rv	I	AAAACCTCCCACTTATTCTACACCTCTC
NSUN1 C4447 28S rRNA Fw	I	TAGTGTTAGGTGGGGAGTTTGATTG
NSUN1 C4447 28S rRNA Rv	I	AAATTCTACTTCACAATAATAAAAAAAACC
NSUN7 CCDC9B Fw	II	CTGGAAGGTCAGGAGGGAAG
NSUN7 CCDC9B Rv	II	GGAAGCCACTGACTCCCC

Expression Analyses

RNA expression was assessed by real time quantitative reverse transcription PCR (qRT-PCR) experiments. Total RNA was first extracted from cell pellets using the SimplyRNA kit (Promega) on a Maxwell RSC device (Promega).1 to 2.5 µg of total RNA were retro-transcribed using ThermoScript RT-PCR system (ThermoFisher) or GoScript Reverse Transcription system (Promega) following manufacturer's instructions. Oligo(dT) primers were used when looking for an enrichment for coding RNAs, while random hexamers when selected when analysing non-coding RNAs. Quantitative PCR was performed using SYBR® Green PCR Master Mix (Applied Biosystems) in a QuantStudio 5 Real-Time PCR System (ThermoFisher). GAPDH was used as the housekeeping gene to enable normalization.

Reactivation treatments with the demethylating agent 5-aza-2′-deoxycytidine or decitabine (AZA; Sigma) were performed at 0.5 μM and 1 μM for 72 hours. RNA extraction and qRT-PCR were carried out as described before.

Protein expression was analyzed by western blotting. Total protein extracts were obtained using RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA and EGTA, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS) containing protease inhibitor cocktail cOmplete™ (Roche). Protein levels were quantified using Pierce™ BCA protein assay kit (Thermo Scientific) following manufacturer's recommendations. Proteins were resolved on SDS-polyacrylamide gels, transferred to nitrocellulose membranes, and blotted against specific antibodies. All primary antibodies used are listed in Table 3.

Table 3. List of the antibodies used and their related information.

Name	Study	Supplier	Reference
Rabbit polyclonal anti-NSUN1	I	ThermoFisher	PA5-34712
Rabbit polyclonal anti-NSUN5	I	Abcam	ab121633
Mouse monoclonal anti-NQO1	I	Cell signaling	3187
Rabbit polyclonal anti-KIDINS220	I	Novus Biologicals	NBP1-31216
Rabbit polyclonal anti-NDRG3	I	Cell signaling	5846
Mouse monoclonal anti-ACO2	I	ThermoFisher	MA1-029
Rabbit polyclonal anti-NSUN7	II	Biorbyt	orb258175
Rabbit polyclonal anti-GAPDH	II	Trevigen	2275-PC-100
Rabbit polyclonal anti-CCDC9B	II	Novus Biologicals	NBP1-93764
Monoclonal anti-β-Actin	I, II	Sigma	A3854
Rabbit polyclonal anti-Lamin B1	I, II	Abcam	ab16048
Mouse monoclonal anti-αTubulin	I, II	Abcam	ab40742

Cellular models

Study I

Overexpression models

The NSUN5 wild-type cDNA sequence was cloned into the pLVX-IRES-ZsGreen1 expression vector (Clontech), and a mutated version coding for a catalytically-inactive NSUN5 protein was cloned into the pLVX-IRES-tdTomato vector (Addgene). DBTRG-05MG cell line total RNA was retrotranscribed and used to amplify the *NSUN5* transcript. Specific adapter primers were designed in order to include the EcoRI restriction site to the 5' region of the transcript; and the XbaI restriction site and a FLAG-Tag in the 3' carboxy-terminal end. NSUN5 cDNA sequence from the DBTRG-05MG cell line was used, as well, to incorporate 4 specific point mutations by a PCR phusion method; 3 at SAM-binding sites (D258L, R263L, D305L) and one at the active site (C359V). For every substituted amino acid, the most frequent nucleotide codon sequence was used. For primer sequences refer to Table 4. Lentiviruses containing the aforementioned constructs (NSUN5 or NSUN5 mut), or the empty vectors (EV) were produced by co-transfecting HEK-293T cells with psPAX2 (Addgene) and pMD2.G (Addgene), using jetPRIME® Transfection Reagent (Polyplus Transfection) and following manufacturer's instructions. After 72 hours, viral supernatants were collected, 0.45 µM-filtered and stored at 4 ºC before infection. The recombinant products were randomly inserted by lentiviral transduction into the genome of A172 and LN229 NSUN5-silenced glioma cell lines. After 5 passages, green and red fluorescent cells were sorted by FACS and cultured as already stated (see Human cell lines and tissues).

Short hairpin interference RNA (shRNA)-mediated depletion models

Four different short hairpin RNA molecules (shRNAs) targeting *NSUN5* mRNA were designed and cloned into the pLVX-shRNA2-ZsGreen plasmid (Clontech). A shRNA against the *MSS2* yeast mRNA (not present in mammals) was used as Scramble control. Lentiviruses containing the aforementioned constructs (shNSUN5 #1-4 or Scramble) were produced by co-transfecting HEK-293T cells with psPAX2 (Addgene) and pMD2.G (Addgene), using jetPRIME® Transfection Reagent (Polyplus Transfection), as above described. NSUN5-expressing glioma cell lines were stably transduced, and after 5 passages, GFP-positive cells were sorted by FACS and maintained in culture. For shRNA sequences refer to Table 4.

Study II

Overexpression model

The cDNA sequence of NSUN7 was cloned into the pLVX-IRES-ZsGreen1 expression plasmid (Clontech), and a mutated cDNA version was cloned into the pLVX-IRES-tdTomato vector (Addgene). A normal peripheral blood mononuclear cells (PBMCs) RNA sample, from a healthy donor, was retrotranscribed and used to amplify NSUN7 coding sequence. Specific adapter primers were designed to add restriction sites for XhoI and NotI to 5' and 3' ends of NSUN7 cDNA sequence. This construct was used as a template to incorporate a specific point mutation at the active site (C439V) by a PCR phusion strategy. Oligonucleotide sequences are listed in Table 4. Lentiviruses containing the aforementioned constructs (NSUN7 or NSUN7 mut), or the empty vectors (EV) were produced by co-transfecting HEK-293T cells with psPAX2 (Addgene) and pMD2.G (Addgene), using jetPRIME® Transfection Reagent (Polyplus Transfection), as above described. The NSUN7-silenced SNU-423 cell line was lentiviral-transduced with the different constructs, GFP or Tomato-FACS sorted and maintained in culture as already stated (see Human cell lines and tissues).

CRISPR-mediated depletion model

The NSUN7-expressing HCC cell line HEP3B2-1-7 was depleted of its expression by using a CRISPR-mediated knock out strategy. The CRISPR-Cas9 strategy makes use of an RNA-guided DNA endonuclease, the CRISPR-associated protein 9 (Cas9) that promotes genome editing by stimulating a double-strand break (DBS) at a target genomic locus. The guide RNA (gRNA) design was made by using the CRISPR Design Tool (http://tools.genome-engineering.org) that takes a genomic sequence of interest and identifies suitable target sites. Using this tool, 3 different gRNAs were designed to target the Cas9 machinery to different NSUN7 exons within the functional domain. For each gRNA, a pair of oligos were purchased (Sigma), annealed and cloned into the pSpCas9-2A-GFP plasmid (Clontech). HEP3B2-1-7 cells were transiently transfected with the pool of guides using jetPRIME® Transfection Reagent (Polyplus Transfection), and green fluorescent cells were sorted by FACS and subjected to clonal proliferation. NSUN7 knock out (KO) was assessed by qRT-PCR and western blot, already described in Expression analyses section. gRNA and primer sequences are listed in Table 4.

Table 4. List of the oligonucleotides employed for the establishment of cellular models.

Name	Study	Sequence
NSUN5 Cloning Fw	I	AAAAAAAAGAATTCGCCGCCACCATGGGGCTGTATGCTGCAGCTGCAGGCGTGTTGGCCGGCGTGGAGAGCCGCCAGGGCTCTATCAAGGGGT
NSUN5 Cloning Rv	I	AAAAAAAATCTAGATCACTTATCGTCGTCATCCTTGTAATCGCCGGAGCCCCTTGGCACCTCGACCCGTTCAATTACAGCAACGAAGAAGCCACTGCTGAGTGTGGTCTC
NSUN5 mut Cloning Fw	I	AAAAAAAAGAATTCGCCGCCACCATGGGGCTGTATGCTGCAGCTGCAGGCGTGTTGGCCGGCGTGGAGAGCC
NSUN5 mut Cloning Rv	I	AAAAAAAACTAGTTCAGGCGTAGTCGGGCACGTCGTAGGGGTAGCCGCTGCCGCCCGGAGCCTCTGCTATGTGCAAGGCGGTGTGCAAGCACCGGCTGCGGCTCTTTGCT
NSUN5 mut PCR phusion A Fw	I	ATGGGGCTGTATGCTGCAGCTG
NSUN5 mut PCR phusion A Rv	I	CATGGATGCCAGCAGCTTGGCATCCAGCAGAAAGGCAAAGATCTTCCCTTGGTTCTTCAGAAGAGCAGCCAAGTGACTGGTCTTATTGCC
NSUN5 mut PCR phusion B Fw	I	GATCTTTGCCTTTCTGCTGGATGCCAAGCTGCTGGCATCCATGGCCACGCTGCTGGCCCGGGCTGGCGTCTCTTGCTGTGAACTGGCTGAG
NSUN5 mut PCR phusion B Rv	I	TGCAGGAAGGCAGCAGCAGGATGTAGTGGACCTCATGGTAGCGTGG
NSUN5 mut PCR phusion C Fw	I	CATCCTGCTGCTGCCTTCCTGCAGTGGCTCGGGTATGCCGAGC
NSUN5 mut PCR phusion C Rv	I	GGCAGAGGGACACCGTGGAGTAGACGAGCCGCTGCAGGGAAGGG
NSUN5 mut PCR phusion D Fw	I	GTCTACTCCACGGTGTCCCTCTGCCAGGAGGAGAATGAAGACGTGGTGC
NSUN5 mut PCR phusion D Rv	I	TTTTTTTTACTAGTTCACTTGTCGTCGTCGTCCTTGTAGTCGCCGCTGCCGCCCGGAGCCTCTGCTATGTGCAAGGCGGTGTGCAAGCACCGGCTGCGGCTCTTTGCTGT
NSUN5 ShRNA 1 Fw	I	GATCCGCTCCGATGATGTAGTTGATTTCAAGAGAATCAACTACATCATCGGAGTTTTTTACGCGTG
NSUN5 ShRNA 1 Rv	I	AATTCACGCGTAAAAAACTCCGATGATGTAGTTGATTCTCTTGAAATCAACTACATCATCGGAGCG
NSUN5 ShRNA 2 Fw	I	GATCCGCCATGAGGTCCACTACATCTTCAAGAGAGATGTAGTGGACCTCATGGTTTTTTACGCGTG
NSUN5 ShRNA 2 Rv	I	AATTCACGCGTAAAAAACCATGAGGTCCACTACATCTCTCTTGAAGATGTAGTGGACCTCATGGCG
NSUN5 ShRNA 3 Fw	I	GATCCGCCGATGATGTAGTTGATTATTCAAGAGATAATCAACTACATCATCGGTTTTTTACGCGTG
NSUN5 ShRNA 3 Rv	I	AATTCACGCGTAAAAAACCGATGATGTAGTTGATTATCTCTTGAATAATCAACTACATCATCGGCG

NSUN5 ShRNA 4 Fw	I	GATCCGTGCTAGTGTATGAGTTGTTTCAAGAGAA CAACTCATACACTAGCACTTTTTTACGCGTG
NSUN5 ShRNA 4 Rv	I	ATTCACGCGTAAAAAAGTGCTAGTGTATGAGTTGT TCTCTTGAAACAACTCATACACTAGCACG
NSUN7 Cloning Fw	II	AAAAAAAACTCGAGCCGCCACCATGCTGAATT
NSUN7 Cloning Rv	II	AAAAAAAAGCGGCCGCTCAAAGCCATCGCCGAG GAGGCCTGAGTAGGG
NSUN7 mut PCR phusion Fw	II	TTACTGCACAGTGTCAGTTTTTCCAGAAGAAAATG AAGCTGTTGT
NSUN7 mut PCR phusion Rv	II	GGAAAAACTGACACTGTGCAGTAAACAACTGCTT GAGCTTTAGTAAATT
NSUN7 gRNA#1 Fw	II	CACCGCTACCTCGTTGTTCAGGAC
NSUN7 gRNA#1 Rv	II	AAACGTCCTGAACAACGAGGTAGC
NSUN7 gRNA#2 Fw	II	CACCGCACTGGAATTTCAAGACCTT
NSUN7 gRNA#2 Rv	II	AAACAAGGTCTTGAAATTCCAGTGC
NSUN7 gRNA#3 Fw	II	CACCGCTCCTTGAGAGTGATCTTTA
NSUN7 gRNA#3 Rv	II	AAACTAAAGATCACTCTCAAGGAGC

RNA Methylation Analyses

RNA extraction

Total RNA from cell pellets was extracted using the SimplyRNA kit (Promega) on a Maxwell RSC device (Promega) or the RNeasy kit (Qiagen), following manufacturer's instructions. Cytoplasmic and nuclear RNA were extracted by using PARIS™ kit (Ambion® PARIS™ system).

RNA bisulfite conversion

Total RNA was bisulfite converted using the Methylamp RNA Bisulfite Conversion Kit (EpiGentek) following a modified version of the manufacturer's protocol, tripling the cycles stated.

RNA bisulfite sequencing

Bisulfite-modified RNA was retrotranscribed using ThermoScript RT-PCR system (ThermoFisher) or GoScript Reverse Transcription system (Promega) following manufacturer's instructions, with random hexamers. PCR primers for the sequences of interest, containing the potentially methylated cytosines, were designed with Primer3web

(https://primer3.ut.ee/), aiming for smaller amplicon sizes in order to reduce amplification of unconverted cytosines due to strong secondary structure. Primer sequences are available in Table 2. PCR amplicons were purified and ligated into pGEM®-T Easy vector (Promega), transformed and sequenced as described previously in Bisulfite genomic sequencing section. Clones with incomplete bisulfite conversion (unconverted cytosines never described to be methylated) were discarded.

Transcriptome-wide bisulfite RNA sequencing (bsRNA-seq)

Study I

A minimum of three independent biological replicates from EV- and NSUN5-transfected LN229 glioma cells were made use of. Total RNA from samples was isolated using the SimplyRNA Maxwell kit, DNA depleted with Turbo DNase (Ambion), phenol/chloroform extracted and ethanol precipitated. RNA quality was estimated on a Bioanalyzer total RNA Nano Chip (Agilent). In order to evaluate the bisulfite conversion efficiency between samples, 10 ng of two Renilla Luciferase *in vitro* transcripts were spiked in 10 µg of each RNA replicate. rRNA was reduced with the Ribo-Zero rRNA Removal Kit (Illumina) prior to bisulfite conversion as previously stated [90], and the efficiency of rRNA depletion was checked using a Bioanalyzer total RNA Pico Chip (Agilent). Bisulfite conversion was carried out as described previously [187]. Libraries were then prepared according to Illumina TruSeq Stranded Total RNA protocol and sequenced on an Illumina Hiseq2500, in paired-end mode and 2x100bp length using the Illumina HiSeq Rapid SBS kit v2. The quality of the raw fastq files was checked with the FastQC software, and low quality reads and adapter sequences were removed with the BBMAP-BBDUK software (https://sourceforge.net /projects/bbmap). The pre-processed reads were mapped with Hisat2-2.0.4 software (https://ccb.jhu.edu/software/hisat2/index.shtml) against the human transcriptome bisulfite converted, corresponding to hg19 build, including the sequences of rRNAs. Only uniquely mapped reads containing less than 3 unconverted cytosines (3C filter) were considered for the following analysis. The m^5C methylation value for each sample in the Watson and Crick strands was inferred using the BS-RNA software (http://bs-rna.big.ac.cn/) from the mapped reads. Only m^5C positions with a coverage greater than 30 in all the samples were considered. Finally, the read counts were normalized with the limma-voom function, and using the R package limma [188] cytosines differentially methylated between conditions (EV- and NSUN5-transfected) were defined as those with a multiple correction adjusted p-value < 0.05, a log fold-change > 0 and a methylation difference over 0.33.

Study II

Three independent biological replicates from EV-transfected and NSUN7-transfected SNU-423 cells, and from NSUN7-expressing and NSUN7-knocked-out HEP3B2-1-7 HCC cells were used. Total RNA was extracted and processed as previously described. In this case, to evaluate bisulfite conversion efficiency across samples, 2 ng of the two Luciferase *in vitro* transcripts were spiked in. Ribosomal RNA (rRNA) was reduced, and RNA was bisulfite-converted as previously stated. Libraries were prepared according to Illumina TruSeq Stranded Total RNA protocol and sequenced on an Illumina Hiseq2500, in paired-end mode and 2x100bp length using the Illumina HiSeq Rapid SBS kit v2. The quality of the raw fastq files was checked with the FastQC software, and low-quality bases and adaptor sequences were removed using Trimmomatic (v0.36). Unmapped reads were mapped to the appropriate converted reference (hg38 build) using the meRanGh tool in MeRanTK. Only uniquely mapped reads containing less than 3 unconverted cytosines (3C filter) were considered for the following analysis. Read counts at each cytosine position in the genome were obtained using the 'mpileup' function in samtools. Non-conversion sites were determined using a custom script. m^5C sites with more than 30 read coverage in at least 1 sample and more than 10 read coverage in all the samples were selected, and a differential methylation analysis between groups (WT vs KO and EV- vs NSUN7-transduced) using Fisher's exact test was performed. Cytosines differentially methylated between conditions were defined as those with a multiple correction adjusted p-value < 0.05 and a methylation difference over 0.1.

RNA Immunoprecipitation and RT-qPCR (RIP-qPCR)

Total extracts from LN229 EV- and NSUN5-transfected cells were overnight immunoprecipitated with 10 µl of Anti-Flag M2 Magnetic Beads (Sigma) in RIP buffer (150 mM KCl, 25 mM Tris pH7.4, 5 mM EDTA, 0.5 mM DTT, 0.5% NP40, protease inhibitors). Beads were then washed three times with RIP buffer. Prior to RNA elution, immunoprecipitation efficiency was assessed by boiling and loading 10% of beads with 1X Laemmli buffer on a polyacrylamide gel. The pulled-down RNA from the remaining 90% of beads was eluted with TRIzol Reagent (ThermoFisher), extracted with phenol/chloroform, DNA-depleted by DNAse treatment and resuspended in water. Equal amounts of each sample were then retrotranscribed and analyzed by qRT-PCR (see Expression analyses). The equivalent amount of input RNA was processed in parallel to estimate pull-down efficiency.

Mung bean nuclease protection assay and reverse phase-high pressure liquid chromatography (RP-HPLC)

Total RNA from glioma cell lines was extracted by phenol/chloroform, and 1 mg per condition was subjected to mung bean nuclease protection assay followed by RP-HPLC, as described previously [189]. The specific sequence of the 28S rRNA containing the NSUN5 target cytosine was protected by hybridization to complementary synthetic deoxyoligonucleotides (covering from the 3765 to the 3812 nucleotide). After digestion with the mung bean nuclease (NEB) and 0.05 mg/ml RNase A (Sigma-Aldrich), protected RNA fragments were resolved in a 8 M urea 13% polyacrylamide gel. The rRNA band was excised and passive eluted overnight with 0.3 M NaAc by rotation at 4 °C. Eluted rRNA fragments were precipitated with 100% EtOH and resuspended in RNAse-free water. Isolated fragments were then digested with nuclease P1 and bacterial alkaline phosphatase (Sigma Aldrich) and single nucleosides were subsequently analyzed by RP-HPLC on a Supelcosil LC-18-S HPLC column (25 cm × 4.6 cm × 5 µm) equipped with a precolumn (4.6 cm × 20 mm) at 30 °C on an Agilent 1200 HPLC system.

Molecular dynamics simulations

The Human 80S ribosome was taken from the high-resolution (2.9 Å) cryoEM structure with PDB id 6EK0 [190]. Since the modelling of the complete 80S ribosome with all the modified bases reported [190] is not accessible with the current state-of-the-art methods available in the field [191], we focused on a region of the ribosome 30 Å around m^5C3782, which comprises part of the 28S subunit (model labelled r30A5mC). The region contained 263 RNA residues, 21 amino acids belonging to the α-helix of the 60S ribosomal protein L41, and 13 Mg^{2+} ions. Several RNA chains of the 28S subunit were cut by our selection, consequently the Cartesian coordinates of all atoms in the mentioned region were fixed in the space as found experimentally by means of a harmonic potential of 5 KcalÅ2. The RNA residues 3750-3793 and 3807-3816 from the 28S subunit, the residues 15 to 24 from protein L41, and 8 Mg^{2+} (labelled 4405, 5186, 5188, 5215, 5252, 5299, 5632, 5639 in 6EK0) were completely free to move during the Molecular Dynamics (MD) simulation. The r30A5mC model also contained 10 different types of modified bases. PSU, 5MC, and 7MG were treated with the available parameters [192], while OMC was converted to C, B8H to 5MU, A2M to M2A, OMG to G, B8W to G, UR3 to 3MU, and MA6 to 6MA. The model was then minimized in vacuo, neutralized (with 164 Na^+ ions), solvated (with explicit waters and 0.4 M of added Na^+Cl^-

), and minimized in solution with positional restraints on the solute using our well-established multi-step protocol [193,194]. To produce the final models, the minimized structures were thermalized to 298ºC at NVT, and then simulated during 60 ns by means of Molecular Dynamics simulations at NPT (P = 1 atm). The first 10 ns of the simulations were considered as an equilibration step and were discarded for further analysis. The system labelled r30AC was built from r30A5mC by replacing m^5C3782 by C3782 (i.e., removing the methyl group).

To treat the L41 protein we used the ff12SB force field [195], which is compatible with the modified RNA bases developed by Santa Lucia and co-workers [192]. The RNA was represented using parmbsc0 [196], and the entire system was surrounded by a truncated octahedral box of ~20,000 TIP3P water molecules [197], applying Dang's parameters on ions [198]. Ions were initially placed randomly, at a minimum distance of 5 Å from the solute and 3.5 Å from one another. All systems were simulated using the Berendsen algorithm [199] to control the temperature and the pressure, with a coupling constant of 5 ps. Center-of-mass motion was removed every 10 ps to limit build-up of the translational kinetic energy of the solute. SHAKE [200] was used to keep all bonds involving hydrogen at their equilibrium values, which allowed us to use a 2 fs step for the integration of Newton equations of motion. Long-range electrostatic interactions were accounted for by using the Particle Mesh Ewald method [201] with standard defaults, and a real-space cut-off of 10 Å. All simulations were carried out using the PMEMD CUDA code module [202] of AMBER 18, and analysed with CPPTRAJ [203], VMD 1.9.3 [204], and PYMOL v2.2.3.

Global determination of protein synthesis

De novo protein synthesis in cell cultures was determined by O-propargyl-puromycin (OP-Puro) and [3H]-leucine incorporation, on basal culture conditions and upon stress induction. Oxidative stress was induced by addition of 100 µM H_2O_2 to the cell culture medium for an average of 16 hours. Nutrient deprivation stress was achieved by culturing cells in low-nutrient medium [DMEM medium with low Glucose, no L-glutamine, Phenol Red and HEPES (ThermoFisher), supplemented with 10% dialyzed FBS, 1X GlutaMAX (ThermoFisher) and 1mM Sodium Pyruvate (ThermoFisher)]. Cells were washed three times with PBS prior to the addition of the nutrient deprivation medium, and were cultured for 16 hours before each below described procedure.

OP-Puro incorporation

Cells were labeled for nascent protein synthesis using the Click-iT® OPP Reagent (Thermo Fisher). Briefly, cells were incubated with OPP Reagent in cell medium for 1 h. Cells were then fixed with 3.7% paraformaldehyde for 15 minutes, and permeabilized in PBS supplemented with 1% fetal bovine serum (Invitrogen) and 0.1% saponin (Sigma-Aldrich) for 5 minutes at room temperature. To conjugate OP-Puro to a fluorochrome, an azide-alkyne cycloaddition was performed using the Click-iT® Plus OPP Protein Synthesis Kit (Thermo Fisher Scientific) with Alexa Fluor-647. Cells were then washed twice with PBS, and further stained for DAPI. Slides were mounted for imaging in glycerol supplemented with Mowiol 4-88 (Calbiochem). Fluorescence images were acquired using a confocal microscope (Leica TCS SP5) at 1,024 × 1,024 dpi resolution. All images were further processed using Fiji (ImageJ) software. To quantify protein synthesis rates, the fluorescence intensity of the OP-Puro signal was thresholded (mean OP-Puro = Integrated density/area) and divided by the total number of cells per image. The mean of OP-Puro incorporation was averaged from a minimum of 15 images per condition.

[3H]-leucine incorporation

Cells at 60-70% confluence were incubated with 10 mCi/mL 3H-leucine (Perkin Elmer) for 30 minutes. Cells were washed with PBS, and lysed in RIPA buffer [50 mM Tris-HCl pH 8, 150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 1 mM dithiothreitol (DTT)]. The lysates were cleared by centrifugation and incubated with cold 10% trichloroacetic acid (TCA) for 10 minutes. TCA-insoluble proteins were then washed twice with 5% TCA, solubilised with 0.1 M NaOH, and analyzed using a liquid scintillation counter (Tri-Carb 2100TR, Perkin Elmer). Results were normalized to protein concentration quantified using Pierce™ BCA protein assay kit (Thermo Scientific).

RNA-sequencing (RNA-seq)

Total RNA was extracted from cell pellets using the SimplyRNA kit (Promega) on a Maxwell RSC device (Promega). 3 µg of total RNA from 2-3 replicates were used for RNA-seq analyses. Libraries were prepared with TruSeq stranded mRNA kit (Illumina) and sequenced in paired-end mode with a read length of 2x100bp on a HiSeq2000 (Illumina). An average of 70 to 100 million pairs of 100bp paired-end read per sample were generated. For alignment and quantification, read quality and adapter sequence

presence was checked with FastQC (v0.11.4) software. Sequencing adapters were removed from the 5' and 3' ends of the reads using BBDUK. Differential expression was analysed using the DESeq2 package (v1.16.1) [205] with default options, in the R programming language (v3.4). RNA-Seq reads were aligned to the human reference transcriptome (GRCh37 build) using HISAT2 v.2.0.5 [206]. The alignments were then assembled with Stringtie v.1.3.3b [207]. The average of mapped reads per sample was 98%. Genes were considered to be differentially expressed with an absolute log2 fold-change > = 2 and an adjusted p-value < 0.01.

Ribosomal profiling (Ribo-seq)

Ribo-seq was performed following a modified version of the original protocol [208]. A minimum of five replicates were performed for each sample. Stress-induction in cell cultures was carried out as previously described in Global determination of protein synthesis section. Briefly, cells were grown until confluence, and then washed twice with PBS, and in-dish lysed [20 mM Tris-Cl pH 7.4, 150 mM NaCl, 5 mM MgCl2, 1 mM DTT, 100 µg/ml cycloheximide, 1% Triton X-100, 25 U/ml of Turbo DNase I (Thermo Fisher Scientific)]. Lysates were further triturated by passing them ten times through a 26-G needle and centrifuged at 20,000 g for 10 minutes to remove nuclei and debris. Supernatants were digested with RNaseI (100 U/µl, Thermo Fisher Scientific) for 45 minutes at room temperature. Digestion was blocked with SUPERase*In (Invitrogen), and lysates were layered on a 1 M sucrose cushion and separated by ultracentrifugation at 80,000 RPM in a TLA110 rotor for 204 minutes at 4ºC. RNA was purified from ultracentrifuged pellets using the miRNeasy kit (Qiagen) according to manufacturer's instructions. RNA footprints were resolved in 15% TBE-urea gels (Invitrogen), and ribosome-protected RNA fragments (26–34 nucleotides) were size-selected. RNA fragments were then 3′-dephosphorylated with a T4 polynucleotide kinase (New England Biolabs). Ribo-Seq libraries were prepared with the TruSeq Small RNA Library Preparation kit (Illumina). 3′-adenylated and 5′-phosphorylated adapters were ligated to the RNA fragments, and RNA was reverse transcribed using the SuperScript III cDNA synthesis kit (Invitrogen). Ribosomal DNA (rDNA) was removed by subtractive hybridization using oligonucleotides (as stated in the original protocol), and cDNAs were PCR amplified using a different primer index per sample. Libraries were sequenced by Macrogen (Seoul, Rep. of Korea). Raw reads were quality assessed with FastQC (v0.11.4) software. Sequencing adapters were removed from the 5' and 3' ends of the

reads using BBDUK. For Ribo-seq analyses, DESeq2 default options were turned off and an independent filtering was applied by using the Riborex R package as well as a minimum count transcript variance greater than 0 or 1, depending on the dataset quality. The cut off for ribosome-associated RNAs was p-value < = 0.01 and a log2 fold-change > = 0.5.

Assessment of Translational Efficiency

Translational efficiency for each RNA was determined by normalizing the Ribo-seq data to transcript length and total transcript abundance according to the corresponding RNA-seq data, as previously described [209]. Count tables for mRNAs (RNA-seq) and ribosome-protected RNAs (Ribo-seq) were calculated from the StringTie output, with PrepDe.py. To analyze differential translation of mRNAs the R package RiboDiff was used. This package assesses translation efficiency (TE) by comparing the relative differences in the ratio of Ribo-seq mRNA fractions to total mRNA expression levels, with a cut off of an absolute fold-change in TE > 0.5 and a p-value < 0.01. The final TE gene candidates were selected from the intersection of transcripts without RNA abundance differences and transcripts with significantly less ribosome occupancy in LN229 NSUN5 cells. For DBTRG-05MG cells, the final TE gene candidates were selected from the intersection of transcripts without RNA abundance differences and transcripts with significantly less ribosome occupancy in DBTRG-05MG Scramble cells.

Gene functional enrichment analysis

Gene sets included in the publicly available GSEA signature database and the DAVID web service (https://david.ncifcrf.gov/) were used to carry out a gene functional enrichment analysis on GO biological processes and KEGG pathways. A hypergeometric test was performed, and the top gene clusters were defined as those with a FDR adjusted p-value < 0.05.

Stable isotope labeling by amino acids in cell culture (SILAC)

SILAC labeling was performed by using SILAC-Lys8- Arg10-Kit media (Silantes). LN229 EV- and NSUN5-transfected cells were cultured in heavy medium (Lys 8- Arg 10) or light medium (Lys 0 – Arg 0), respectively, for 10 days. After labeling, protein extraction was carried out by using RIPA buffer. Protein levels were quantified by the Lowry method

using the DC Protein Assay (BioRad) and mixed in 1:1 ratio. Proteins were next processed by the Filter aided sample preparation (FASP) method, using disposable centrifugal ultrafiltration units that allow for detergent depletion, protein digestion, and isolation of peptides. Samples were further subjected to alkylation with iodoacetamide and overnight digestion with Lys-c and trypsin at 37°C in 50 mM ammonium bicarbonate. Peptide mixes were analyzed using an OrbitrapPhusion Lumos mass spectrometer (Thermo Scientific) coupled to an EasyLC (Thermo Scientific). All data were acquired with Xcalibur software v3.0.63. Proteome Discoverer software suite v2.0 (Thermo Fisher Scientific) and the Mascot search engine v2.5 (Matrix Science) were used for peptide identification and quantification. Samples were searched against a SwissProt database containing entries corresponding to Human (April 2016 version), a list of common contaminants and all the corresponding decoy entries. Resulting data files were filtered for FDR < 1%. Only peptide intensities corresponding to master protein category were used. The ratios from average peptide intensities were calculated and normalized for each protein. Significantly upregulated and downregulated proteins between EV- and NSUN5-transfected groups were calculated by a Student's t-test, p-value < 0.05.

Polysomal profiling

Prior to polysome fractionation, cells were incubated with 100 mg/ml cycloheximide for 15 minutes. Plates were then washed with ice-cold PBS containing 100 mg/ml cycloheximide, and cells were scraped and collected in ice-cold PBS with cycloheximide. Cells were lysed in polysome lysis buffer [20 mM Tris-HCl pH 7.4, 5 mM MgCl2, 150 mM NaCl, 1% Triton X-100, 2.5 mM DTT, 1% deoxycholate, 200 U/ml RNasin, 100 mg/ml cycloheximide, protease inhibitor cocktail cOmplete (Roche), α1-antitrypsin (EMD Biosciences)]. Lysates were assessed for protein concentration by Bradford assay, and equals amounts of protein were loaded on a 15%–50% sucrose gradient containing 100 mg/ml cycloheximide, 1 mM DTT and 0.2 mg/ml heparin. Gradients were separated by ultracentrifugation at 36,000 RPM in a SW40 rotor for 3 hours at 4ºC, subsequently fractionated using an ISCO-Foxy Jr. fraction collector and ordered by absorbance at 254 nm using a UA-6 absorbance detector connected to the fraction collector. RNA was extracted with TRIzol Reagent (ThermoFisher), and equal amounts of each sample were then retrotranscribed and analyzed by qRT-PCR (see Expression analyses).

In vitro cell growth

Cell proliferation was assessed by the sulforhodamine B (SRB) assay. 1,000 to 3,000 cells per well were seeded in 96-well plates. During 7 consecutive days, cells were fixed with 10% trichloroacetic acid, washed twice with distilled water and stained with 0.4% SRB in 1% acetic acid for 30 minutes. SRB was then solubilized in 10mM Tris-HCl pH 10.0 and 540 nm-optical densities were determined using a microplate reader (Perkin Elmer Viktor 3). Data was analyzed with GraphPad Prism 5 software.

Cell viability

For *in vitro* IC_{50} studies, 5,000 cells per well were seeded in 96-well plates, and maintained overnight to allow cell attachment prior to drug exposure. Cell viability was assessed by the SRB assay (as previously described) 48h after drug treatment. For *in vivo* treatment with isobutyl-Deoxynyboquinone (IB-DNQ), 12 mice were inoculated with $2x10^6$ A172 or DBTRG-05MG cells into the brain, and 7 mice of each group were randomly chosen to be treated, while five mice remained untreated. Mice were orally administered every two days, with either IB-DNQ (12 mg/g) or vehicle (2-hydroxypropyl-beta-cyclodextrin, HP-CD).

Electronic microscopy

For transmission electron microscopy, cells at 80% of confluence were first fixed in 2.5% glutaraldehyde in 0.1 M phosphate buffer, and next fixed in 1.5% osmium tetroxide. Samples were processed using the tEPON 812 embedding Kit (Tousimis®). Ultrathin sections were stained with uranyl acetate and lead citrate, and assessed at 80 kv with a JEM-1011 (JEOL) transmission electron microscope. Ribosome number was quantified from different randomly selected regions of several images (n = 8) by three different researchers, and normalized by the area.

Autoradiographic Analysis of rRNA Synthesis

To label newly synthesized RNA, cells were incubated for 2 hours with 1.2 µCi/ml of [^{3}H]-uridine (Perkin Elmer). RNA was then extracted with TRIzol reagent (Invitrogen). 2 µg of total RNA were resolved on a formaldehyde-containing 1.2 % agarose gel and transferred to Hybond N+ membrane (GE Healthcare). Membranes were UV-crosslinked, sprayed with EN3HANCE (Perkin Elmer) and exposed to Kodak BioMax MS film (Kodak) at -80^{o}C for 1 week.

Animal models

Orthotopic tumor xenografts

IVIS images: 1x10^6 cells were stereotactically inoculated into the corpus striatum of right brain hemispheres of 9-week-old Nude-Foxn1nu athymic mice (Charles River Laboratories). Tumor size was estimated by luciferase activity of inoculated glioma cells, quantified in a Xenogen-CCD camera from IVIS. All mouse experiments were approved by and performed in accordance with the guidelines of the Institutional Animal Care Committee of the Vall d'Hebron Research Institute in agreement with the European Union and national directives.

Xenograft's growth*:* 3x10^5 cells for LN229 (EV or NSUN5), and 1,5x10^6 cells for DBTRG-05MG (Scramble or shNSUN5) were inoculated into the corpus striatum of 9 (Scramble and shNSUN5) or 10 (EV and NSUN5) mice for each cell line individually. All mouse experiments were approved and performed in accordance with the guidelines of the Institutional Animal Care Committee of IDIBELL.

Subcutaneous tumor xenografts

In vivo tumor growth was analysed on five-week-old male athymic nu/nu mice (Charles River, Wilmington, MA, USA). Xenografts were initiated by subcutaneously inoculating 2x10^6 LN229 EV and NSUN5 (n = 8) or 4x10^6 DBTRG-05MG Scramble and DBTRG-05MG-shNSUN5 glioma cells (n = 8), into the left and right flanks, respectively. Tumor growth was measured in two dimensions using calipers, and tumor volume (V) was estimated applying the formula: V (mm^3) = (L × W^2)/2, being L (length) > W (width). Mice were euthanized 28 days after inoculation, and tumors were excised and weighed.

Clinical data

Study I

Gene mutation, copy number variation (CNV), DNA methylation and mRNA expression data in cell lines were obtained from COSMIC cell line database and the TCGA Data Portal. DNA methylation and mRNA expression data in the TCGA set of glioma patients were obtained from the TCGA Data Portal, and values for LGG and GBM TCGA primary tumor samples were obtained from the NCI's Genomic Data Commons (GDC) (https://gdc.cancer.gov/). To assemble the glioma validation cohort, 115 formalin-fixed paraffin-embedded (FFPE) glioma tumors were retrospectively collected from four different centers from 1989 to 2018 (Hospital Germans Trias i Pujol, Hospital Universitari de Bellvitge, Hospital Clínic de Barcelona and University Hospital Basel) and were histologically reviewed. Most of these patients received temozolamide as first-line treatment, in combination or not with radiotherapy. All cases were provided with patient-informed consent under the ethical approval from each Institution Boards/Ethics Committees.

Statistical analysis

The associations between variables were assessed by Fisher’s exact test, Welch's t-test, Chi-squared test, Mann-Whitney test, Wilcoxon paired test or Spearman correlation as appropriate. Kaplan-Meier plots and log-rank test were used to estimate Progression-Free Survival (PFS) and Overall Survival (OS). Univariate Cox regression analysis was performed in which hazard ratio with a 95% of confidence interval was evaluated. Statistical analyses were carried out by using SPSS (Armonk, NY) and GraphPad Prism 5 (La Jolla, CA). P-values lower than 0.05 were considered statistically significant.

RESULTS

RESULTS

Identification of epigenetically altered m^{5}C RNA methylation-related genes in human cancer

To identify epigenetic defects affecting m^{5}C RNA methylation in human tumorigenesis, we first decided to perform a pan-cancer analysis of our Illumina 450K methylation array data collection, as one of the major epigenetic mechanisms causing transcriptional silencing is gene promoter hypermethylation. From all the CpG sites interrogated by the 450K methylation array, promoter-associated CpG islands were defined as those CpGs located in DNA regions spanning 200 base pairs upstream to the Transcription Start Site (TSS) of a gene. Promoter-associated CpG islands of all m^{5}C-related genes described so far, including writers (*NSUN1-7*, *DNMT2*), readers (*ALYREF*, *YBX1*) and one eraser (*ALKBH1*) were assessed for their promoter methylation status across more than one thousand human cancer cell lines, belonging to the COSMIC Cell Lines project panel [210].

Out of the eleven m^{5}C-related genes assessed, two were found hypermethylated in a relevant percentage of the cancer cell lines studied: the m^{5}C RNA methyltransferase genes *NSUN5* and *NSUN7* (**Figure 8**).

We also wondered if other mechanisms could be affecting the expression of our candidate genes in cancer. The COSMIC Cell Line panel also counts with exome mutational and copy number (CN) data, allowing us to rule out potential point mutations or gene amplifications. Mining of the available CN data revealed practically no gene amplifications in the analyzed cell lines, while point mutations of some m^{5}C-related genes appeared specifically enriched in endometrium, cervix and gastric cancer cell lines, although not at elevated percentages (**Figure 9A, B**). Despite our focus was directed towards epigenetic mechanisms regulating m^{5}C RNA methylation, it would also be interesting to characterize the effect of m^{5}C regulators point mutations on tumorigenesis.

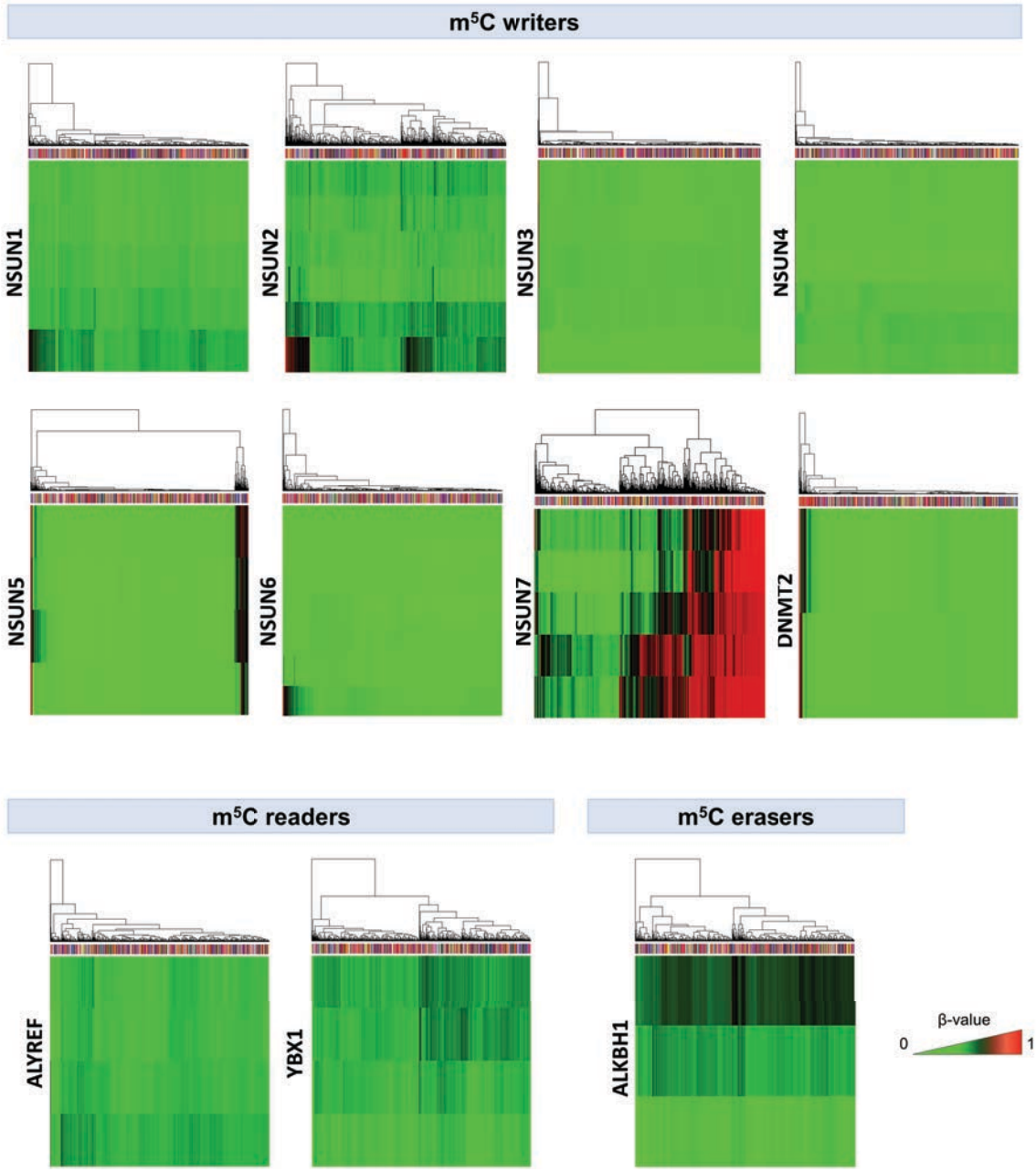
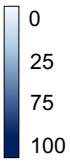


Figure 8. Promoter-associated CpG island methylation status of m^{5}C-related genes across cancer cell lines. Heatmap representation of the methylation levels of the CpGs considered. Every single CpG has an absolute methylation level (β-value), ranging from 0 to 1. Green, unmethylated; red, methylated. Columns correspond to cell lines and rows to CpGs.

A

	NSUN1	NSUN2	NSUN3	NSUN4	NSUN5	NSUN6	NSUN7	DNMT2	ALYREF	YBX1	ALKBH1
Autonomic Ganglia	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Biliary Tract	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bone	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Breast	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CNS	0.0	0.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Cervix	0.0	7.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Colon	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Endometrium	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Haematopoietic & Lymphoid Tissue	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Kidney	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Liver	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Lung	0.6	4.1	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0
Oesophagus	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Ovary	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Pancreas	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Skin	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Soft Tissue	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	0.0
Stomach	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Thyroid	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Urinary Tract	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0



B

	NSUN1	NSUN2	NSUN3	NSUN4	NSUN5	NSUN6	NSUN7	DNMT2	ALYREF	YBX1	ALKBH1
Autonomic Ganglia	3.0	6.1	0.0	0.0	6.1	6.1	0.0	0.0	0.0	3.0	0.0
Biliary Tract	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Bone	0.0	2.9	2.9	0.0	2.8	2.9	0.0	0.0	0.0	0.0	2.9
Breast	0.0	4.0	0.0	4.0	8.0	6.1	0.0	4.0	0.0	2.0	2.0
CNS	1.8	0.0	0.0	1.9	5.5	0.0	1.9	0.0	0.0	3.7	1.9
Cervix	0.0	7.1	0.0	0.0	14.3	7.1	7.1	14.3	0.0	7.1	0.0
Colon	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Endometrium	16.7	33.3	8.3	0.0	25.0	0.0	25.0	8.3	0.0	0.0	8.3
Haematopoietic & Lymphoid Tissue	4.4	2.5	1.9	0.6	9.4	1.9	2.5	5.1	1.3	4.4	1.9
Kidney	3.3	3.3	0.0	0.0	6.7	0.0	3.3	3.3	0.0	0.0	0.0
Liver	0.0	0.0	0.0	0.0	0.0	7.1	0.0	7.1	0.0	0.0	0.0
Lung	2.3	4.1	3.0	2.3	2.9	5.3	0.6	1.2	1.8	4.1	1.8
Oesophagus	0.0	8.3	2.9	0.0	5.6	5.7	0.0	0.0	5.7	11.1	0.0
Ovary	0.0	8.1	2.7	8.1	0.0	0.0	2.7	2.7	0.0	5.4	5.4
Pancreas	0.0	0.0	3.6	0.0	3.4	3.6	0.0	0.0	0.0	7.1	0.0
Skin	3.6	9.3	1.9	3.7	0.0	3.7	7.4	1.9	0.0	1.9	5.6
Soft Tissue	17.4	4.3	4.3	0.0	4.3	4.3	4.3	0.0	0.0	0.0	0.0
Stomach	0.0	7.4	0.0	0.0	0.0	0.0	3.7	14.8	0.0	7.4	3.7
Thyroid	6.3	0.0	0.0	0.0	0.0	0.0	0.0	6.3	0.0	6.3	0.0
Urinary Tract	0.0	0.0	0.0	0.0	5.3	5.3	0.0	0.0	0.0	0.0	0.0

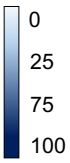


Figure 9. Frequency of genetic alterations on m^{5}C-related genes across cancer cell lines. Heatmap representation of **(A)** gene amplifications and **(B)** somatic mutations of the candidate gene set in the COSMIC Cell Line panel. Color intensities represent the percentage of cell lines of every tumor type affected. A gene was considered amplified with a total copy number >= 8.

STUDY I: Insights into the epigenetic loss of the m^5C RNA methyltransferase NSUN5, and its contribution to gliomagenesis

Characterization of NSUN5 promoter hypermethylation in glioma

The aforementioned data mining of the DNA methylation array data across cell lines showed that NSUN5 promoter CpG island was hypermethylated in different malignancies, with the highest percentages in central nervous system (CNS) cancer (38%, 21 of 55 cell lines), thyroid cancer (25%, 4 of 16 cell lines), biliary tract cancer (20%, 1 of 5 lines), autonomic ganglia cancer (18%, 6 of 33 lines) and soft-tissue sarcoma (13%, 3 of 23 cell lines) (**Figure 10A**).

To confirm that NSUN5 epigenetic alteration takes place in human cancer samples, ruling out a potential cell line event, NSUN5 promoter CpG island was assessed for its methylation status in The Cancer Genome Atlas (TCGA) human primary tumors collection. This analysis revealed NSUN5 promoter hypermethylation happening mainly in human primary gliomas from all grades (28%, 190 of 681 tumor samples), and to a minor extent, in gastric tumors (5%,18 of 361 tumors) (**Figure 10B**). Importantly, NSUN5 promoter CpG island was unmethylated in all the normal tissues studied (Data not shown), suggesting a cancer-specific epigenetic event.

To take a deeper look inside NSUN5 promoter methylation in human glioma, we subdivided the TCGA glioma sample collection according to the histological tumor grade based on the degree of anaplasia, following the classification established by the World Health Organization (WHO) on Tumors of the Central Nervous System (CNS) [211]. The TCGA glioma collection counts with diffuse low-grade gliomas (LGGs; grades II and III), including astrocytomas, oligoastrocytomas and oligodendrogliomas, and high-grade gliomas (grade IV), also known as glioblastoma multiforme (GBM). NSUN5 promoter CpG island was found hypermethylated in 34% of LGG tumors (180 of 527), and in 6% of glioblastomas (10 of 154) (**Figure 10C**).

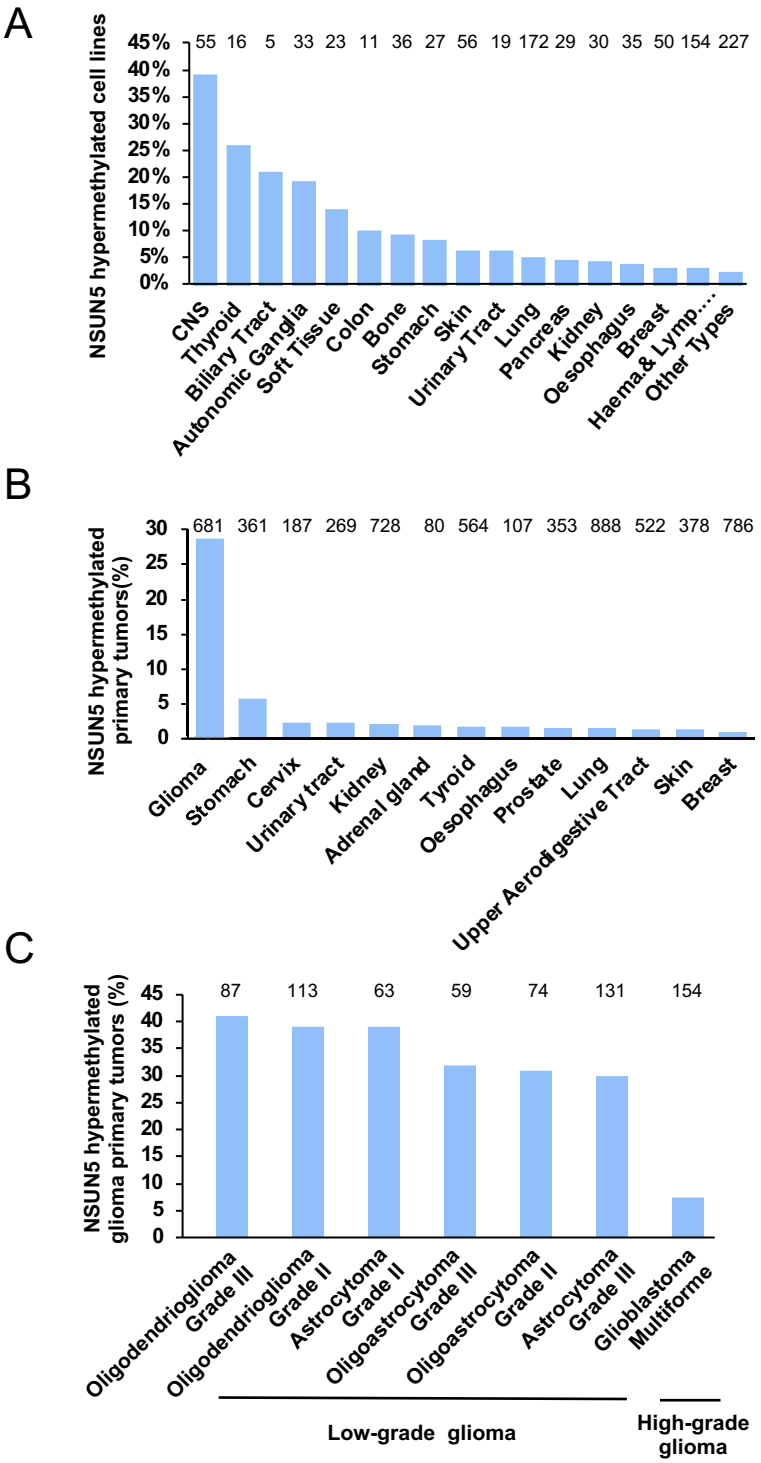


Figure 10. NSUN5 promoter hypermethylation in glioma. (A) Percentage of human cancer cell lines with a hypermethylated NSUN5 promoter. **(B)** Percentage of NSUN5 promoter methylation in primary tumors from the TCGA collection, classified by tumor type. **(C)** Percentage of primary glioma tumors with a hypermethylated NSUN5 promoter, classified according to the histological tumor grade as established by the WHO. Total number of cell lines or primary tumors is shown on top of each bar.

NSUN5 promoter hypermethylation is associated with its transcriptional silencing

Having identified a cancer-specific NSUN5 epigenetic defect in glioma, we sought to determine if NSUN5 promoter hypermethylation was responsible for a transcriptional inactivation of the *NSUN5* gene. A pan-cancer analysis of the available transcriptome profiles of the COSMIC Cell Line collection showed that NSUN5 hypermethylation was, indeed, associated with transcript downregulation (Spearman's rank correlation, rho = −0.336; p-value < 0.0001) (**Figure 11A**), being NSUN5 significantly more expressed in hypomethylated cells (**Figure 11B**). The negative correlation between promoter methylation and *NSUN5* mRNA expression was magnified when analyzing only glioma cell lines (Spearman's rank correlation, rho = − 0.756; p-value < 0.0001) (**Figure 11C**), with hypermethylated cells significantly reduced of *NSUN5* mRNA expression (**Figure 11D**).

RNA-sequencing (RNA-seq) data from the TCGA primary glioma cohort confirmed the association between NSUN5 promoter methylation and transcript downregulation across all glioma tumors (rho = − 0.7; p-value < 0.0001), and after tumor grade classification into low- and high-grade gliomas (rho = − 0.65; p-value < 0.0001 and rho = − 0.37; p-value = 0.0044, respectively) (**Figure 12**).

To study this epigenetic mechanism in greater detail we moved to an *in vitro* setting. We selected two subsets of glioma cell lines, a LGG subset, comprising three cell lines, and a glioblastoma subset, composed of 6 cell lines, based on NSUN5 promoter methylation status. According to the DNA methylation array, the glioblastoma cell lines DBTRG-05MG, MO59J and CAS-1 have a NSUN5 unmethylated promoter, while the A172, LN229, and KS-1 cell lines display a hypermethylated promoter (**Figure 13A**). In order to validate the array data, we performed bisulfite genomic sequencing of multiple clones of those cell lines, using primers that spanned the TSS-associated CpG island. We confirmed our previous results, finding NSUN5 promoter-associated CpG island hypomethylated in the DBTRG-05MG, CAS-1, and MO59J cells, while it was hypermethylated in the LN229, A172, and KS-1 glioma cell lines (**Figure 13C**). Importantly, according to the methylation array data, NSUN5 promoter was hypomethylated in normal brain cells and samples, including commercial astrocytes, fetal brain, and different adult brain regions (frontal cortex, posterior cingulate cortex, hippocampus and white matter) (**Figure 13B**), and the array results were validated by bisulfite genomic sequencing in the case of white matter and astrocytes samples (**Figure 13C**).

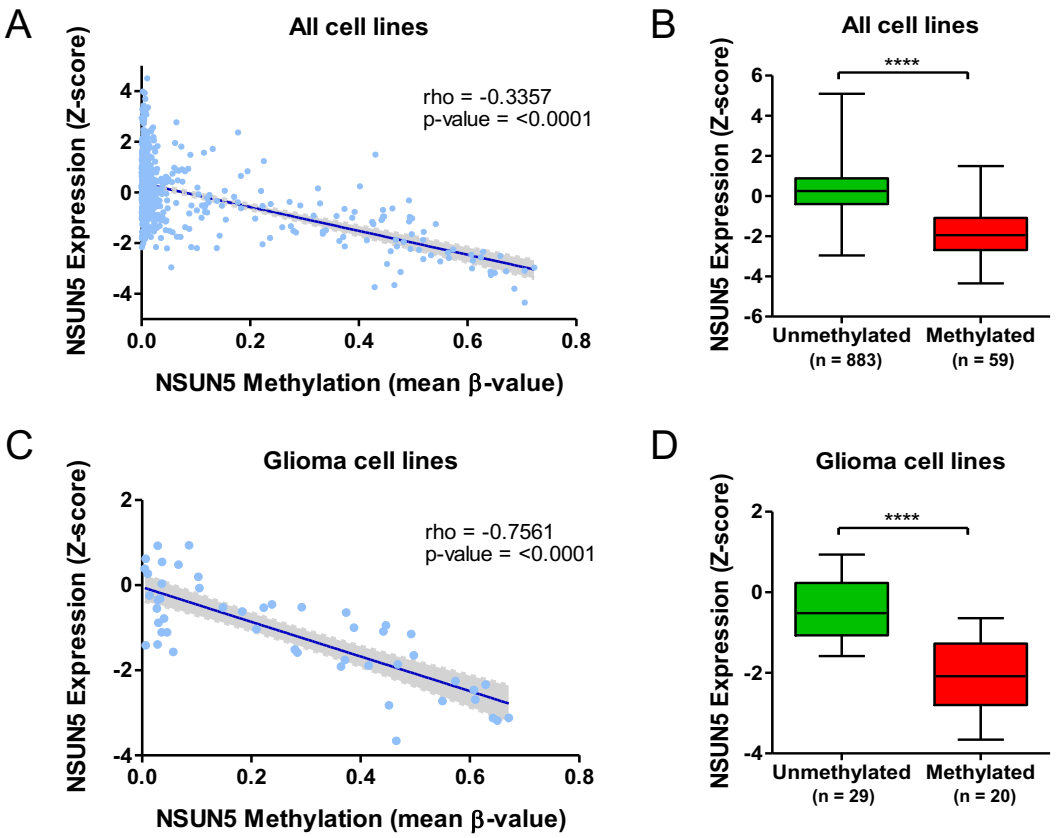


Figure 11. *In silico*-detected NSUN5 transcriptional silencing by promoter CpG island hypermethylation in cancer cell lines. (A) NSUN5 methylation is inversely correlated with *NSUN5* transcript expression in human cell lines belonging to the COSMIC panel (n = 942). Spearman's rank correlation test with its p-value and the associated rho coefficient are shown. **(B)** Box plots displaying *NSUN5* transcript expression (Z-Score) in unmethylated and methylated cell lines from all cancer types. **(C)** Correlation analysis between NSUN5 promoter methylation (mean β-value) and *NSUN5* transcript expression (Z-Score) in glioma cell lines (n = 49). Spearman's rank correlation test p-value and rho coefficient are shown. **(D)** Box plots representing *NSUN5* transcript abundance (Z-Score) in glioma cell lines show significantly reduced expression in methylated lines.

NSUN5 expression on the cell line panel was analyzed both at the RNA and protein levels, by quantitative real-time PCR and western blot, respectively. NSUN5-hypermethylated glioblastoma cell lines LN229, A172 and KS-1 minimally expressed NSUN5, whereas the unmethylated cells (DBTRG-05MG, CAS-1, and MO59J) expressed NSUN5 at both levels (**Figure 14A**). In order to validate if NSUN5 silencing is caused by a DNA methylation-mediated epigenetic mechanism, hypermethylated glioblastoma cell lines were treated with 5-aza-2-deoxycytidine or decitabine (AZA). This molecule is a cytosine analogue that inhibits DNA methyltransferases (DNMTs), promoting a passive demethylation through cell divisions. AZA treatment restored NSUN5 expression both at the RNA and protein levels (**Figure 14B**).

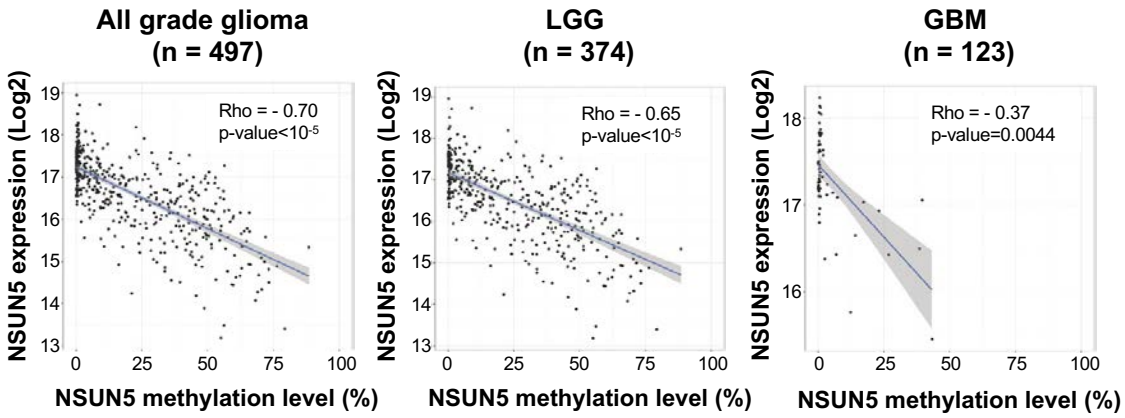


Figure 12. *In silico*-detected NSUN5 transcriptional silencing by promoter CpG island hypermethylation in human primary gliomas. NSUN5 promoter methylation is associated with its transcript downregulation in primary gliomas from all grades, in LGG and in GBM from the TCGA cohort. Spearman's rank correlation tests with their p-values and associated rho coefficients are shown.

The previous *in vitro* characterization was carried out, as well, in the LGG cellular panel. It was composed of the grade III astrocytoma cell lines SW1088 and MOG-G-CCM, both wild-type for IDH1, and the BT142 mut/− cell line that harbors a homozygous IDH1 mutation (R132H). According to the methylation array, the BT142 mut/− cell line is hypermethylated, whereas the SW1088 and MOG-G-CCM cell lines are hypomethylated for the NSUN5 promoter (**Figure 15A**). Cells were subjected to bisulfite genomic sequencing, and results showed the same promoter methylation pattern obtained with the DNA microarray (**Figure 15B**). Moreover, BT142 mut/− cell line, with proved hypermethylated promoter had minimal expression of NSUN5, analyzed by quantitative real-time PCR and western blot, compared to the hypomethylated SW1088 and MOG-G-CCM cell lines (**Figure 15C**). Importantly, treatment with the demethylating molecule AZA led to a recovery of NSUN5 expression in the BT142 mut/− hypermethylated cell line (**Figure 15D**). All these data indicate that NSUN5 promoter CpG island hypermethylation leads to its silencing in glioma.

A

Target_ID	Location	Glioma cell lines						
		DBTRG-05MG	M059J	CAS-1	A172	LN-229	KS-1	Astrocytes
cg27330627	1stExon	0.00	0.06	0.22	0.81	0.75	0.75	0.00
cg09232499	TSS200	0.20	0.19	0.33	0.79	0.69	0.73	0.00
cg21802401	TSS200	0.20	0.17	0.20	0.65	0.58	0.59	0.00
cg01251255	TSS200	0.19	0.30	0.09	0.68	0.56	0.59	0.00

B

TargetID	Location	Fetal Brain			Frontal Cortex					Hippocampus			White matter											
		FB1	FB2	FB3	FC1	FC2	FC3	FC4	FC5	H1	H2	H3	WM1	WM2	WM3	WM4	WM5	WM6	WM7	WM8	WM9	WM10	WM11	WM12	WM13
cg27330627	1stExon	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01
cg09232499	TSS200	0.00	0.01	0.00	0.01	0.01	0.08	0.00	0.02	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01
cg21802401	TSS200	0.01	0.01	0.00	0.00	0.00	0.07	0.02	0.01	0.01	0.00	0.00	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01
cg01251255	TSS200	0.00	0.00	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.01

TargetID	Location	Posterior cingulate cortex																							
		PCC1	PCC2	PCC3	PCC4	PCC5	PCC6	PCC7	PCC8	PCC9	PCC10	PCC11	PCC12	PCC13	PCC14	PCC15	PCC16	PCC17	PCC18	PCC19	PCC20	PCC21	PCC22	PCC23	PCC24
cg27330627	1stExon	0.00	0.02	0.00	0.00	0.03	0.02	0.02	0.00	0.02	0.00	0.03	0.03	0.01	0.02	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
cg09232499	TSS200	0.05	0.00	0.03	0.02	0.03	0.05	0.01	0.05	0.05	0.01	0.04	0.02	0.04	0.04	0.01	0.03	0.03	0.01	0.03	0.02	0.02	0.05	0.02	0.03
cg21802401	TSS200	0.02	0.03	0.04	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.02	0.04	0.04	0.03	0.02	0.02	0.02	0.03	0.04	0.02	0.03
cg01251255	TSS200	0.02	0.01	0.02	0.01	0.00	0.02	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.03	0.01	0.00	0.03	0.03	0.03	0.01	0.03

C

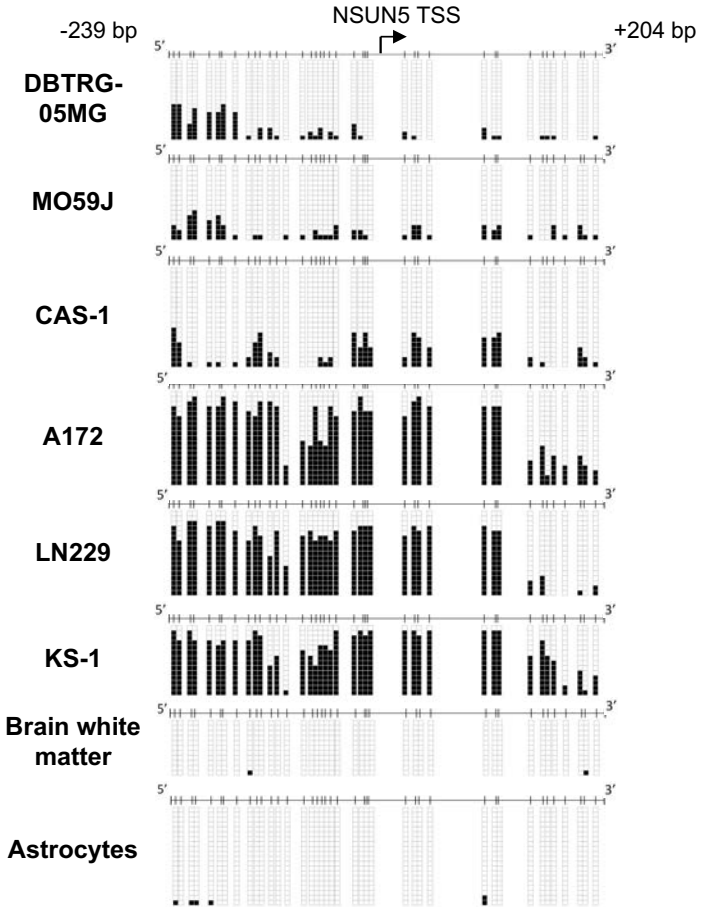


Figure 13. Characterization of NSUN5 promoter methylation in glioma cell lines. DNA methylation levels (absolute β-values) of the NSUN5 promoter-associated CpG island analyzed by the 450K methylation array. Green, unmethylated; red, methylated. Data from the selected GBM cellular set and normal astrocyte cells **(A)**, and from normal fetal and adult brain samples from different regions **(B)**. **(C)** Bisulfite genomic sequencing of NSUN5 promoter CpG Island in GBM cells, plus two normal samples (brain white matter and astrocyte cells). CpG dinucleotides are represented as short vertical lines, NSUN5 TSS is indicated by a black arrow, and unmethylated or methylated cytosines are represented as white or black squares, respectively. Single clones are shown for each sample (n > 10).

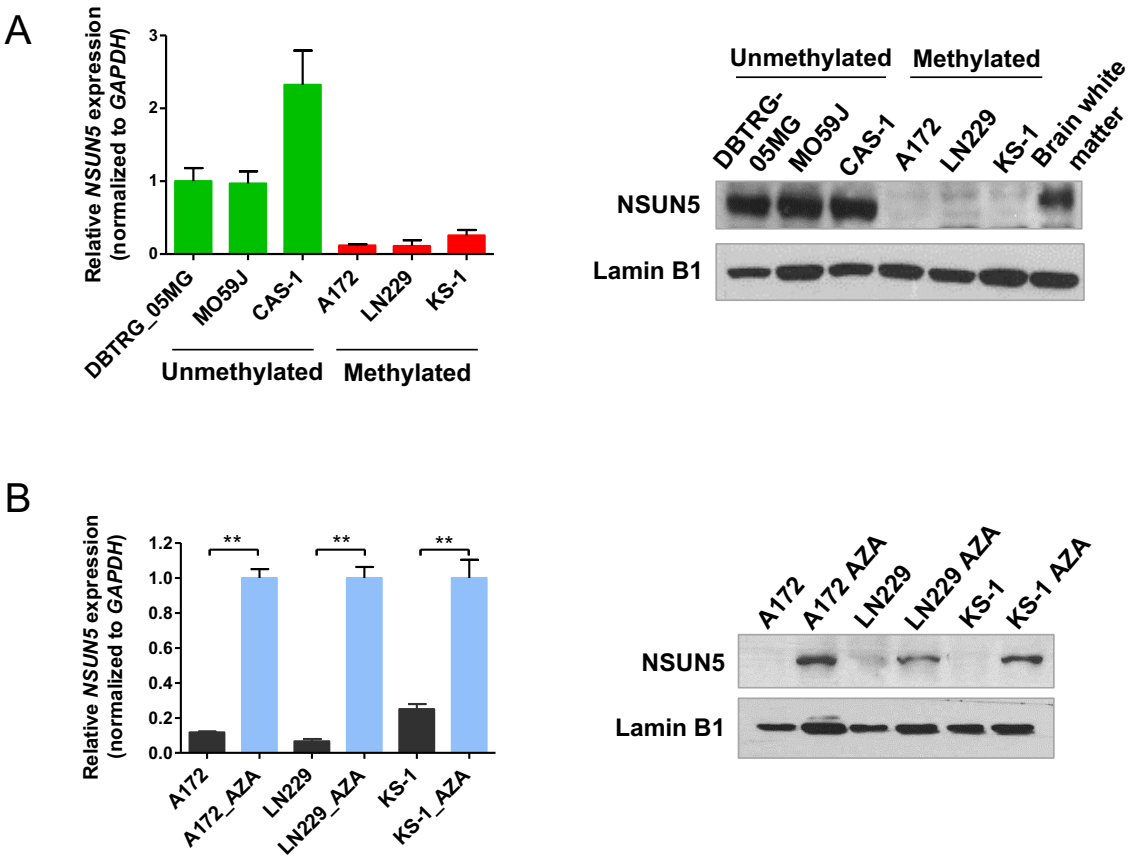


Figure 14. NSUN5 epigenetic loss in GBM cell lines. (A) NSUN5 expression levels in the glioblastoma set of cell lines, plus a brain white matter sample, determined by real-time PCR (*left*) and western blot (*right*). **(B)** NSUN5 expression at both, RNA (*left*) and protein levels (*right*), was restored in the three hypermethylated cell lines (A172, LN229 and KS-1) after treatment with the demethylating drug AZA. RNA expression data shown represent the mean ± S.D. of biological triplicates, and p-values were calculated by the Mann–Whitney test. **p-value < 0.01

A

Target_ID	Location	SW1088	MOG-G-CCM	BT142 mut/-
cg27330627	1stExon	0.03	0.03	0.62
cg09232499	TSS200	0.02	0.02	0.63
cg21802401	TSS200	0.02	0.02	0.54
cg01251255	TSS200	0.03	0.02	0.51

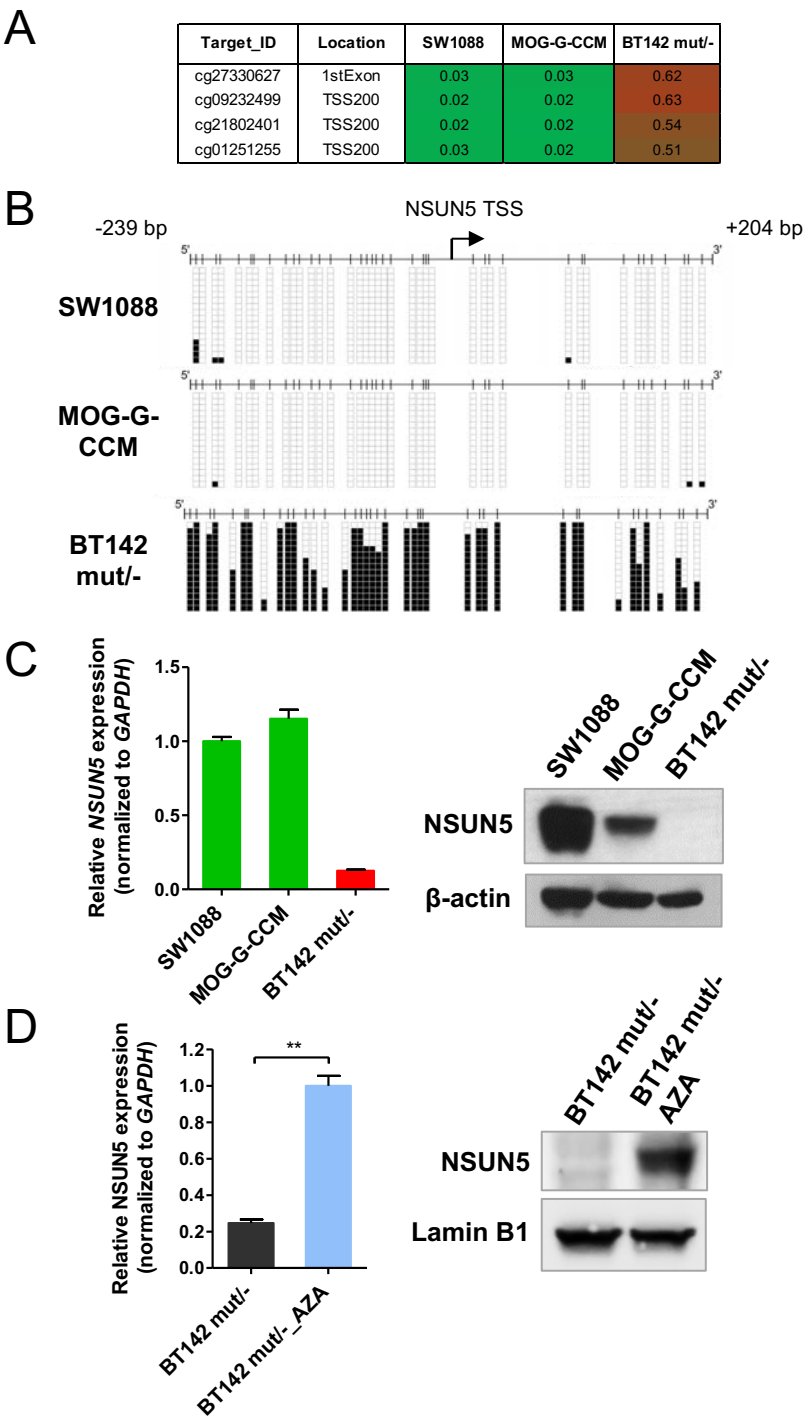


Figure 15. NSUN5 epigenetic silencing by promoter hypermethylation in LGG cell lines. (A) Methylation β-values of the NSUN5 promoter-associated CpG island in the selected LGG cellular panel, analyzed by the 450K methylation array. Green, unmethylated; red, methylated. **(B)** In vitro validation of the array data by bisulfite genomic sequencing of NSUN5 promoter CpG Island in LGG cells. CpG dinucleotides are represented as short vertical lines, NSUN5 TSS is indicated by a black arrow, and unmethylated or methylated cytosines are represented as white or black squares, respectively. Single clones are shown for each sample (n > 10). **(C)** Expression levels of the NSUN5 transcript and protein, analyzed by real-time PCR (left), and western blot (right). **(D)** NSUN5 expression at both, RNA (*left*) and protein levels (*right*), was restored in the BT142 mut/− NSUN5 promoter hypermethylated cell line by treatment with the demethylating agent AZA. RNA expression data shown represent the mean ± S.D. of biological triplicates, and p-values were calculated by the Mann–Whitney test. **p-value < 0.01.

NSUN5 shows tumor suppressor properties both *in vitro* and *in vivo*

Having demonstrated the epigenetic loss of NSUN5 in glioma cell lines, we studied its biological contribution to the growth of these cells. To this end, we developed both NSUN5 gain-of-function and loss-of-function models. NSUN5 expression was restored in the NSUN5-hypermethylated glioblastoma cell lines LN229 and A172 by stable transduction of an integrative expression vector (**Figure 16A**). On the contrary, NSUN5-hypomethylated glioblastoma cell lines DBTRG-05MG and CAS-1 were depleted of NSUN5 expression by means of short hairpin interference RNA (shRNA) (**Figure 16B**). NSUN5 recovery led to a significant reduction of cellular growth *in vitro*, compared with empty-vector-transduced (EV) cells (**Figure 16C**), and, conversely, NSUN5 depletion increased cell viability (**Figure 16D**).

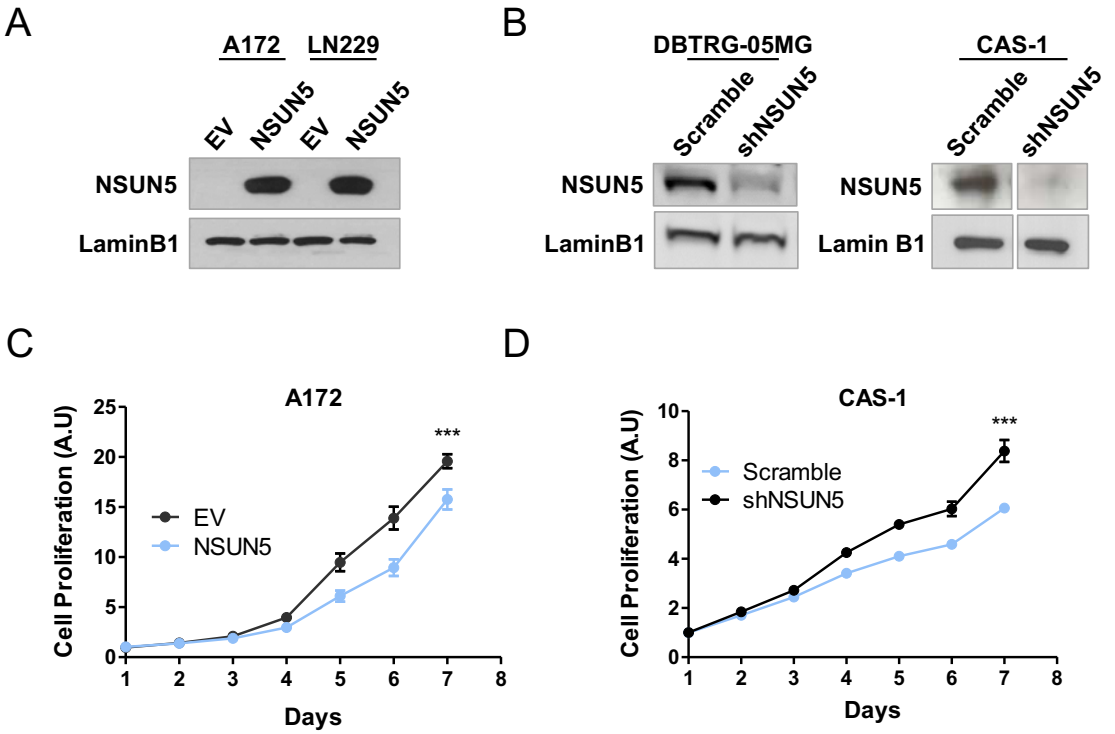


Figure 16. NSUN5 role in glioma cell proliferation. (A) Restoration of NSUN5 protein expression by stable transduction in A172 and LN299 glioblastoma cells, analyzed by western blot. **(B)** Efficient shRNA-mediated depletion of NSUN5 protein expression in the hypomethylated DBTRG-05MG and CAS-1 glioma cell lines. **(C)** SRB assay shows that NSUN5 restoration decreases cell viability in the NSUN5 hypermethylated cell line A172. **(D)** NSUN5 depletion increases cell viability in the hypomethylated CAS-1 glioblastoma cell line. Viability data shown represent the mean ± S.D. of biological triplicates, and p-values were calculated by the Mann–Whitney test. ***p-value < 0.001.

Based on these results, we decided to test the tumorigenic capacity of the cells *in vivo*, by carrying out subcutaneous mouse xenotransplantation experiments. We observed a reduced volume of the tumors derived from NSUN5-transduced LN229 cells, compared to EV-transduced cells (**Figure 17A**), whereas tumors arising from NSUN5-depleted DBTRG-05MG cells had a significantly higher volume (**Figure 17C**). Moreover, tumor samples obtained at the experiment endpoint had a lower weight when derived from NSUN5-transduced LN229 cells, compared to EV-transduced cells, and conversely, tumors originated from NSUN5-depleted DBTRG-05MG cells weighed more than their Scramble-transduced counterparts (**Figures 17B, D**).

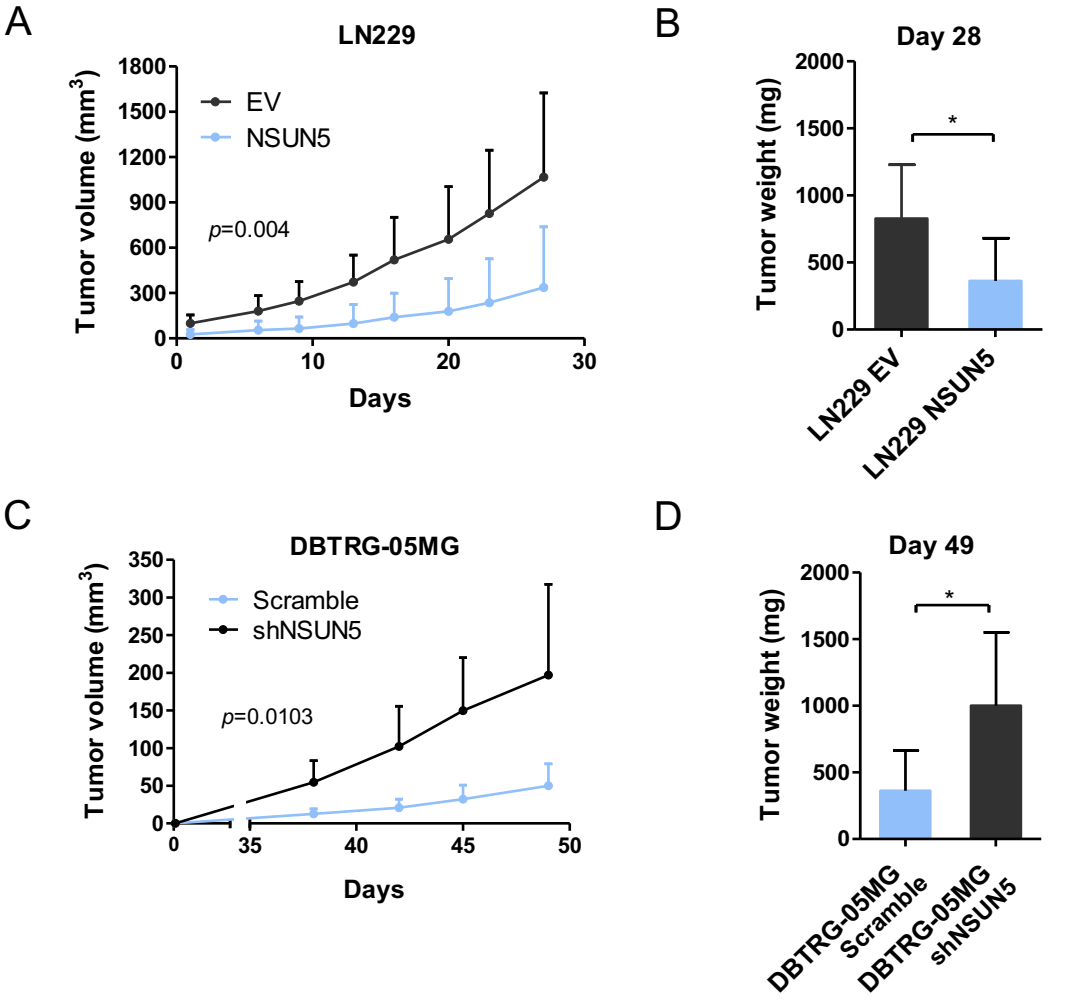


Figure 17. NSUN5 role in subcutaneous tumor xenograft growth. (A) LN229 EV- and NSUN5-transduced cells were subcutaneously injected in the left and right flank of 10 mice, respectively, and tumor volume was measured over time. **(B)** Tumors originated from EV- and NSUN5-transduced LN229 cells were weighed at the experiment endpoint. **(C)** DBTRG-05MG cells Scramble and NSUN5-shRNA-depleted were injected in the left and right flank of 10 mice, respectively. Tumor volume, measured over time, is shown. **(D)** Weight of tumors derived from Scramble and NSUN5-depleted DBTRG-05MG cells, analyzed at the experiment endpoint. Error bars show means ± s.d. P-values were calculated by Student's T test. *p-value < 0.05, **p-value < 0.01.

To better characterize NSUN5 role on the tumorigenic cell growth of gliomas we also performed orthotopic mouse xenotransplantation experiments, which allow to create a tumor microenvironment that resembles more to the actual tumor. LN229 and A172 EV- and NSUN5-transduced cells were Luciferase-transduced prior to their mouse brain inoculation, in order to quantify tumor growth over time by analyzing luciferase activity. The study of orthotopically injected glioma cells into athymic mice showed similar results: tumors deriving from NSUN5-transfected LN229 and A172 glioblastoma cells had a significantly reduced size in comparison with EV-transfected tumors at day 10 post-inoculation (**Figure 18A**). On the contrary, tumors originating from DBTRG-05MG cells depleted of NSUN5 expression had a significantly bigger size to those deriving from Scramble-transduced cells, at day 6 post-inoculation (**Figure 18B**).

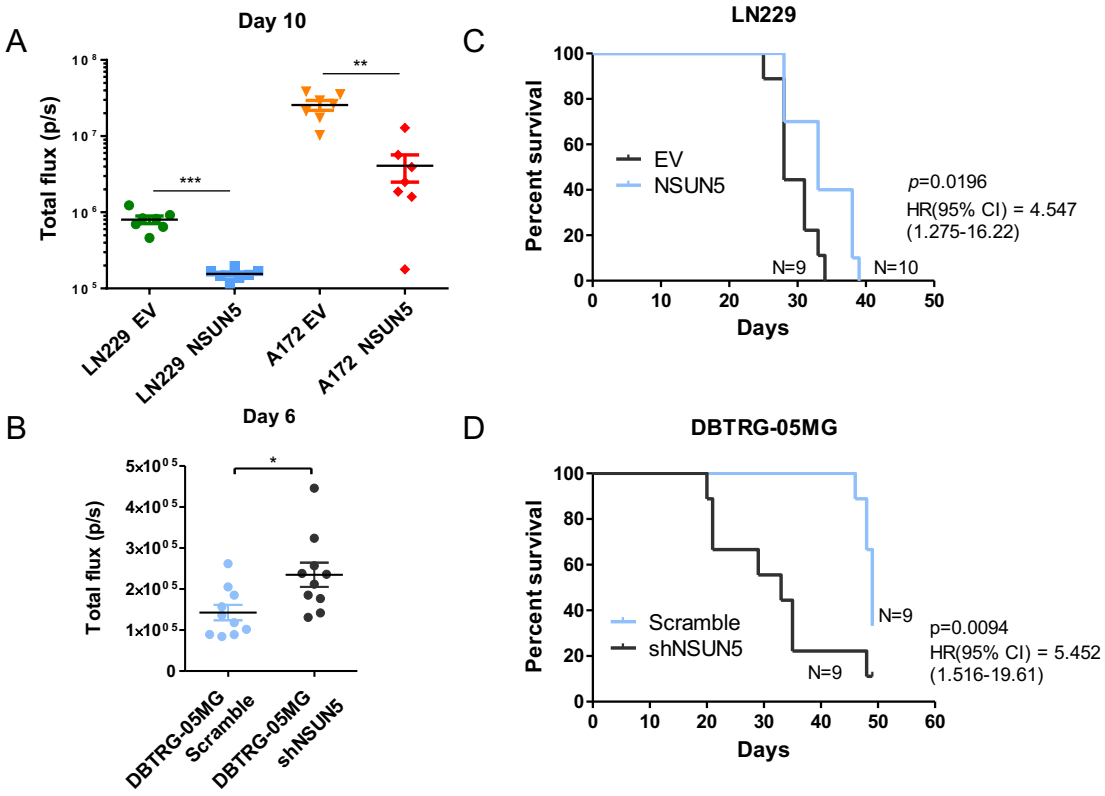


Figure 18. NSUN5 role in orthotopic tumor xenograft growth. (A) LN229 and A172 cells, EV- and NSUN5-transduced were stereotactically injected into the brain of athymic mice, and tumor growth was analyzed at day 10 post-inoculation by quantifying luciferase activity of tumoral cells. **(B)** DBTRG-05MG Scramble and shNSUN5 cells were stereotactically inoculated into the brain of athymic mice. Scatter plots show the size of individual tumors at day 6 post-inoculation. Error bars show means ± s.d. P-values were calculated by Student's T test. *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001. **(C)** Kaplan-Meier curves showing the survival of mice with intracranial implanted tumors originated from the hypermethylated LN229 cell line, EV- and NSUN5-transduced. **(D)** Survival of mice intracranially-inoculated with tumors derived from Scramble and NSUN5-depleted DBTRG-05MG cells. P-values were calculated by the log-rank test, and results of the univariate Cox regression analysis are shown by hazard ratio (HR) and 95% confidence interval (CI). *p-value < 0.05, **p-value < 0.01.

We also analyzed survival of orthotopically-injected mice, and observed that NSUN5 recovery into the LN229 cell line was associated with a longer overall survival of animals (log-rank; p-value=0.0196; hazard ratio [HR] = 4.547, 95% CI = 1.275–16.22) (**Figure 18C**), whereas NSUN5 downregulation by shRNA-mediated depletion in DBTRG-05MG cells decreased the survival of mice (log-rank; p-value=0.0094; hazard ratio [HR] = 5.452, 95% CI = 1.516–19.61) (**Figure 18D**). Overall, these observations indicate that NSUN5 displays tumor suppressor properties in glioma cells.

NSUN5 m^5C RNA methyltransferase is responsible for methylating the C3782 position of the 28S rRNA

Next, we wondered if NSUN5 tumor suppressor properties might be better understood by characterizing NSUN5 role in human cells. As explained at the Introduction chapter, NSUN5 is a putative m^5C RNA methyltransferase whose human targets have not been formally described yet. However, the function of its orthologues in lower eukaryotes is well described: to methylate a single cytosine on the corresponding 28S orthologue RNAs. Thus, we sought to identify NSUN5 target(s) in human cells, and how the loss of their methylated cytosines due to NSUN5 epigenetic loss could result in an advantage for cancer cells. For this reason, we first performed an alignment of the regions of the 25S and 26S rRNAs containing the known methylated cytosines, the C2278 of the yeast 25S rRNA, methylated by Rcm1 (*S. cerevisiae* orthologue) and the C2381 of the worm 26S rRNA, methylated by *nsun-5* (*C. elegans* orthologue), and included the mouse and human 28S rRNAs. What we first observed is that this rRNA region is extremely conserved across eukaryotes, which might imply a critical function on protein synthesis (**Figure 19A**). Next, that C2381 and C2278 positions correspond to the cytosine 3782 of the human 28S rRNA (**Figure 19A**). We also carried out a phylogenetic analysis by aligning the protein sequences of human NSUN5, with its previously described orthologues in *C. elegans*, *S. cerevisiae* and *M. musculus*, and observed a highly conserved RNA-methyltransferase domain between species (**Figure 19B**).

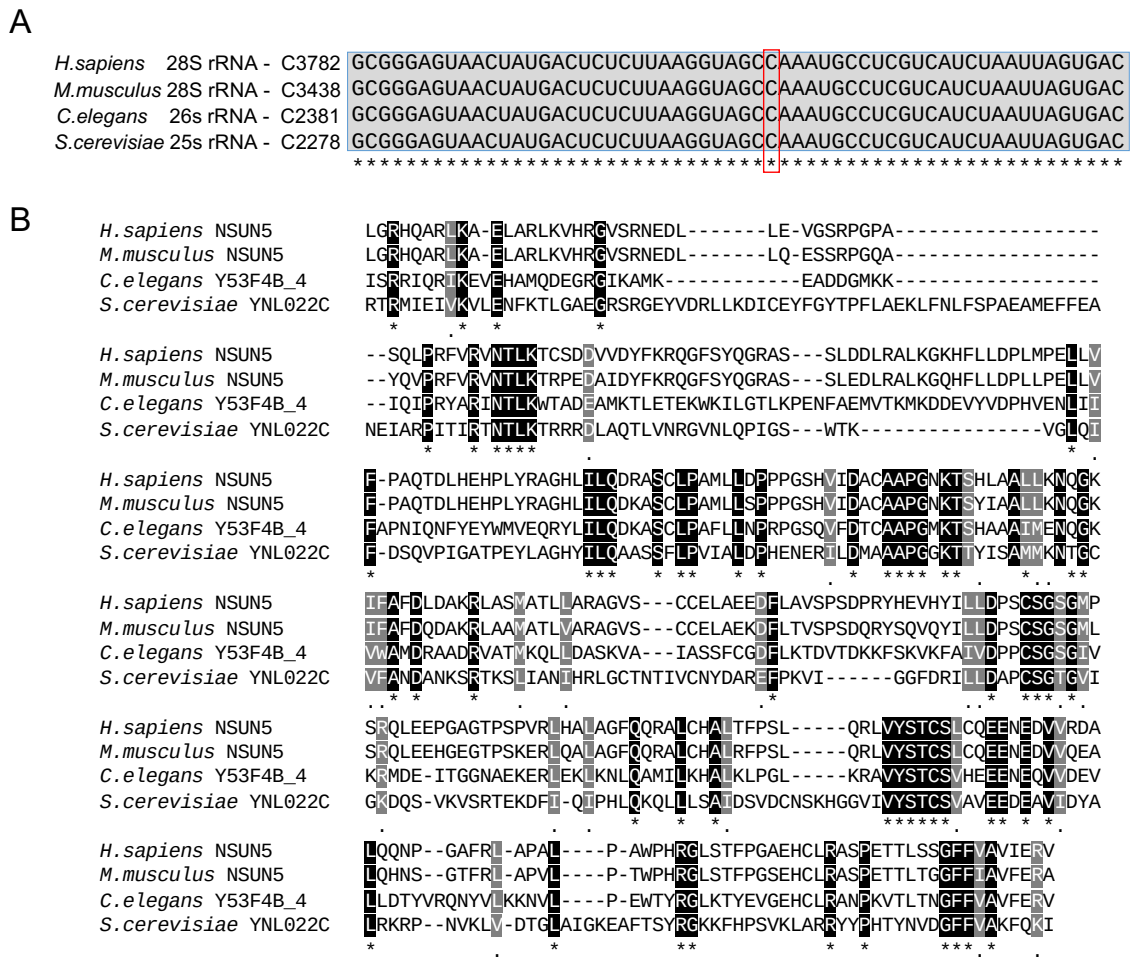


Figure 19. NSUN5 protein and putative target are evolutionary conserved between eukaryotes. (A) Sequence alignment of a region of the human 28S rRNA and their corresponding 28S, 26S and 25S rRNAs orthologues from *M. musculus*, *C. elegans* and *S. cerevisiae*, respectively, containing the already described target cytosines. **(B**) Sequence alignment of a region of the human NSUN5 protein, including the RNA-methyltransferase domain, with its *M. musculus*, *C. elegans* and *S. cerevisiae* protein orthologues. Regions in black depict identical residues, while in grey physicochemically similar residues (scoring > 0.5 in the Gonnet PAM 250 matrix).

Our first step towards validating this putative NSUN5 human target was to demonstrate if NSUN5 binds to the 28S rRNA. To this end we carried out an RNA pull-down assay in LN229 EV- and NSUN5-transduced cells, and observed that upon specific immunoprecipitation of NSUN5, quantitative RT-PCR of the retrieved RNA showed an enrichment of 28S rRNA transcripts (**Figure 20**).

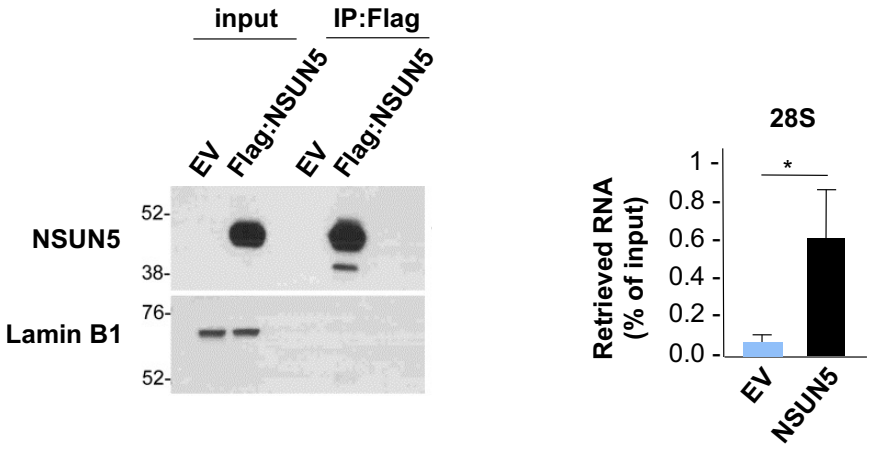


Figure 20. NSUN5 binds to the 28S rRNA. Total extracts from EV- and NSUN5-transduced LN229 cells were efficiently immunoprecipitated with an anti-Flag antibody (*left panel*), and 28S rRNA levels were measured by quantitative RT-PCR from the retrieved RNA (*right panel*).

Next, to properly validate NSUN5-mediated m^5C-methylation we performed RNA bisulfite sequencing of multiple clones of 28S rRNAs extracted from all our cell line collections. Regarding the GBM panel, NSUN5 unmethylated and expressing cell lines DBTRG-05MG, MO59J, and CAS-1 displayed high levels of C3782 28S rRNA methylation, whereas NSUN5-silenced LN229, A172, and KS-1 cell lines showed a lack of methylation at the 3782 cytosine (**Figure 21A**). Importantly, treatment of NSUN5 promoter hypermethylated cell lines LN229, A172, and KS-1 with the DNA-demethylating agent AZA, apart from restoring NSUN5 expression as previously shown (**Figure 14B**), recovered rRNA methylation at the 3782 target cytosine (**Figure 21B**). We next carried out rRNA bisulfite sequencing on the cell lines belonging to the LGG panel, as well, observing similar results: SW1088 and MOG-G-CCM cell lines, with a hypomethylated NSUN5 promoter, showed high levels of C3782 28S rRNA methylation, compared to the NSUN5 methylated and silenced BT142 mut/− cell line, which completely lacks the methyl group on the C3782 (**Figure 21C**).

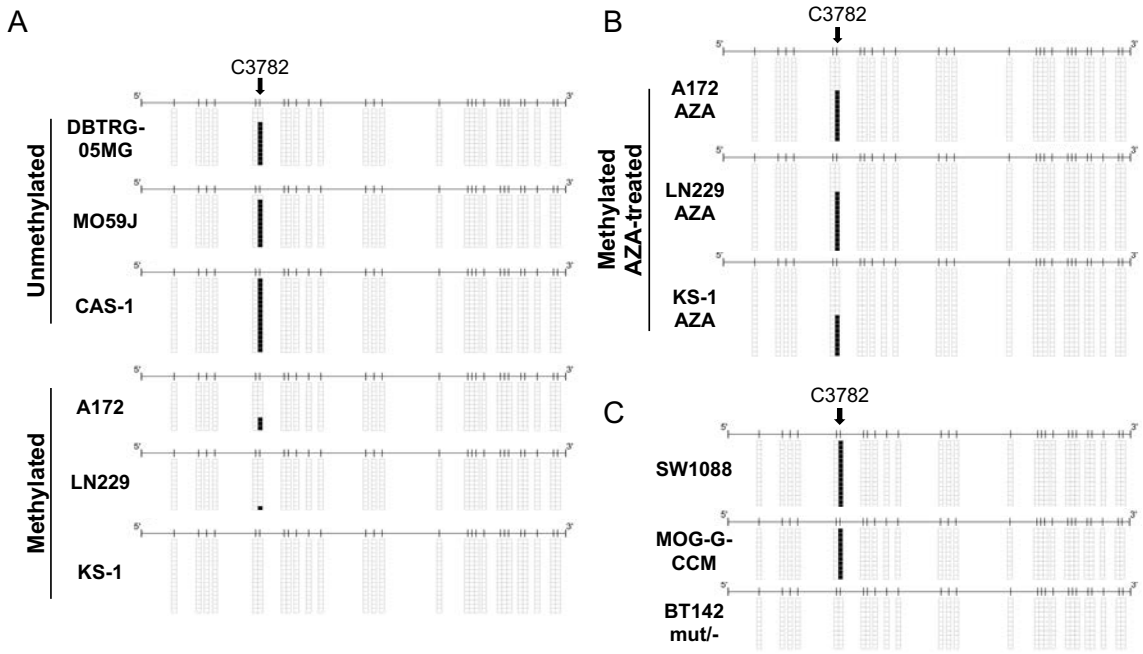


Figure 21. NSUN5 epigenetic loss abolishes the methyl group on the C3782 position of the 28S rRNA. (A) RNA bisulfite sequencing of multiple clones of the 28S rRNA from the GBM cell panel, according to NSUN5 methylation status. **(B)** RNA bisulfite sequencing of the 28S rRNA from the three NSUN5 hypermethylated glioblastoma cell lines, after treatment with the demethylating drug AZA. **(C)** RNA bisulfite sequencing of the 28S rRNA extracted from the three LGG cell lines. Cytosines are represented as short vertical lines, the C3782 site is indicated by a black arrow, and presence of a methyl group is represented as a black square.

In order to confirm these results, we decided to carry out an orthogonal method, able to detect 5-methylcytosine RNA methylation without the need for bisulfite conversion of the RNA. We decided to perform a mung bean nuclease protection assay, for isolation of the rRNA region including the C3782 position, coupled to RP-HPLC. This experiment confirmed our previous RNA bisulfite sequencing data for C3782 28S rRNA methylation: NSUN5 hypermethylated cell lines LN229, A172, and KS-1 had minimum C3782 methylation levels compared to the NSUN5 unmethylated and expressing cell lines DBTRG-05MG, MO59J, and CAS-1 (**Figure 22**).

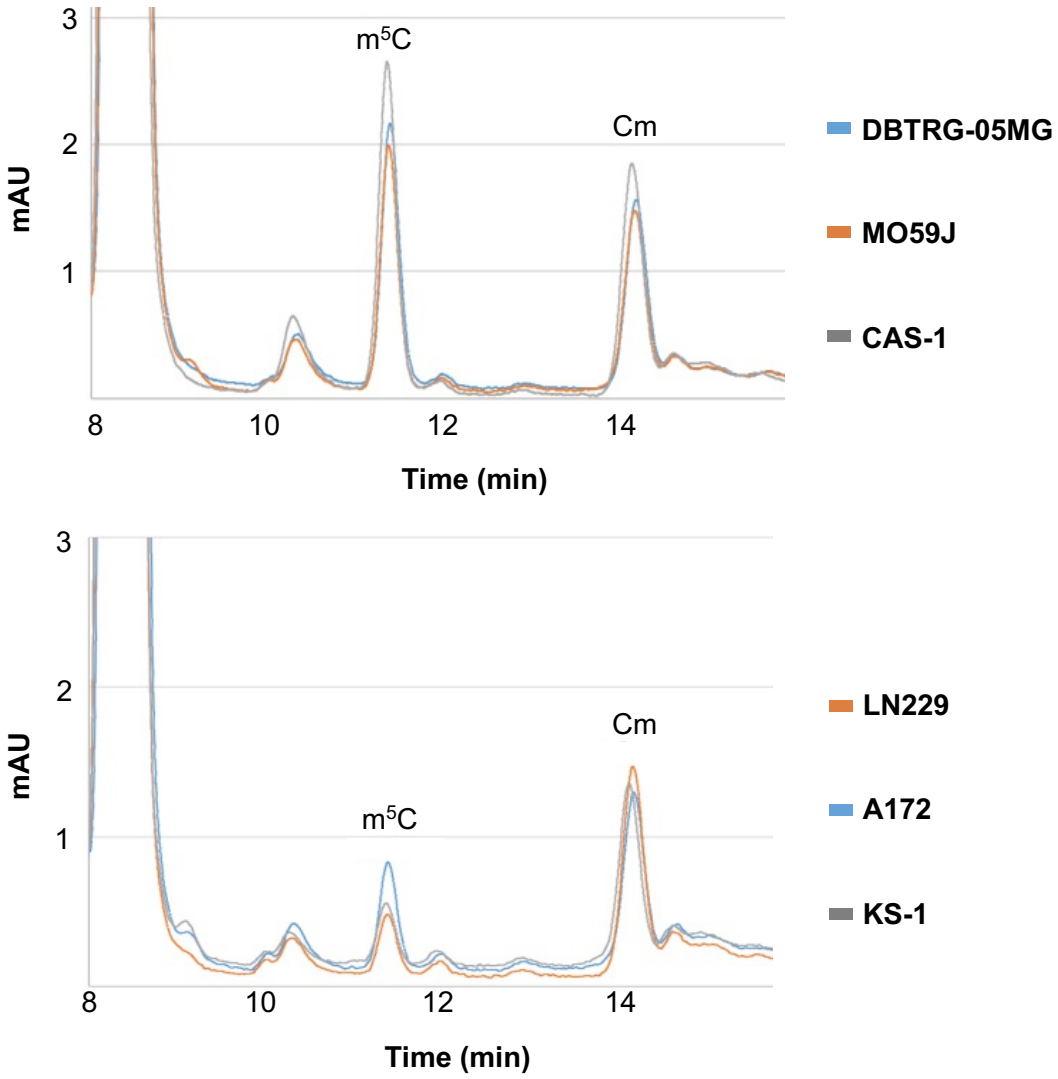


Figure 22. Methylation status of the C3782 site of the 28S rRNA by mung bean nuclease protection assay-RP-HPLC. Region of the RP–HPLC chromatogram with the nucleosides of interest. A specific region of 28S rRNAs from NSUN5 hypomethylated (*top panel*) and hypermethylated (*bottom panel*) glioma cells were isolated by hybridization to complementary deoxyoligonucleotides, digested by the mung bean nuclease and analyzed by RP–HPLC. The peak corresponding to m^5C shows a retention time of 11,5 minutes.

Having demonstrated the robustness of our validation method, we analyzed C3782 methylation levels upon NSUN5 restoration and depletion in our *in vitro* models. NSUN5 recovery in the LN229 and A172 silenced cells resulted in a recovery of C3782 28S rRNA methylation (**Figure 23A**). Conversely, NSUN5 downregulation in the DBTRG-05MG and CAS-1 expressing cell lines resulted in a decrease of methylation at the rRNA site (**Figure 23B**).

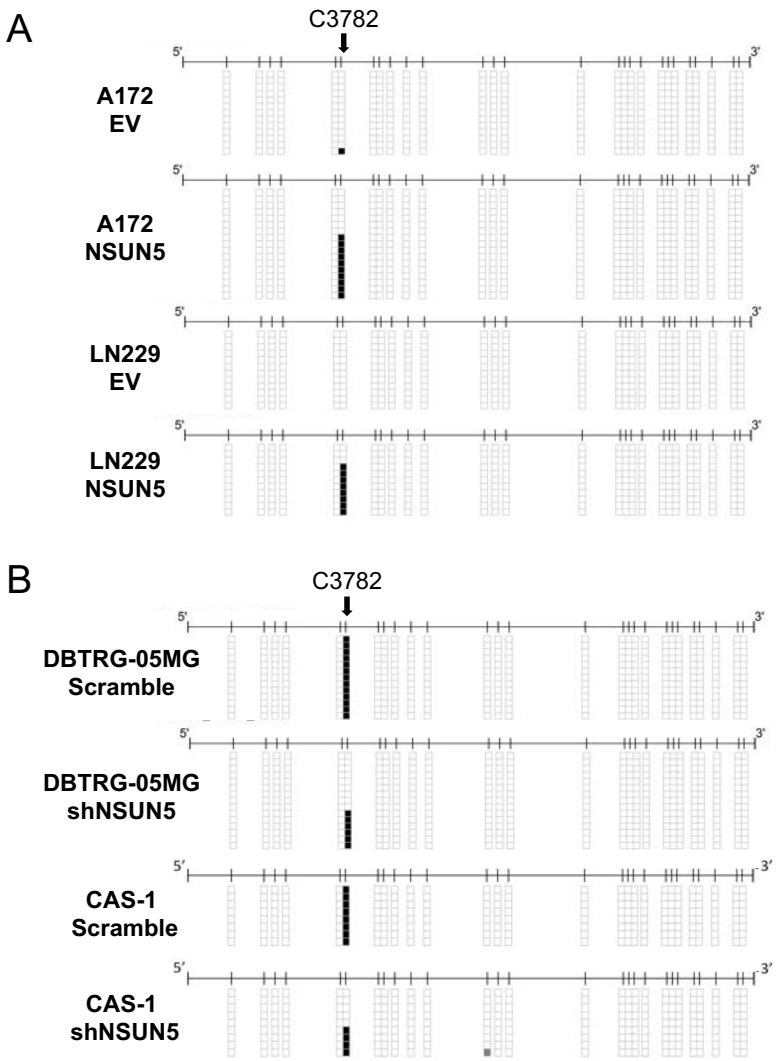


Figure 23. NSUN5 methylates the C3782 position of the 28S rRNA in glioma cells. RNA bisulfite sequencing of multiple clones of 28S rRNAs from EV- and NSUN5-transduced LN229 and A172 glioblastoma cells **(A)**, and Scramble and NSUN5-shRNA-depleted DBTRG-05MG and CAS-1 glioblastoma cells **(B)**. Cytosines are represented as short vertical lines, the C3782 site is indicated by a black arrow, and presence of a methyl group is represented as a black square.

Finally, we developed a NSUN5 catalytic defective protein (mut) by introducing point mutations at key residues for its enzymatic activity, all of them at the methyltransferase domain (**Figure 24A**). The mutant NSUN5 form was transduced into the NSUN5-hypermethylated LN229 glioblastoma cell line, together with the wild-type version, and we observed similar expression levels of both constructs, both at the RNA and protein levels (**Figure 24B**). Importantly, when looking at the methylation level of the 28S rRNA C3782 site, both by RNA bisulfite sequencing (**Figure 24C**) and mung bean nuclease protection assay-RP-HPLC (**Figure 24D**), we showed that the NSUN5 enzymatically dead version was unable to restore C3782 methylation as its wild-type counterpart.

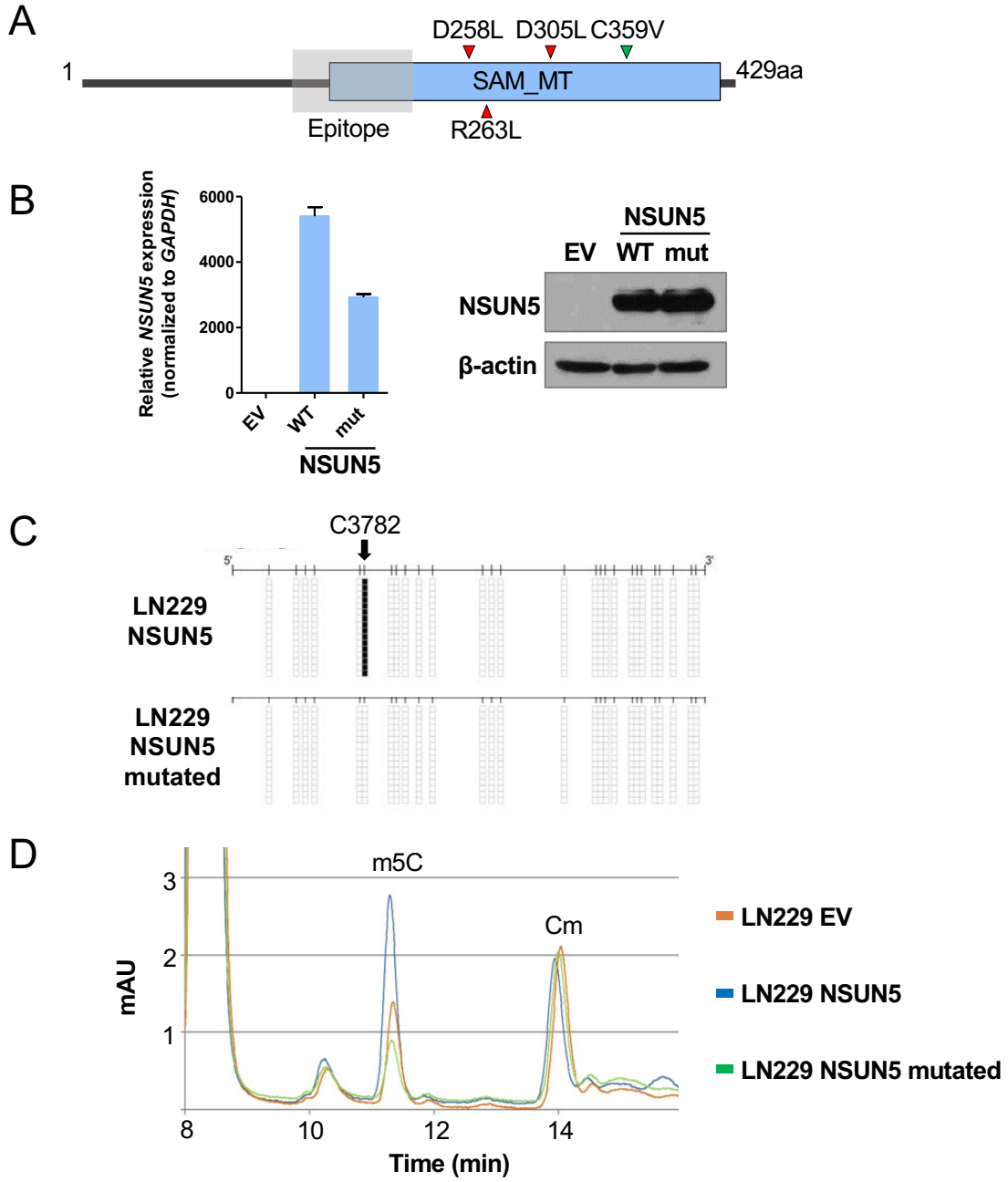


Figure 24. A catalytically-inactive NSUN5 protein is not able to restore the C3782 methylation status. (A) Schematic representation of the mutated NSUN5 protein. Triangles mark the position of substituted amino acids, with red triangles representing SAM-binding sites, and in green, the active site. The major functional domain (SAM_MT: SAM-dependent MT NOL1/NOP2/SUN domain) is displayed in blue, and the antibody epitope is marked by a grey square. **(B)** NSUN5 expression levels in the LN229 cell line, EV- and NSUN5-wild-type (WT) and mutated (mut)-transduced, determined by real-time PCR (*left*) and western blot (*right*). **(C)** RNA bisulfite sequencing of the 28S rRNA in NSUN5-silenced LN229 cells transduced with the wild-type version, and the mutated NSUN5 protein. **(D)** Methylation status of the 28S rRNA C3782 upon EV, NSUN5 wild-type and NSUN5-mutated transduction in the epigenetically silenced LN229 cell line, analyzed by mung bean nuclease protection assay-RP-HPLC.

To confirm the specificity of our approach, we analyzed the expression and methylation target of the second 28S rRNA methyltransferase, NSUN1, in our cell lines and models. First, we observed NSUN1 protein expression in the glioblastoma panel (**Figure 25A**). Next, we performed RNA bisulfite sequencing of the 28S rRNA extracted from these glioma cell lines, and analyzed the methylation status of the known NSUN1 target, the 28S 4447 cytosine. The methyl group at this position was present in all cell lines, regardless of NSUN5 expression (**Figure 25B**). Finally, we showed that recovery of NSUN5 in the silenced glioblastoma cell lines LN229 and A172 neither affect the protein level of the other 28S rRNA methyltransferase nor the methylation status of the C4447 position of the 28S rRNA (**Figure 25C, D**).

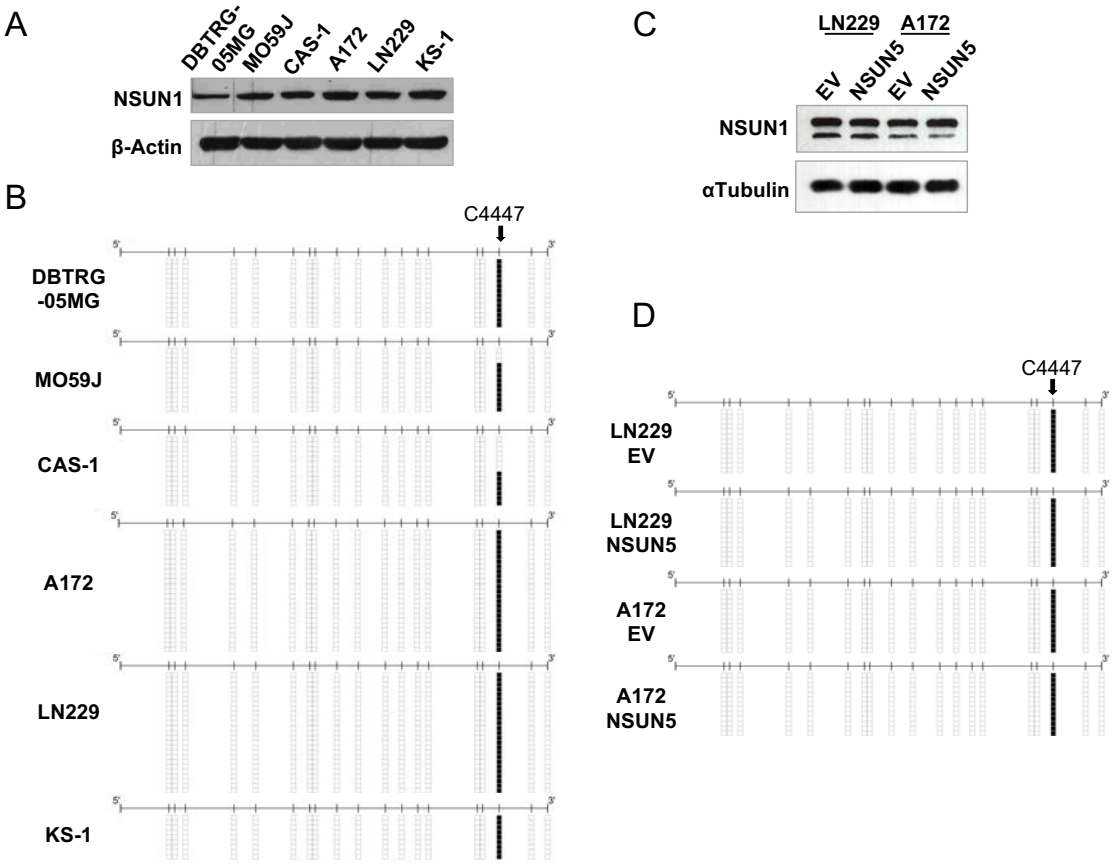


Figure 25. Ribosomal RNA methyltransferase NSUN1 expression and catalytic activity in glioma cells. (A) Western blot showing NSUN1 expression levels in the six glioblastoma cell lines. **(B)** RNA bisulfite sequencing of a region of the 28S rRNA containing NSUN1 target cytosine, in the GBM cellular panel. **(C)** Western blot showing how the restoration of NSUN5 expression in A172 and LN299 glioma cells does not alter the protein levels of the other ribosomal RNA methyltransferase, NSUN1. **(D)** RNA bisulfite sequencing of multiple clones of 28S rRNAs from EV- and NSUN5-transduced LN229 and A172 glioblastoma cells, displaying NSUN1-mediated C4447 methylation. Cytosines are represented as short vertical lines, the C4447 site is indicated by a black arrow, and presence of a methyl group is represented as a black square.

Therefore, all these results allow us to establish that the human NSUN5 RNA methyltransferase is responsible for catalyzing the incorporation of a methyl group onto the 3782 cytosine of the 28S rRNA, and that its epigenetic loss in glioma cells leads to an unmethylated status at that position.

NSUN5 m^5C RNA methyltransferase methylates mRNAs involved in the translation machinery

As previously disclosed at the Introduction chapter, many members of the NSUN family of RNA methyltransferases have numerous targets, some even from different RNA species. We wondered if this could be the case for NSUN5 as well. For this reason, we moved from our candidate approach to an unbiased setting, in order to potentially widen the spectrum of NSUN5 targets. We performed transcriptome-wide bisulfite sequencing (bsRNA-seq), by bisulfite converting total RNA from EV- and NSUN5-transduced LN229 glioma cells, and subjecting it to RNA sequencing. This unbiased approach confirmed the NSUN5-dependent methylation of the C3782 site (**Figure 26A**).

Surprisingly, we found several cytosines differentially methylated between EV- and NSUN5-transduced conditions (multiple correction adjusted p-value < 0.05 and methylation fold-change > 0.33), in mRNAs. Interestingly, the list of putative novel NSUN5 targets was especially enriched in mRNAs coding for ribosomal proteins (RPs) from the large ribosomal subunit (LSU), where the C3782 position is located (**Figure 26B**).

In order to validate these results, we first demonstrated that some of the putative targets, *RPLP0*, *RPL10* and *RPL41* mRNAs are binding targets of NSUN5, by means of immunoprecipitation of NSUN5 followed by analysis of the retrieved RNA by RT-PCR (**Figure 26C**).

We next performed RNA bisulfite sequencing of multiple clones of the candidate mRNAs from EV- and NSUN5-transduced LN229 glioma cells, but due to technical reasons, we were only capable of analyzing *RPL41* mRNA. Importantly, we observed that upon NSUN5 restoration in the silenced LN229 cell line, the C524 position on the *RPL41* mRNA recovers the methyl group, suggesting C524 as a novel cytosine methylated by NSUN5 (**Figure 27A**).

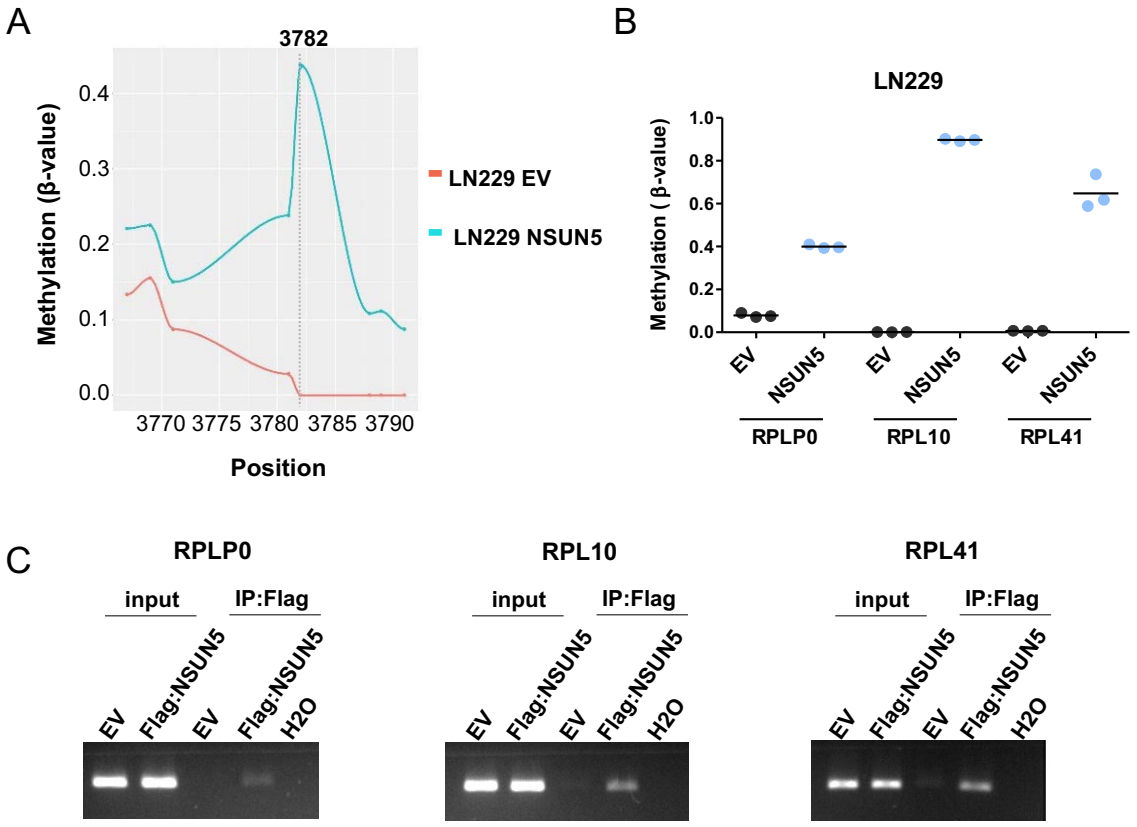


Figure 26. New putative NSUN5 targets. (A) Methylation β-value results from bsRNA-seq for the C3782 position of human 28S RNA in EV- and NSUN5-transduced LN229 glioma cells. **(B)** Methylation β-values for some RP mRNAs in EV- and NSUN5-transduced LN229 glioma cells. **(C)** NSUN5 interaction with ribosomal protein mRNAs, analyzed by RT-PCR from RNA retrieved after efficient NSUN5 immunoprecipitation with and anti-FLAG antibody.

To evaluate the effect of the m^5C524 loss upon NSUN5 epigenetic silencing, we analyzed RPL41 expression in LN229 cells, upon EV and NSUN5 transduction. We showed an increased RPL41 expression both at the RNA and the protein level upon NSUN5 restoration, and therefore, when the mark was present (**Figure 27B**), which may suggest a role for the methyl group in the mRNA stability. This hypothesis was further tested by an Actinomycin D chase assay, which showed a higher *RPL41* mRNA stability in NSUN5-recovered C524-methylated LN229 cells compared to NSUN5-lacking cells (**Figure 27C**).

Taken all together, our results may suggest a new layer of ribosomal function regulation exerted by NSUN5, which appears to not only modulate components of the RNA fraction of ribosomes, but the protein fraction as well.

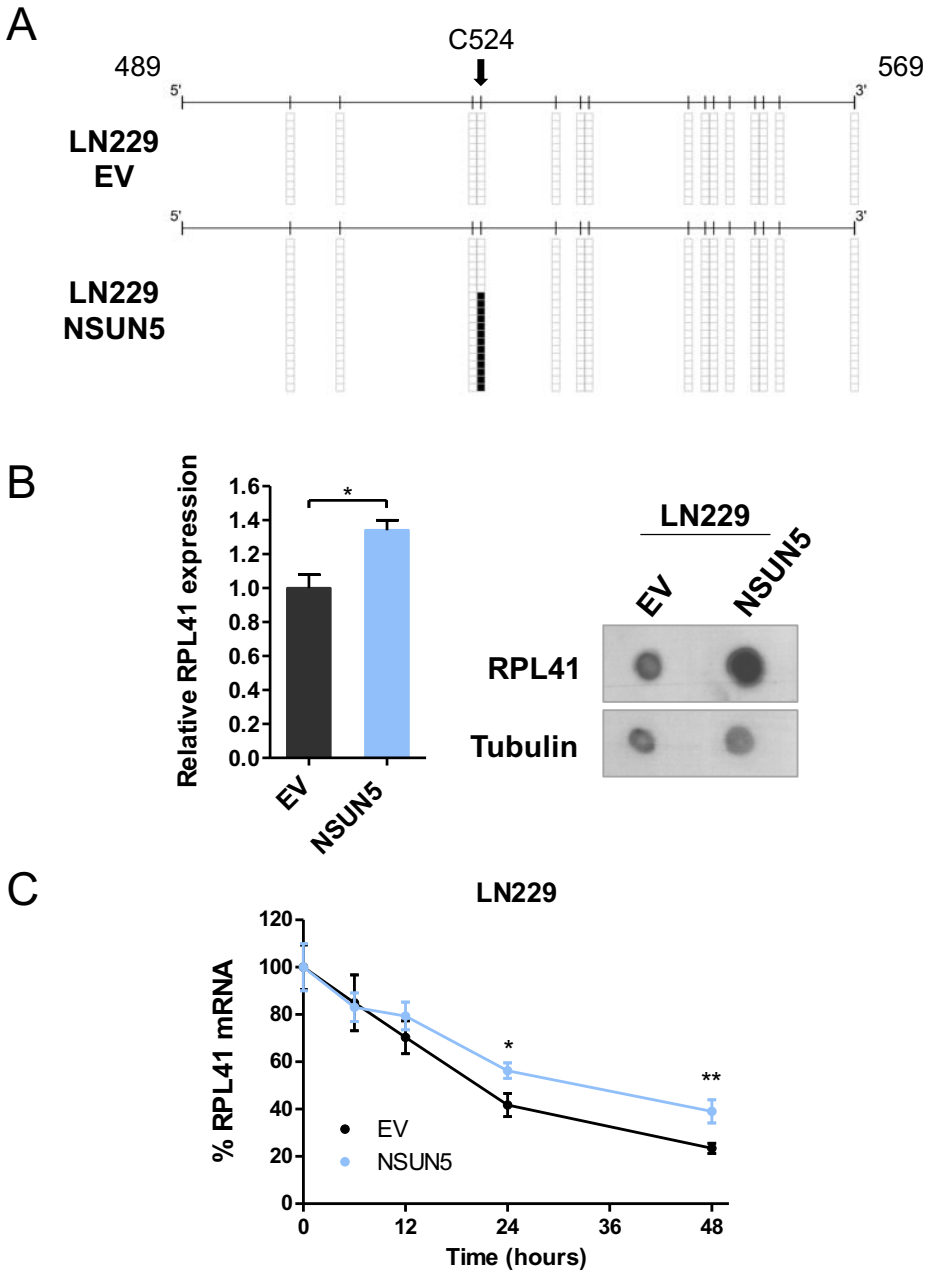


Figure 27. Effect of NSUN5-mediated C524 methylation on *RPL41* mRNA. (A) RNA bisulfite sequencing of multiple clones of *RPL41* mRNA from EV- and NSUN5-transduced LN229 glioblastoma cells. Cytosines are represented as short vertical lines, the C524 site is indicated by a black arrow, and presence of a methyl group is represented as a black square. **(B)** RPL41 expression levels in EV- and NSUN5-transduced LN229 cells, analyzed by quantitative RT-PCR (*left*) and western blot (*right*). **(C)** *RPL41* transcript half-life depending on NSUN5 expression, analyzed by Actinomycin D chase assay.

Loss of 28S rRNA m^{5}C3782 does not affect rRNA biogenesis nor ribosome distribution, but impacts ribosomal conformation

Having demonstrated NSUN5 role in methylating different RNA components of the translation machinery, we next wondered how its epigenetic loss could impact ribosomal biogenesis and function.

First of all, we analyzed rRNA biogenesis upon NSUN5 restoration in the LN229 glioblastoma cell line by northern blot, under basal conditions and under oxidative stress, and did not observe differences in rRNA levels between conditions (**Figure 28A**). Next, we evaluated if NSUN5 epigenetic loss could affect the distribution of ribosomes within cells, by means of transmission electron microscopy (TEM) imaging. Comparing hypermethylation-mediated NSUN5-silenced LN229 cells EV- and NSUN5-transduced, no major differences in ribosome distribution were observed (**Figure 28B**).

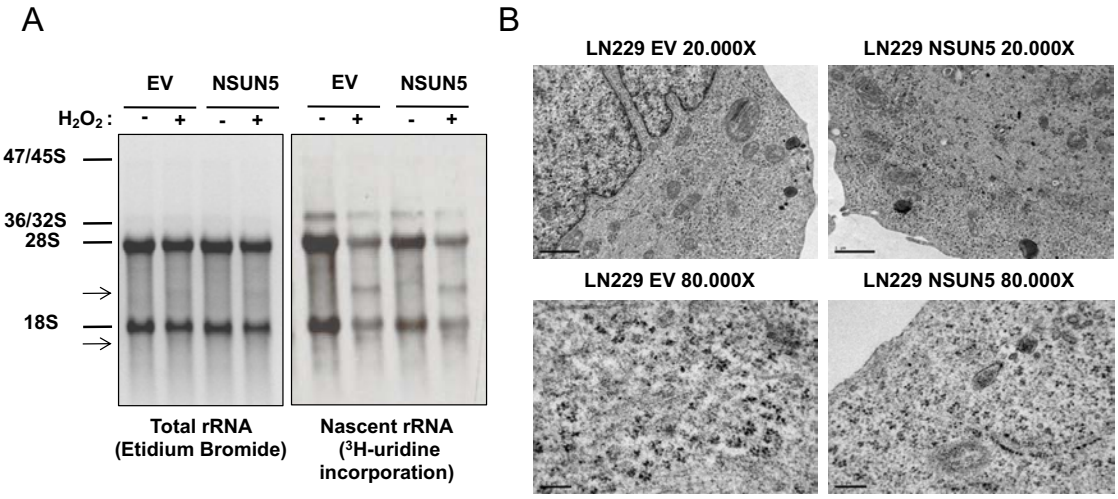


Figure 28. NSUN5 restoration does not affect rRNA biogenesis, maturation or ribosome distribution. (A) Northern blot of total RNA extracted from EV- or NSUN5-transduced LN229 cells, under basal conditions and upon H_2O_2 treatment. Total rRNA was analyzed by ethidium bromide staining, and nascent rRNA was visualized by 3H-uridine incorporation. Mature rRNA species are indicated by hyphens on the left of the pictures, and major degradation products by arrows. **(B)** TEM images show no differences in ribosome distribution upon NSUN5 restoration in hypermethylated LN229 cells.

Finally, we evaluated if the loss of methylation at the 28S rRNA C3782 site upon NSUN5 epigenetic silencing could alter human ribosomal structure, via Molecular Dynamics (MD) simulations (**Figure 29A-E**). A sub-region of the 28S subunit, comprising C3782 microenvironment, was simulated in two scenarios: basal m^{5}C3782 methylated

conditions (labelled r30A5mC, in blue), and with an unmethylated cytosine (labelled r30AC, in red). The methylated system (r30A5mC) showed a stable methyl-π interaction between m^5C3782 and the C3781 base, which was not present upon m^5C loss (r30AC) (**Figure 29D**). Moreover, loss of the 5-methylcytosine mark significantly altered the base pair between C3781 and G3810 sites, altering the conformation of a bulge located at position C3809 (**Figure 29D**). This bulge interacts with a small hairpin (comprising residues 3754–3778) that is exposed to the peptidyl site (P-site) of the ribosome (**Figure 29D, F**), and whose function is to directly interact with the tRNA and the mRNA at the P-site [212]. In the simulated non-methylated scenario (r30AC), we observed that this hairpin undergoes a conformational change that leads to a 2.8 Å displacement compared to the m^5C3782 methylated system (r30A5mC) (**Figure 29E**). This displacement moves away the hairpin from the P-site, potentially impairing its function of ensuring the stability of the tertiary complex rRNA–tRNA–mRNA (**Figure 29F**) at the P-site. Therefore, loss of a single methyl group at one cytosine of the 28S rRNA may be sufficient to alter normal protein synthesis, through impairing the tertiary complex structural stability at the P-site of the ribosome.

NSUN5 epigenetic silencing impairs global protein synthesis

To experimentally validate the aforementioned results, suggesting a link between the presence of a chemical mark in one base of the 28S rRNA, and the structural integrity of the second tRNA binding site of the ribosome, potentially affecting normal translation, we analyzed *de novo* global protein synthesis by O-propargyl-puromycin (OP-Puro) incorporation into nascent proteins and fluorescence-mediated detection. First of all, we assessed protein synthesis in the GBM cell panel, comprising, as already explained, 3 hypermethylation-mediated NSUN5 silenced cells (A172, LN229 and KS-1) and 3 NSUN5 hypomethylated and expressing cell lines (DBTRG-05MG, MO59J and CAS-1). However, under basal unstressed conditions we did not observe significant differences in global translation (Data not shown). Nevertheless, as the NSUN5-orthologue in yeast Rcm1, and its mediated methylation of C2278 of 25S rRNA has been shown to play a role in stress resistance, we repeated our approach under oxidative stress, through H_2O_2 treatment. Interestingly, we observed a significant reduction in overall protein synthesis in NSUN5-silenced C3782-hypomethylated cells, compare to those cells expressing the RNA methyltransferase (**Figure 30**).

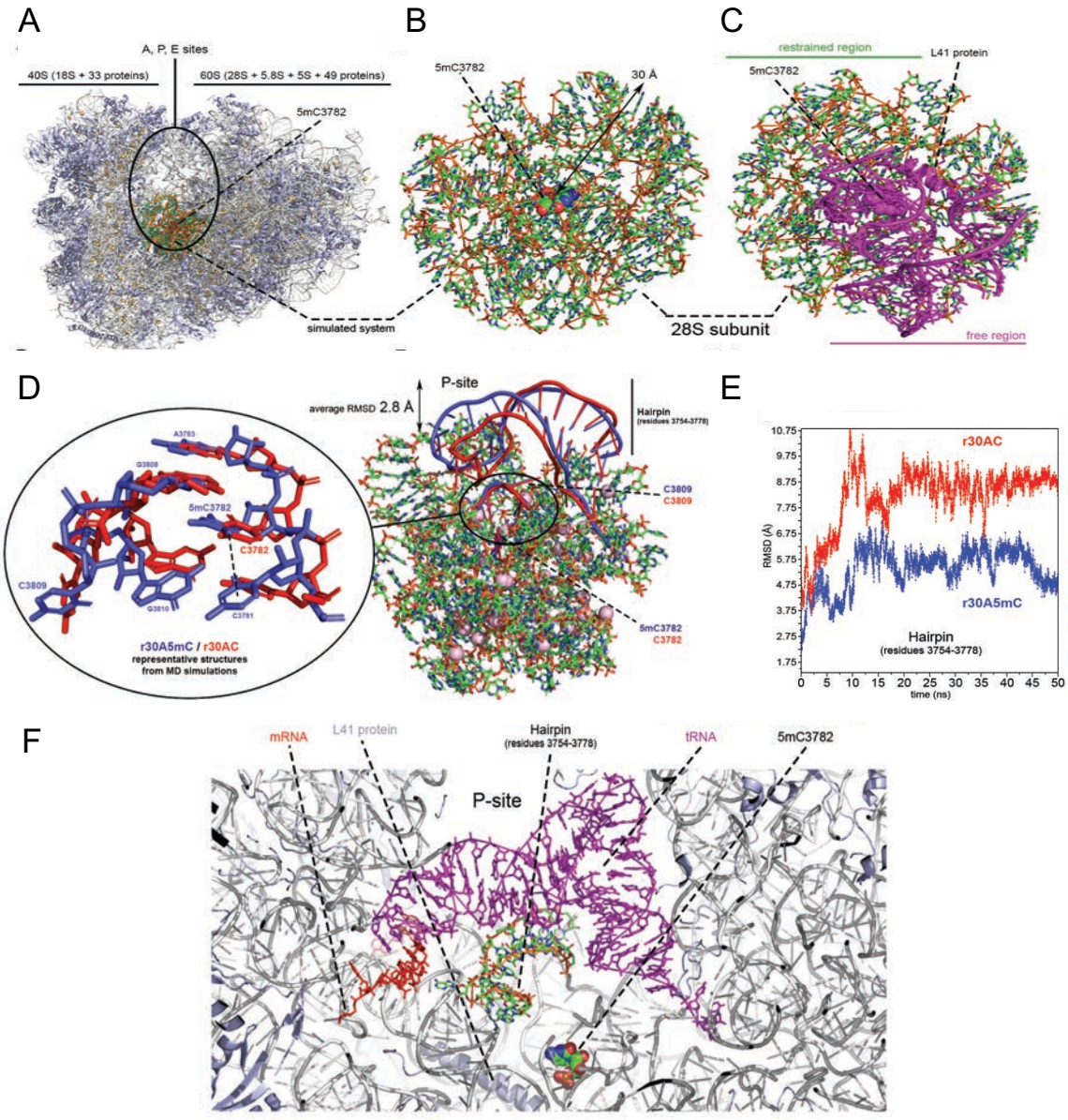


Figure 29. Loss of m^{5}C3782 on the 28S rRNA impairs the tertiary complex (rRNA–tRNA–mRNA) structural stability at the P-site of the ribosome. (A) Human 80S ribosome structure taken from PDB code 6EK0. Proteins are displayed in light purple, and RNA molecules in white. **(B)** System considered in the MD simulations: 30 Å around the C3782 position, which is located in the center and depicted by its van der Waals radii. Nucleobases are depicted in green, while backbone is shown in orange. **(C)** Image shows the residues free to move (in pink), and the residues constrained during the simulation. **(D)** Local (*left*) and "long-range" (*right*) structural distortions observed in MD simulations after loss of the methyl group at the C3782 position (labelled r30AC, in red), compared to the methylated system (labelled r30A5mC, in blue). **(E)** Root-Mean-Squared-Deviation (RMSD) of the heavy atoms in the hairpin exposed to the P-site (residues 3754–3778) along time. Compared to the cryo-electron microscopy (cryoEM) structure, the hairpin in the methylated system experiments a deviation of 5.4 ± 0.8 Å in average, while in the unmethyated system the average deviation is 8.2 ± 1.2 Å. **(F)** CryoEM structure (5.5 Å-resolution) of an eukaryotic 80S ribosome (PDB code 4V6I). Interaction between the P-site hairpin (nucleobases depicted in green and backbone in orange), the mRNA (depicted in red) and the tRNA (in pink) is shown.

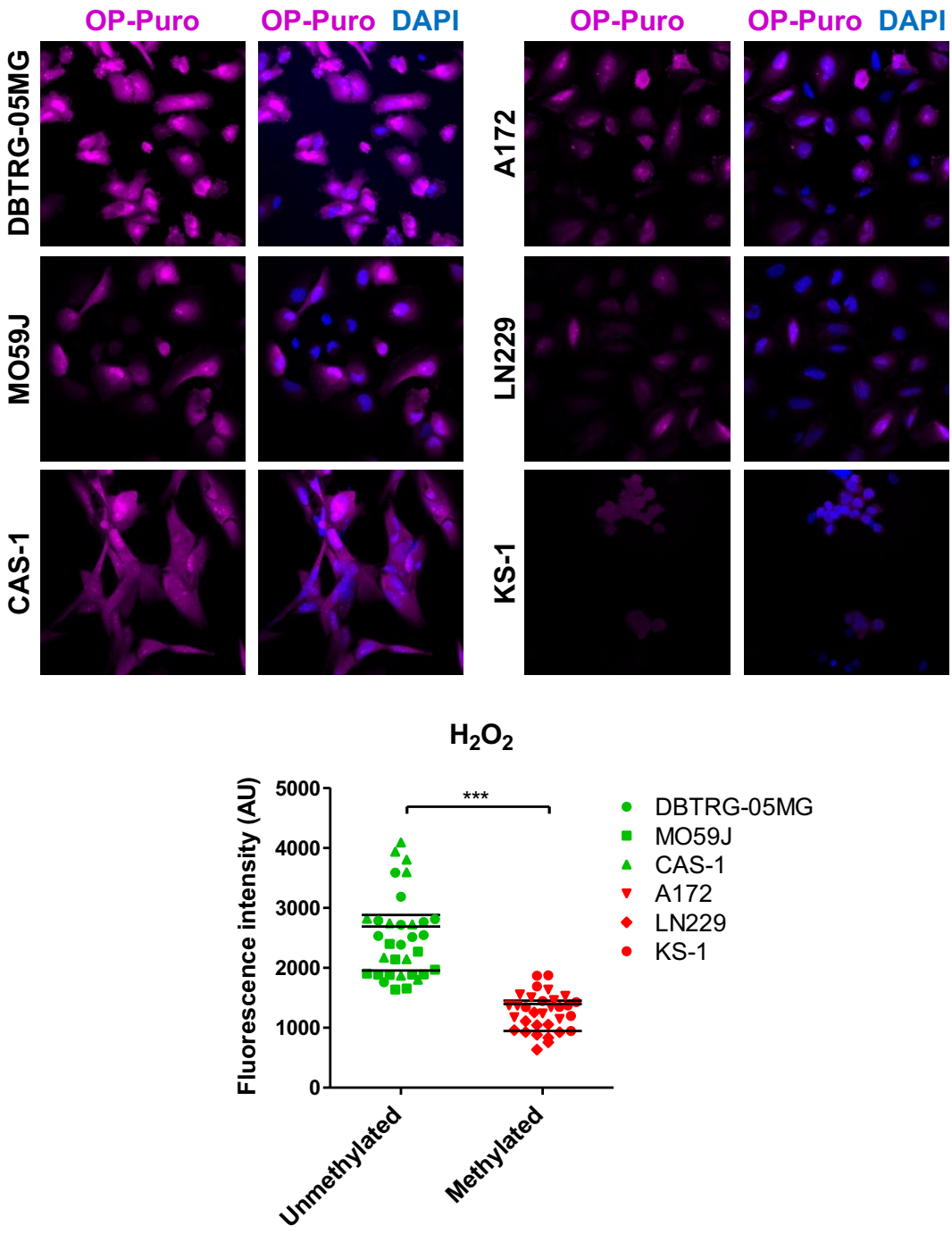


Figure 30. NSUN5 epigenetic loss leads to a global depletion of protein synthesis. NSUN5 hypermethylated and silenced cell lines (A172, LN229 and KS-1) show significantly lower overall protein synthesis than NSUN5 hypomethylated and expressing glioma cell lines (DBTRG-05MG, MO59J, and CAS-1) under oxidative stress (100 µM H_2O_2). Protein synthesis is assessed by OP-Puro incorporation into cells and visualized by confocal microscopy. Representative confocal images (*left panel*) and fluorescence intensity quantification (*right panel*) are shown.

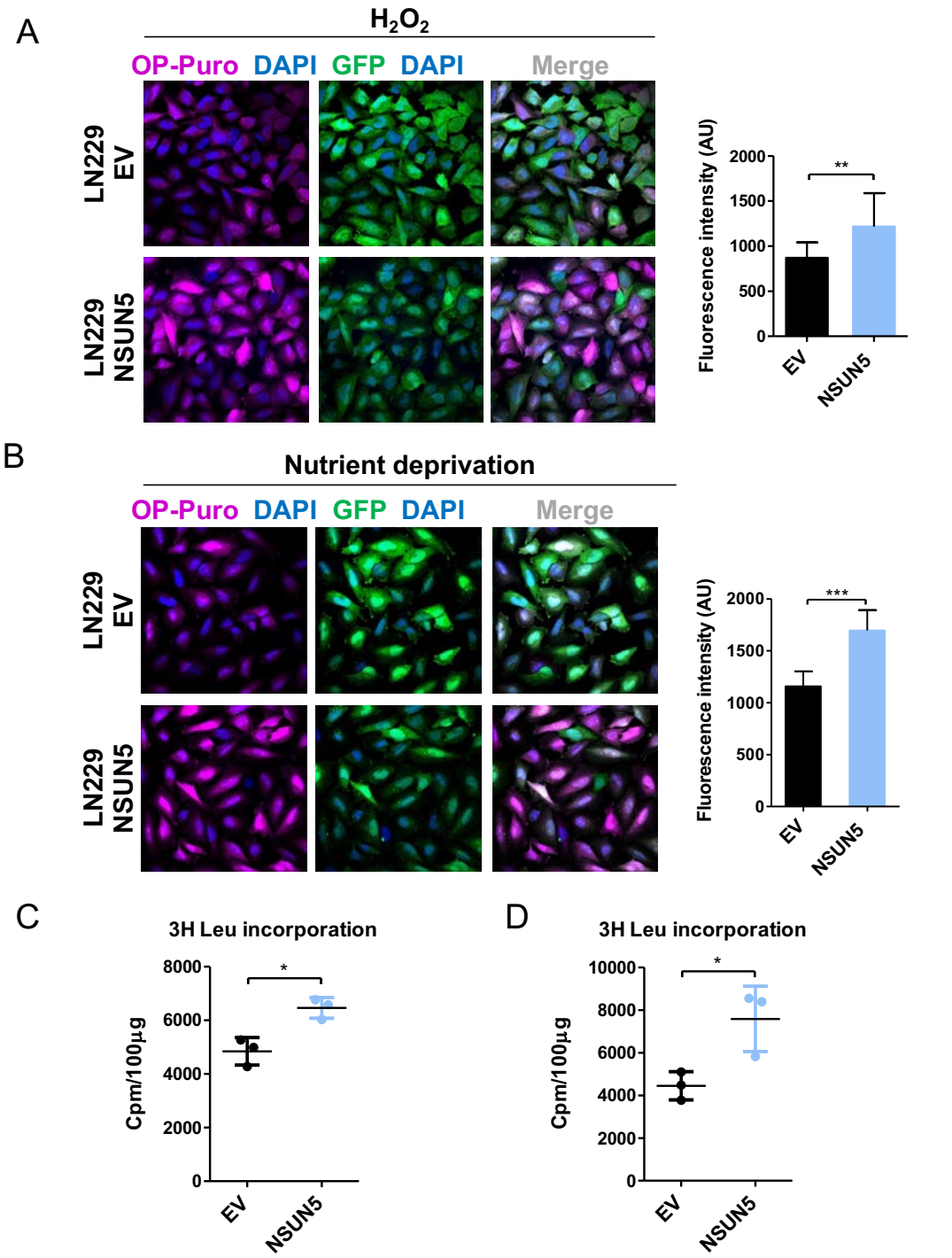


Figure 31. NSUN5 recovery restores global protein synthesis under different stress conditions. (A) Under oxidative stress (100 µM H_2O_2) and **(B)** nutrient deprivation, NSUN5-transduced LN229 cells show increased overall protein synthesis compared to their EV-transduced counterparts. Protein synthesis is assessed by OP-Puro incorporation into cells and visualized by confocal microscopy. Representative confocal images (*upper panel*) and fluorescence intensity quantification (*bottom panel*) are shown. **(C)** Increased global protein synthesis upon NSUN5 restoration in epigenetically-silenced LN229 cells under oxidative stress (100 µM H_2O_2) is also observed by a second technique, [3H] leucine incorporation and radioactivity quantification. **(D)** [3H] leucine incorporation into proteins shows similar results upon nutrient deprivation.

Moreover, we showed that NSUN5 recovery in LN229 hypermethylated cells significantly increased overall protein synthesis, compared to the EV-transduced counterpart cells (**Figure 31A**). Analysis of *de novo* translation under a second stress condition, nutrient deprivation, rendered similar results: NSUN5-restored LN229 cells significantly synthesized more proteins than NSUN5-silenced cells (**Figure 31B**).

With the aim of confirming these results we analyzed global protein synthesis by an orthogonal method, [3H] leucine incorporation into nascent proteins, followed by radioactivity quantification. We observed the same pattern: under oxidative stress (**Figure 31C**) and nutrient deprivation (**Figure 31D**), epigenetically-inactive-NSUN5 LN229 cells had lower global protein synthesis in comparison to NSUN5-restored cells.

Overall, our results indicate that epigenetic loss of NSUN5 in glioma cells, and its consequently lack of the methyl group on the C3782 of the 28S rRNA, is associated with an impairment of global protein synthesis under stress conditions.

NSUN5 epigenetic silencing induces the translation of a specific mRNA subset for stress survival

Apart from global protein synthesis inhibition, some human tumors have been shown to implement additional methods for coping with cellular stress, such as the activation of alternative translational programs that are involved in the adaptive response to stress [213–215]. In order to characterize if NSUN5-epigenetically inactive cells promote the translation of specific transcripts to cope with cellular stress we followed the next strategy: to compare total RNA-seq data to ribosome-protected mRNA sequencing (Ribo-seq) data from EV- and NSUN5-transduced LN229 cells.

First of all, we observed that restoration of NSUN5 expression into silenced LN229 cells had almost no effect on the overall coding transcriptome, as only 2% of all transcripts showed differences in abundance between conditions, being 255 upregulated and 244 downregulated (adjusted p-value < 0.01 and absolute log2 fold-change > 2) (**Figure 32A**). However, when looking at the mRNA fraction being translated for the same conditions, we observed higher differences: up to 7.2% of transcripts showed differences in ribosome occupancy, with 1,028 mRNAs overrepresented and 2,009 downrepresented upon NSUN5 restoration in LN229 cells (p-value < 0.01 and absolute log2 fold-change > 0.5) (**Figure 32A**).

In order to decipher if NSUN5 epigenetic silencing in glioma promotes the emergence of specific translational programs despite the global impairment of protein synthesis, we decided to better characterize the fraction of transcripts with enhanced translation upon NSUN5 epigenetic loss. This fraction was defined as those mRNAs unaltered in the RNA-seq data regardless of NSUN5 expression, but downregulated in the Ribo-seq data upon NSUN5 recovery in LN229 cells, or what is the same, with higher ribosome occupancy in EV-transduced cells (NSUN5-silenced). We observed a total of 1987 transcripts with enhanced translational efficiency in NSUN5 epigenetically silenced LN229 cells EV-transduced (**Figure 32B**), and to confirm these observations we applied a statistical tool that calculates translational efficiency by using a generalized linear model (GLM) over mRNA abundance (RNA-seq data) and ribosome occupancy (Ribo-seq data) (**Figure 32C**). To take a deeper look into these 1987 mRNAs with higher translational efficiency upon NSUN5 loss in glioma cells, we carried out a gene set enrichment analysis (GSEA) using GO biological process and KEGG pathway signature collections. Interestingly, we observed an overrepresentation of pathways related to cancer and the ribosome, and gene ontologies related to cellular and metabolic adaptation to stress (FDR adjusted p-value < 0.05) (**Figure 32D**).

To better characterize the potential translational program arising upon NSUN5 epigenetic loss in glioma cells, we carried out Stable Isotopic Labeling of Amino acids in Cell culture (SILAC), comparing the hypermethylated LN229 cells EV- and NSUN5-recovered. We observed 128 proteins differentially expressed based on NSUN5 expression *(*p-value < 0.05 and absolute log2 fold-change > 0.5) (**Figure 33A**). The decreased protein expression upon NSUN5 recovery was validated by western blot for three of them (**Figure 33B**). Moreover, we carried out a gene functional annotation of differentially expressed proteins by GSEA, using, as previously described, GO biological process and KEGG pathway signature collections. We observed pathways and gene ontologies related to response to stress by metabolic adaptation, mismatch repair and RNA processing (FDR adjusted p-value < 0.05) (**Figure 33C**). Importantly, overlapping SILAC results to ribosome occupancy data (Ribo-seq) (**Figure 32C**), we observed that 26% of the downregulated proteins upon NSUN5 recovery (9 of 34), derived from transcripts with higher translational efficiency in the EV condition (**Figure 33A**).

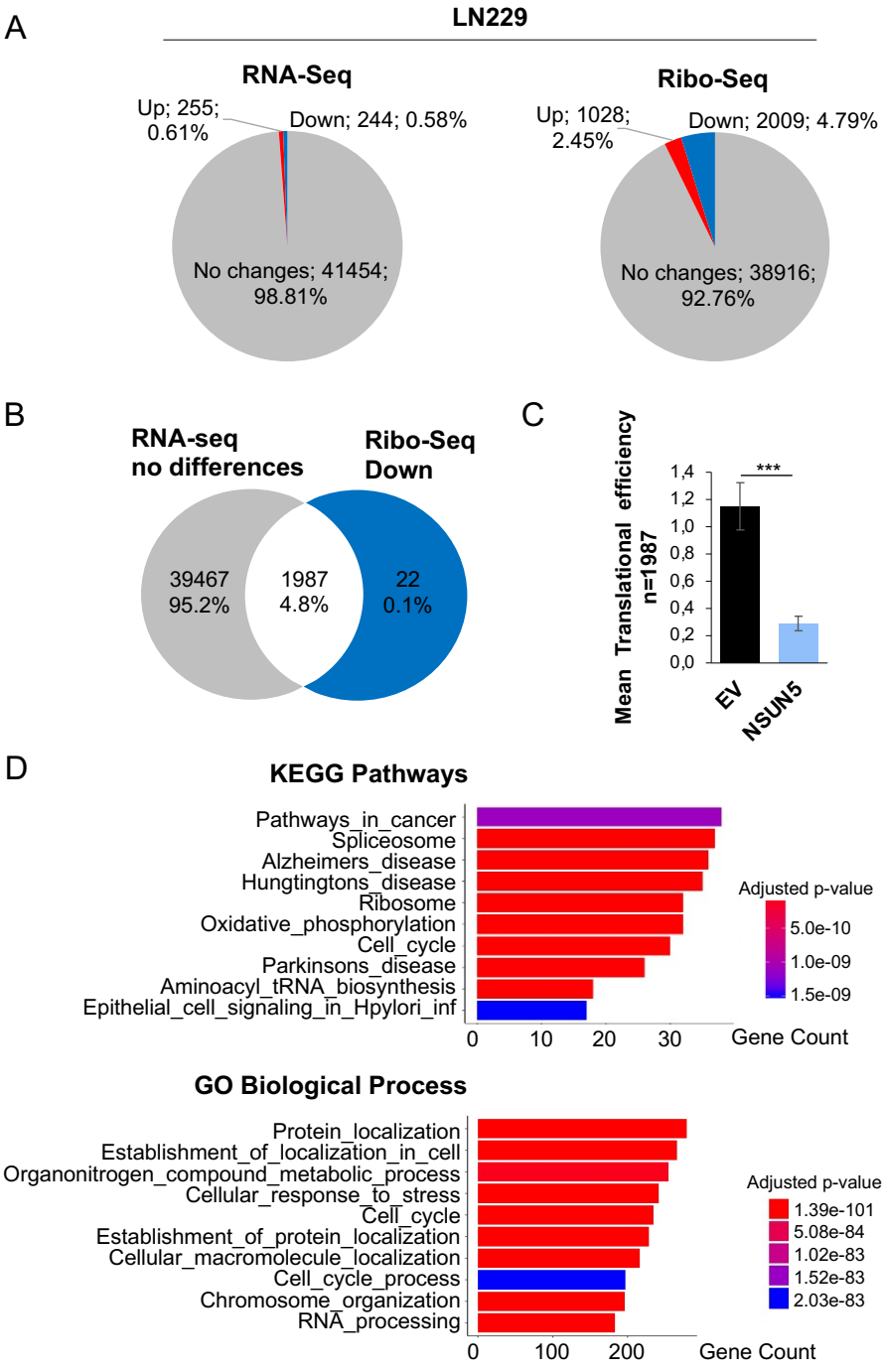


Figure 32. NSUN5 epigenetic silencing increases the translational efficiency of specific transcripts. (A) Pie charts summarizing RNA-seq and Ribo-seq deep-sequencing profiles upon NSUN5 recovery into epigenetically-silenced LN229 glioblastoma cells. **(B)** Merge of transcripts unaltered in RNA-seq data with transcripts downregulated in Ribo-seq data upon NSUN5 restoration allowed to identify 1987 RNAs with potentially higher translational efficiency upon NSUN5 epigenetic silencing (EV condition). **(C)** Assessment of translational efficiency for the identified set of 1987 transcripts. A GLM is applied, which estimates the over-dispersion of RNA-Seq and Ribo-seq measurements separately, and then performs a statistical test to identify differential translational efficiency. **(D)** Gene set enrichment analysis (GSEA) of the RNAs with enhanced translational efficiency in NSUN5 depleted cells (hypergeometric test with FDR adjusted p-value < 0.05).

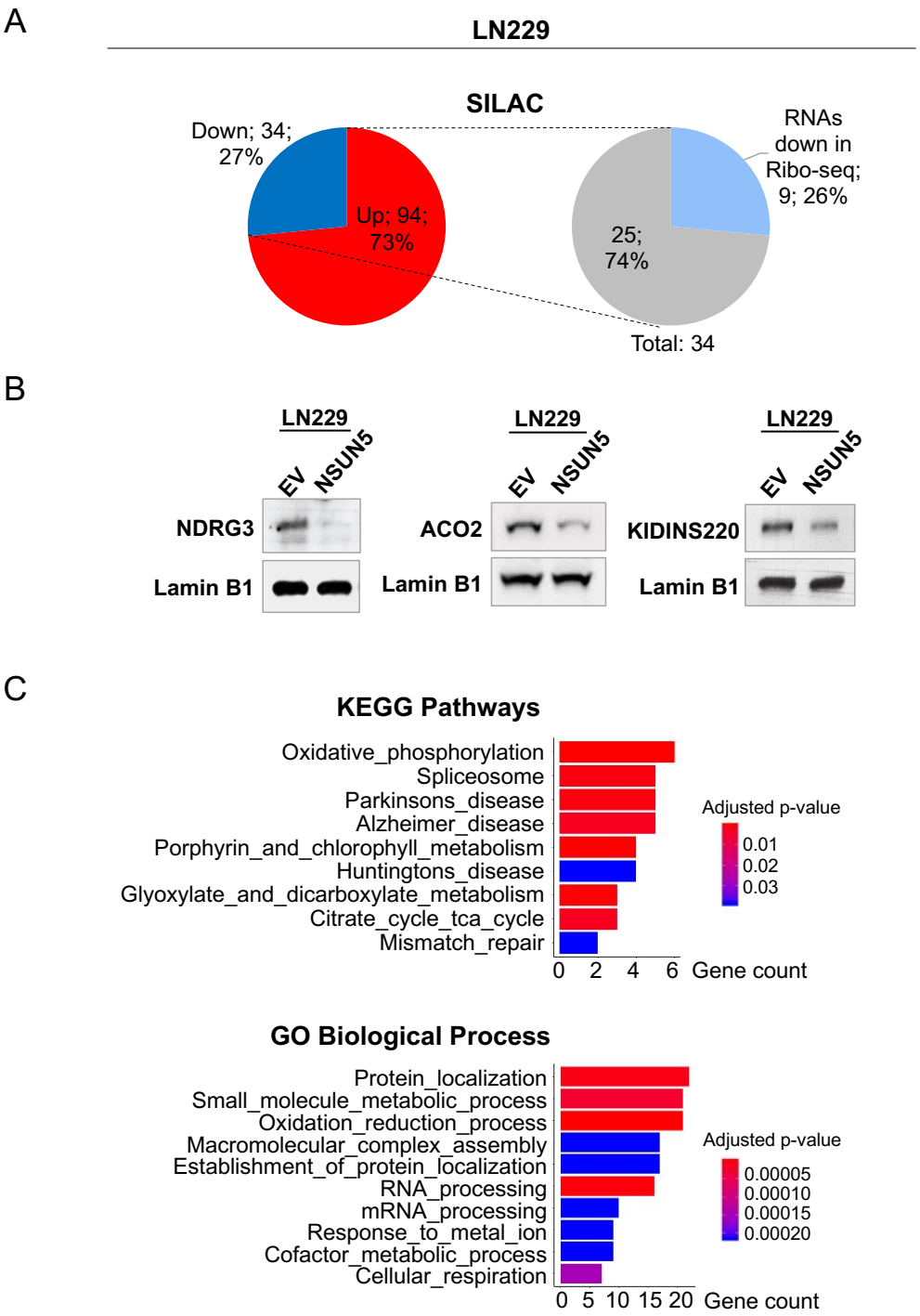


Figure 33. Proteome changes derived from NSUN5 recovery. (A) Pie charts summarizing differentially expressed proteins between LN229 cells EV- and NSUN5-recovered, assessed by SILAC approach. 26% of the downregulated proteins upon NSUN5 recovery derived from transcripts with higher translational efficiency in NSUN5-silenced LN229 cells (EV condition). **(B)** Western blot validation of some of the proteins downregulated upon NSUN5 restoration in epigenetically-silenced LN229 cells, according to SILAC results. **(C)** Gene set enrichment analysis (GSEA) of differentially expressed proteins upon NSUN5 recovery in LN229 glioma cells, according to SILAC results (hypergeometric test with FDR adjusted p-value < 0.05).

Moreover, we carried out additional RNA-seq experiments, first, for the same EV- and NSUN5-transduced LN229 cell model, but upon oxidative stress, and second, for the other GBM restoration model. Both experiments confirmed that NSUN5 restoration into epigenetically-silent cells had minimal effect on the overall transcriptome: only 0.5% of all transcripts showed differential abundance between EV- and NSUN5-transduced LN229 cells in the presence of H_2O_2 (**Figure 34A**), and recovery of NSUN5 in the A172 glioblastoma cell line affected the expression of just 1% of all mRNAs (**Figure 34C**). A further ribosomal profiling of the LN229 model under oxidative stress confirmed the existence of a candidate gene set with an enhanced translation upon NSUN5 epigenetic silencing, as we observed a total of 137 transcripts with unaltered expression between conditions in the RNA-seq data, but downregulated in the Ribo-seq data upon NSUN5 restoration (**Figure 34B**).

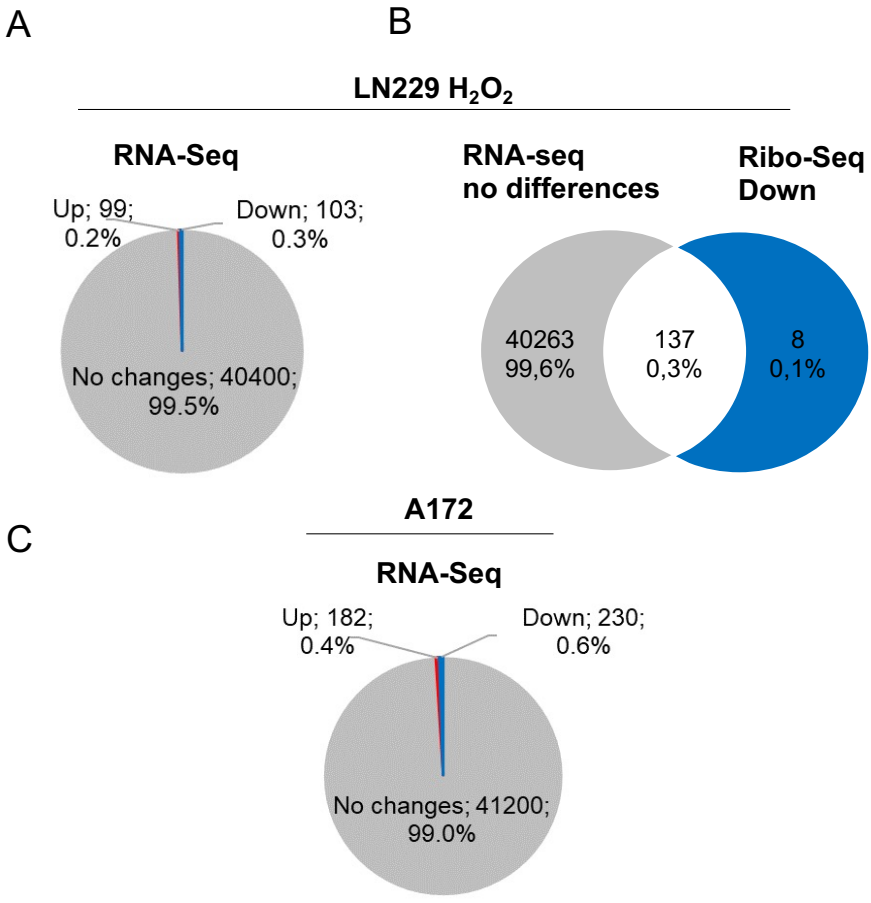


Figure 34. NSUN5 recovery does not induce major changes in the overall transcriptome. **(A)** Pie chart summarizing RNA-seq results upon NSUN5 restoration in epigenetically-silenced LN229 cells under oxidative stress (100 µM H_2O_2). **(B)** Merging of transcripts without abundance differences according to RNA-seq data, and transcripts downregulated in Ribo-seq data upon NSUN5 recovery, identified 137 RNAs with potentially higher translational efficiency upon NSUN5 epigenetic silencing (EV condition). **(C)** Pie chart showing RNA-seq results upon NSUN5 restoration in a second glioblastoma cell line, A172.

The previously disclosed strategy was also validated in the loss-of-function DBTRG-05MG model. Firstly, we carried out both RNA-seq and Ribo-seq experiments in hypomethylated and expressing DBTRG-05MG cells, Scramble- and shNSUN5-transduced. We observed that depletion of NSUN5 expression in those cells did not affect the overall coding transcriptome, as only 0.5% of all transcripts had abundance differences, and the remaining 99.5% were unchanged (adjusted p-value < 0.01 and absolute log2 fold-change > 2). However, for the same conditions, we observed that 2.2% transcripts showed differences in ribosome occupancy (p-value < 0.01 and absolute log2 fold-change > 0.5) (**Figure 35A**). As aforementioned, to better characterize the fraction of transcripts with enhanced translation upon NSUN5 epigenetic loss, we merged mRNAs unaltered between conditions in the RNA-seq data, with those upregulated in the Ribo-seq data upon shRNA-mediated loss of NSUN5. We observed a total of 787 transcripts with enhanced translational efficiency in NSUN5-depleted DBTRG-05MG cells compared to their Scrambled counterparts, and to confirm these observations we calculated the translational efficiency of those transcripts by applying the GLM as previously described (**Figure 35A**). A further characterization of this gene set with higher translational efficiency by means of GSEA showed, as described for the gain-of-function LN229 model, an overrepresentation of ontologies related to response to stress and metabolic adaptation (**Figure 35B**).

Taken all together, our results suggest that loss of NSUN5 through an epigenetic mechanism in glioma cells leads to a global depletion of protein synthesis, but at the same time, it promotes a selective synthesis of proteins to help cope with cellular stress.

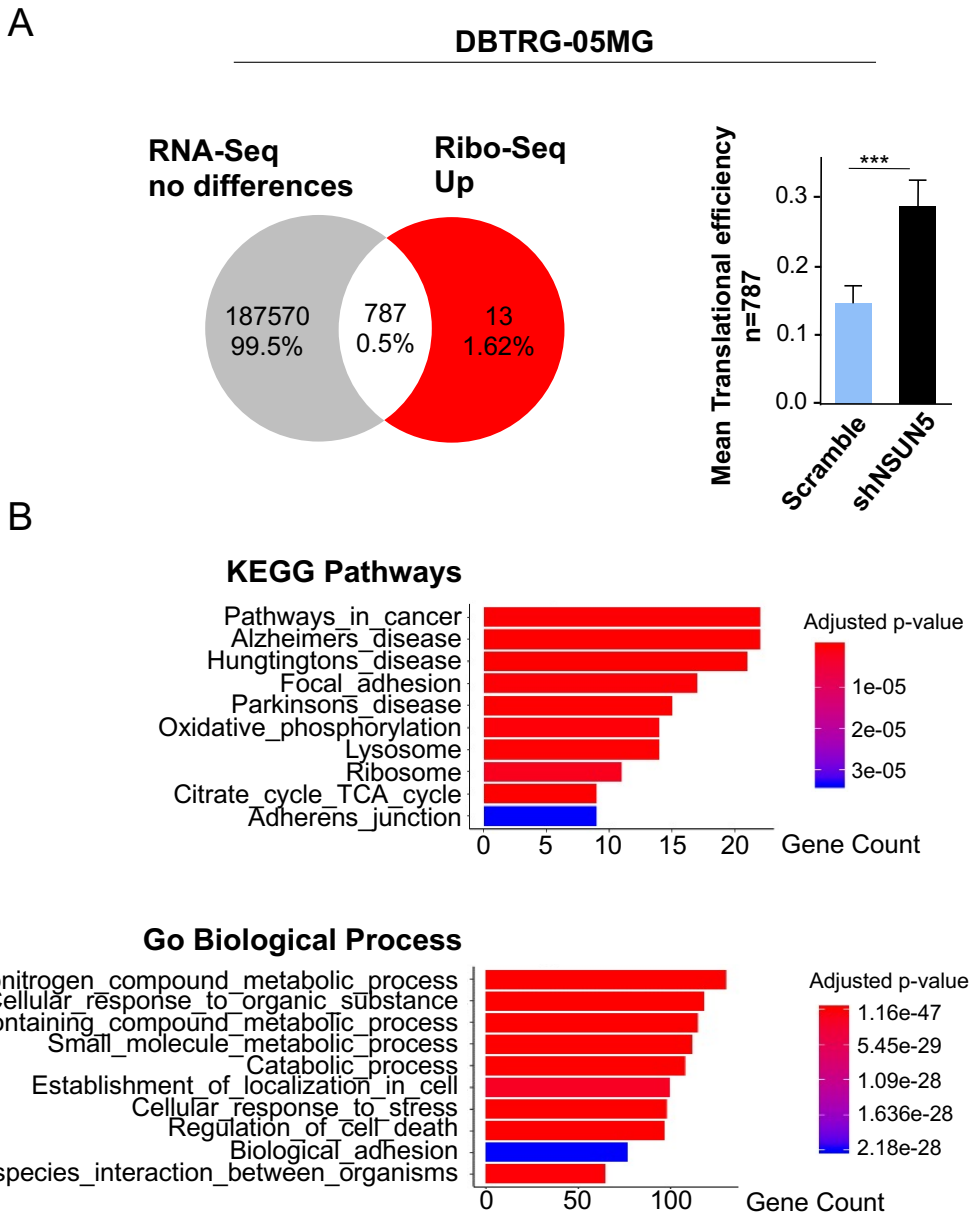


Figure 35. ShRNA-mediated depletion of NSUN5 increases the translational efficiency of specific transcripts. (A) Merge of transcripts unaltered in RNA-seq data with transcripts upregulated in Ribo-seq data upon NSUN5 depletion in DBTRG-05MG cells led to the identification of 787 RNAs with potentially higher translational efficiency upon NSUN5 loss (*left panel*). To confirm these results, translational efficiency for these transcripts was estimated from over-dispersion analysis of both RNA-seq and Ribo-seq measurements by applying a GLM (*right panel*). **(B)** Gene set enrichment analysis (GSEA) of the mRNAs with higher translational efficiency after shRNA-mediated depletion of NSUN5 in DBTRG-05MG cells (hypergeometric test with FDR adjusted p-value < 0.05).

NSUN5 epigenetic silencing as a hallmark of sensitivity to NQO1-targeting compounds

Taking a deeper insight into the translational program arising upon loss of NSUN5 in glioma cells we found an interesting candidate for further validation: NQO1, a multifunctional antioxidant enzyme regulated by the Keap1/Nrf2/ARE pathway [216]. *NQO1* mRNA levels in the NSUN5-restoration LN229 model were unaffected regardless of NSUN5 expression, according to RNA-seq data, and it was further confirmed by quantitative real-time PCR (**Figure 36A**). However, western blot analysis of NQO1 showed higher protein levels in EV-transduced LN229 cells, confirming that NSUN5-mediated NQO1 regulation occurs at the translational level (**Figure 36B**). Furthermore, analysis of the additional RNA-seq experiments showed that *NQO1* mRNA neither showed abundance differences in LN229 EV- and NSUN5-transduced cells upon H_2O_2 stress nor upon NSUN5 recovery into A172 glioblastoma cells (adjusted p-value < 0.01 and absolute log2 fold-change > 2). And importantly, NQO1 was one of the transcripts with greater translational efficiency in EV-transduced LN229 cells undergoing oxidative stress, according to Ribo-seq data (p-value < 0.01 and absolute log2 fold-change > 0.5) (**Figure 34B**).

In order to validate NQO1 translational regulation we quantified *NQO1* mRNA abundance in polysome fractionations (polysome profiling) from EV- and NSUN5-transduced LN229 glioblastoma cells. *NQO1* mRNAs from EV-transduced LN229 cells were mostly found bound to polysomes rather than to single ribosomes, meaning that NQO1 is being highly translated (**Figure 36C, *left***). And importantly, the polysome fraction extracted from NSUN5-silenced LN229 EV cells was significantly more enriched in *NQO1* mRNAs, compared to its NSUN5-transduced counterpart (**Figure 36C, *right***).

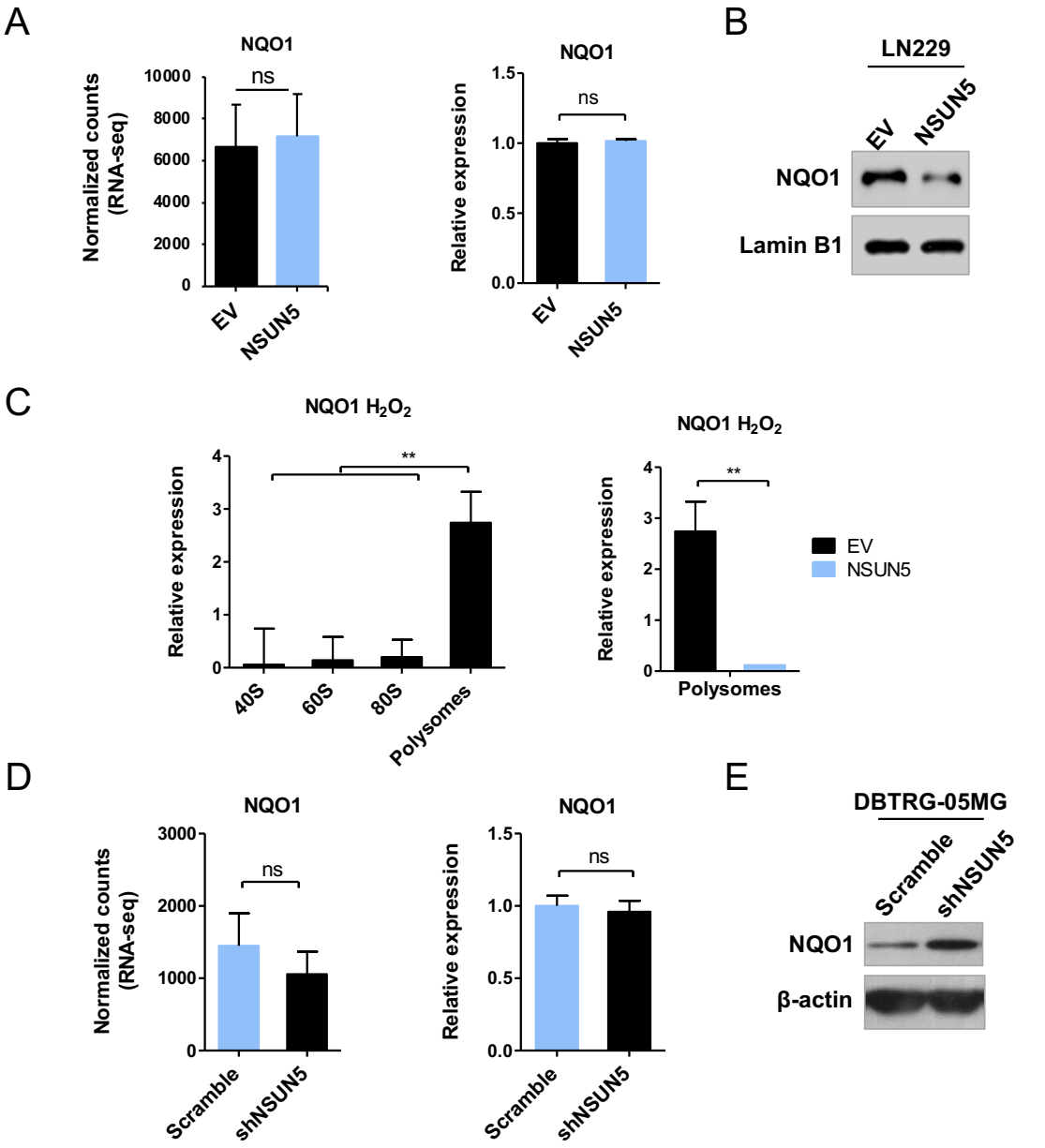


Figure 36. NSUN5 loss induces a higher translation of the stress-related enzyme NQO1. **(A)** NQO1 expression at the RNA level, determined by RNA-seq counts (left) and quantitative RT-PCR (right), does not change upon NSUN5 recovery in LN229 glioma cells. **(B)** NQO1 protein levels decrease in LN229 cells upon NSUN5 restoration, determined by western blot. **(C)** Quantitative RT-PCR shows significantly higher abundance of the *NQO1* mRNA in the polysome fraction of EV-transduced LN229 cells, compared to NSUN5-recovered LN229 cells. **(D)** *NQO1* transcript expression, determined by RNA-seq counts (*left*) and quantitative RT-PCR (*right*), does not change between Scramble and NSUN5-depleted DBTRG-05MG cells. **(E)** Western blot analysis shows increased NQO1 protein levels in DBTRG-05MG cells depleted of NSUN5 expression.

NSUN5-mediated NQO1 translational regulation was characterized in the loss-of-function model, as well. Ribo-seq data mining confirmed *NQO1* as one of the specific transcripts with greater translational efficiency in DBTRG-05MG cells depleted of NSUN5 expression (shNSUN5) (p-value < 0.01 and absolute log2 fold-change > 0.5). Similar to the LN229 gain-of-function model, *NQO1* mRNA levels were unchanged between conditions (Scramble vs shNSUN5), both looking at the RNA-seq counts and analyzed by quantitative real-time PCR (**Figure 36D**), while its expression at the protein level was higher in DBTRG-05MG cells with shRNA-mediated depletion of NSUN5, in comparison to Scramble cells. (**Figure 36E**).

All the results confirmed the selective translation of the stress-related enzyme NQO1 upon epigenetic loss of NSUN5 in glioma cells, despite the overall inhibition of protein synthesis. This finding led us to wonder if NQO1 overexpression could be therapeutically exploited. There are bioactivatable compounds such as deoxynyboquinone (DNQ) and isobutyl–deoxynyboquinone (IB-DNQ) that, upon reduction by this stress-related enzyme, cause the rapid production of large quantities of ROS that push cells into their death [217]. We thus wondered if NSUN5-hypermethylated tumors, with an enhanced NQO1 expression, may be more susceptible to targeting with NQO1 bioactivatable compounds.

First of all, we tested DNQ growth inhibitory effect on the LN229 gain-of-function model, by analyzing cell viability through the SRB assay. We observed that upon restoration of NSUN5, and therefore downregulation of NQO1 protein expression (**Figure 36B**), LN229 glioma cells were significantly more resistant to the inhibitory effect of the bioactivatable NQO1 substrate than EV-transduced cells (Student's T test, p-value = 0,0056) (**Figure 37A**). Next, with the aim of broadening the significance of our findings, we extended the assay to the GBM cellular set. First, NQO1 protein expression was assessed by western blot, and we observed that the three NSUN5 hypermethylated cells lines (A172, LN229, and KS-1) showed greater levels of expression compared to the three NSUN5 hypomethylated cell lines (DBTRG-05MG, MO59J, and CAS-1), confirming the role exerted by NSUN5 in the translational regulation of the stress-related enzyme (**Figure 37B**). Importantly, treatment of these cells with both, DNQ and IB-DNQ drugs, showed that NSUN5 hypermethylated cell lines were significantly more sensitive to the inhibitory growth effect of NQO1 bioactivatable substrates (z test, p-value < 0.0001) (**Figure 37C**).

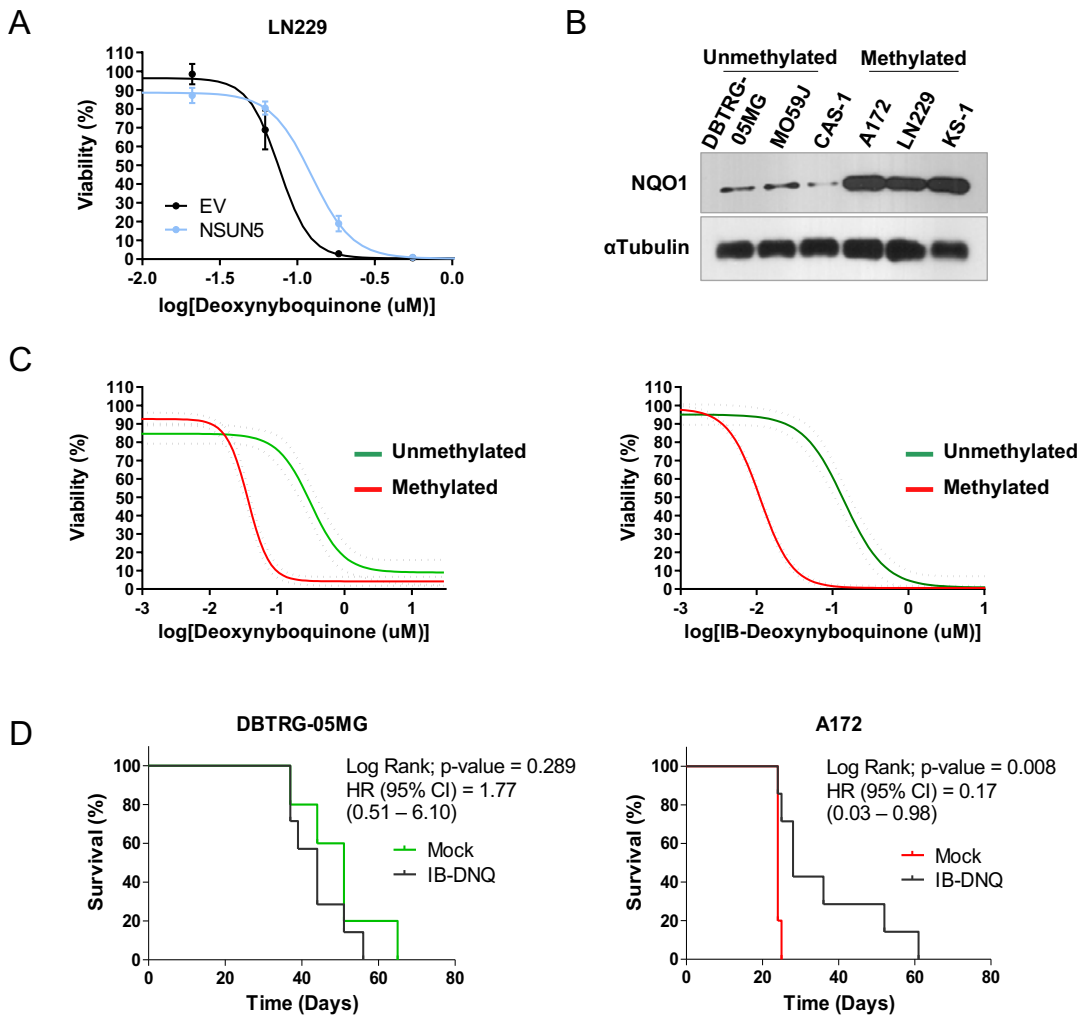


Figure 37. NSUN5 epigenetic silencing confers sensitivity to NQO1-targeting compounds. **(A)** Representative experiment employing SRB assay for analyzing DNQ growth inhibitory effect on EV- and NSUN5-transduced LN229 glioma cells. IC_{50} values were calculated from the slopes in triplicates, and differences between curves were statistically estimated by Student's T test. **p-value = 0,0056. **(B)** Western blot showing NQO1 protein expression levels in the GBM cellular panel, according to NSUN5-promoter methylation status. **(C)** DNQ (left panel) and IB-DNQ (*right panel*) growth inhibitory effect on glioblastoma cells grouped by NSUN5-promoter methylated status, analyzed by SRB assay. NSUN5 hypermethylated cell lines (A172, LN229 and KS-1), with proven NQO1 overexpression, show significantly higher sensitivity to DNQ and IB-DNQ compared to NSUN5 unmethylated cell lines (DBTRG-05MG, MO59 J, and CAS-1). Dashed curves represent the 95% confidence interval (CI) for each group. Differences between curves were statistically estimated by z test. ****p-value < 0.0001. **(D)** Survival of mice intracranially-inoculated with tumors derived from unmethylated DBTRG-05MG cells (*left*) and methylated A172 (*right*) cells, according to treatment conditions (mock vs IB-DNQ treated groups). P-values were calculated by the log-rank test, and results of the univariate Cox regression analysis are shown by hazard ratio (HR) and 95% confidence interval (CI).

Finally, we moved to an *in vivo* setting, and studied the survival of orthotopically-injected nude mice undergoing IB-DNQ treatment. While we did not observe survival differences between mock and treated mice with tumors originating from the NSUN5 hypomethylated and expressing DBTRG-05MG cell line, with minimal expression of NQO1, we showed that IB-DNQ treatment significantly increased the survival of mice carrying tumors derived from the A172 glioma cell line, hypermethylated for NSUN5 and with higher NQO1 protein expression (**Figure 37D**).

Overall, these observations suggest that NSUN5 hypermethylation could be used to distinguish glioma cells that are more sensitive to NQO1-targeting compounds.

NSUN5 epigenetic silencing is associated with good clinical outcome

Finally, we wondered if NSUN5 promoter methylation in human glioma could have any prognostic value. To address this question, we first analyzed the clinical data from the TCGA glioma cohort (n = 497) based on NSUN5 methylation status. Without taking tumor grade into account, we showed that NSUN5 hypermethylation was associated with longer overall survival (OS) (log-rank; p-value $<10^{-5}$; hazard ratio [HR] = 5.17, 95% CI = 2.78 – 9.61) (**Figure 38A**). A further analysis subdividing glioma samples based on tumor grade, as previously explained, provided similar results: NSUN5 hypermethylation was a predictor of better outcome both in LGG (log-rank; p-value = 0.044; hazard ratio [HR] = 2.05, 95% CI = 1.01 – 4.18) (Figure.) and GBM (log-rank; p-value = 0.012; hazard ratio [HR] = 8.26, 95% CI = 1.15 – 59.5) (**Figure 38B, C**). Similar results were obtained when analyzing TCGA glioma clinical data based on NSUN5 transcript expression: low *NSUN5* expression was associated with higher OS in all glioma cases (log-rank; p-value $<10^{-5}$; hazard ratio [HR] = 4.57, 95% CI = 2.40 – 9.04) (**Figure 38D**). Low levels of *NSUN5* transcripts were also significantly associated with longer OS in LGG (log-rank; p-value = 0.014; hazard ratio [HR] = 2.37, 95% CI = 1.16 – 4.84) (**Figure 38E**) and a similar fashion was observed for GBM (**Figure 38F**), although no statistical significance could be inferred due to the small number of cases with low *NSUN5* expression (n = 2).

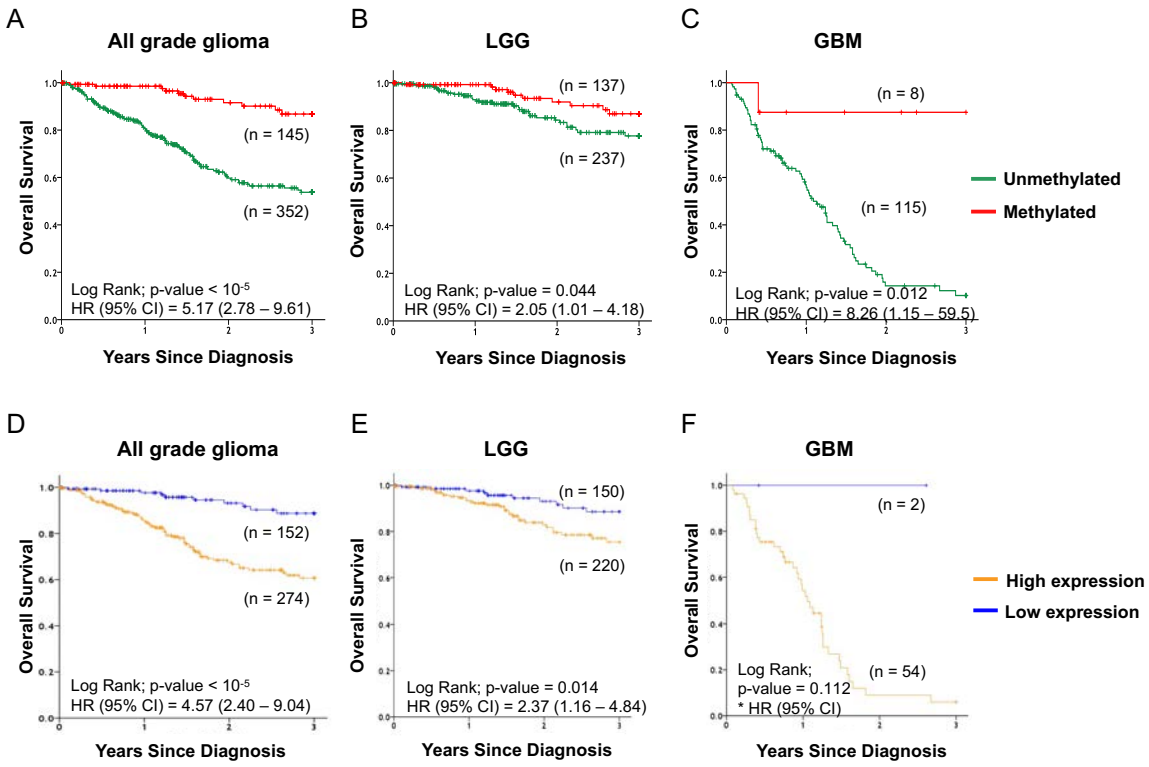


Figure 38. Transcriptional silencing of NSUN5 is associated with good clinical outcome in glioma. Kaplan-Meier analysis of OS in primary glioma tumors from the TCGA, from all grades **(A)**, LGG **(B)**, and GBM **(C)** according to NSUN5-promoter methylation status. Similar results were obtained by analyzing OS according to NSUN5 transcript expression levels in primary glioma tumors from the TCGA, from all grades **(D)**, LGG **(E)**, and GBM **(F)**. P-values were assessed by the log-rank test, and results of the univariate Cox regression analysis are shown by hazard ratio (HR) and 95% confidence interval (CI).

To confirm the prognostic value of NSUN5 methylation in glioma patients, we generated a validation cohort by obtaining new glioma samples (n = 115) and determining their NSUN5 methylation status by pyrosequencing. The validation cohort confirmed our findings: NSUN5 hypermethylation was associated with longer OS (log-rank; p-value = 0.043; HR = 1.90; 95% CI = 1.01 – 3.58) and extended progression-free survival (PFS) (log-rank; p-value = 0.010; HR = 1.95; 95% CI = 1.16 – 3.27) (**Figure 39**). Clinicopathological features of both cohorts can be found on the Annexes section (**Supplementary Table S1**).

Moreover, we observed that NSUN5 methylation was associated with glioma biomarkers of clinical benefit such as IDH1 mutation, co-deletion of 1p/19q and MGMT methylation (**Supplementary Table S1**).

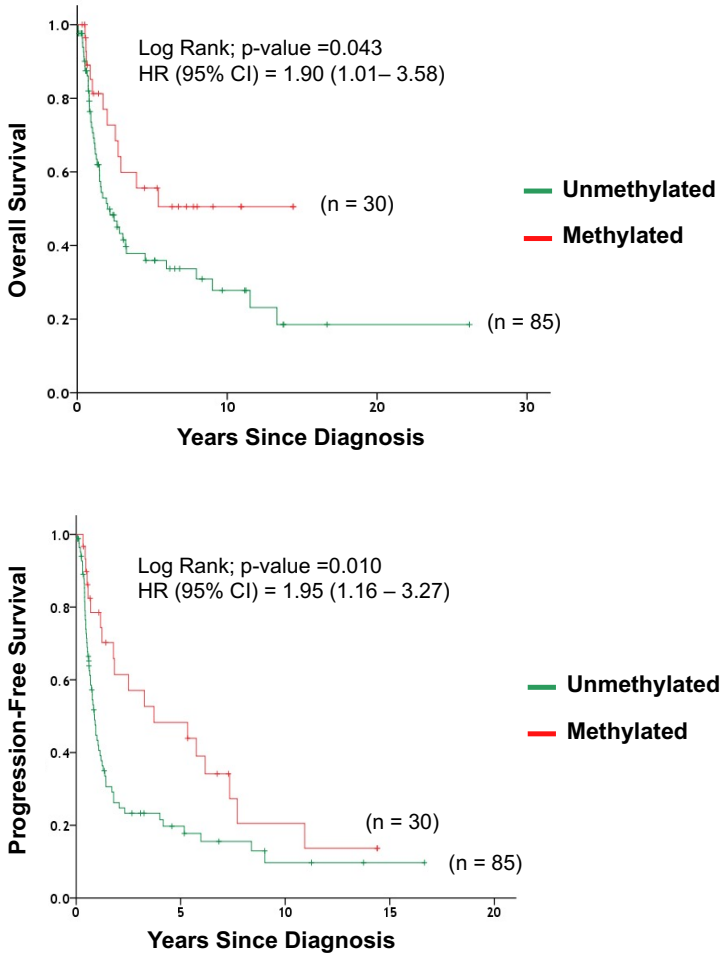


Figure 39. A validation cohort confirms the prognostic value of NSUN5 methylation status in glioma. Kaplan-Meier analysis of OS (left) and PFS (right) in primary glioma tumors from the validation cohort (n = 115), according to NSUN5-promoter methylation status. Log-rank test p-values are displayed, and results of the univariate Cox regression analysis are shown by hazard ratio (HR) and 95% confidence interval (CI).

Finally, with the aim of determining if NSUN5 methylation is an independent biomarker of clinical outcome, we carried out a multivariate analysis for the TCGA cohort, including described biomarkers of glioma outcome (IDH1 mutational, 1p/19q co-deletion and MGMT methylation status) [218]. Importantly, in the interrogated TCGA cohort, the multivariate Cox regression analysis showed that NSUN5 hypermethylation was a significant covariate (HR = 0.50; 95% CI = 0.26 – 0.94; p-value = 0.032), meaning that it is an independent predictor of longer OS in glioma (**Figure 40**).

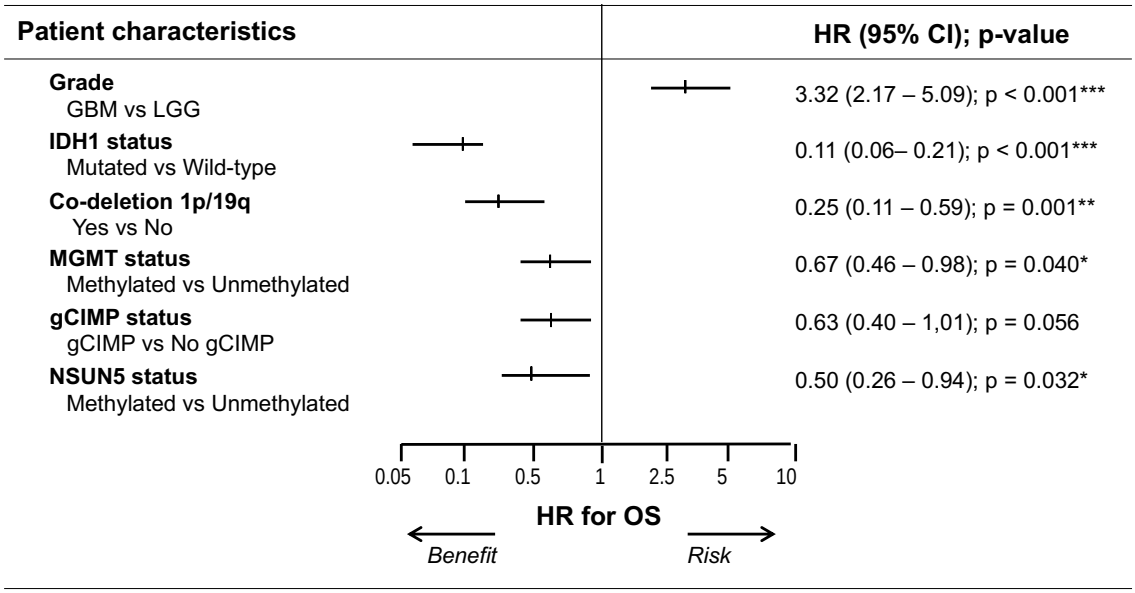


Figure 40. NSUN5-promoter methylation is an independent predictor of longer OS in glioma. Forest plot showing results of the multivariate Cox regression analysis of OS in the TCGA primary glioma cohort (n = 497). Analysis includes different glioma prognostic factors together with NSUN5 methylation status. P-values and hazard ratios (HR) with a 95% of confidence interval (95%CI) are shown. Co-variables with p-value < 0.05 were considered as independent prognostic factors. *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001.

STUDY II: Epigenetic loss of the RNA methyltransferase NSUN7 in Hepatocellular Carcinoma (HCC)

Identification of NSUN7 promoter hypermethylation in hepatocellular carcinoma

Our aforementioned 450K methylation array data mining aiming to find epigenetically altered expression of m^5C RNA methylation-related genes gave another candidate besides NSUN5, a second member of the NSUN family, NSUN7. NSUN7 promoter CpG island was found hypermethylated in several malignancies, and especially, in 74% (32 of 43) of the melanoma cell lines studied, in 58% (7 of 12) of the hepatocellular carcinoma (HCC) cell lines, and 55% (83 of 149) of the haematological cancer cell lines, at the top rank (**Figure 41A**). Importantly, NSUN7 promoter CpG island was unmethylated in all the normal tissues studied (Data not shown).

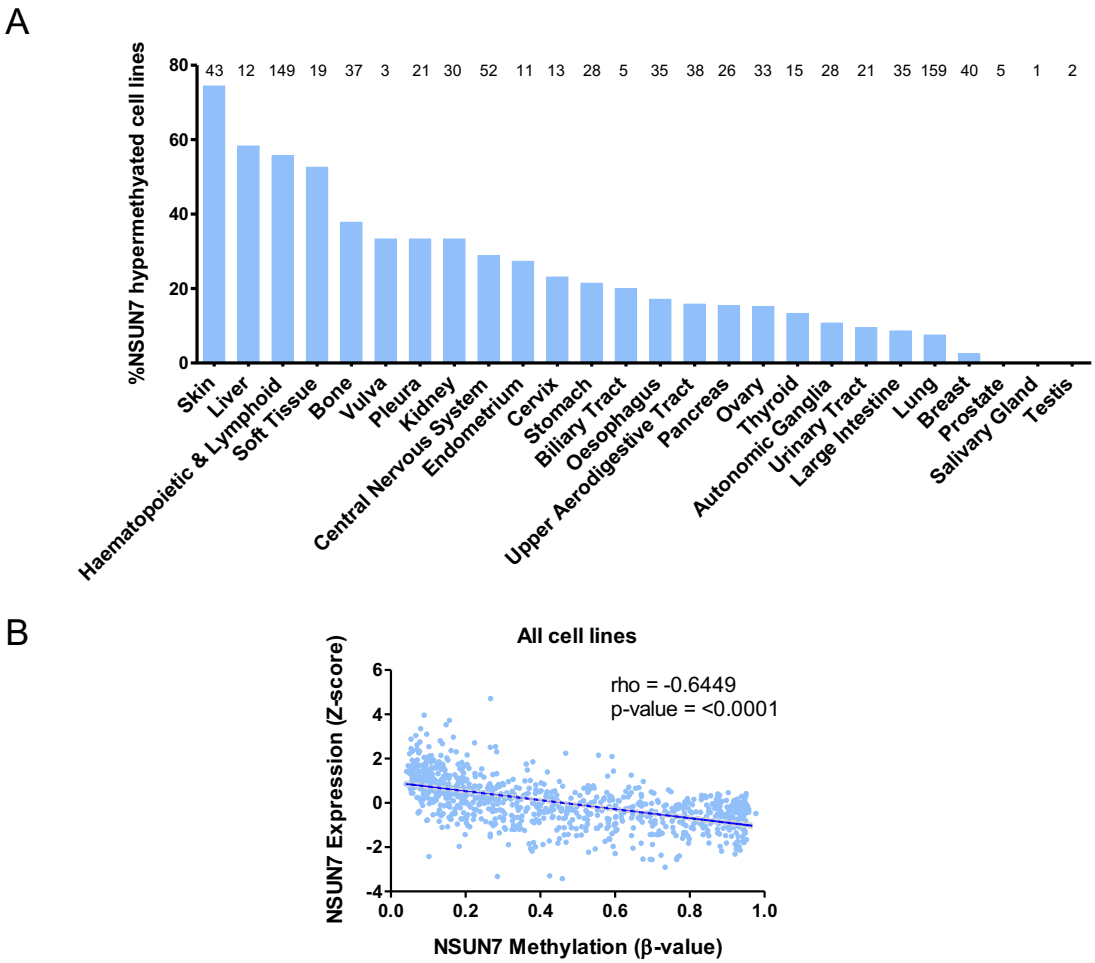


Figure 41. NSUN7 promoter hypermethylation in human cancer cell lines. (A) Percentage of human cancer cell lines, classified by primary tumor site, with a hypermethylated NSUN7 promoter. Total number of cell lines is shown on top of each bar (n = 861). **(B)** Correlation analysis between NSUN7 promoter methylation (mean β-value) and NSUN7 transcript expression (Z-score) in cancer cell lines (n = 861). Spearman's rank correlation test with its p-value and the associated rho coefficient are shown.

Moreover, there seemed to be a correlation between promoter CpG island methylation and NSUN7 expression (rho= -0.65, p-value < 0.0001) (**Figure 41B**), what was pointing out to a cancer-mediated epigenetic silencing of this gene. To validate the mechanism of NSUN7-transcriptional silencing we decided to focus on one of the highest methylated malignancies, hepatocellular carcinoma (HCC).

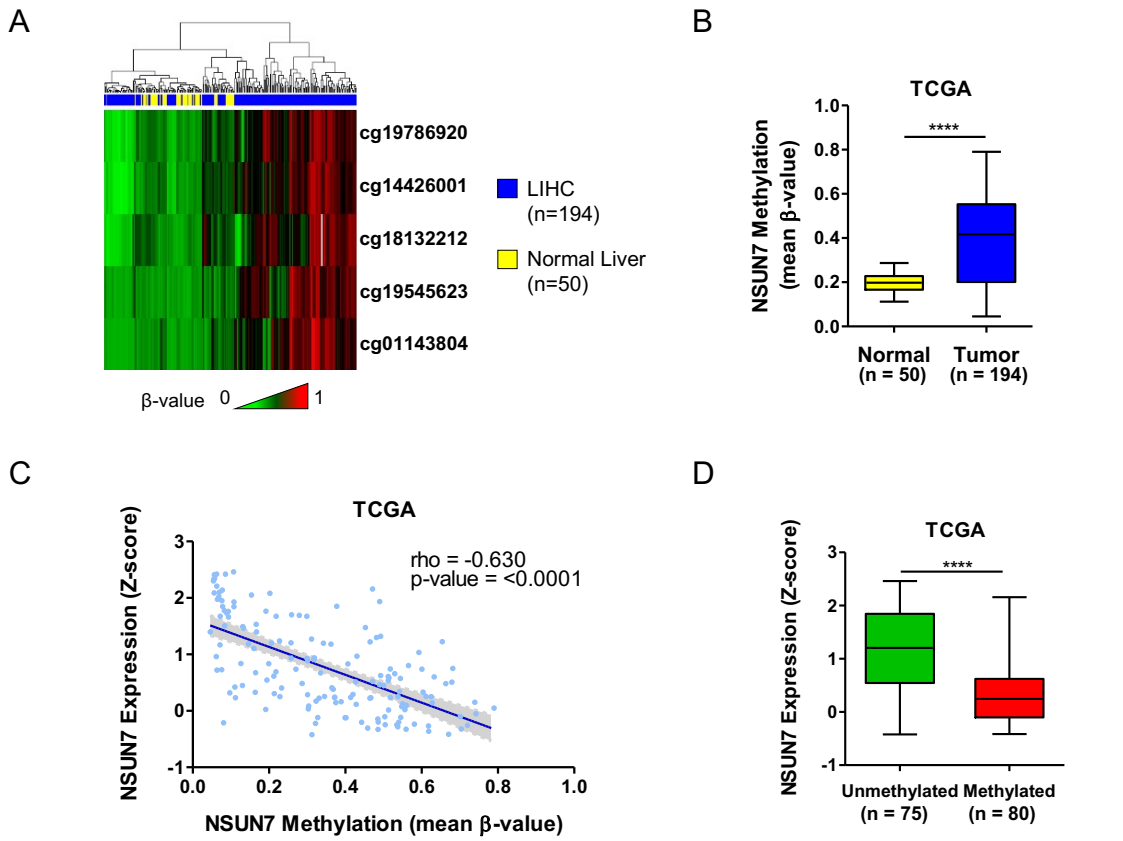


Figure 42. *In silico*-detected NSUN7 transcriptional silencing by promoter CpG island hypermethylation in the TCGA HCC cohort. (A) Heatmap representation of the methylation status of NSUN7 promoter-associated CpGs in human HCC tumors and normal liver samples. Every CpG has an absolute methylation level (β-value), ranging from 0 to 1. Green, unmethylated; red, methylated. **(B)** NSUN7-promoter methylation status (mean β-value) in TCGA HCC tumors and normal liver samples. **(C)** NSUN7-promoter methylation is inversely correlated with *NSUN7* transcript expression in TCGA HCC tumors (n = 194). Spearman's rank correlation test p-value and rho coefficient are shown. **(D)** Box plots showing *NSUN7* transcript expression (Z-Score) in unmethylated and methylated liver tumors from the TCGA cohort.

To confirm that NSUN7 epigenetic alteration takes place in human cancer samples, two publicly available HCC tumor cohorts were data mined: the TCGA HCC cohort, counting with 450K methylation array data from 194 tumor samples, and a second cohort, belonging to the HEPTROMIC Consortium, for which our lab collaborated establishing the methylome of the samples (n=229) [219]. Analyzing CpG island promoter methylation,

TCGA HCC tumors displayed a significantly higher NSUN7 promoter methylation, compared to normal liver samples (**Figure 42A, B**), and TCGA RNA sequencing data showed that NSUN7 hypermethylation was associated with transcript downregulation across tumor samples (rho = -0.63, p-value < 0.0001) (**Figure 42C**), being *NSUN7* significantly more expressed in hypomethylated tumors (β-value < 0.33) (**Figure 42D**).

We observed similar results for the second HCC cohort: NSUN7 promoter displayed significantly higher methylation levels in liver tumors, in comparison to normal samples (**Figure 43A, B**). Moreover, expression microarray data from the HEPTROMIC cohort showed that NSUN7 promoter hypermethylation was inversely correlated to NSUN7 expression (**Figure 43C**), being the *NSUN7* transcript significantly more expressed in unmethylated tumors (β-value < 0.33) (**Figure 43D**).

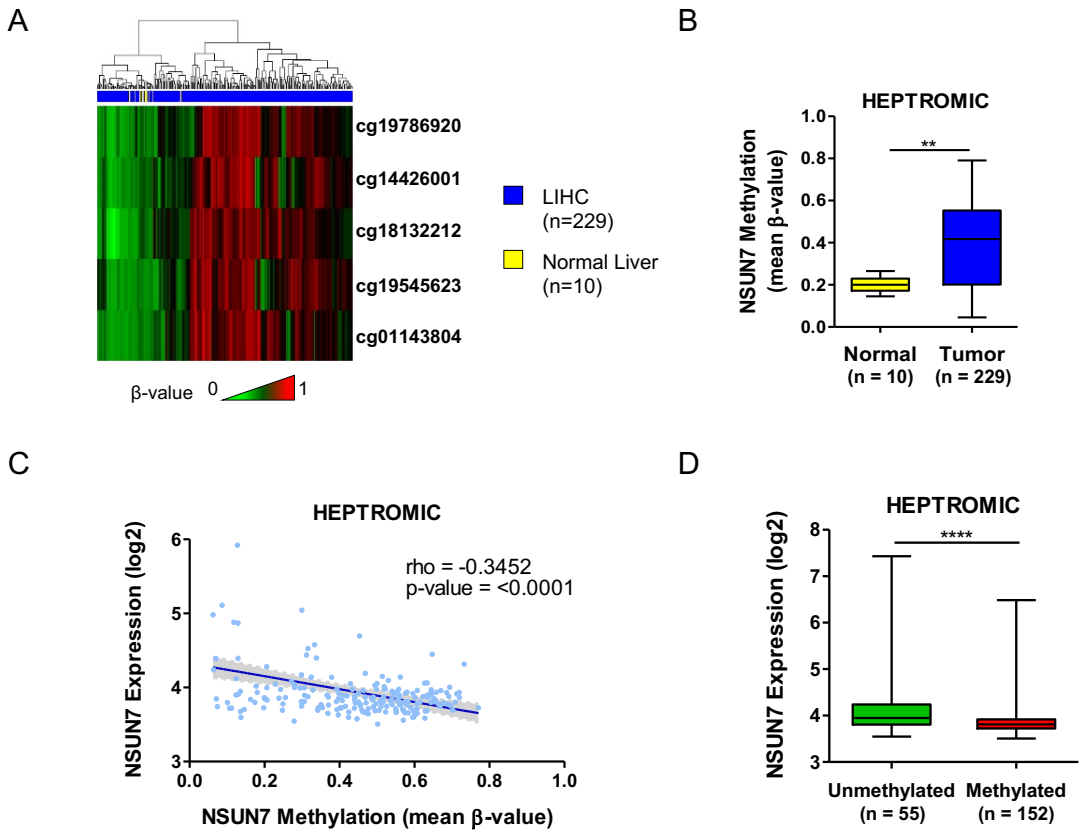


Figure 43. *In silico*-detected NSUN7 transcriptional silencing by promoter CpG island hypermethylation in the HEPTROMIC HCC cohort. (A) Heatmap representation of the methylation status of the CpGs considered, in primary HCC tumors and normal liver samples. Every CpG displays a methylation β-value that ranges from 0 to 1. Green, unmethylated; red, methylated. **(B)** NSUN7-promoter methylation status (mean β-value) is significantly higher in HCC tumors, compared to normal liver samples. **(C)** Correlation analysis between NSUN7 promoter methylation (mean β-value) and *NSUN7* transcript expression (log2) in HEPTROMIC HCC tumors (n = 229). Spearman's rank correlation test with its p-value and the associated rho coefficient are shown. **(D)** Box plots showing significantly higher *NSUN7* transcript expression (log2) in unmethylated liver tumors from the HEPTROMIC cohort.

Functional characterization of NSUN7 promoter methylation-mediated silencing in HCC cells

To experimentally validate the above-mentioned promoter methylation of NSUN7 in liver tumors, we moved to an *in vitro* setting. According to the methylation array data (**Figure 44A**), three NSUN7 hypomethylated lines (HEP3B2-1-7, SNU-387, SNU-475) and three NSUN7 hypermethylated cell lines (SNU-423, SNU-398, HUH-7) were subjected to bisulfite genomic sequencing, using primers specifically designed to analyze the TSS-associated CpG island. Results confirmed the promoter methylation pattern obtained with the DNA microarray (**Figure 44C**). Moreover, as aforementioned, NSUN7 promoter was found hypomethylated in normal liver tissues (**Figure 44B**), and the array results were validated by bisulfite genomic sequencing of one human liver sample (**Figure 44C**).

Importantly, SNU-423, SNU-398 and HUH-7 HCC cell lines, with proved hypermethylated promoter, showed minimal expression of NSUN7, analyzed by quantitative real-time PCR and western blot, whereas the hypomethylated HEP3B2-1-7, SNU-475 and SNU-387 cell lines showed NSUN7 expression, both at the RNA and protein levels (**Figure 45A**). Finally, treatment with the DNA-demethylating agent AZA restored *NSUN7* transcript expression in the three hypermethylated cell lines (**Figure 45B**), confirming DNA methylation as the epigenetic mechanism responsible for NSUN7 silencing.

Overall, these results confirmed that the *NSUN7* gene undergoes a cancer-specific transcriptional silencing by promoter CpG island hypermethylation in HCC cell lines.

A

		HCC cell lines					
Target ID	Location	HEP3B2-1-7	SNU-475	SNU-387	SNU-423	SNU-398	HUH-7
cg14426001	TSS200	0.32	0.53	0.29	0.76	0.86	0.92
cg19786920	TSS200	0.17	0.40	0.39	0.72	0.91	0.95
cg18132212	TSS200	0.06	0.34	0.37	0.71	0.85	0.90
cg01143804	TSS200	0.01	0.12	0.33	0.52	0.99	0.99
cg19545623	5'UTR;1stExon	0.12	0.15	0.27	0.71	0.97	0.98

B

		Normal Liver					
Target ID	Location	NL1	NL2	NL3	NL4	NL5	NL6
cg14426001	TSS200	0.19	0.08	0.23	0.26	0.20	0.19
cg19786920	TSS200	0.26	0.17	0.23	0.24	0.20	0.19
cg18132212	TSS200	0.17	0.16	0.19	0.26	0.19	0.16
cg01143804	TSS200	0.05	0.05	0.03	0.05	0.03	0.02
cg19545623	5'UTR;1stExon	0.08	0.11	0.05	0.11	0.03	0.06

C

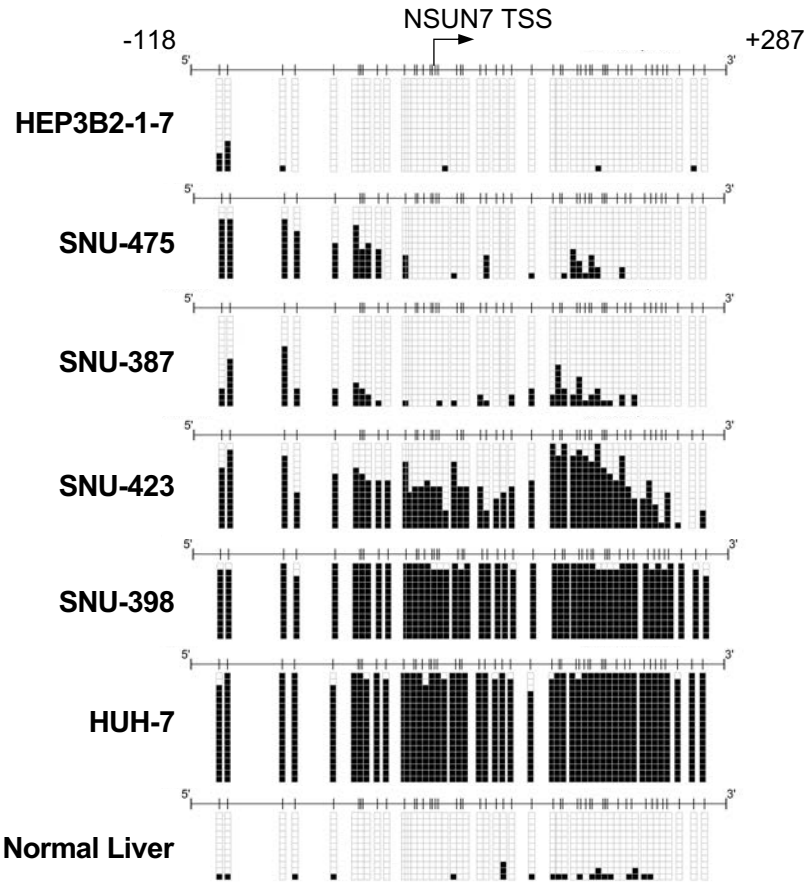


Figure 44. *In vitro* validation of NSUN7 promoter methylation in HCC cell lines. (A) Absolute methylation β-values of NSUN7 promoter-associated CpGs analyzed by the 450K methylation array. Green, unmethylated; red, methylated. Data from the six HCC cell lines, and **(B)** from representative normal liver samples. **(C)** Bisulfite genomic sequencing of NSUN7-promoter CpG Island in HCC cells, plus a normal liver sample. CpG dinucleotides are represented as short vertical lines, NSUN7 TSS is indicated by a black arrow, and unmethylated or methylated cytosines are represented as white or black squares, respectively. Single clones are shown for each sample (n > 10).

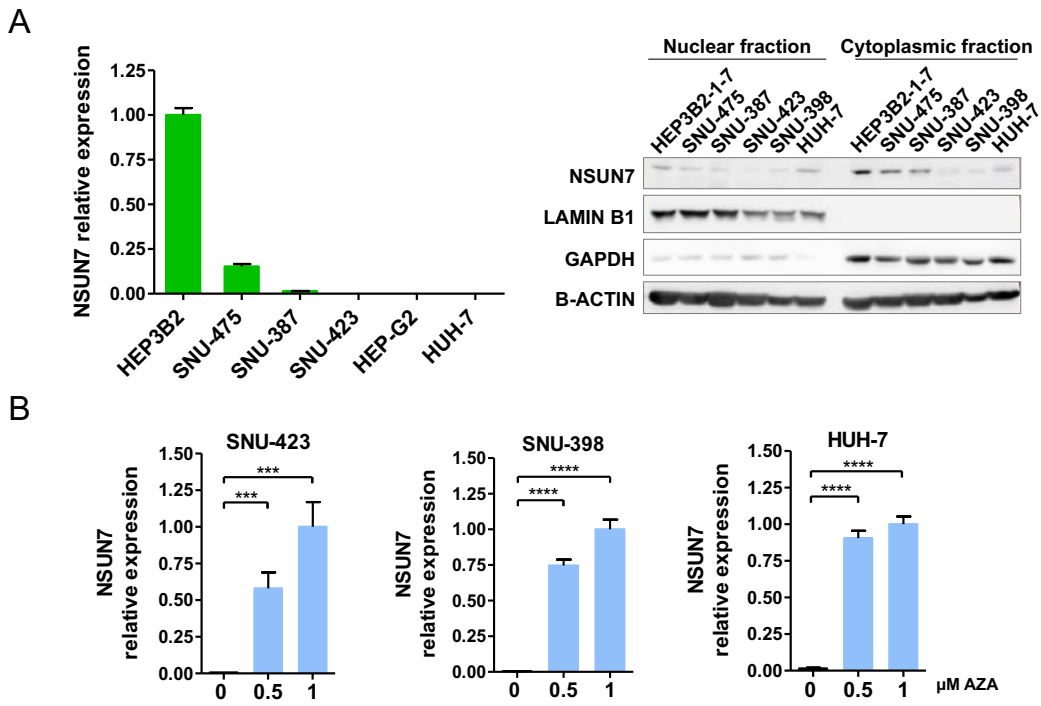


Figure 45. NSUN7 epigenetic loss in HCC cell lines. (A) NSUN7 expression at the RNA (*left*) and protein levels (*right*), analyzed by real-time PCR and western blot, respectively. **(B)** *NSUN7* transcript expression is restored in the three NSUN7 hypermethylated cell lines (SNU-423, SNU-398 and HUH-7) by treatment with the demethylating agent AZA. RNA expression data shown represent the mean ± S.D. of biological triplicates, and p-values were calculated by a Student's T test. ***p-value < 0.001, ****p-value < 0.0001.

Development of cellular tools for investigating NSUN7 molecular function

Having demonstrated NSUN7 loss in HCC, we aimed to study its contribution to the biology of these tumors. To this end, we developed loss- and gain-of-function models.

To develop a NSUN7 loss-of-function model we knocked out its expression in the NSUN7-hypomethylated HEP3B2-1-7 cell line, via the CRISPR-Cas9 system. Three different sgRNAs were designed to target the Cas9 machinery to different NSUN7 exons within the SAM methyltransferase functional domain (**Figure 46A**). After obtaining a clonal population, we analyzed NSUN7 protein expression by western blot, and observed a NSUN7 knockout (KO) (**Figure 46B**). Conversely, NSUN7 wild-type expression was restored in the NSUN7-hypermethylated SNU-423 cell line (**Figure 46E**) by stable transduction of an integrative expression vector. Moreover, we developed a NSUN7 enzymatically-defective protein (MUT) by introducing a point mutation at one of the cytosines involved in the catalysis (C439V) (**Figures 46C, D**), and SNU-423 HCC cells were transduced with the NSUN7 mutant version, as well (**Figure 46E**).

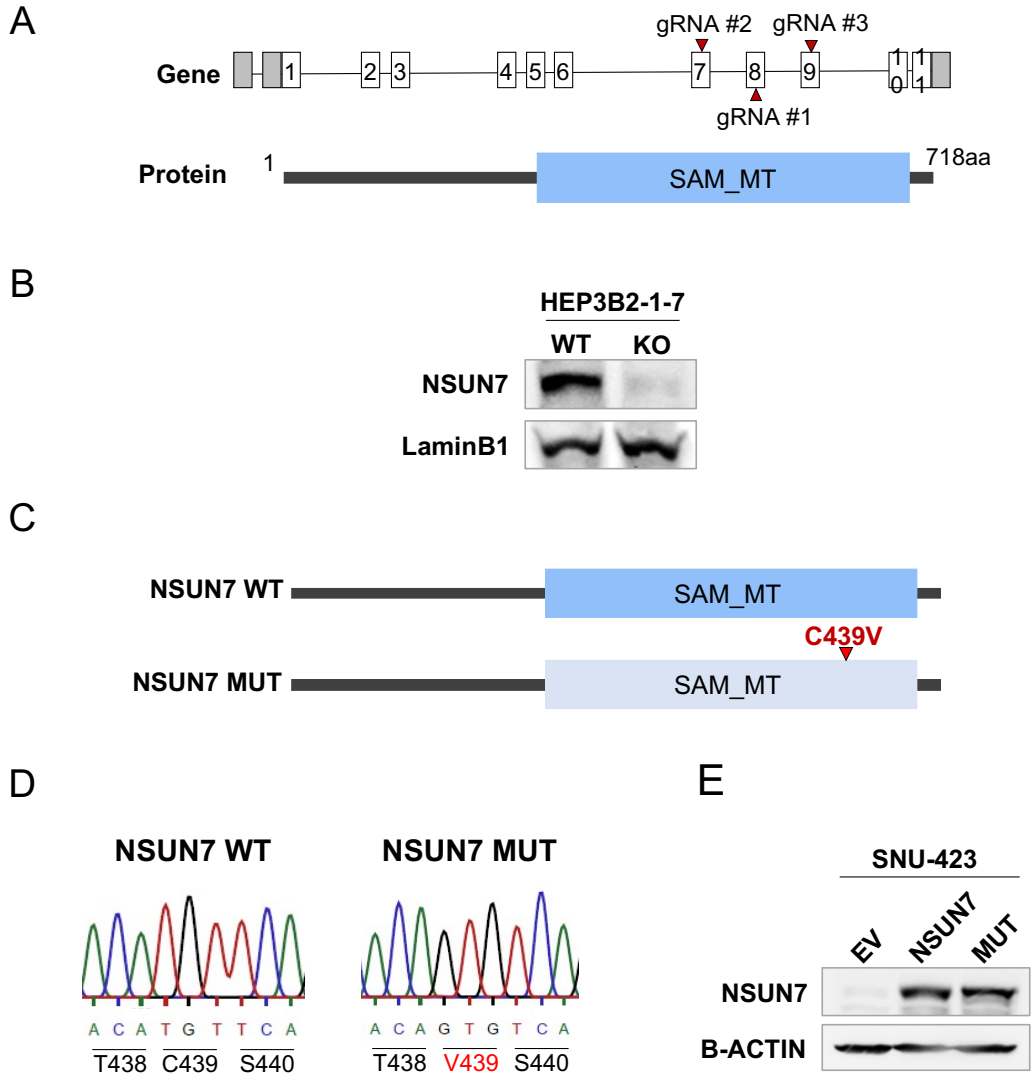


Figure 46. Development of NSUN7 loss- and gain-of-function models. (A) Schematic representation of the *NSUN7* gene and its correspondent translation into protein. Triangles depict the target site of each gRNA. **(B)** Efficient CRISPR-mediated knockout of NSUN7 protein expression in the hypomethylated and expressing HEP3B2-1-7 HCC cell line. **(C)** Schematic representation of wild-type and mutated NSUN7 proteins. Triangle marks the point mutation introduced. The major functional domain (SAM_MT: SAM-dependent MT NOL1/NOP2/SUN domain) is displayed in blue. **(D)** Sanger sequencing of the NSUN7 mutant construct, confirming the substitution of the cysteine 439 by a valine. **(E)** Restoration of NSUN7 wild-type and mutated protein expression by stable transduction in epigenetically-silenced SNU-423 HCC cells, analyzed by western blot.

Insights into NSUN7 role in sperm motility

We next wondered about the molecular pathways potentially altered upon NSUN7 loss. To this end, we performed RNA sequencing (RNA-seq) of the NSUN7 KO model. Strikingly, a clustering of the RNA-seq results by the top 30 most variable genes showed

that two sperm-expressed transcripts were highly depleted in NSUN7-knocked out HEP3B2-1-7 cells, *CATSPERG* and *ATP2B4* (PMCA4) (marked with an asterisk) (**Figure 47A**). We carried out quantitative real-time PCR analysis and validated the RNA-seq data: upon NSUN7 knockout in HEP3B2-1-7 cells, *CATSPERG* and *ATP2B4* transcript levels were dramatically reduced (**Figure 47B**).

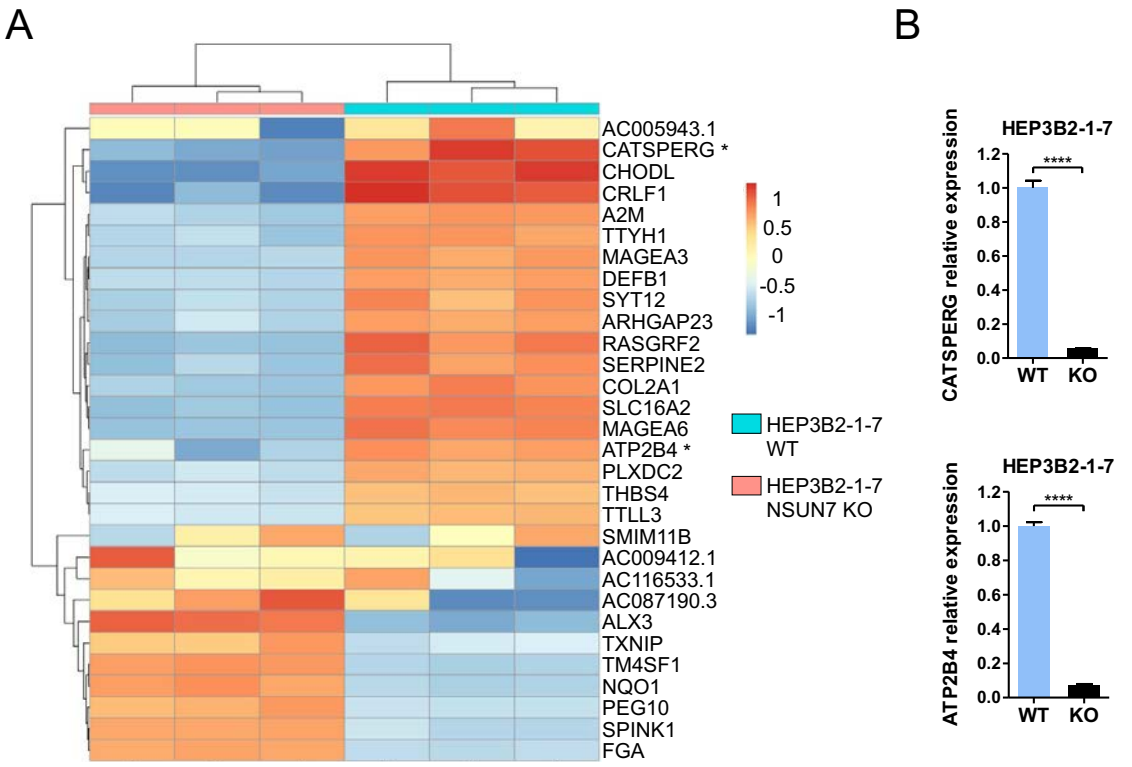


Figure 47. NSUN7 loss is associated with depletion of two flagellar plasma membrane channels. (A) Heatmap representation of the top 30 most variable genes upon NSUN7 knockout in HEP3B2-1-7 HCC cell, according to the RNA-seq results. *CATSPERG* and *ATP2B4* transcripts are marked with an asterisk. **(B)** Validation of *CATSPERG* and *ATP2B4* transcript expression levels in NSUN7 wild-type and knocked-out HEP3B2-1-7 cells by quantitative RT-PCR. RNA expression data shown represent the mean ± S.D. of biological triplicates, and p-values were calculated by a Student's T test. ****p-value < 0.0001.

Our data, thus, showed that NSUN7 loss results in a dramatic reduction of the flagellar plasma membrane channels CATSPERG and PMCA4, both with described roles in sperm motility. Further research would be needed in order to connect the loss of these two flagellar channels to the sperm motility defects found in NSUN7-mutation-bearing infertile men. However, as it is not a cancer-related process, we moved our focus towards other biological effects of NSUN7 loss.

Unraveling NSUN7 catalytic role in HCC cells

In order to uncover NSUN7's target spectrum we carried out transcriptome-wide bisulfite sequencing (bsRNA-seq) in both loss- and gain-of-function models. We observed that lack of NSUN7 did not cause a major change in m^5C sites in 'well-recognised' RNA types. However, we were still able to filter out around 39 mRNA sites that were differentially methylated depending on NSUN7 presence (≥ 20% change in non-conversion rate), although we did not observe overlap between the two sets of transcripts (HEP3B2-1-7 NSUN7 WT vs KO and SNU-423 EV vs NSUN7). It is worth mentioning that putative NSUN7 target transcripts appeared methylated at low stoichiometries.

In order to validate bsRNA-seq results, we next performed RNA bisulfite sequencing of those candidate mRNAs with a methylation fold-change > 0.30, from both the NSUN7 KO and overexpression models. Interestingly, by this technique we identified a NSUN7 methylation target conserved between both models, the *CCDC9B* mRNA, which codes for a putative RNA-binding protein. According to bsRNA-seq results, the C1324 of the *CCDC9B* mRNA, on the CDS, undergoes m^5C methylation in a NSUN7-dependent manner. However, by RNA bisulfite sequencing not only C1324 but also C1322 and C1323 sites appeared to be methylated in the three NSUN7 unmethylated cell lines (HEP3B2-1-7, SNU-475 and SNU-387 cells), while NSUN7-silenced SNU-423, SNU-398 and HUH-7 cells showed lack of methylation at these three sites (**Figure 48**). Strikingly, while the stoichiometry of C1322-1324 *CCDC9B* mRNA methylation in the NSUN7 unmethylated cancer lines was low, all *CCDC9B* transcripts were fully methylated in the normal liver sample (**Figure 48**). To further confirm NSUN7 role in the deposition of m^5C1322-1324 on the *CCDC9B* mRNA, we tested their methylated status in our *in vitro* models. We observed that upon NSUN7 KO in NSUN7-expressing HEP3B2-1-7 cells the methyl group was lost in the three positions (**Figure 49A**). Conversely, the NSUN7-silenced SNU-423 cell line, already proven C1322-1324 *CCDC9B* mRNA unmethylated, recovered m^5C mRNA methylation after restoration of the wild-type NSUN7 protein, and, importantly, restoration of the mutant NSUN7 version was unable to recover methylation at the three cytosines (**Figure 49B**).

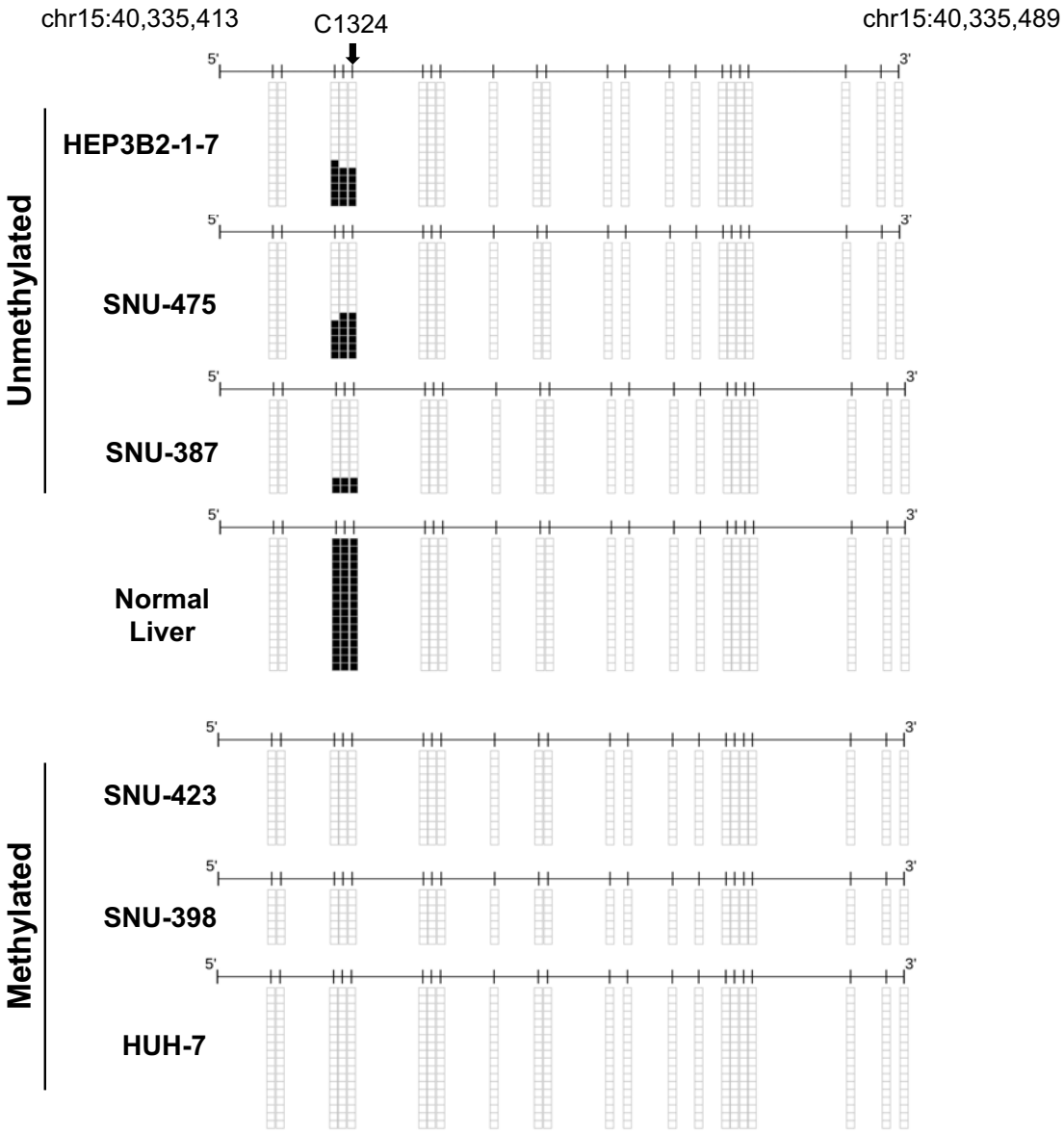


Figure 48. NSUN7 epigenetic loss abolishes the methyl group on the C1322, 1323 and 1324 positions of the *CCDC9B* mRNA. RNA bisulfite sequencing of multiple clones of *CCDC9B* transcripts from the HCC cell panel plus a normal liver sample, according to NSUN7-promoter methylation status. Cytosines are represented as short vertical lines, the bsRNA-seq target C1324 is indicated by a black arrow, and presence of a methyl group is represented as a black square.

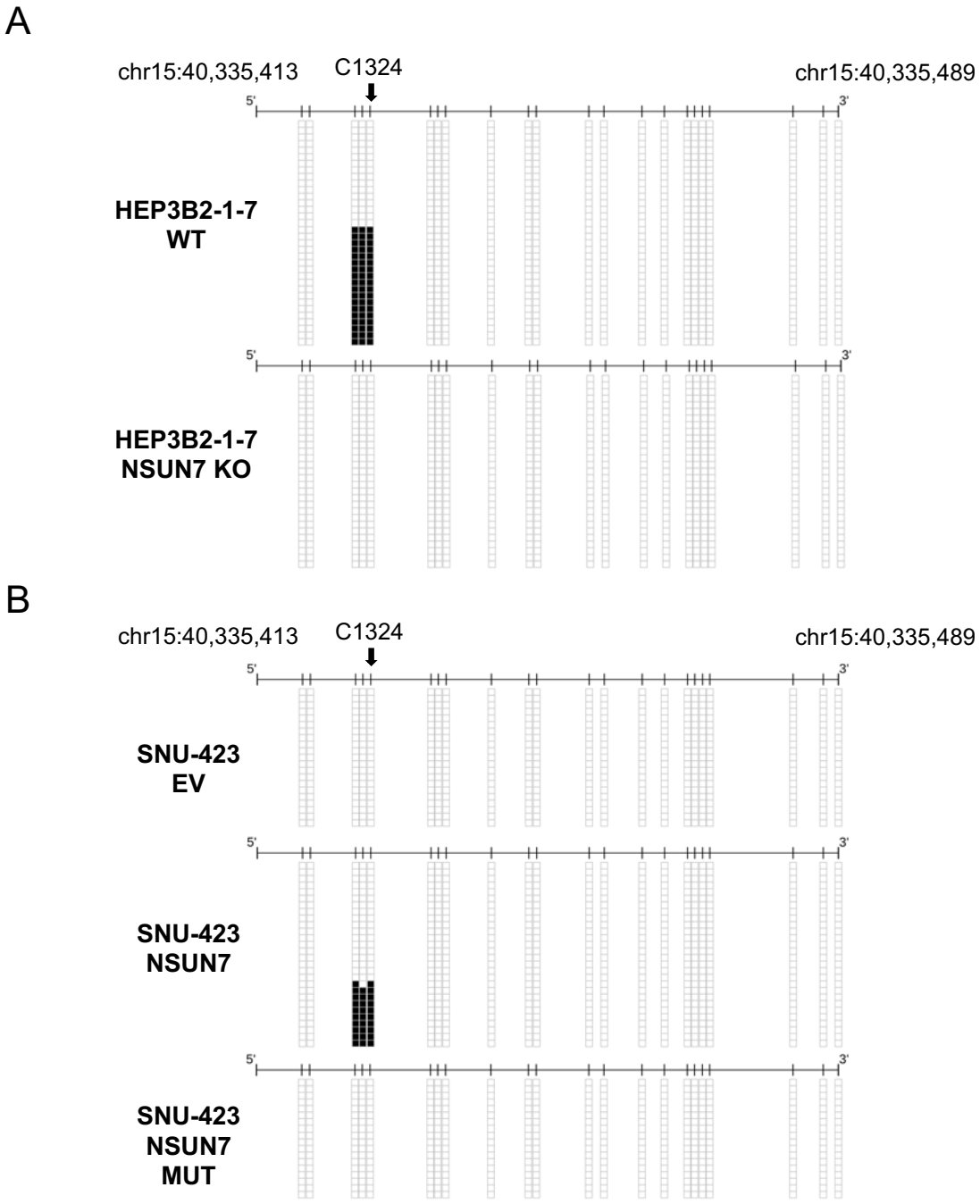


Figure 49. NSUN7 methylates the C1322, 1323 and 1324 positions of the *CCDC9B* mRNA in HCC cells. RNA bisulfite sequencing of multiple clones of *CCDC9B* transcripts from NSUN7 WT and NSUN7 KO HEP3B2-1-7 cells **(A)**, and EV- and NSUN7 wild-type and mutated-transduced SNU-423 cells **(B)**. While loss of NSUN7 in HEP3B2-1-7 cells depletes m^5C methylation on C1322-24 positions, its restoration into SNU-423 epigenetically silent cells restores the three methyl groups. Importantly, a catalytically-inactive NSUN7 protein is not able to restore the m^5C1322-24 methylation status. Cytosines are represented as short vertical lines, the bsRNA-seq target C1324 is indicated by a black arrow, and presence of a methyl group is represented as a black square.

To evaluate the effect of the m^5C1322-24 loss upon NSUN7 epigenetic silencing, we analyzed CCDC9 expression in the HEP3B2-1-7 KO model. We observed a decreased CCDC9B expression both at the RNA and the protein level upon NSUN7 knockout (**Figure 50A**). To test if the methyl group plays a role in the stability of the *CCDC9B* mRNA, we carried out an Actinomycin D chase assay, to measure mRNA half-life. We observed longer *CCDC9B* transcript half-life in NSUN7-expressing and m^5C1322-24 methylated HEP3B2-1-7 cells compared to their NSUN7 KO counterparts (**Figure 50B**).

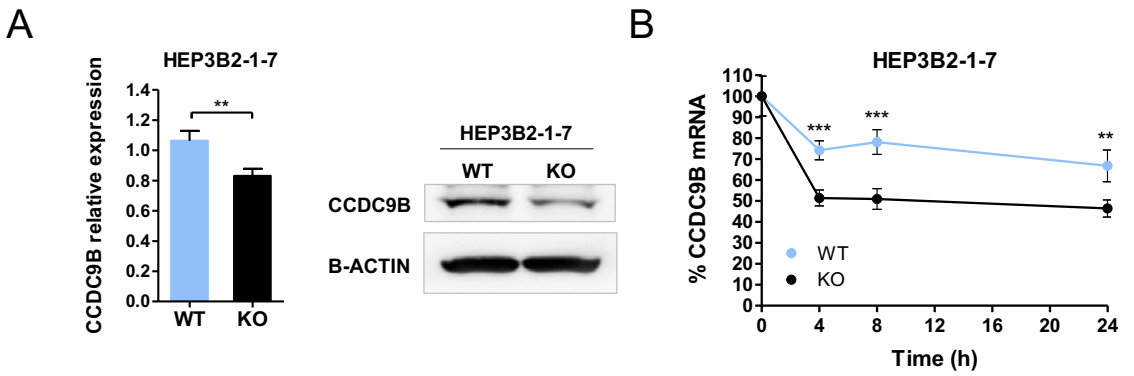


Figure 50. Effect of NSUN7-mediated C1322-24 methylation on *CCDC9B* mRNA. (A) CCDC9B expression levels in NSUN7 wild-type and knocked out HEP3B2-1-7 cells, analyzed by quantitative RT-PCR (left) and western blot (right). **(B)** CCDC9B transcript half-life depending on NSUN7 expression, analyzed by Actinomycin D chase assay.

All the above results allow us to establish that the human NSUN7 RNA methyltransferase is responsible for catalyzing the incorporation of a methyl group onto the 1322, 1323 and 1324 cytosines of the *CCDC9B* mRNA, and these chemical marks contribute to transcript stability. Moreover, NSUN7 epigenetic silencing in HCC cells results in the loss of methylation at the three positions, leading to a decreased CCDC9B expression.

DISCUSSION

DISCUSSION

RNA modifications have long been known by the scientific community, and yet, the absence of proper methods to measure or map them made the field unattended until the last decade. Recent technological and methodological advances, especially the development of transcriptome-wide sequencing techniques, have led to the discovery of several previously unknown RNA modifications in all domains of life, promoting, as well, research to uncover the molecular players depositing, decoding and removing those marks. This boost on the epitranscriptomic field brought to light the crucial role for RNA modifications along all steps of RNA life cycle and functions, and how disruptions in RNA modification-related enzymes and, therefore, altered single epitranscriptomes are associated with numerous diseases, including cancer. In fact, distinct epitranscriptomic mechanisms have been described to exert a role in the acquisition of different hallmarks of cancer, mainly by genetic mutations and aberrant expression of epitranscriptomic enzymes [183,184] .

However, although the contribution of epigenetics to tumorigenesis has been vastly described [186,220], by the time this thesis was started, no epigenetic mechanisms affecting the expression of epitranscriptomic players had been described in cancer. This fact encouraged us to contribute to the cancer epitranscriptomics field by potentially identifying epigenetic alterations in epitranscriptomic players occurring in cancer.

As the study of RNA modifications has been, all along its history, hampered by the existence of methods to detect every of them, we decided to focus our first dive into the epitranscriptomic field on a mark already familiar to the laboratory, 5-methylcytosine. m^5C epitranscriptomic players described so far include 8 writers (NSUN1-7, DNMT2), 2 readers (ALYREF, YBX1) and one eraser (ALKBH1). As the main epigenetic mechanism causing transcriptional silencing is gene promoter hypermethylation, we decided to assess the methylation status of the 11 m^5C-related genes across about 1,000 cancer cell lines. Surprisingly, we found that two m^5C methyltransferase genes, *NSUN5* and *NSUN7*, were hypermethylated in several cancer cell lines from different origins (**Figure 8**), what became the starting point of two independent studies.

STUDY I: Insights into the epigenetic loss of the m^5C RNA methyltransferase NSUN5, and its contribution to gliomagenesis

A closer look into *NSUN5* promoter hypermethylation in cancer cell lines showed that methylation was enriched in tumoral cells from the CNS (**Figure 10**). CNS tumors are a heterogeneous group of tumors that arise from CNS cells, and can be benign or malignant, being malignant brain tumors among the deathliest cancer types, with a 5-year OS not higher than 35% [221]. Thus, identification of a cancer-specific epigenetic silencing of NSUN5 appeared to us a great opportunity to better understand brain tumorigenesis, and potentially contribute to its clinical management.

The most common adult malignant CNS tumors are gliomas, constituting 75% of all cases, with the remaining types including meningiomas, lymphomas, choroid plexus tumors, neuronal and mixed-glial tumors, tumors of the pineal region, or malignant peripheral nerve sheath tumors, among others. Gliomas have a neuroectodermal origin, arise from glial or precursor cells, and can be divided into astrocytomas, oligodendrogliomas and oligoastrocytomas (grades I-III) and glioblastomas (grade IV). The 2016 WHO classification of CNS tumors underwent a major restructuring in glioma nomenclature and grading from the previous 2007 version, by including for the first-time molecular markers besides the traditional histological and immunohistochemistry parameters, improving patient diagnosis and prognosis [211]. The most commonly used glioma molecular markers are IDH mutational status, 1p/19q codeletion and MGMT promoter methylation [218]. Mutations in Isocitrate dehydrogenase 1 and 2 (*IDH1* and *IDH2*) genes are an early event in gliomagenesis [222]. Circumscribed low-grade gliomas (astrocytomas grade I) do not show IDH mutations, but they are frequently found in diffuse low-grade gliomas (> 70% in astrocytomas grades II and III; 100% oligodendrogliomas) and in 10% glioblastomas. Gliomas harboring IDH mutations show better prognosis, compared to their IDH wild-type counterparts [222]. Likewise, the non-balanced centromeric translocation t(1:19) (q10:p10) resulting in codeletion of 1p and 19q chromosomes, is associated with better prognosis [223–225], and predicts sensitivity to alkylating agents [226]. On the other hand, the DNA repair protein O^6-methylguanine-DNA methyltransferase (MGMT) that removes mutagenic alkyl adducts from guanine bases, is commonly found silenced by promoter hypermethylation in glioblastomas (30-50% of IDH wild-type tumors) and oligodendrogliomas (>90%). *MGMT* promoter methylation confers sensitivity to alkylating agents such as temozolomide, and is associated with longer survival of patients [227–229].

We proved the existence of NSUN5 promoter hypermethylation in primary glioma tumor samples of all grades from the TCGA collection (**Figure 10**), and importantly, its unmethylated status in normal tissue samples. Our study, thus, confirmed NSUN5 epigenetic alteration as a cancer-specific mechanism, being the first lesion described for this RNA methyltransferase in cancer.

Next, by *in silico* mining *NSUN5* transcript expression both in the COSMIC cell line collection and in the TCGA human glioma cohort, we observed that, as expected, NSUN5 hypermethylation was associated with transcript downregulation (**Figures 11, 12**). Moreover, we experimentally validated NSUN5 translational silencing due to promoter hypermethylation both in a LGG cellular cohort and in a GBM cohort (**Figures 13-15**). The reason to stratify glioma cells is because tumor grade profoundly impacts patient prognosis, with GBM showing an extremely unfavorable outcome (median OS for IDH-mutant glioblastomas is lower than 3 years, and for IDH-wild-type tumors lower than 1 year) [211]. Management of those patients with single-agent therapies has proven disappointing so far, and combination of different therapies is required, what leaves plenty of room for investigating new strategies, such as biomarker-targeted compounds. Studying NSUN5 implication in GBM tumorigenesis, may give rise to new drug target candidates for this devastating disease.

To unveil which role the previously unknown epigenetic silencing of NSUN5 was playing in glioma, we evaluated its effect in tumorigenesis both *in vitro* and *in vivo*. While NSUN5 restoration into silenced cells reduced cell viability, downregulation of NSUN5 by shRNA led to an increment of cell growth *in vitro* (**Figure 16**). Moreover, NSUN5-depleted tumors in xenograft mice, and later in orthotopic models, showed significantly higher growth and weight compared to tumors derived from glioma cells expressing the methyltransferase, and, consequently, mice harboring NSUN5-depleted tumors showed decreased OS (**Figures 17, 18**). These observations pointed out that NSUN5 displays tumor suppressor properties in glioma cells, a reason explaining its epigenetic silencing in glioma tumors.

We then hypothesized that NSUN5 tumor suppressor role may be mediated by its methylation targets, at that point, still uncharacterized in human. An alignment of protein sequences of NSUN5 orthologues in yeast (Rcm1), worm (*nsun-5*) and mice (Nsun5) showed high conservation, specially within the RNA methyltransferase catalytic domain (**Figure 19**), and, most strikingly, alignment of the described targets for those orthologues (C2278 of the yeast 25S rRNA and C2381 of the worm 26S rRNA) [107,128] to the human and mouse 28S rRNA, showed that the rRNA region is extremely conserved across lower and higher eukaryotes (**Figure 19**). Moreover, it informed us that the C2381

and C2278 positions correspond to the cytosine 3782 of the human 28S rRNA, and to the 3438 of the mouse 28S rRNA. In order to validate if NSUN5 is responsible for methylating the C3782 of 28S rRNA in glioma cells, we employed two orthogonal m^5C RNA methylation detecting methods, a bulk quantification of m^5C levels by LC-MS and RNA bisulfite sequencing. Results showed how, both, LGG and GBM cell lines expressing NSUN5 displayed m^5C3782 28S rRNA status, while NSUN5-depleted cell lines exhibited lack of methylation at this position (**Figures 21-23**). Moreover, to undoubtedly link NSUN5 methyltransferase activity to the presence of a methyl group on the C3782, we developed and ectopically expressed a NSUN5 catalytic defective protein and observed how it was unable to restore C3782 methylation status in NSUN5-epigenetically silent cells (**Figure 24**). Therefore, our results described, for the first time, the human target for the m^5C writer NSUN5. Importantly, not long after our discovery, another study confirmed NSUN5-mediated methylation of C3782 28S rRNA in HeLa cells, and Nsun5-mediated methylation of C3438 of the mouse 28S rRNA [230], confirming our predicted target position in the mouse transcriptome (**Figure 19**).

The discovery of NSUN5 catalytic activity raised the question of how the loss of a methyl group on a cytosine on the 28S rRNA, due to NSUN5 epigenetic loss, could be of any advantage for the glioma tumorigenic process. Literature research proved that rRNA modifications play significant roles in ribosome structure and function. In fact, rRNAs undergo vast modifications mainly consisting of ribose (N_m) and base methylation, and pseudouridylations (Ψ). The modified sites cluster in functionally important ribosomal regions and are conserved from prokaryotes to eukaryotes, increasing their numbers with organismal complexity (36 in *E. coli*, 112 in yeast, and > 200 in human) [231]. In fact, there are several reports proving the importance of rRNA modifications for proper ribosome function. As an example, a cluster of 10 Ψs in a helix located above the A-site, on the yeast LSU positively affects translational efficiency and fidelity [232]. Moreover, removal of several 2'-O-methylations and Ψs in the yeast SSU decoding center, through depletion of their snoRNA guides, reduced translation fidelity impairing stop codon termination and reading frame maintenance [233] . Even the function of the peptidyl-transferase center (PTC) is affected by rRNA modifications, as removal of Ψs in the PTC of yeast 25S rRNA resulted in rRNA structural changes that increased sensitivity to translation inhibitors and affected translational fidelity [234]. Strikingly, in multicellular organisms, loss of rRNA modifications can have widespread consequences. Some instances are the developmental defects observed in zebrafish upon loss of a single N_m modification in the 28S rRNA [235], and the effects over lifespan and oogenesis in *C. elegans* caused by loss of *nsun-1*, and therefore, loss of m^5C2381 on the 26S rRNA [108].

Altered rRNA modification patterns have also been proposed to contribute to tumorigenesis [236]. For instance, changes in rRNA 2'-O-methylation patterns through p53-mediated control of FBL expression leads to changes in stop codon read-through and enhanced translation of Internal Ribosome Entry Sites (IREs)-containing oncogenic mRNAs [237]. Moreover, patients of X-linked dyskeratosis congenita (X-DC), a syndrome characterized by mutations in the rRNA pseudouridine synthase *DKC1* gene, exhibit higher risks of developing cancer [238]. Molecular insights revealed how the consequent hypopseudouridylation of rRNA in X-DC patients specifically impairs the translation of IREs-containing mRNAs, such as those of the tumor suppressor genes *TP53* and *CDKN1B* [239,240].

Regarding our specific C3782 base methylation, its conserved position in the yeast 25S rRNA, C2278, is also described to exert a phenotype upon loss of its methyl group. Sharma et al., described that depletion of NSUN5-orthologue Rcm1 increased yeast sensitivity to anisomycin, potentially through alterations in the structure of the 25S rRNA [107], and Gigova et al., confirmed that loss of a cluster of methylations including m^5C2278 induces structural changes causing ribosome instability under high-salt conditions [241]. Therefore, we next sought to evaluate if loss of methylation at the C3782 site upon NSUN5 epigenetic silencing could impact ribosomal biogenesis and/or ribosomal structure. As we did not observe differences in the levels of mature rRNAs depending on NSUN5 presence or absence, we ruled out m^5C3782 implication in pre-rRNA processing (**Figure 28**), and decided to analyze potential structural changes on the ribosome. To this end we carried out MD simulations of a sub-region of the 28S rRNA comprising C3782 microenvironment under two scenarios: with an unmethylated site and with the presence of the methyl group. Simulations showed how loss of the methyl group may impair several base pairs potentially leading to a conformational change of a hairpin expose to the P-site of the LSU, whose function is interacting with the tRNA and the mRNA, ensuring the stability of the tertiary complex rRNA–tRNA–mRNA [212] (**Figure 29**). Therefore, lack of m^5C3782 through NSUN5 epigenetic silencing might be causing a structural distortion in the P-site cavity that would impair physiological protein synthesis.

To experimentally validate the suggested impairment of protein synthesis, we carried out OP-Puro and [3H] leucine incorporation for assessing *de novo* global translation. However, under basal conditions, we did not observe significant differences in overall protein synthesis between NSUN5-expressing (m^5C3782 methylated) and NSUN5-epigenetically silent (C3782 unmethylated) glioma cells (Data not shown). These results are in agreement with two independent reports in yeast by Sharma and Schosserer et

al., where comparisons between polysomal profiles of wild-type and Rcm1 knock-out cells did not reveal differences in the distribution of polysomes, 80S monosomes and free 60S and 40S subunits [107,128]. Polysome profiles provide snapshots of translationally active ribosomal populations, with shifts from polysome peaks to monosomes or even free subunits meaning decreased active translation and even disturbed ribosome assembly [242]. Therefore, unaltered profiles rule out translation differences. Nonetheless, Schosserer et al., did observe a shift of actively transcribed polysomes towards the 80S monosome peak under oxidative stress treatment of Rcm1-deficient yeast cells compared to wild-type cells, meaning a reduction of global translation happening upon loss of m^5C2278 25S rRNA methylation. For that reason, we decided to assess *de novo* global translation under two different stress conditions, oxidative stress through H_2O_2 treatment, and nutrient deprivation. Surprisingly, under both stress scenarios, we observed a significant reduction in overall protein synthesis in NSUN5-silenced C3782-hypomethylated glioma cells, compared to NSUN5-expressing cells, both endogenously and by ectopic restoration (**Figures 30, 31**).

Since it has been described that oxidative stress affects translational fidelity, promoting stop codon read-through and codon misreading [215], Schosserer et al., estimated the translational efficiency of individual transcripts in Rcm1 wild-type and knock-out cells, with and without oxidative stress. They observed that, while wild-type cellular translatomes were only affected by the stress treatment, for Rcm1 knock-out cells even under unstressed conditions, the absence of m^5C2278 was leading to a decreased translational fidelity and an increase in the presence of stress-related mRNAs in actively translating polysomes, despite they had not observed major differences in global translation. Then, under stress, they observed fewer differentially translated mRNAs between conditions. In fact, a third of the transcripts translationally regulated by H_2O_2 treatment in wild-type cells where already differentially expressed in basal Rcm1-depleted cells. Therefore, loss of Rcm1 was leading to a pre-activated translational stress response for an improved counteraction of cellular stress conditions.

As our results indicated that epigenetic loss of NSUN5 in glioma cells, and its consequently lack of m^5C3782 methylation is associated with an impairment of global protein synthesis under stress conditions, suggesting a conserved phenotype between yeast and human cells, we decided to test if NSUN5-silenced cells promote, as well, the translation of specific transcripts to cope with cellular stress despite their global inhibition of protein synthesis. To this end, we estimated total RNA expression (RNA-seq data) of the epigenetically-silent LN229 glioma cell line upon ectopic restoration of NSUN5, and

compared it to the fraction of those transcripts being actively translated within polysomes (Ribo-seq data), under basal conditions. We observed how, while the LN229 transcriptome was mainly unaltered regardless NSUN5 presence, more than three thousand transcripts showed ribosome occupancy differences between conditions (**Figure 32**). In order to characterize the putative translational program arising upon NSUN5 epigenetic loss, we took a deeper look into those transcripts with higher ribosome occupancy in the absence of NSUN5, and validated how up to 1987 mRNAs had a higher translational efficiency in NSUN5-epigenetically silent LN229 cells (**Figure 32**). Importantly, GSEA clustering of those transcripts revealed an overrepresentation of pathways related to cancer and the ribosome, and gene ontologies related to cellular and metabolic adaptation to stress (**Figure 32**). The same experimental strategy was followed for the NSUN5-loss-of-function model, and similar results were obtained: while depletion of NSUN5 in unmethylated DBTRG-05MG glioma cells did not affect the overall coding transcriptome, it affected the ribosome occupancy of more than 2% of mRNAs. Indeed, almost eight hundred transcripts showed enhanced translational efficiency upon loss of NSUN5 (**Figure 35**), and similarly, we found an overrepresentation of ontologies related to response to stress and metabolic adaptation among them (**Figure 35**).

Next off, we evaluated the translational behavior of cells under oxidative stress, carrying out the same battery of experiments already disclosed in the LN229 gain-of-function model. As we had already observed under basal cellular conditions (**Figure 32**), restoration of NSUN5 had minimal effect on the overall transcriptome (**Figure 34**). However, Ribo-seq data revealed fewer differentially translated transcripts between conditions upon H_2O_2 treatment, an order of magnitude smaller (**Figure 34**). All our results, then, are in agreement with Schosserer et al., described phenotype in yeast. Loss of NSUN5 through an epigenetic mechanism in glioma cells leads to an overall depletion of protein synthesis, but at the same time, it enhances the selective synthesis of stress-response mRNAs. Even in the absence of stress, loss of NSUN5, and therefore, lack of m^5C3782-modified ribosomes drives the pre-activation of a translational stress-response program for a rapid modulation of cytoplasmic translation upon stress. As the exchange of modified rRNAs to unmodified ones is a slow process, requiring turnover of ribosomes, it is reasonable that NSUN5 activity could be modulated by long-term chronic stress, what would explain how its epigenetic loss is more frequently found in LGG patients (**Figure 10**), as it would help the tumorigenic process from its early stages. Tumors at initial stages remain avascular, with associated consequences ranging from low oxygen availability (hypoxic stress) to nutrient starvation [243,244], and such unfavorable

microenvironments may exert a selective pressure that favors the survival of stress-reluctant NSUN5-silent cells.

To deepen our knowledge into the specific translational program driven by NSUN5 epigenetic silencing, we next carried out a proteomic approach to assess differentially expressed proteins between epigenetically-silent LN229 cells, and those same cells restored of NSUN5 expression. We detected 128 proteins differentially expressed between conditions, and a gene functional annotation analysis showed pathways and gene ontologies related to response to stress, mismatch repair and RNA processing (**Figure 33**). Importantly, a closer look into the proteins downregulated upon NSUN5 restoration (meaning they are greater expressed in the absence of the methyltransferase) revealed that 26% of them were deriving from transcripts undergoing higher translational efficiency upon NSUN5 loss, according to the Ribo-seq results (**Figures 32, 33**). These results, thus, confirmed the 'activated' cell status driven by NSUN5 silencing in glioma, allowing cells to better respond to stress environments.

Nonetheless, it is worth mentioning that, although the discovery that the rRNA was the ribosome component responsible for catalyzing the peptide bond formation relegated RPs to a scaffolding role, recent evidence suggests that almost half of RPs are essential for ribosomal function, and some even required for its catalytic activity [245–248]. Therefore, our discovery that NSUN5 may have a wider target spectrum and could be mediating m^5C modification of a number of RP mRNAs from the LSU (**Figure 26**) should be considered in line with our results. In an attempt to validate the putative, never described, NSUN5 mRNA targets, we were only able to confirm NSUN5 role in methylating the m^5C524 of *RPL41* mRNA. Importantly, we showed higher *RPL41* mRNA stability in NSUN5-expressing C524-methylated cells compared to NSUN5-lacking cells (**Figure 27**). Hence, if NSUN5 epigenetic loss could be impairing some RPs expression in a m^5C-dependent manner, we cannot rule out its implication in the phenotype observed. Strikingly, eL41, the yeast RPL41 orthologue, has been described to directly contact the Rcm1-mediated m^5C2278 modification in the LSU, and to form a bridge between the LSU and the SSU decoding center, acting as an axis for SSU rotation during translation [249,250]. This mechanism would imply a rRNA modification checkpoint in order to maintain ribosomal translational efficiency. If conserved in humans, it may be an explanation of the observed unaltered translation upon NSUN5 loss under basal conditions: impaired RPL41 expression through loss of m^5C524 of *RPL41* mRNA could bypass the lack of m^5C3782 28S rRNA methylation. However, substantial experimental work should be performed in order to rise further conclusions.

Despite the new avenues of research opened up, we next decided to keep interrogating our described stress-response program, in this case for potential drug target candidates. An illustrative example of the translational program in NSUN5-silenced cells is provided by the NAD(P)H quinone oxidoreductase 1 (NQO1), a multitarget antioxidant enzyme. While *NQO1* mRNA level is unchanged regardless NSUN5 expression, its translational efficiency is enhanced upon loss of the methyltransferase, resulting in a higher NQO1 protein expression in NSUN5-depleted cells (**Figure 36**). Validation of NSUN5-hypermethylation-mediated NQO1 overexpression represented a milestone in our study, as it opened up a new therapeutic opportunity for glioma management. NQO1 is a 2-electron reductase in charge of detoxification of a variety of xenobiotics, including quinones, quinoneimines, nitroaromatics and azo dyes [216]. Its expression is regulated through the Keap1/Nrf2/ARE pathway, one of the major signaling cascades for stress response [251]. NQO1 has been found overexpressed in several solid tumors, including breast, lung, colon and pancreatic cancers, while its expression is normally quite low in normal tissues [217], fact that made the enzyme a promising target for selective anticancer strategies. However, different NQO1 inhibitors failed to exert cancer cell growth inhibition [252]. Nevertheless, another targeting strategy is showing promising results: to exploit NQO1 catalytic function to induce a ROS-mediated cancer cell death. The 2-electron reduction of quinones carried out by NQO1 is typically considered to detoxify them, as hydroquinones are stable and can be eliminated through conjugation to glutathione or glucuronic acid. However, some hydroquinones are not stable and can rapidly be re-oxidized by molecular oxygen, giving rise to the generation of superoxide. To date, three NQO1 bioactivatable substrates have been described to generate large ROS quantities upon reduction: streptonigrin, β-lapachone and deoxynyboquinone (DNQ) [217], being the last one the most efficient NQO1 substrate. In fact, DNQ has been described to potently induce cancer cell death *in vitro*, mediated by the generation of > 60 mol of ROS per mole of drug each minute, and validated through the fact that its cytotoxic action is strongly reduced by cotreatment with antioxidants or by treatment in a reduced oxygen atmosphere [253].

Luckily, we were kindly provided of DNQ by Dr. Paul Hergenrother (University of Illinois, USA), as well as of a DNQ derivative developed by his laboratory, isobutyl-DNQ (IB-DNQ), with increased solubility and superior pharmacokinetics for *in vivo* experiments [254,255]. First, we evaluated DNQ and IB-DNQ growth inhibitory effects in an *in vitro* setting. As expected, NSUN5-silenced cells, with higher NQO1 protein levels, were significantly more sensitive to the inhibitory growth effect of NQO1 bioactivatable substrates than NSUN5 endogenously or ectopically-expressing cell lines (**Figure 37**). Importantly, we

were able to validate our findings in an *in vivo* setting, showing how mice harboring NSUN5-depleted subcutaneous tumors displayed an increased OS when treated with IB-DNQ, compared to the mock group, while no survival differences were observed between mock and treated mice with tumors originating from a NSUN5-expressing glioma cell line, with minimal expression of NQO1 **(Figure 37**). Therefore, our results described how NSUN5-promoter hypermethylation pinpoints glioma tumors that could benefit from the treatment with NQO1 bioactivated small molecules.

Finally, we wondered if NSUN5 promoter methylation may show any translational implication for glioma patients. Analysis of both, the TCGA glioma cohort (n = 497) and a newly obtained glioma validation cohort (n = 115) revealed how NSUN5 methylation, and its associated transcript silencing, are associated with better patient outcome (**Figures 38, 39**). Moreover, multivariate Cox regression analysis proved NSUN5-promoter hypermethylation to be an independent prognostic biomarker, as valuable as the currently validated molecular markers for glioma management, already explained (*IDH* mutational status, 1p/19q codeletion and MGMT promoter methylation) (**Figure 40**).

The apparent contradiction between the observed *in vitro* phenotype, assigning a tumor suppressor role for the RNA methyltransferase in glioma cells, and the clinical data, suggesting longer survival of those glioma patients with silenced NSUN5 expression, resembles another feature of glioma tumors: the occurrence of *IDH* mutations. As already disclosed, *IDH* mutations are especially prevalent in human glioma, observed in more than 80% of WHO grade II/III cases [222]. Aside CNS malignancies, *IDH* mutations are also found in 16% of acute myeloid leukemia cases [256], 23% of intrahepatic cholangiocarcinomas [257], and more than half chondrosarcomas [258]. Cancer-associated *IDH* mutations are heterozygous, and remarkably specific to codons encoding the arginine residue responsible for substrate recognition (R132 in IDH1, and R140 or R172 in IDH2) [259]. IDH enzymes form homodimers and carry out the oxidative decarboxylation of isocitrate to α-ketoglutarate (α-KG), reducing $NADP^+$ to NADPH [260]. However, heterozygous *IDH1* mutations results in tumoral IDH1 heterodimers, with a wild-type monomer converting isocitrate to α-KG and producing NADPH, whereas the mutant component, with a decreased affinity for isocitrate, exhibits a neomorphic activity converting α-KG into D-2-hydroxyglutarate (D-2-HG) in an NADPH-dependent manner [261]. This neomorphic activity of IDH1 mutant enzymes has been shown to give rise to a vast metabolic reprogramming in cancer cells: It depletes the Krebs cycle by draining α-KG through production of D-2-HG. To compensate for the metabolic impact, non-Krebs-cycle sources of carbohydrates such as glutamine and/or glutamate are recruited to

maintain metabolic homeostasis [262,263]. In addition, IDH mutant consumption of NADPH for the generation of D-2-HG impairs *de novo* lipogenesis, turning IDH mutant cells dependent on exogenous lipid sources [264]. Furthermore, lactate dehydrogenase A (LDHA), a critical enzyme for the Warburg phenotype is commonly found downregulated by promoter hypermethylation in IDH-mutant gliomas [265]. IDH-mutant glioma tumors, thus, instead of displaying the prevalent cancer-associated increased glycolysis to sustain high proliferation rates, exhibit a limited glycolytic capacity. This unique metabolic pattern, compared to other solid tumors, explains the slow-growing nature of IDH mutant gliomas, and therefore, the better prognosis of the patients.

NSUN5-promoter hypermethylation might constitute another example of the particular biological nature of glioma tumors, as well. According to our results, the epigenetic loss of NSUN5 leads to a hypomethylation event at a specific position of the 28S rRNA, at the large ribosomal subunit, that depletes global protein synthesis but prearranges cells to cope with cellular stress through the specific translation of stress-related transcripts. Cancer cells normally exhibit enhanced protein synthesis rates to sustain, as already mentioned, their oncogenic proliferation, typically through gene amplification of translation initiation factors (EIF) or oncogenic alterations affecting MYC, MAPK or PI3K-AKT-mTOR signaling pathways, which converge in activation of translation [266,267]. Nonetheless, protein synthesis is a double-edged sword that can, eventually, turn around them. Exacerbated protein synthesis rates increase the likelihood of generation of misfolded or damaged proteins that will accumulate and aggregate within the endoplasmic reticulum (ER) and even in the cytosol. The consequent proteotoxic stress can activate the unfolded protein response (UPR) that might end up triggering cell death if homeostasis is not achieved [268,269]. Therefore, and especially in the context of facing unfriendly environments, such as those characterized by oxidative, hypoxic, genotoxic, proteotoxic, or metabolic stress, a strategy followed by cancer cells is to shut down overall protein synthesis, blocking this way a highly energy-consuming process and avoiding UPR-triggered apoptosis. But at the same time, the translation of specific mRNAs, destined to cope with such unfriendly conditions is enhanced. The fact that the translation process, and therefore, the cellular proteome could adapt to environmental cues is in great part achieved through ribosome heterogeneity. The classical view of the ribosome as a static ‘molecular machine’ is currently outdated, and there is mounting evidence of heterogeneous ribosome populations able to target specific mRNAs to fine-tune translation in a tissue-specific, context-specific, and even within the same cell, in a spatially-dependent manner, what has been called ‘specialized ribosomes’ [231,250,270]. Ribosome heterogeneity can be achieved through variant rRNA alleles, incorporation of

RP paralogs or changes in RP stoichiometry, post-translational RP modifications, and, as the example provided by this thesis, altered or substoichiometric rRNA modifications [231]. Several further examples of specialized ribosomes fine-tuning translation under stress scenarios have been found in all domains of life. For instance, nutrient-deprived conditions induce the expression of the specific 16S rRNA variant rrnH in *E. coli*, and rrnH-containing ribosomes promote the translation of RpoS, a master regulator of the stress-response cascade [271]. In *C. elegans*, loss of Metl-5, and its mediated m^6A1717 on the 18S rRNA, promotes the specific translation of cyp-29A3, rendering knockout worms more resistant to a variety of stress conditions than their Metl-5 wild-type counterparts [272]. Likewise, exposure of yeast cells to osmotic stress induces RP paralog switching [273], while other study shows how it also increases the abundance of Rps26-deficient ribosomes, that selectively translate stress-related transcripts [274]. Another report by Shi and colleagues showed how heterogenous ribosomes with different RP composition preferentially translate specific pools of mRNAs in mouse embryonic stem cells (mESCs) [275]. Importantly, one of the substoichiometrically expressed RP was Rpl10, whose transcript is putatively m^5C-modified by NSUN5 in glioma cells, according to our results (**Figure 26**). Hence, a potential explanation of how C3782-hypomodified ribosomes upon NSUN5 loss specifically promote the translation of stress-related transcripts could be their substoichiometric expression of some RPs, due to lack of NSUN5-mediated m^5C-methylation of their transcripts. If validated, it would postulate NSUN5 as a major regulator of the cellular translational machinery and would explain its conserved presence from unicellular to highly complex superior eukaryotes.

To conclude, it is imperative to keep in mind how, nowadays, the prognosis for advanced forms of glioma tumors is extremely poor, and therapeutic options beyond surgical resection and first-line chemotherapy are limited [211]. Despite the substantial progress made in the last decade to uncover the molecular basis of brain tumors, it has not been materialized into effective therapies. Some of the reasons are the poor blood-brain barrier drug permeability, the redundancy of intracellular signaling pathways or the lack of targetable biomarkers [221]. Therefore, the identification of NSUN5 DNA methylation-associated silencing in human gliomagenesis, if validated in larger prospective studies, may serve for the identification of those glioma patients who, in the context of a devastating disease, are likely to have a good clinical outcome. Besides, patients harboring NSUN5 hypermethylation might benefit from an alternative therapeutic approach, consisting of a targeted therapy with NQO1 bioactivated small molecules.

STUDY II: Epigenetic loss of the RNA methyltransferase NSUN7 in Hepatocellular Carcinoma (HCC)

The screening of 450K array-derived DNA methylation data comprising the promoter regions of 11 m^5C methyltransferase genes from around 1,000 cancer cell lines allowed the identification of two candidate enzymes that could potentially exert tumor suppressive roles. Besides NSUN5, another member of the NSUN family was, as well, hypermethylated in a number of cancer cell lines. A closer look into NSUN7 promoter hypermethylation revealed its enrichment in melanoma, HCC and haematological cancer cell lines (**Figure 41**). We then decided to direct the focus of our study towards HCC, as the only published paper concerning NSUN7 molecular function pointed out to a role in fine-tuning energy metabolism in liver cells [153].

The observed robust negative correlation between NSUN7 promoter methylation status, and its transcript levels in TCGA and HEPTROMIC HCC tumor samples (**Figures 42, 43**), on the one hand, and NSUN7 expression recovery in HCC cancer cell lines upon DNA demethylation treatment on the other (**Figure 45**), demonstrated the functional role of promoter methylation in mediating NSUN7 silencing in HCC.

HCC is the most common type of primary liver tumor, accounting for almost 90% of all cases [276]. Liver cancer remains a global health threat, as its incidence continues increasing, especially in some parts of Europe, and the USA [277]. To date, liver cancer is considered the sixth most frequent cancer type, and the fourth leading cause of cancer-related death worldwide [278], but if this trend is maintained it is proposed to reach the third death-causing position by 2030 [279]. The reason lies in HCC-associated risk factors. The main risk factors for development of HCC tumors are the infection by hepatitis B (HBV) and C virus (HCV) [280,281], that explains why the highest incidence and mortality of this cancer type are described in Africa and East Asia [277], where there are not universal HBV vaccination programs implemented or surveillance for HCV infection [282,283]. Nonetheless, as mentioned, HCC incidence is becoming more frequent in the West, owing to an increased alcohol consumption, to the higher prevalence of obesity-driven non-alcoholic steatohepatitis (NASH), and to a NASH associated with metabolic syndrome or diabetes mellitus [284,285].

Due to the asymptomatic disease progression, and the lack of early diagnostic biomarkers, only 40% to 50% of patients with HCC are diagnosed at early stages, when

ablative therapies, such as liver resection or transplantation, are feasible [286,287]. Yet, up to 40% of patients are diagnosed at an advanced stage, when are no longer eligible for ablative treatments, and depend on systemic therapies [287]. At present, there are only six systemic therapies approved based on phase III trials (sorafenib, lenvatinib, combination of atezolizumab and bevacizumab, regorafenib, cabozantinib and ramucirumab) [278]. However, multikinase inhibitors such as sorafenib and lenvatinib have proven to extend patient overall survival by less than one year [286], what makes imperative to continue investigating new avenues of treatment. The main obstacle remains HCC biology itself, as the principal mutational drivers promoting tumorigenesis (*TERT*, *TP53*, *CTNNB1*) are so far undruggable [286,288]. In fact, only around 25% of HCC tumors display targetable mutations with available anticancer therapies, and the majority of those, show a prevalence lower than 10% [289]. Hence, further understanding of HCC tumor biology is needed in order to improve its clinical management. While epigenetic research has already proven its utility when translated into clinics [38,290,291], the progressive understanding of epitranscriptomic mechanisms, RNA modifications and their modifiers may, as well, eventually become valuable for clinical purposes. Therefore, our identification of an epitranscriptomic lesion, highly prevalent in HCC tumors, may contribute to any of the unmet clinical needs.

Consequently, we decided to address NSUN7 putative function, hypothesizing it would be the cause of its silencing in HCC tumors. In mice, Nsun7 has been described to colocalize with PGC-1α at enhancer regions of the genome, where it mediates m^5C modification of eRNAs, which affects the expression of their cognate genes. Some of the genes whose expression is modulated through Nsun7-mediated eRNA-methylation are Pfkl, Sirt5, Idh3b, and Hmox2 [153]. In our initial attempt to decipher NSUN7 role in human cells, we assessed if those genomic distal regulatory regions encoding Pfkl, Sirt5, Idh3b, and Hmox2 eRNAs were conserved between mice and human, and observed lack of conservation (Data not shown). Accordingly, an *in silico* analysis of PFKL, SIRT5, IDH3B and HMOX2 expression in the TCGA and HEPTROMIC cohorts revealed no differences in transcript expression regardless both, NSUN7-promoter methylation status and *NSUN7* transcript expression (Data not shown). These results point out that Nsun7 role in mediating m^5C methylation of eRNAs might not be conserved in humans, or at least not of those specific target genes.

Hence, in order to uncover NSUN7's target spectrum we carried out transcriptome-wide bisulfite sequencing (bsRNA-seq) in both loss- and gain-of-function models, as we did for NSUN5. However, while with NSUN5 overexpression we observed a clear increase

in m^5C sites across the transcriptome, with NSUN7 manipulation we did not detect such 'mass effect'. Some reasons explaining our modest results could be that NSUN7 does not really target many RNAs that are detectable at a moderate sequencing depth, or that the majority of NSUN7 target sites occur at low stoichiometry, being their detection hampered by technical limitations. Even so, we were still able to filter out around 39 mRNA sites that were differentially methylated depending on NSUN7 presence (absolute methylation fold-change > 0.2). Despite unsatisfactory, our results go in line with recent top-down studies aiming to characterize m^5C RNA writers through unbiased transcriptome-wide methods [135,136]. These studies reported that almost all m^5C sites in mRNA are deposited either by NSUN2 or NSUN6. Likewise, we cannot rule out NSUN7 role in methylating eRNAs, as not many of these molecules were detectable at the sequencing depth employed.

Next off, we experimentally validated one of the previously disclosed mRNAs as a NSUN7 target, the *CCDC9B* transcript. Through RNA bisulfite sequencing we demonstrated how HCC lines endogenously or ectopically expressing NSUN7 displayed m^5C1322-1324 status, how NSUN7-depleted cell lines exhibited lack of methylation at the three positions, and importantly, that restoration of a catalytically-dead NSUN7 version was unable to recover methylation at the three cytosines (**Figures 48, 49**). Moreover, we observed how loss of m^5C modifications reduced *CCDC9B* transcript half-life (**Figures 50**). The subsequent step after validating a human NSUN7 target was to wonder how loss of *CCDC9B* m^5C-methylation and consequent reduced CCDC9B expression through NSUN7 epigenetic loss could be of any profit for HCC tumors. However, a literature review revealed that little is known about CCDC9B (also known as C15ORF52). Nevertheless, a study published in 2012 by Castello et al., establishing the HeLa mRNA interactome by RNA interactome capture (RIC), identified CCDC9B as a novel RNA binding protein (RBP) that specifically binds mRNAs [292]. Moreover, a posterior work by Hein and colleagues, determining the HeLa protein interactome in three quantitative dimensions, suggested CCDC9B physical interaction with THOC1 and THOC7 [293]. These proteins are members of the transcription-and-export (TREX) complex, in charge of recognizing mature mRNAs at the nucleus, and exporting them to the cytoplasm [294,295]. The TREX complex can be found across all eukaryotes, and it is composed by the multi-subunit THO complex (including THOC1, -2, -3, -5, -6 and -7), the RNA helicase DDX39B, and an RNA export adapter, oftentimes constituted by the m^5C reader protein ALYREF [296]. Additional TREX interactors have been described under different cellular contexts [297], influencing the core machinery and promoting the export of specific transcripts [298,299]. All the previous evidence, hence, may reveal CCDC9B as a

novel mRNA export factor through interaction with TREX components. And what is more, NSUN7-mediated CCDC9B altered expression in HCC cells may be affecting the translation of specific mRNAs through impairment of their nuclear export. In order to unveil CCDC9B role in HCC we are currently carrying out two high-throughput approaches: CCDC9B-IP coupled to MS, to identify the interactome of the putative RBP, and individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP), to uncover RNA binding targets of CCDC9B. Results of both experiments will hopefully shed light into NSUN7 role in human cells, and why its epigenetic deregulation is frequently observed in cancer.

Additionally, although not cancer-related, it is worth mentioning our preliminary insights into NSUN7 potential novel biological roles. Interestingly, mining our RNA-seq data we observed that upon NSUN7 knockout, two transcripts coding for flagellar plasma membrane channels (*CATSPERG* and *ATP2B4*) were dramatically depleted. As already stated in the Introduction section, several studies have described that NSUN7 point mutations are associated with sperm motility defects both in mice and human, and are found in infertile men [149–152]. Sperm movement is activated by changes in intracellular ion concentration, achieved by an increase in pH, calcium ion and cAMP. The sperm membrane specific, calcium-selective ion channel CatSper is located in the sperm flagellum and is required for the hyperactivation of sperm motility in the oviduct. CatSper is composed by four main α subunits and three auxiliary subunits, being CatSperγ (*CATSPERG*) one of the latest. Importantly, knockout experiments in mice have suggested that all CatSper subunits are required for proper channel assembly, with the absence of a single subunit leading to the degradation of the remaining CatSper proteins [300]. Moreover, the Plasma membrane Ca^{2+}-ATPase PMCA4 (*ATP2B4*), also found in the principal piece of the sperm flagellum, has, as well, being described to be essential for proper hyperactivation of sperm flagellum movement [301]. Therefore, loss of both flagellar channels upon NSUN7 loss might represent the molecular basis behind the infertility phenotype observed in NSUN7-mutation-bearing men. In fact, some of the characterized mutations result in an impairment of NSUN7 protein function, such as the Ste5Jcs1 mutation, that generates a premature stop codon leading to a truncated catalytic domain [149], or the C26232T-transition and T26248G-transversion mutations that affect protein folding and ligand binding sites [151]. Such evidence would suggest that NSUN7 catalytic activity is needed for proper sperm motility. Further experiments are then required in order to link NSUN7 RNA methyltransferase function to CATSPERG and PMC4 expression [151]. So far, we can exclude both *CATSPERG* and *ATP2B4* transcripts as direct NSUN7 methylation targets, as bsRNA-seq results did not prove the existence of

differentially methylated cytosines in those transcripts, based on NSUN7 expression. However, we cannot rule out an indirect effect over CATSPERG and PMC4 expression mediated by a NSUN7-methylated target.

CONCLUSIONS

CONCLUSIONS

STUDY I: Insights into the epigenetic loss of the m^5C RNA methyltransferase NSUN5, and its contribution to gliomagenesis

1. NSUN5 expression is epigenetically regulated through CpG-island promoter methylation in human cancer.
2. Human glioma cells undergo promoter hypermethylation-mediated silencing of NSUN5.
3. NSUN5 shows tumor suppressor properties both *in vitro* and *in vivo.*
4. NSUN5 is a m^5C RNA methyltransferase whose loss generates an unmethylated status at the C3782 position of human 28S rRNA.
5. NSUN5 epigenetic loss drives an overall depletion of protein synthesis, along with the emergence of an adaptive translational program for survival under conditions of cellular stress.
6. NSUN5 epigenetic inactivation promotes NQO1 overexpression, which in turn renders gliomas sensitive to NQO1-targeted substrates.
7. NSUN5 epigenetic inactivation is a hallmark of glioma patients with longer overall survival.

This work was published in:

Janin, M. & Ortiz-Barahona, V. *et al*. Epigenetic loss of RNA-methyltransferase NSUN5 in glioma targets ribosomes to drive a stress adaptive translational program. *Acta Neuropathologica* **138**, 1053-1074 (2019).

For all the supplemental information related to this study and not included in this thesis, please refer to the paper.

STUDY II: Epigenetic loss of the RNA methyltransferase NSUN7 in Hepatocellular Carcinoma (HCC)

1. NSUN7 expression is epigenetically regulated through CpG-island promoter methylation in human cancer.

2. Human HCC cells undergo promoter hypermethylation-mediated silencing of NSUN7.

3. NSUN7 is a m^5C RNA methyltransferase responsible for catalyzing the incorporation of a methyl group onto the 1322, 1323 and 1324 cytosines of the *CCDC9B* mRNA.

4. NSUN7 epigenetic loss drives a dramatic reduction of the flagellar plasma membrane channels CATSPERG and PMCA4, with described roles in sperm motility.

REFERENCES

REFERENCES

1. Alberts, B. *Molecular Biology of the Cell*. (Garland Science, 2014).
2. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
3. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
4. Waddington, C. H. Preliminary Notes on the Development of the Wings in Normal and Mutant Strains of Drosophila. *Proceedings of the National Academy of Sciences of the United States of America* **25**, 299 (1939).
5. Holliday, R. The inheritance of epigenetic defects. *Science* **238**, 163–170 (1987).
6. Zhang, G. & Pradhan, S. Mammalian epigenetic mechanisms. *IUBMB Life* **66**, 240–256 (2014).
7. Wyatt, G. R. Occurrence of 5-methylcytosine in nucleic acids. *Nature* **166**, 237–238 (1950).
8. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
9. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
10. Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H. & Tajima, S. DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a and Dnmt3b through a Direct Interaction. *Journal of Biological Chemistry* **279**, 27816–27823 (2004).
11. Song, J., Teplova, M., Ishibe-Murakami, S. & Patel, D. J. Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. *Science* **335**, 709–712 (2012).
12. Franks, J. L. *et al.* In silico APC/C substrate discovery reveals cell cycle-dependent degradation of UHRF1 and other chromatin regulators. *PLOS Biology* **18**, e3000975 (2020).
13. Wu, X. & Zhang, Y. TET-mediated active DNA demethylation: Mechanism, function and beyond. *Nature Reviews Genetics* **18**, 517–534 (2017).
14. Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes and Development* **25**, 1010–1022 (2011).
15. Klose, R. J. & Bird, A. P. Genomic DNA methylation: the mark and its mediators. *Trends in biochemical sciences* **31**, 89–97 (2006).

16. Ferguson-Smith, A. C. Genomic imprinting: The emergence of an epigenetic paradigm. *Nature Reviews Genetics* **12**, 565–575 (2011).
17. Payer, B. & Lee, J. T. X chromosome dosage compensation: How mammals keep the balance. *Annual Review of Genetics* **42**, 733–772 (2008).
18. Mohn, F. *et al.* Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Molecular Cell* **30**, 755–766 (2008).
19. Gopalakrishnan, S., Sullivan, B. A., Trazzi, S., della Valle, G. & Robertson, K. D. DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. *Human Molecular Genetics* **18**, 3178–3193 (2009).
20. Zhang, G. *et al.* N6-Methyladenine DNA Modification in Drosophila. *Cell* **161**, 893–906 (2015).
21. Greer, E. L. *et al.* DNA Methylation on N6-Adenine in C. elegans. *Cell* **161**, 868–878 (2015).
22. Fu, Y. *et al.* N6-methyldeoxyadenosine marks active transcription start sites in Chlamydomonas. *Cell* **161**, 879–892 (2015).
23. Arber, W. & Dussoix, D. Host specificity of DNA produced by Escherichia coli: I. Host controlled modification of bacteriophage λ. *Journal of Molecular Biology* **5**, 18–36 (1962).
24. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
25. Petty, E. & Pillus, L. Balancing chromatin remodeling and histone modifications in transcription. *Trends in Genetics* **29**, 621–629 (2013).
26. Simó-Riudalbas, L. & Esteller, M. Targeting the histone orthography of cancer: drugs for writers, erasers and readers. *British journal of pharmacology* **172**, 2716–2732 (2015).
27. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* **21**, 381–395 (2011).
28. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Experimental & Molecular Medicine* **49**, e324 (2017).
29. Feinberg, A. P. & Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**, 89–92 (1983).
30. Karpf, A. R. & Matsui, S. I. Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. *Cancer Research* **65**, 8635–8639 (2005).

31. Soejima, H. & Higashimoto, K. Epigenetic and genetic alterations of the imprinting disorder Beckwith-Wiedemann syndrome and related disorders. *Journal of Human Genetics* **58**, 402–409 (2013).
32. Jirtle, R. L. IGF2 Loss of Imprinting: A Potential Heritable Risk Factor for Colorectal Cancer. *Gastroenterology* **126**, 1190–1193 (2004).
33. Bjornsson, H. T. *et al.* Epigenetic specificity of loss of imprinting of the IGF2 gene in wilms tumors. *Journal of the National Cancer Institute* **99**, 1270–1273 (2007).
34. Lee, E. *et al.* Landscape of somatic retrotransposition in human cancers. *Science* **337**, 967–971 (2012).
35. Fu, X. *et al.* Distinct expression patterns of hedgehog ligands between cultured and primary colorectal cancers are associated with aberrant methylation of their promoters. *Molecular and cellular biochemistry* **337**, 185–192 (2010).
36. Vizoso, M. *et al.* Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. *Nature Medicine* **21**, 741–750 (2015).
37. Baylin, S. B. & Jones, P. A. Epigenetic determinants of cancer. *Cold Spring Harbor Perspectives in Biology* **8**, a019505 (2016).
38. Ortiz-Barahona, V., Joshi, R. S. & Esteller, M. Use of DNA methylation profiling in translational oncology. *Seminars in cancer biology* **20**, 30271–30276 (2020).
39. Audia, J. E. & Campbell, R. M. Histone Modifications and Cancer. *Cold Spring Harbor Perspectives in Biology* **8**, (2016).
40. Peng, Y. & Croce, C. M. The role of MicroRNAs in human cancer. *Signal Transduction and Targeted Therapy* **1**, 1–9 (2016).
41. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature reviews. Genetics* **13**, 227–232 (2012).
42. Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Research* **46**, D303–D307 (2018).
43. Davis, F. F. & Allen, F. W. Ribonucleic acids from yeast which contain a fifth nucleotide. *The Journal of biological chemistry* **227**, 907–915 (1957).
44. Cohn, W. E. Pseudouridine, a Carbon-Carbon Linked Ribonucleoside in Ribonucleic Acids: Isolation, Structure, and Chemical Characteristics. *The Journal of Biological Quemistry* **235**, 1488–1498 (1960).
45. Edmons, M. & Abrams, R. Polynucleotide biosynthesis: formation of a sequence of adenylate units from adenosine triphosphate by an enzyme from thymus nuclei. *The Journal of biological chemistry* **235**, 1142–1149 (1960).

46. Furuichi, Y., Muthukrishnan, S. & Shatkin, A. J. 5'-Terminal m-7G(5')ppp(5')G-m-p in vivo: identification in reovirus genome RNA. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 742–745 (1975).
47. Perry, R. P., Kelley, D. E., Friderici, K. & Rottman, F. The methylated constituents of L cell messenger RNA: evidence for an unusual cluster at the 5' terminus. *Cell* **4**, 387–394 (1975).
48. Jia, G. *et al.* N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature chemical biology* **7**, 885–887 (2011).
49. Saletore, Y. *et al.* The birth of the Epitranscriptome: deciphering the function of RNA modifications. *Genome Biology 2012 13:10* **13**, 1–12 (2012).
50. Novoa, E. M., Mason, C. E. & Mattick, J. S. Charting the unknown epitranscriptome. *Nature Reviews Molecular Cell Biology 2017 18:6* **18**, 339–340 (2017).
51. Helm, M. & Motorin, Y. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nature Reviews Genetics* **18**, 275–291 (2017).
52. Grozhik, A. V. *et al.* Antibody cross-reactivity accounts for widespread appearance of m1A in 5'UTRs. *Nature communications* **10**, 5126 (2019).
53. Borchardt, E. K., Martinez, N. M. & Gilbert, W. V. Regulation and Function of RNA Pseudouridylation in Human Cells. *Annual review of genetics* **54**, 309–336 (2020).
54. Rintala-Dempsey, A. C. & Kothe, U. Eukaryotic stand-alone pseudouridine synthases - RNA modifying enzymes and emerging regulators of gene expression? *RNA biology* **14**, 1185–1196 (2017).
55. Penzo, M. & Montanaro, L. Turning Uridines around: Role of rRNA Pseudouridylation in Ribosome Biogenesis and Ribosomal Function. *Biomolecules* **8**, 38 (2018).
56. Spenkuch, F., Motorin, Y. & Helm, M. Pseudouridine: still mysterious, but never a fake (uridine)! *RNA biology* **11**, 1540–1554 (2014).
57. Schwartz, S. *et al.* Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148–162 (2014).
58. Karijolich, J. & Yu, Y. T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* **474**, 395–399 (2011).
59. Ayadi, L., Galvanin, A., Pichot, F., Marchand, V. & Motorin, Y. RNA ribose methylation (2′-O-methylation): Occurrence, biosynthesis and biological functions. *Biochimica et Biophysica Acta* **1862**, 253–269 (2018).
60. Nichols, J. L. & Lane, B. G. N-4-methyl-2'-O-methyl cytidine and other methyl-substituted nucleoside constituents of Escherichia coli ribosomal and soluble RNA. *Biochimica et biophysica acta* **119**, 649–651 (1966).

61. Motorin, Y. & Marchand, V. Detection and Analysis of RNA Ribose 2′-O-Methylations: Challenges and Solutions. *Genes* **9**, 642 (2018).
62. Lin, J. *et al.* Structural basis for site-specific ribose methylation by box C/D RNA protein complexes. *Nature* **469**, 559–563 (2011).
63. Dai, Q. *et al.* Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. *Nature Methods* **14**, 695–698 (2017).
64. Ji, L. & Chen, X. Regulation of small RNA stability: methylation and beyond. *Cell Research* **22**, 624–636 (2012).
65. Monaco, P. lo, Marcel, V., Diaz, J.-J. & Catez, F. 2′-O-Methylation of Ribosomal RNA: Towards an Epitranscriptomic Control of Translation? *Biomolecules* **8**, (2018).
66. Höfler, S. & Carlomagno, T. Structural and functional roles of 2'-O-ribose methylations and their enzymatic machinery across multiple classes of RNAs. *Current opinion in structural biology* **65**, 42–50 (2020).
67. Li, X., Xiong, X. & Yi, C. Epitranscriptome sequencing technologies: decoding RNA modifications. *Nature Methods* **14**, 23–31 (2016).
68. Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201–206 (2012).
69. Linder, B. *et al.* Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nature methods* **12**, 767–772 (2015).
70. Ke, S. *et al.* A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes & development* **29**, 2037–2053 (2015).
71. Meyer, K. D. *et al.* 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **163**, 999–1010 (2015).
72. Yang, Y., Hsu, P., Chen, Y. & Yang, Y. Dynamic transcriptomic m6A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Research* **28**, 616–624 (2018).
73. Warda, A. S. *et al.* Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO reports* **18**, 2004–2014 (2017).
74. Xiao, W. *et al.* Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Molecular cell* **61**, 507–519 (2016).
75. Roundtree, I. A. *et al.* YTHDC1 mediates nuclear export of N 6-methyladenosine methylated mRNAs. *eLife* **6**, (2017).
76. Shi, H. *et al.* YTHDF3 facilitates translation and decay of N 6-methyladenosine-modified RNA. *Cell research* **27**, 315–328 (2017).

77. Du, H. *et al.* YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. *Nature communications* **7**, (2016).
78. Alarcón, C. R. *et al.* HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events. *Cell* **162**, 1299–1308 (2015).
79. Dunn, D. B. The occurrence of 1-methyladenine in ribonucleic acid. *Biochimica et Biophysica Acta* **46**, 198–200 (1961).
80. Zhang, C. & Jia, G. Reversible RNA Modification N1-methyladenosine (m1A) in mRNA and tRNA. *Genomics, Proteomics & Bioinformatics* **16**, 155–161 (2018).
81. Oerum, S., Dégut, C., Barraud, P. & Tisné, C. m1A Post-Transcriptional Modification in tRNAs. *Biomolecules* **7**, 20 (2017).
82. Schwartz, S. m1A within cytoplasmic mRNAs at single nucleotide resolution: a reconciled transcriptome-wide map. *RNA* **24**, 1427–1436 (2018).
83. Safra, M. *et al.* The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature* **551**, 251–255 (2017).
84. Ozanick, S., Krecic, A., Andersland, J. & Anderson, J. T. The bipartite structure of the tRNA m1A58 methyltransferase from S. cerevisiae is conserved in humans. *RNA* **11**, 1281–1290 (2005).
85. Chujo, T. & Suzuki, T. Trmt61B is a methyltransferase responsible for 1-methyladenosine at position 58 of human mitochondrial tRNAs. *RNA* **18**, 2269–2276 (2012).
86. Liu, F. *et al.* ALKBH1-Mediated tRNA Demethylation Regulates Translation. *Cell* **167**, 816–828 (2016).
87. Li, X. *et al.* Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome. *Nature Chemical Biology 2016 12:5* **12**, 311–316 (2016).
88. Amos, H. & Korn, M. 5-Methyl cytosine in the RNA of Escherichia coli. *Biochimica et Biophysica Acta* **29**, 444–445 (1958).
89. Desrosiers, R., Friderici, K. & Rottman, F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 3971–3975 (1974).
90. Squires, J. E. *et al.* Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Research* **40**, 5023–5033 (2012).
91. Hussain, S., Aleksic, J., Blanco, S., Dietmann, S. & Frye, M. Characterizing 5-methylcytosine in the mammalian epitranscriptome. *Genome biology* **14**, (2013).

92. Huang, T., Chen, W., Liu, J., Gu, N. & Zhang, R. Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nature structural & molecular biology* **26**, 380–388 (2019).
93. Khoddami, V. *et al.* Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 6784–6789 (2019).
94. Helm, M. & Motorin, Y. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nature Reviews Genetics 2017 18:5* **18**, 275–291 (2017).
95. Edelheit, S., Schwartz, S., Mumbach, M. R., Wurtzel, O. & Sorek, R. Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. *PLoS genetics* **9**, (2013).
96. Khoddami, V. & Cairns, B. R. Transcriptome-wide target profiling of RNA cytosine methyltransferases using the mechanism-based enrichment procedure Aza-IP. *Nature protocols* **9**, 337–361 (2014).
97. Hussain, S. *et al.* NSun2-Mediated Cytosine-5 Methylation of Vault Noncoding RNA Determines Its Processing into Regulatory Small RNAs. *Cell Reports* **4**, 255–261 (2013).
98. Frommer, M. *et al.* A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 1827–1831 (1992).
99. Schaefer, M., Pollex, T., Hanna, K. & Lyko, F. RNA cytosine methylation analysis by bisulfite sequencing. *Nucleic Acids Research* **37**, (2009).
100. Garalde, D. R. *et al.* Highly parallel direct RNA sequencing on an array of nanopores. *Nature methods* **15**, 201–206 (2018).
101. Smith, A. M., Jain, M., Mulroney, L., Garalde, D. R. & Akeson, M. Reading canonical and modified nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing. *PLOS one* **14**, (2019).
102. Anreiter, I., Mir, Q., Simpson, J. T., Janga, S. C. & Soller, M. New Twists in Detecting mRNA Modification Dynamics. *Trends in biotechnology* **39**, 72–89 (2021).
103. Bohnsack, K. E., Höbartner, C. & Bohnsack, M. T. Eukaryotic 5-methylcytosine (m5C) RNA Methyltransferases: Mechanisms, Cellular Functions, and Links to Disease. *Genes* **10**, 102 (2019).

104. Fonagy, A. *et al.* Cell cycle regulated expression of nucleolar antigen P120 in normal and transformed human fibroblasts. *Journal of cellular physiology* **154**, 16–27 (1993).
105. Jhiang, S. M., Yaneva, M. & Busch, H. Expression of Human Proliferation-associated Nucleolar Antigen p120. *Cell Growth & Differentiation* **1**, 319–324 (1990).
106. Hong, J., Lee, J. H. & Chung, I. K. Telomerase activates transcription of cyclin D1 gene through an interaction with NOL1. *Journal of cell science* **129**, 1566–1579 (2016).
107. Sharma, S., Yang, J., Watzinger, P., Kötter, P. & Entian, K.-D. Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. *Nucleic Acids Research* **41**, 9062–9076 (2013).
108. Heissenberger, C. *et al.* The ribosomal RNA m5C methyltransferase NSUN-1 modulates healthspan and oogenesis in Caenorhabditis elegans. *eLife* **9**, (2020).
109. Hong, B., Wu, K., Brockenbrough, J. S., Wu, P. & Aris, J. P. Temperature sensitive nop2 alleles defective in synthesis of 25S rRNA and large ribosomal subunits in Saccharomyces cerevisiae. *Nucleic Acids Research* **29**, 2927–2937 (2001).
110. Bourgeois, G. *et al.* Eukaryotic rRNA Modification by Yeast 5-Methylcytosine-Methyltransferases and Human Proliferation-Associated Antigen p120. *PLOS one* **10**, e0133321 (2015).
111. Frye, M. & Watt, F. M. The RNA Methyltransferase Misu (NSun2) Mediates Myc-Induced Proliferation and Is Upregulated in Tumors. *Current Biology* **16**, 971–981 (2006).
112. Van Haute, L. *et al.* NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs. *Nucleic Acids Research* **47**, 8720–8733 (2019).
113. Shinoda, S. *et al.* Mammalian NSUN2 introduces 5-methylcytidines into mitochondrial tRNAs. *Nucleic Acids Research* **47**, 8734–8745 (2019).
114. Blanco, S. *et al.* Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. *The EMBO Journal* **33**, 2020–2039 (2014).
115. Sibbritt, T., Patel, H. R. & Preiss, T. Mapping and significance of the mRNA methylome. *Wiley Interdisciplinary Reviews: RNA* **4**, 397–422 (2013).
116. Sajini, A. A. *et al.* Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation. *Nature Communications 2019 10:1* **10**, 1–13 (2019).
117. Abbasi-Moheb, L. *et al.* Mutations in NSUN2 Cause Autosomal- Recessive Intellectual Disability. *The American Journal of Human Genetics* **90**, 847–855 (2012).

118. Martinez, F. J. *et al.* Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *Journal of Medical Genetics* **49**, 380–385 (2012).
119. Luo, Y., Feng, J., Xu, Q., Wang, W. & Wang, X. NSun2 Deficiency Protects Endothelium From Inflammation via mRNA Methylation of ICAM-1. *Circulation Research* **118**, 944–956 (2016).
120. Guo, G. *et al.* Disease Activity-Associated Alteration of mRNA m5 C Methylation in CD4+ T Cells of Systemic Lupus Erythematosus. *Frontiers in Cell and Developmental Biology* **0**, 430 (2020).
121. Nakano, S. *et al.* NSUN3 methylase initiates 5-formylcytidine biogenesis in human mitochondrial tRNAMet. *Nature Chemical Biology* **12**, 546–551 (2016).
122. van Haute, L. *et al.* Deficient methylation and formylation of mt-tRNAMet wobble cytosine in a patient carrying mutations in NSUN3. *Nature Communications* **7**, 1–10 (2016).
123. Haag, S. *et al.* NSUN3 and ABH1 modify the wobble position of mt-tRNAMet to expand codon recognition in mitochondrial translation. *The EMBO Journal* **35**, 2104–2119 (2016).
124. Cantara, W. A., Murphy, F. V., Demirci, H. & Agris, P. F. Expanded use of sense codons is regulated by modified cytidines in tRNA. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 10964–10969 (2013).
125. Paramasivam, A., Meena, A. K., Venkatapathi, C., Pitceathly, R. D. S. & Thangaraj, K. Novel Biallelic NSUN3 Variants Cause Early-Onset Mitochondrial Encephalomyopathy and Seizures. *Journal of Molecular Neuroscience* **70**, 1962–1965 (2020).
126. Metodiev, M. D. *et al.* NSUN4 Is a Dual Function Mitochondrial Protein Required for Both Methylation of 12S rRNA and Coordination of Mitoribosomal Assembly. *PLOS Genetics* **10**, e1004110 (2014).
127. Cámara, Y. *et al.* MTERF4 Regulates Translation by Targeting the Methyltransferase NSUN4 to the Mammalian Mitochondrial Ribosome. *Cell Metabolism* **13**, 527–539 (2011).
128. Schosserer, M. *et al.* Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. *Nature Communications* **6**, 1–17 (2015).
129. Doll, A. & Grzeschik, K.-H. Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome. *Cytogenetic and Genome Research* **95**, 20–27 (2001).

130. Zhang, T. *et al.* Cognitive deficits in mice lacking Nsun5, a cytosine-5 RNA methyltransferase, with impairment of oligodendrocyte precursor cells. *Glia* **67**, 688–702 (2019).
131. Chen, P., Zhang, T., Yuan, Z., Shen, B. & Chen, L. Expression of the RNA methyltransferase Nsun5 is essential for developing cerebral cortex. *Molecular Brain 2019 12:1* **12**, 1–14 (2019).
132. Yuan, Z., Chen, P., Zhang, T., Shen, B. & Chen, L. Agenesis and Hypomyelination of Corpus Callosum in Mice Lacking Nsun5, an RNA Methyltransferase. *Cells* **8**, 552 (2019).
133. Haag, S. *et al.* NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs. *RNA* **21**, 1532–1543 (2015).
134. Long, T. *et al.* Sequence-specific and Shape-selective RNA Recognition by the Human RNA 5-Methylcytosine Methyltransferase NSun6. *Journal of Biological Chemistry* **291**, 242993–24303 (2016).
135. Fang, L. *et al.* CIGAR-seq, a CRISPR/Cas-based method for unbiased screening of novel mRNA modification regulators. *Molecular Systems Biology* **16**, e10025 (2020).
136. Selmi, T. *et al.* Sequence- and structure-specific cytosine-5 mRNA methylation by NSUN6. *Nucleic Acids Research* **49**, 1006–1022 (2021).
137. Brzezicha, B. *et al.* Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic Acids Research* **34**, 6034–6043 (2006).
138. Auxilien, S., Guérineau, V., Szweykowska-Kulińska, Z. & Golinelli-Pimpaneau, B. The human tRNA m5C methyltransferase Misu is multisite-specific. *RNA biology* **9**, 1331–1338 (2012).
139. Khoddami, V. & Cairns, B. R. Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nature Biotechnology* **31**, 458–464 (2013).
140. Xing, J. *et al.* NSun2 Promotes Cell Growth via Elevating Cyclin-Dependent Kinase 1 Translation. *Molecular and Cellular Biology* **35**, 4043–4052 (2015).
141. Tang, H. *et al.* NSun2 delays replicative senescence by repressing p27 (KIP1) translation and elevating CDK1 translation. *Aging* **7**, 1143–1155 (2015).
142. Cai, X. *et al.* RNA methyltransferase NSUN2 promotes stress-induced HUVEC senescence. *Oncotarget* **7**, 19099–19110 (2016).
143. Li, Q. *et al.* NSUN2-Mediated m5C Methylation and METTL3/METTL14-Mediated m6A Methylation Cooperatively Enhance p21 Translation. *Journal of Cellular Biochemistry* **118**, 2587–2598 (2017).

144. Wang, N., Tang, H., Wang, X., Wang, W. & Feng, J. Homocysteine upregulates interleukin-17A expression via NSun2-mediated RNA methylation in T lymphocytes. *Biochemical and Biophysical Research Communications* **493**, 94–99 (2017).
145. Chen, X. *et al.* 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nature Cell Biology* **21**, 978–990 (2019).
146. Xu, X., Zhang, Y., Zhang, J. & Zhang, X. NSun2 promotes cell migration through methylating autotaxin mRNA. *Journal of Biological Chemistry* **295**, 18134–18147 (2020).
147. Jurkowski, T. P. *et al.* Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *RNA* **14**, 1663–1670 (2008).
148. Begik, O. *et al.* Integrative analyses of the RNA modification machinery reveal tissue- And cancer-specific signatures. *Genome Biology* **21**, (2020).
149. Harris, T., Marquez, B., Suarez, S. & Schimenti, J. Sperm Motility Defects and Infertility in Male Mice with a Mutation in Nsun7, a Member of the Sun Domain-Containing Family of Putative RNA Methyltransferases. *Biology of Reproduction* **77**, 376–382 (2007).
150. Ren, H., Zhong, R., Ding, X., Chen, Z. & Jing, Y. Investigation of polymorphisms in exon7 of the NSUN7 gene among Chinese Han men with asthenospermia. *Genetics and Molecular Research* **14**, 9261–9268 (2015).
151. Khosronezhad, N. *et al.* T26248G-transversion mutation in exon7 of the putative methyltransferase Nsun7 gene causes a change in protein folding associated with reduced sperm motility in asthenospermic men. *Reproduction, Fertility and Development* **27**, 471–480 (2015).
152. Khosronezhad, N., Hosseinzadeh Colagar, A. & Mortazavi, S. M. The Nsun7 (A11337)-deletion mutation, causes reduction of its protein rate and associated with sperm motility defect in infertile men. *Journal of Assisted Reproduction and Genetics* **32**, 807–815 (2015).
153. Aguilo, F. *et al.* Deposition of 5-Methylcytosine on Enhancer RNAs Enables the Coactivator Function of PGC-1α. *Cell Reports* **14**, 479–492 (2016).
154. Goll, M. G. *et al.* Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science* **311**, 395–398 (2006).
155. Schaefer, M. *et al.* RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes & Development* **24**, 1590–1595 (2010).
156. Shanmugam, R. *et al.* Cytosine methylation of tRNA-Asp by DNMT2 has a role in translation of proteins containing poly-Asp sequences. *Cell discovery* **1**, 15010 (2015).

157. Yang, X. *et al.* 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Research* **27**, 606–625 (2017).
158. Tran, H., Maurer, F. & Nagamine, Y. Stabilization of Urokinase and Urokinase Receptor mRNAs by HuR Is Linked to Its Cytoplasmic Accumulation Induced by Activated Mitogen-Activated Protein Kinase-Activated Protein Kinase 2. *Molecular and Cellular Biology* **23**, 7177–7188 (2003).
159. Doller, A. *et al.* Posttranslational Modification of the AU-Rich Element Binding Protein HuR by Protein Kinase Cδ Elicits Angiotensin II-Induced Stabilization and Nuclear Export of Cyclooxygenase 2 mRNA. *Molecular and Cellular Biology* **28**, 2608–2625 (2008).
160. Huber, S. M. *et al.* Formation and Abundance of 5-Hydroxymethylcytosine in RNA. *ChemBioChem* **16**, 752–755 (2015).
161. Fu, L. *et al.* Tet-mediated formation of 5-hydroxymethylcytosine in RNA. *Journal of the American Chemical Society* **136**, 11582–11585 (2014).
162. Perlaky, L. *et al.* Increased Growth of NIH/3T3 Cells by Transfection with Human p120 Complementary DNA and Inhibition by a p120 Antisense Construct. *Cancer Research* **52**, 428–436 (1992).
163. Freeman, J. W. *et al.* Identification and Characterization of a Human Proliferation-associated Nucleolar Antigen with a Molecular Weight of 120,000 Expressed in Early G1 Phase. *Cancer Research* **48**, 1244–1251 (1988).
164. Freeman, J. W. *et al.* Prognostic Significance of Proliferation Associated Nucleolar Antigen P120 in Human Breast Carcinoma. *Cancer Research* **51**, 1973–1978 (1991).
165. Bantis, A. *et al.* Expression of p120, Ki-67 and PCNA as proliferation biomarkers in imprint smears of prostate carcinoma and their prognostic value. *Cytopathology* **15**, 25–31 (2004).
166. Saijo, Y. *et al.* Expression of nucleolar protein p120 predicts poor prognosis in patients with stage I lung adenocarcinoma. *Annals of Oncology* **12**, 1121–1125 (2001).
167. Frye, M. *et al.* Genomic gain of 5p15 leads to over-expression of Misu (NSUN2) in breast cancer. *Cancer Letters* **289**, 71–80 (2010).
168. Okamoto, M. *et al.* Frequent increased gene copy number and high protein expression of tRNA (cytosine-5-)-methyltransferase (NSUN2) in human cancers. *DNA and Cell Biology* **31**, 660–671 (2012).

169. Lu, L., Zhu, G., Zeng, H., Xu, Q. & Holzmann, K. High tRNA Transferase NSUN2 Gene Expression is Associated with Poor Prognosis in Head and Neck Squamous Carcinoma. *Cancer Investigation* **36**, 246–253 (2018).
170. Rosace, D., López, J. & Blanco, S. Emerging roles of novel small non-coding regulatory RNAs in immunity and cancer. *RNA biology* **17**, 1196–1213 (2020).
171. Ivanov, P., Emara, M. M., Villen, J., Gygi, S. P. & Anderson, P. Angiogenin-induced tRNA fragments inhibit translation initiation. *Molecular Cell* **43**, 613–623 (2011).
172. Blanco, S. *et al.* Stem cell function and stress response are controlled by protein synthesis. *Nature* **534**, 335–340 (2016).
173. Tuorto, F. *et al.* The tRNA methyltransferase Dnmt2 is required for accurate polypeptide synthesis during haematopoiesis. *The EMBO Journal* **34**, 2350–2362 (2015).
174. Subramaniam, D., Thombre, R., Dhar, A. & Anant, S. DNA Methyltransferases: A Novel Target for Prevention and Therapy. *Frontiers in Oncology* **0**, 80 (2014).
175. Elhardt, W., Shanmugam, R., Jurkowski, T. P. & Jeltsch, A. Somatic cancer mutations in the DNMT2 tRNA methyltransferase alter its catalytic properties. *Biochimie* **112**, 66–72 (2015).
176. Cheng, J. X. *et al.* RNA cytosine methylation and methyltransferases mediate chromatin organization and 5-azacytidine response and resistance in leukaemia. *Nature Communications* **9**, 1–16 (2018).
177. Kar, S. P. *et al.* Genome-Wide Meta-Analyses of Breast, Ovarian, and Prostate Cancer Association Studies Identify Multiple New Susceptibility Loci Shared by at Least Two Cancer Types. *Cancer Discovery* **6**, 1052–1067 (2016).
178. He, Y. *et al.* Role of m5C-related regulatory genes in the diagnosis and prognosis of hepatocellular carcinoma. *American Journal of Translational Research* **12**, 912 (2020).
179. Yang, R. *et al.* The RNA methyltransferase NSUN6 suppresses pancreatic cancer development by regulating cell proliferation. *EBioMedicine* **63**, 103195 (2021).
180. Domínguez-Sánchez, M. S., Sáez, C., Japón, M. A., Aguilera, A. & Luna, R. Differential expression of THOC1 and ALY mRNP biogenesis/export factors in human cancers. *BMC cancer* **11**, (2011).
181. Wang, J. Z. *et al.* The role of the HIF-1α/ALYREF/PKM2 axis in glycolysis and tumorigenesis of bladder cancer. *Cancer Communications* **41**, 560–575 (2021).
182. Archer, N. P. *et al.* Family-based exome-wide assessment of maternal genetic effects on susceptibility to childhood B-cell acute lymphoblastic leukemia in hispanics. *Cancer* **122**, 3697–3704 (2016).

183. Barbieri, I. & Kouzarides, T. Role of RNA modifications in cancer. *Nature Reviews Cancer* **20**, 303–322 (2020).
184. Esteve-Puig, R., Bueno-Costa, A. & Esteller, M. Writers, readers and erasers of RNA modifications in cancer. *Cancer Letters* **474**, 127–137 (2020).
185. Esteller, M. Epigenetics in Cancer. *New England Journal of Medicine* **358**, 1148–1159 (2008).
186. Cossío, F. P., Esteller, M. & Berdasco, M. Towards a more precise therapy in cancer: Exploring epigenetic complexity. *Current Opinion in Chemical Biology* **57**, 41–49 (2020).
187. Sibbritt, T., Shafik, A., Clark, S. J. & Preiss, T. Nucleotide-level profiling of m5C RNA methylation. *Methods in Molecular Biology* **1358**, 269–284 (2016).
188. Parker, B. J. Statistical methods for transcriptome-wide analysis of RNA methylation by bisulfite sequencing. *Methods in Molecular Biology* **1562**, (2017).
189. Yang, J. *et al.* Mapping of Complete Set of Ribose and Base Modifications of Yeast rRNA by RP-HPLC and Mung Bean Nuclease Assay. *PLOS one* **11**, (2016).
190. Natchiar, S. K., Myasnikov, A. G., Kratzat, H., Hazemann, I. & Klaholz, B. P. Visualization of chemical modifications in the human 80S ribosome structure. *Nature* **551**, 472–477 (2017).
191. Sponer, J. *et al.* RNA structural dynamics as captured by molecular simulations: A comprehensive overview. *Chemical reviews* **118**, 4177–4338 (2018).
192. Aduri, R. *et al.* AMBER force field parameters for the naturally occurring modified nucleosides in RNA. *Journal of Chemical Theory and Computation* **3**, 1464–1475 (2007).
193. Pérez, A., Luque, F. J. & Orozco, M. Dynamics of B-DNA on the microsecond time scale. *Journal of the American Chemical Society* **129**, 14739–14745 (2007).
194. Dans, P. D. *et al.* Long-timescale dynamics of the Drew-Dickerson dodecamer. *Nucleic Acids Research* **44**, 4052–4066 (2016).
195. Hornak, V. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function and Genetics* **65**, 712–725 (2006).
196. Pérez, A. *et al.* Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ conformers. *Biophysical Journal* **92**, 3817–3829 (2007).
197. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926–935 (1983).
198. Smith, D. E. & Dang, L. X. Computer simulations of NaCl association in polarizable water. *The Journal of Chemical Physics* **100**, 3757–3766 (1994).

199. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Dinola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **81**, 3684–3690 (1984).
200. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**, 327–341 (1977).
201. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98**, (1993).
202. Salomon-Ferrer, R., Götz, A. W., Poole, D., le Grand, S. & Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *Journal of Chemical Theory and Computation* **9**, 3878–3888 (2013).
203. Roe, D. R. & Thomas E. Cheatham, I. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **9**, 3084–3095 (2013).
204. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).
205. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, (2014).
206. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* **11**, 1650–1667 (2016).
207. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology* **33**, 290–295 (2015).
208. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols* **7**, 1534–1550 (2012).
209. Wolfe, A. L. *et al.* RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **513**, 65–70 (2014).
210. Iorio, F. *et al.* A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **166**, 740–754 (2016).
211. Louis, D. N. *et al.* The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* **131**, 803–820 (2016).
212. Noller, H. F. RNA Structure: Reading the Ribosome. *Science* **309**, 1508–1514 (2005).

213. Leprivier, G., Rotblat, B., Khan, D., Jan, E. & Sorensen, P. H. Stress-mediated translational control in cancer cells. *Biochimica et Biophysica Acta* **1849**, 845–860 (2015).
214. Sulima, S. O., Hofman, I. J. F., Keersmaecker, K. de & Dinman, J. D. How Ribosomes Translate Cancer. *Cancer Discovery* **7**, 1069–1087 (2017).
215. Gerashchenko, M. v., Lobanov, A. v. & Gladyshev, V. N. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proceedings of the National Academy of Sciences* **109**, 17394–17399 (2012).
216. Dinkova-Kostova, A. T. & Talalay, P. NAD(P)H:quinone acceptor oxidoreductase 1 (NQO1), a multifunctional antioxidant enzyme and exceptionally versatile cytoprotector. *Archives of Biochemistry and Biophysics* **501**, 116–123 (2010).
217. Parkinson, E. I. & Hergenrother, P. J. Deoxynyboquinones as NQO1-Activated Cancer Therapeutics. *Accounts of Chemical Research* **48**, 2715–2723 (2015).
218. Louis, D. N. *et al.* The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. *Neuro-Oncology* **23**, 1231–1251 (2021).
219. Villanueva, A. *et al.* DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. *Hepatology* **61**, 1945–1956 (2015).
220. Nebbioso, A., Tambaro, F. P., Dell'Aversana, C. & Altucci, L. Cancer epigenetics: Moving forward. *PLOS Genetics* **14**, e1007362 (2018).
221. Lapointe, S., Perry, A. & Butowski, N. A. Primary brain tumours in adults. *The Lancet* **392**, 432–446 (2018).
222. Yan, H. *et al.* IDH1 and IDH2 Mutations in Gliomas. *New England Journal of Medicine* **360**, 765–773 (2009).
223. van den Bent, M. J. *et al.* Chromosomal anomalies in oligodendroglial tumors are correlated with clinical features. *Cancer* **97**, 1276–1284 (2003).
224. Okamoto, Y. *et al.* Population-based study on incidence, survival rates, and genetic alterations of low-grade diffuse astrocytomas and oligodendrogliomas. *Acta Neuropathologica* **108**, 49–56 (2004).
225. Jenkins, R. B. *et al.* A t(1;19)(q10;p10) Mediates the Combined Deletions of 1p and 19q and Predicts a Better Prognosis of Patients with Oligodendroglioma. *Cancer Research* **66**, 9852–9861 (2006).
226. Bromberg, J. E. C. & van den Bent, M. J. Oligodendrogliomas: Molecular Biology and Treatment. *The Oncologist* **14**, 155–163 (2009).
227. Esteller, M. *et al.* Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *New England Journal of Medicine* **343**, 1350–1354 (2000).

228. Hegi, M. E. *et al.* Correlation of O6-methylguanine methyltransferase (MGMT) promoter methylation with clinical outcomes in glioblastoma and clinical strategies to modulate MGMT activity. *Journal of Clinical Oncology* **26**, 4189–4199 (2008).
229. Mansouri, A. *et al.* MGMT promoter methylation status testing to guide therapy for glioblastoma: refining the approach based on emerging evidence and current challenges. *Neuro-Oncology* **21**, 167–178 (2019).
230. Heissenberger, C. *et al.* Loss of the ribosomal RNA methyltransferase NSUN5 impairs global protein synthesis and normal growth. *Nucleic Acids Research* **47**, 11807–11825 (2019).
231. Gay, D. M., Lund, A. H. & Jansson, M. D. Translational control through ribosome heterogeneity and functional specialization. *Trends in Biochemical Sciences* **0**, (2021).
232. Piekna-Przybylska, D., Przybylski, P., Baudin-Baillieu, A., Rousset, J.-P. & Fournier, M. J. Ribosome Performance Is Enhanced by a Rich Cluster of Pseudouridines in the A-site Finger Region of the Large Subunit. *The Journal of Biological Chemistry* **283**, 26026 (2008).
233. Baudin-Baillieu, A. *et al.* Nucleotide modifications in three functionally important regions of the Saccharomyces cerevisiae ribosome affect translation accuracy. *Nucleic Acids Research* **37**, 7665 (2009).
234. Baxter-Roshek, J. L., Petrov, A. N. & Dinman, J. D. Optimization of Ribosome Structure and Function by rRNA Base Modification. *PLOS one* **2**, e174 (2007).
235. Higa-Nakamine, S. *et al.* Loss of ribosomal RNA modification causes developmental defects in zebrafish. *Nucleic Acids Research* **40**, 391–398 (2012).
236. Janin, M., Coll-SanMartin, L. & Esteller, M. Disruption of the RNA modifications that target the ribosome translation machinery in human cancer. *Molecular Cancer* **19**, 1–13 (2020).
237. Marcel, V. *et al.* p53 Acts as a Safeguard of Translational Control by Regulating Fibrillarin and rRNA Methylation in Cancer. *Cancer Cell* **24**, 318–330 (2013).
238. Heiss, N. S. *et al.* X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nature Genetics* **19**, 32–38 (1998).
239. Yoon, A. *et al.* Impaired Control of IRES-Mediated Translation in X-Linked Dyskeratosis Congenita. *Science* **312**, 902–906 (2006).
240. Bellodi, C., Kopmar, N. & Ruggero, D. Deregulation of oncogene-induced senescence and p53 translational control in X-linked dyskeratosis congenita. *The EMBO Journal* **29**, 1865–1876 (2010).

241. Gigova, A., Duggimpudi, S., Pollex, T., Schaefer, M. & Koš, M. A cluster of methylations in the domain IV of 25S rRNA is required for ribosome stability. *RNA* **20**, 1632–1644 (2014).
242. Chassé, H., Boulben, S., Costache, V., Cormier, P. & Morales, J. Analysis of translation using polysome profiling. *Nucleic Acids Research* **45**, e15–e15 (2017).
243. Nagy, J. A., Chang, S.-H., Dvorak, A. M. & Dvorak, H. F. Why are tumour blood vessels abnormal and why is it important to know?. *British Journal of Cancer* **100**, 865–869 (2009).
244. Pries, A. R. *et al.* Structural Adaptation and Heterogeneity of Normal and Tumor Microvascular Networks. *PLOS Computational Biology* **5**, e1000394 (2009).
245. Schulze, H. & Nierhaus, K. H. Minimal set of ribosomal components for reconstitution of the peptidyltransferase activity. *The EMBO Journal* **1**, 609–613 (1982).
246. Volarević, S. *et al.* Proliferation, But Not Growth, Blocked by Conditional Deletion of 40S Ribosomal Protein S6. *Science* **288**, 2045–2047 (2000).
247. Matsson, H. *et al.* Targeted Disruption of the Ribosomal Protein S19 Gene Is Lethal Prior to Implantation. *Molecular and Cellular Biology* **24**, 4032–4037 (2004).
248. Kirn-Safran, C. B. *et al.* Global growth deficiencies in mice lacking the ribosomal protein HIP/RPL29. *Developmental Dynamics* **236**, 447–460 (2007).
249. Ben-Shem, A. *et al.* The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. *Science* **334**, 1524–1529 (2011).
250. Sloan, K. E. *et al.* Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biology* **14**, 1138–1152 (2016).
251. Leinonen, H. M., Kansanen, E., Pölönen, P., Heinäniemi, M. & Levonen, A. L. Role of the Keap1–Nrf2 Pathway in Cancer. *Advances in Cancer Research* **122**, 281–320 (2014).
252. Reigan, P. *et al.* Development of Indolequinone Mechanism-Based Inhibitors of NAD(P)H:Quinone Oxidoreductase 1 (NQO1): NQO1 Inhibition and Growth Inhibitory Activity in Human Pancreatic MIA PaCa-2 Cancer Cells. *Biochemistry* **46**, 5941–5959 (2007).
253. Bair, J. S., Palchaudhuri, R. & Hergenrother, P. J. Chemistry and Biology of Deoxynyboquinone, a Potent Inducer of Cancer Cell Death. *Journal of the American Chemical Society* **132**, 5469–5478 (2010).
254. Lundberg, A. P. *et al.* Pharmacokinetics and derivation of an anticancer dosing regimen for the novel anti-cancer agent isobutyl-deoxynyboquinone (IB-DNQ), a NQO1 bioactivatable molecule, in the domestic felid species. *Investigational New Drugs* **35**, 134–144 (2016).

255. Lee, H. Y. *et al.* Reactive Oxygen Species Synergize To Potently and Selectively Induce Cancer Cell Death. *ACS Chemical Biology* **12**, 1416–1424 (2017).
256. Paschka, P. *et al.* IDH1 and IDH2 Mutations Are Frequent Genetic Alterations in Acute Myeloid Leukemia and Confer Adverse Prognosis in Cytogenetically Normal Acute Myeloid Leukemia With NPM1 Mutation Without FLT3 Internal Tandem Duplication. *Journal of Clinical Oncology* **28**, 3636–3643 (2010).
257. Borger, D. R. *et al.* Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *The Oncologist* **17**, 72–79 (2012).
258. Amary, M. F. *et al.* IDH1 and IDH2 mutations are frequent events in central chondrosarcoma and central and periosteal chondromas but not in other mesenchymal tumours. *The Journal of Pathology* **224**, 334–343 (2011).
259. Han, S. *et al.* IDH mutation in glioma: molecular mechanisms and potential therapeutic targets. *British Journal of Cancer* **122**, 1580–1589 (2020).
260. Reitman, Z. J. & Yan, H. Isocitrate Dehydrogenase 1 and 2 Mutations in Cancer: Alterations at a Crossroads of Cellular Metabolism. *JNCI: Journal of the National Cancer Institute* **102**, 932–941 (2010).
261. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744 (2009).
262. Ohka, F. *et al.* Quantitative metabolome analysis profiles activation of glutaminolysis in glioma with IDH1 mutation. *Tumor Biology* **35**, 5911–5920 (2014).
263. Maus, A. & Peters, G. J. Glutamate and α-ketoglutarate: key players in glioma metabolism. *Amino Acids* **49**, 21–32 (2016).
264. Badur, M. G. *et al.* Oncogenic R132 IDH1 Mutations Limit NADPH for De Novo Lipogenesis through (D)2-Hydroxyglutarate Production in Fibrosarcoma Cells. *Cell Reports* **25**, 1018-1026.e4 (2018).
265. Chesnelong, C. *et al.* Lactate dehydrogenase A silencing in IDH mutant gliomas. *Neuro-Oncology* **16**, 686–695 (2014).
266. Ruggero, D. Translational Control in Cancer Etiology. *Cold Spring Harbor Perspectives in Biology* **5**, a012336 (2013).
267. Truitt, M. L. & Ruggero, D. New frontiers in translational control of the cancer genome. *Nature Reviews Cancer* **16**, 288–304 (2016).
268. Tabas, I. & Ron, D. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nature Cell Biology* **13**, 184–190 (2011).
269. Yadav, R. K., Chae, S.-W., Kim, H.-R. & Chae, and H. J. Endoplasmic Reticulum Stress and Cancer. *Journal of Cancer Prevention* **19**, 75–88 (2014).

270. Dannfald, A., Favory, J. J. & Deragon, J. M. Variations in transfer and ribosomal RNA epitranscriptomic status can adapt eukaryote translation to changing physiological and environmental conditions. *RNA Biology* (2021) doi:10.1080/15476286.2021.1931756.
271. Kurylo, C. M. *et al.* Endogenous rRNA Sequence Variation Can Regulate Stress Response Gene Expression and Phenotype. *Cell Reports* **25**, 236–248 (2018).
272. Liberman, N. *et al.* N6-adenosine methylation of ribosomal RNA affects lipid oxidation and stress resistance. *Science Advances* **6**, (2020).
273. Ghulam, M. M., Catala, M. & Abou Elela, S. Differential expression of duplicated ribosomal protein genes modifies ribosome composition in response to stress. *Nucleic Acids Research* **48**, 1954–1968 (2020).
274. Ferretti, M. B., Ghalei, H., Ward, E. A., Potts, E. L. & Karbstein, K. Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements. *Nature Structural & Molecular Biology* **24**, 700–707 (2017).
275. Shi, Z. *et al.* Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. *Molecular Cell* **67**, 71–83 (2017).
276. Villanueva, A. Hepatocellular Carcinoma. *New England Journal of Medicine* **380**, 1450–1462 (2019).
277. McGlynn, K. A., Petrick, J. L. & London, W. T. Global Epidemiology of Hepatocellular Carcinoma: An Emphasis on Demographic and Regional Variability. *Clinics in Liver Disease* **19**, 223–238 (2015).
278. Llovet, J. M. *et al.* Hepatocellular carcinoma. *Nature Reviews Disease* **7**, 1–28 (2021).
279. Rahib, L. *et al.* Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. *Cancer Research* **74**, 2913–2921 (2014).
280. Fitzmaurice, C. *et al.* The Burden of Primary Liver Cancer and Underlying Etiologies From 1990 to 2015 at the Global, Regional, and National Level: Results From the Global Burden of Disease Study 2015. *JAMA Oncology* **3**, 1683–1691 (2017).
281. Wang, J., Chenivesse, X., Henglein, B. & Bréchot, C. Hepatitis B virus integration in a cyclin A gene in a hepatocellular carcinoma. *Nature* **343**, 555–557 (1990).
282. Chang, M. H. *et al.* Long-term Effects of Hepatitis B Immunization of Infants in Preventing Liver Cancer. *Gastroenterology* **151**, 472–480 (2016).
283. Jain, M. K. *et al.* Evaluation of a Multifaceted Intervention to Reduce Health Disparities in Hepatitis C Screening: A Pre-Post Analysis. *Hepatology* **70**, 40–50 (2019).

284. Jepsen, P., Ott, P., Andersen, P. K., Sørensen, H. T. & Vilstrup, H. Risk for hepatocellular carcinoma in patients with alcoholic cirrhosis: A Danish nationwide cohort study. *Annals of Internal Medicine* **156**, 841–847 (2012).
285. Estes, C., Razavi, H., Loomba, R., Younossi, Z. & Sanyal, A. J. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology* **67**, 123–133 (2018).
286. Llovet, J. M., Montal, R., Sia, D. & Finn, R. S. Molecular therapies and precision medicine for hepatocellular carcinoma. *Nature Reviews Clinical Oncology* **15**, 599–616 (2018).
287. Marrero, J. A. *et al.* Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology* **68**, 723–750 (2018).
288. Schulze, K. *et al.* Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nature Genetics* **47**, 505–511 (2015).
289. Zucman-Rossi, J., Villanueva, A., Nault, J. C. & Llovet, J. M. Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology* **149**, 1226–1239 (2015).
290. Berdasco, M. & Esteller, M. Clinical epigenetics: seizing opportunities for translation. *Nature Reviews Genetics* **20**, 109–127 (2019).
291. Heyn, H. & Esteller, M. DNA methylation profiling in the clinic: applications and challenges. *Nature Reviews Genetics* **13**, 679–692 (2012).
292. Castello, A. *et al.* Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149**, 1393–1406 (2012).
293. Hein, M. Y. *et al.* A Human Interactome in Three Quantitative Dimensions Organized by Stoichiometries and Abundances. *Cell* **163**, 712–723 (2015).
294. Masuda, S. *et al.* Recruitment of the human TREX complex to mRNA during splicing. *Genes & Development* **19**, 1512–1517 (2005).
295. Pühringer, T. *et al.* Structure of the human core transcription-export complex reveals a hub for multivalent interactions. *eLife* **9**, 1–65 (2020).
296. Sträßer, K. *et al.* TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**, 304–308 (2002).
297. Heath, C. G., Viphakone, N. & Wilson, S. A. The role of TREX in gene expression and disease. *Biochemical Journal* **473**, 2911–2935 (2016).
298. Farny, N. G., Hurt, J. A. & Silver, P. A. Definition of global and transcript-specific mRNA export pathways in metazoans. *Genes & Development* **22**, 66–78 (2008).

299. Mancini, A. *et al.* THOC5/FMIP, an mRNA export TREX complex protein, is essential for hematopoietic primitive cell survival in vivo. *BMC Biology* **8**, 1–17 (2010).

300. Lishko, P. v. *et al.* The control of male fertility by spermatozoan ion channels. *Annual Review of Physiology* **74**, 453–475 (2012).

301. Okunade, G. W. *et al.* Targeted ablation of plasma membrane Ca2+-ATPase (PMCA) 1 and 4 indicates a major housekeeping function for PMCA1 and a critical role in hyperactivated sperm motility and male fertility for PMCA4. *Journal of Biological Chemistry* **279**, 33742–33750 (2004).

ANNEXES

ANNEXES

Supplementary Table 1: Association between prognostic factors and NSUN5 methylation status in TCGA dataset and validation cohort of glioma patients.

CHARACTERISTICS	Glioma TCGA dataset NSUN5 methylation (N = 497)				Glioma VALIDATION cohort NSUN5 methylation (N = 115)			
	N (%)	Unmethylated N (%)	Methylated N (%)	P*	N (%)	Unmethylated N (%)	Methylated N (%)	P*
Age (years)								
<50	282 (56%)	192 (68%)	90 (32%)	0.090	47 (41%)	30 (64%)	17 (36%)	0.052
>50	213 (43%)	160 (75%)	53 (25%)		68 (59%)	55 (81%)	13 (19%)	
Unknown	2 (1%)	0 (0%)	2 (100%)		0 (0%)	0 (0%)	0 (0%)	
Gender								
Male	267 (54%)	187 (70%)	80 (30%)	0.074	68 (59%)	51 (75%)	17 (25%)	0.830
Female	228 (45%)	165 (73%)	63 (27%)		47 (41%)	34 (72%)	13 (28%)	
Unknown	2 (1%)	0 (0%)	2 (100%)		0 (0%)	0 (0%)	0 (0%)	
Grade of glioma								
Low-grade glioma	374 (75%)	237 (63%)	137 (37%)	$<10^{-5}$***	41 (36%)	25 (61%)	16 (39%)	0.026*
Glioblastoma	123 (25%)	115 (93%)	8 (7%)		74 (64%)	60 (81%)	14 (19%)	
Treatment								
Radiotherapy	221 (44%)	140 (63%)	81 (37%)	$<10^{-5}$***	8 (7%)	4 (50%)	4 (50%)	0.187
Temozolamide	9 (2.5%)	9 (100%)	0 (0%)		49 (43%)	39 (70%)	10 (30%)	
Radiotherapy + Temozolamide	82 (16%)	74 (90%)	8 (10%)		33 (29%)	24 (73%)	9 (27%)	
Others	9 (2.5%)	9 (100%)	0 (0%)		11 (9%)	6 (54%)	5 (46%)	
No treatment	169 (34%)	113 (67%)	56 (33%)		6 (5%)	6 (100%)	0 (0%)	
Unknown	7 (1%)	7 (100%)	0 (0%)		8 (7%)	6 (75%)	2 (25%)	
IDH1 mutational status								
Wild-type	203 (41%)	187 (92%)	16 (8%)	$<10^{-5}$***	54 (47%)	47 (87%)	7 (13%)	10^{-4}***
Mutated	294 (50%)	165 (56%)	129 (44%)		29 (25%)	13 (45%)	16 (55%)	
Unknown	0 (0%)	0 (0%)	0 (0%)		32 (28%)	25 (78%)	7 (22%)	
Co-deletion 1p/19q								
No co-deletion	349 (65%)	261 (75%)	88 (25%)	0.002**	17 (15%)	10 (59%)	7 (41%)	0.999
Co-deletion 1p/19q	131 (26%)	77 (59%)	54 (41%)		16 (14%)	9 (56%)	7 (44%)	
Unknown	17 (4%)	14 (82%)	3 (18%)		82 (71%)	66 (80%)	16 (20%)	
MGMT methylation status								
Unmethylated	240 (48%)	194 (81%)	46 (19%)	$<10^{-5}$***	25 (22%)	19 (76%)	6 (24%)	0.432
Methylated	257 (52%)	158 (61%)	99 (39%)		56 (49%)	45 (80%)	11 (37%)	
Unknown	0 (0%)	0 (0%)	0 (0%)		34 (29%)	21 (62%)	13 (38%)	

Footnote: *P-value represents Fisher's exact test or Chi-square function whenever required: statistical significance under 0.05*, 0.01** and 0.001***

Las estatuas surgen de la dura y fría roca
solo a base de martillo y cincel.
Saurom