



SAPIENZA
UNIVERSITÀ DI ROMA

SAPIENZA UNIVERSITY OF ROME

FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI

Ph.D. in Cell and Developmental Biology XXXV Cycle

AA 2022-2023

Ph.D. Thesis

Relevance of KDM5B isoforms in breast cancer

Ph.D. Student: Elena Di Nisio

Supervisor: Prof. Rodolfo Negri

Internal reviewer: Prof. Giuseppe Lupo

Ph.D. course Coordinator: Prof.ssa Giulia De Lorenzo

Department of Biology and Biotechnologies "Charles Darwin"

“The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when one contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries to comprehend only a little of this mystery every day.”

Albert Einstein

INDEX

Summary	5
Synopsis of the whole work.....	6
1. INTRODUCTION.....	9
2. AIM OF THE WORK	15
3. RESULTS AND DISCUSSION	16
Part 1 - Characterization of a new KDM5B isoform including exon-6	16
3.1. The frequency of exon-6 inclusion in KDM5B splicing variants is similar in melanoma and breast cancer cell lines	16
3.2. Although the KDM5B transcripts level is significantly higher in MCF7 compared to MDA-MB-231, the fraction of exon-6 including transcripts is similar in the two cell lines	18
3.3. Western blot analyses suggest that a predicted N-terminal truncated KDM5B isoform including exon-6 is expressed in breast cancer cell lines.....	19
3.4. 5'-RLM-RACE demonstrates the presence of a new downstream transcriptional start site, compatible with the expression of a truncated and catalytically inactive isoform of KDM5B.....	21
3.5. MCF7 and MDA-MB-231 breast cancer cell lines show a different relative amount of KDM5B-NTT with respect to the PLU-1 isoform.....	24
3.6. Cycloheximide treatment highlighted the greater instability of PLU-1 compared to the KDM5B-NTT isoform	26
3.7. The proteasomal degradation activity regulates PLU-1 turn-over	28
Part 2 – Investigating the putative role of KDM5B-NTT isoform in breast cancer	31
3.8. RNA interference and transient over-expression as a tool to investigate the function of KDM5B-NTT in breast cancer cells	32
3.9. KDM5B-NTT affects the global H3K4me3 level	35

3.10. KDM5B-NTT regulates gene expression.....	37
3.11. The putative regulatory action of KDM5B-NTT.....	39
4. CONCLUSIONS AND FUTURE PERSPECTIVES.....	42
5. MATERIAL AND METHODS.....	44
6. REFERENCES.....	59
7. SUPPLEMENTAL FILES.....	64
8. SUPPLEMENTAL TABLES AND FIGURES.....	65
9. SUPPLEMENTAL REFERENCES	76
ACKNOWLEDGMENTS.....	78
APPENDIX.....	79

Summary

KDM5B (also known as JARID1B or PLU-1) is a histone lysine demethylase involved in differentiation processes, gene expression regulation, DNA damage repair, cancer genesis, and drug resistance. Aberrant KDM5B expression has been observed in many human cancers. Intriguingly, growing evidence suggests that KDM5B can have a dual role, acting as an oncogene or as a tumor suppressor in a cell-specific context. KDM5B sustains cell proliferation in the luminal breast cancer cell line MCF7, while it inhibits cell migration and invasion in the basal breast cancer cell line MDA-MB-231.

Recently, it has been reported that an isoform of KDM5B including the variant exon-6, namely RBP2H1, might contribute to tumor progression in melanoma.

However, if different KDM5B isoforms could play a role in breast cancer subtypes it has not been investigated.

In this study, we are reporting the characterization of a new KDM5B isoform which is N-terminal truncated (KDM5B-NTT) and catalytically inactive. KDM5B-NTT is a splicing variant having a transcriptional start site downstream to that known for the canonical PLU-1 isoform and including also 36 additional residues encoded by the variant exon-6, previously associated with the RBP2H1 isoform in melanoma. We observed that KDM5B-NTT accumulates in breast cancer cells due to higher protein stability compared to KDM5B-PLU-1 and appears relatively more expressed in MDA-MB-231 than in MCF7. The exogenous over-expression of KDM5B-NTT in MCF7 cells correlates with a global increase of H3K4me3 levels and with the up-regulation of the tumor suppressor Caveolin-1 and genes related to the interferon-alpha and -gamma response. These results give rise to the possibility of further exploring the regulatory roles of KDM5B independent of its catalytic activity.

Synopsis of the whole work

The histone demethylase KDM5B is a master regulator of H3K4 methylome and its aberrant expression correlates with many human cancers. KDM5B often behaves as an oncogene. However, its role in cancer seems multifaceted and complex, depending on the cellular context. A paradigm of this opposite behavior is highlighted by the different roles of KDM5B in gene expression regulation comparing the luminal breast cancer cell line MCF7 versus the basal breast cancer cell line MDA-MB-231. The complexity of KDM5B roles in cancer might be the result of the expression of different isoforms. So far, the expression of KDM5B isoforms in breast cancer cell lines was studied considering only the total pool of proteins, therefore we questioned the relevance of KDM5B isoforms in breast cancer. Recently, it has been proposed that the increase of a KDM5B isoform including the variant exon-6 drives the tumour progression in melanoma. Considering that, our investigation started by analyzing the inclusion frequency of this alternative exon-6 in KDM5B transcripts in breast cancer cell lines compared to melanoma cell lines. The bioinformatic analysis revealed a similar frequency of exon-6 inclusion in melanoma and breast cancer cell lines. RT-qPCR showed that, despite the higher level of total KDM5B transcripts in MCF7 cells than MDA-MB-231, the level of KDM5B transcripts with exon-6 compared with the total KDM5B transcripts level was similar in the two cell lines. Since transcription and translation don't have a linear and simple relationship, we decided to explore the expression of KDM5B isoforms through Western blot analyses. Using two different primary antibodies (Ab1 for the C-terminal region present in all the isoforms of KDM5B and Ab2 specific for the exon-6 variant encoded residues), we identified two major isoforms of KDM5B in breast cancer cells. We observed the expression of an isoform with an apparent molecular weight of approximately 175kDa (corresponding to KDM5B-PLU-1, the canonical isoform) and another isoform at 162kDa, whereas no band was detected around 180kDa (corresponding to RBP2-H1 isoform with exon-6 previously reported in melanoma). Surprisingly, the isoform with the apparent molecular weight of 162kDa is detected using both antibodies, indicating that this is smaller than PLU-1 and that includes the insertion of 36 residues encoded by the exon-6. Focusing our attention on the characterization of the transcriptional start site of KDM5B transcripts which include the exon-6, we found a new transcript compatible with the

expression of a shorter KDM5B isoform N-terminal truncated (hereafter called KDM5B-NTT). This isoform is very fascinating because could potentially act both as loss of function, being catalytically inactive, and as a gain of function, including the residues with still unknown function encoded by the exon-6. Questioning if the relative abundance of the two main KDM5B isoforms (PLU-1 and NTT) might correlate with tumour breast cancer progression, we analyze the relative ratio NTT/PLU-1 in different breast cancer cell lines. Although no clear correlation was found with the different cancer subtypes, the NTT/PLU-1 ratio was significantly higher in MDA-MB-231 cells than in MCF7 cells, suggesting a putative contributing to explain the divergent role of KDM5B previously observed in these two cell lines. Integrating the RT-qPCR and the Western blot analyses, we questioned how it was possible that a transcript representing less than 10% of the KDM5B total transcripts could generate a much higher percentage of protein (approximately 40% in MCF7 and 57% in MDA-MB-231). Testing the hypothesis of a different post-translational regulation of the two isoforms, we found that NTT isoform is much more stable than PLU-1 isoform in both the analyzed breast cancer cell lines. For time being, we don't know why PLU-1 and NTT isoforms have so remarkably different stability. However, our hypothesis is that in the N-terminal region of PLU-1, missing in NTT, there could be a degron motif that targets PLU-1 to proteasome-mediated degradation. The faster turnover of PLU-1 compared to NTT, which is stable for at least 12 hours, suggests that KDM5B-NTT can accumulate in cells. To investigate the putative function of KDM5B-NTT in breast cancer cells we decided to use two complementary approaches: downregulate specifically KDM5B-NTT in MDA-MB-231 cells, where the NTT/PLU-1 ratio was greater than in MCF7 cells; at the opposite, over-express NTT in MCF7 cells, where the NTT/PLU-1 ratio was lower compared to MDA-MB-231 cells. Unfortunately, we observed that the interference was not specific for NTT, so we decided to temporarily pause the downregulation approach, meanwhile designing new strategies. The over-expression of NTT determined a significant increase of global H3K4 trimethylation level in MCF7, reaching the trimethylation basal level of MDA-MB-231 cells. Moreover, KDM5B-NTT overexpression modulated the genes belonging to the interferon-alpha and gamma response and inflammation in MCF7 cells. It is plausible that KDM5B-NTT could interfere with PLU-1 at some target genes, by competing with its repressive activity. Supporting this hypothesis, many of the induced genes upon KDM5B-NTT over-expression in MCF7 cells were previously shown regulated by the direct action

of PLU-1 in MCF7 or other breast cancer cell lines. However, on the putative dominant negative role of this isoform in breast cancer progression many questions are still open and further investigations about the mechanistic details are needed.

1. INTRODUCTION

The wrapping of DNA around the octamer of histone proteins brings to the formation of the nucleosome, the structural unit of the chromatin in eukaryotic cells¹. Post-translational modifications of histones modulate the accessibility to genetic information², regulating many cellular processes highly influenced by the chromatin state, such as replication, transcription and genome stability. Epigenetic players include *writers*, *readers* and *erasers* which respectively add, recognize and remove epigenetic marks. Histone lysine methylation is a dynamic and reversible mark, mainly associated with gene expression regulation, and tightly controlled during early development and cellular differentiation states by the coordinated action of lysine methyltransferases and demethylases^{3–6}. Histone lysine (K) demethylases (KDMs) can act as both readers and erasers⁷. In human, there are two enzymatic classes of histone demethylases: the flavin-adenine-dinucleotide (FAD) dependent amine oxidases, belonging to the Lysine Specific histone Demethylase (LSD) family (or KDM1 subfamily), and the Fe(II) and 2-oxoglutarate-dependent oxygenases with a conserved Jumonji C domain (JmjC), belonging to the Jumonji Histone Demethylases (JHDMs) family (or KDM2–KDM7 subfamilies). The reader domains confer the substrate specificity, recognizing specific lysine residues with different methylation degrees. Histone methylation may lead to different transcriptional outcomes, such as gene activation or repression, depending on the modified residue, the degree of methylation and the location of the mark in the genome^{10,11}. Among KDMs, the KDM5 (or JARID1) enzymes remove di- and tri-methylation of lysine 4 of histone H3 (H3K4me₂/me₃), playing thus an important role in transcriptional regulation. The H3K4me₃ mark in gene promoter regions furthers the recruitment of basal transcription factors^{12–14}. Accordingly, KDM5 enzymes can act as epigenetic repressors in a demethylase-activity-dependent manner, by removing the H3K4me₃ at gene promoter regions^{15–17}. The demethylase activity of KDM5s is also known to be involved in the mechanism of DNA double-strand breaks (DSBs) repair (DDR), by silencing transcription *in cis* to DNA DSBs and by favoring the recruitment of DNA repair factors (Fig.1)^{18–23}. KDM5 proteins can also act as transcriptional repressors in a demethylase-activity-independent manner, by recruiting other repressive complexes^{24,25}. Conversely, they can act as transcriptional activators by removing the H3K4me₃ mark in non-promoter regions. The removal of H3K4me₃ at the body

of actively transcribed genes supports productive transcriptional elongation through the repression of spurious intragenic transcription¹¹. The conversion of H3K4me3 to H3K4me1, combined with acetylated H3K27, is predictive of active enhancers²⁶.

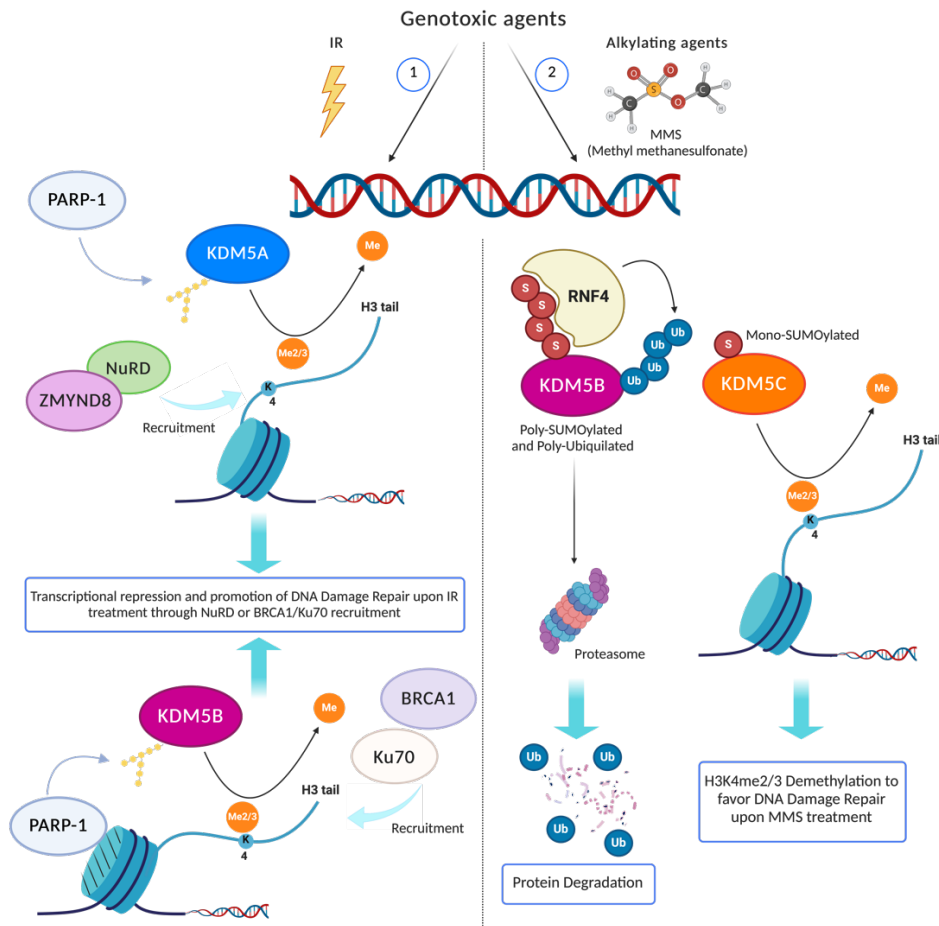


Fig.1 Roles of KDM5 demethylases in the DNA damage response. KDM5 enzymes can play different roles in the DNA Damage Response (DDR) depending on the genotoxic agent type. KDM5A and KDM5B are recruited on the DNA damage site upon ionizing radiation (IR) in a PARP1-dependent manner (as shown in (1), left panel). Their demethylase activity is involved in local transcription repression by lowering the H3K4me2/3 levels and by favoring the recruitment of repair factors, including BRCA1 and Ku70. On the other hand, in response to alkylation damage (as shown in (2), right panel) poly-SUMOylated KDM5B is poly-ubiquitylated by the E3 SUMO-dependent ubiquitin ligase RNF4 and targeted to proteasome degradation, while KDM5C is mono-SUMOylated and recruited to the chromatin to demethylate histone H3K4me2/3. (From Di Nisio et al. 2021)²¹.

In human, the KDM5 subfamily consists of four proteins: KDM5A (also known as JARID1A or RBP2), KDM5B (also known as JARID1B or PLU1), KDM5C (also known as JARID1C or SMCX), and KDM5D (also known as JARID1D or SMCY)²⁷. The KDM5 members show a high degree of sequence homology and domain organization (Fig.2).

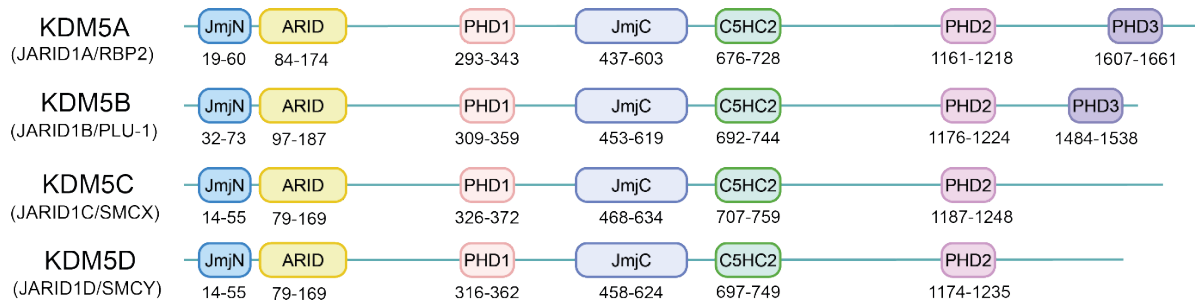


Fig.2 KDM5 demethylases family. Schematic representation of protein domains organization (not on scale) in the human KDM5 histone demethylases family members. Human KDM5 demethylases are highly similar in domain architecture.

This is the only group of histone demethylases with an atypical insertion that separates the Jumonji (Jmj) catalytic domain into two regions (JmjN and JmjC). The inserted domains are the AT-rich interaction (ARID) domain and the plant homeodomain (PHD) finger 1 (PHD1). The ARID domain binds DNA sequences with CCGCCC or GCACA/C motif^{17,28} and the PHD1 recognizes the unmethylated H3K4 histone tail²⁹⁻³¹. The C₅HC₂ zinc finger domain is also required for efficient catalytic activity³². The PHD2 domain is present in all the members and its function is still unknown, whereas KDM5A and KDM5B contain also the PHD3 domain that binds the H3K4me₃^{29,33-35}.

Intriguingly, KDM5s can sometimes directly favor the transcription independently of their demethylase activity, functioning as a scaffold for transcription factors in specific cell contexts: KDM5A in a complex with CLOCK-BMAL can inhibit the activity of the histone deacetylase HDAC1, enhancing histone acetylation and transcriptional activation at the *Per2* promoter; KDM5B orchestrates a biphasic regulation of retinoic acid (RA)-dependent genes. In absence of RA, KDM5B cooperates in complex with Polycomb Repressive Complex 2 (PRC2) to repress transcription, whereas in presence of RA KDM5B mediates the PRC2 dissociation from chromatin and a consequent transcriptional activation³⁶. Gene expression regulation by KDM5 proteins plays a role in many cellular processes, such as the regulation of stem cell function, cell differentiation, early embryonic development, and pathological conditions, including cancer³⁷⁻⁴⁰. Aberrant expression of KDM5s contributes significantly to tumor initiation and progression in many human cancers. Specifically, KDM5B was initially identified as a markedly up-regulated gene in breast cancer⁴² and later found over-expressed in prostate⁴³, bladder and lung cancer⁴⁴, and in stem-

like subpopulations in melanomas⁴⁵. In normal adult tissues, KDM5B expression is mainly restricted to testis, where contributes to transcriptional control during spermatogenesis, suggesting it could belong to the testis/cancer antigens class. KDM5B shows a copy number gain associated with overexpression, especially in luminal breast cancer subtypes⁴¹. In the luminal MCF-7 (estrogen-responsive, ER+) breast cancer cell line, KDM5B favors cell proliferation and is involved in the repression of several tumor-suppressor genes such as BRCA1 and Caveolin-1¹⁵. On the contrary, in triple-negative breast cancer cell line MDA-MB-231, KDM5B inhibits cell proliferation, inflammatory response, adhesion and migration through the interaction with LSD1/NuRD complex²⁹. KDM5B plays a dual role also in melanoma: re-expression of KDM5B causes a decrease in the proliferative rate, suggesting it may act as a tumor suppressor⁴⁸, however, a small subpopulation of slow-cycling cells retains the over-expression of KDM5B, which is necessary for the long-term tumour growth maintenance⁴⁵. Recent observations show that the malignant transformation of melanocytes correlates with a change in the expression pattern of KDM5B isoforms⁴⁹. A significantly increased expression of the RBP2-H1 isoform, which includes the alternative exon-6, was found in melanomas compared to naevi, suggesting a role of this isoform in tumor progression⁵⁰. Exon-6 shows strong homology with Alu sequences, but its function is still unknown⁴⁶.

So far, the expression of KDM5B isoforms in breast cancer cell lines was studied considering only the total pool of proteins, thus we questioned the relevance of KDM5B isoforms in different breast cancer lines. The history of KDM5B isoforms begins in 1999^{42,51,52}, but no clarity has been made yet, and their functional relevance is largely unknown.

Four KDM5B transcripts are reported in NCBI Reference Sequence (RefSeq) database (Fig.3A). They differ for the 5'-end and the inclusion of exon-5 and exon-6. RBP2-H1 is the longest transcript and includes the alternative exon-6, absent in PLU-1, KDM5B-215 and KDM5B-240. KDM5B-215 uses an alternative splice site in exon-1 and KDM5B-240 doesn't include exon-5. The resulting reported protein isoforms are the following: RBP2-H1 (1580 aa, 179.44 kDa) having 36 additional residues (encoded by the variant exon-6) after the Glu237 of PLU-1 isoform (1544 aa, 175.14 kDa), isoform-3 (1499 aa, 170.45 kDa) lacking residues encoding by exon-5 and exon-6 (from KDM5B-240 transcript) and isoform-4 (1539 aa, 175.14 kDa), that differs from PLU-1 for the lacking of 5 residues after Lys63 (Fig.3B).

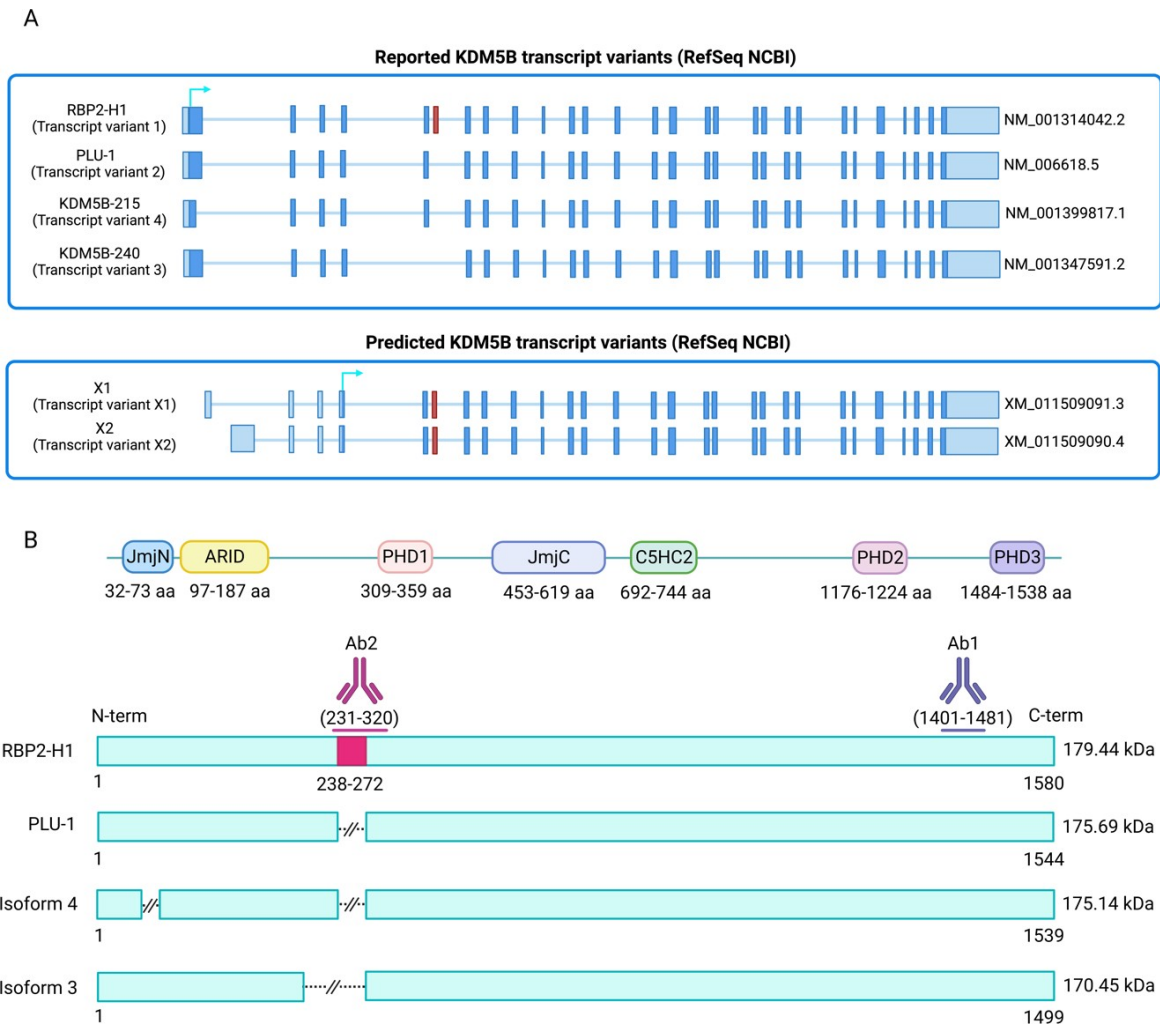


Fig.3 Schematic of KDM5B transcript variants and protein isoforms. A. KDM5B transcripts as reported in RefSeq NCBI database. RBP2-H1, PLU-1, KDM5B-215 and KDM5B-240 are the reported (NM) variants; X1 and X2 are the predicted (XM) variants. Transcripts differ for the 5'-end and the presence of alternative exons. RBP2-H1 represents the longest transcript and includes the exon-6 (in red). The arrows indicate the translation start site (ATG site); the translation of reported variants starts in exon-1; the translation of the predicted variants starts instead in exon-4. **B** Protein domains of human KDM5B protein isoforms that arise from the translation of the reported transcript variants are shown; PLU-1 is the canonical isoform. RBP2-H1 is the longest isoform with exon-6 (dark pink). The antibodies used to study the KDM5B isoforms are represented as Ab1 and Ab2. The Ab1 primary antibody detects the C-terminal region present in all the isoforms; the Ab2 primary antibody is specific for the region of the residues encoded by exon-6.

Here we describe a previously predicted KDM5B isoform that is N-Terminal Truncated (KDM5B-NTT) and catalytically inactive, includes the variant exon-6 and appears expressed in different breast cancer cell lines. The endogenous expression of KDM5B-NTT compared to KDM5B-PLU-1 is significantly lower in the luminal breast cancer cell line MCF7 versus the basal breast cancer cell line MDA-MB-231.

This different relative abundance seems related to the higher protein stability of KDM5B-NTT compared to PLU-1, combined with the faster turnover of KDM5B-PLU-1 in MDA-MB-231. Over-expressing KDM5B-NTT in MCF7 cells, we observed a global increase of H3K4me3 and significant transcriptomic changes, supporting the hypothesis this newly characterized isoform may have regulatory roles in breast cancer cells, independently of its catalytic activity.

2. AIM OF THE WORK

The histone demethylase KDM5B is a master regulator of H3K4 methylome and its aberrant expression correlates with many human cancers, behaving usually as an oncogene. Although growing evidence of the involvement of KDM5B in oncogenesis, its role in cancer seems multifaceted and complex, depending on the cellular context. A paradigm of the opposite behavior of KDM5B in breast cancer is highlighted by its antagonistic role in gene expression regulation comparing luminal versus basal breast cancers. Elsewhere, conflicting roles have been reported also in melanoma. However, recently it has been proposed that the complexity of KDM5B roles in melanoma may be the result of a different relative expression of its isoforms. The KDM5B splicing variant 1 (RBP2-H1) with the alternative exon-6 correlates to tumour progression in melanoma, and the variant 2 (PLU-1) without the alternative exon-6 contributes to long-time tumour maintenance in a slow-cycling subpopulation of melanoma cells. KDM5B is a known luminal-driven oncogene in breast cancer, but so far, the relevance of KDM5B isoforms in breast cancer remains elusive. Therefore, the aim of this study is to explore the putative role of KDM5B isoforms in breast tumour progression.

3. RESULTS AND DISCUSSION

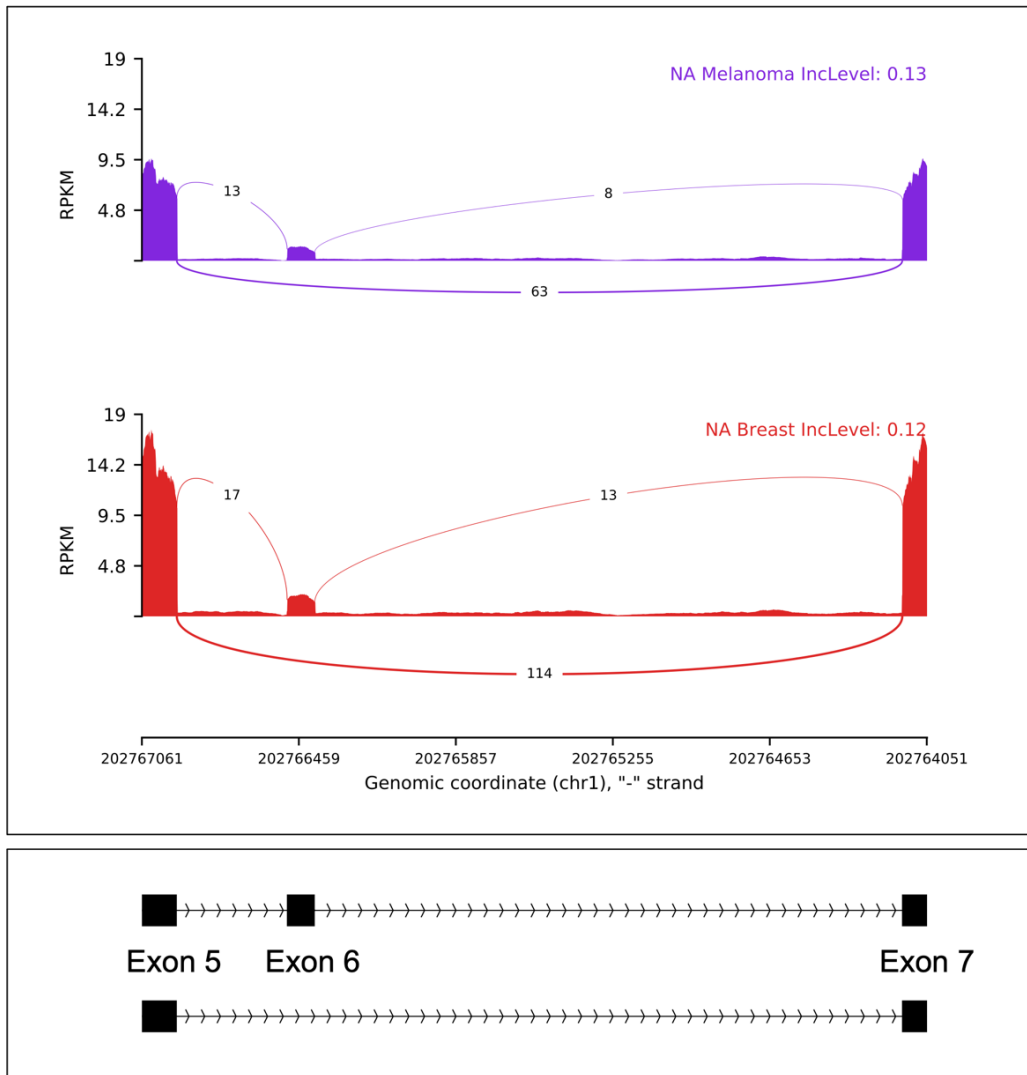
Part 1 - Characterization of a new KDM5B isoform including exon-6

3.1. The frequency of exon-6 inclusion in KDM5B splicing variants is similar in melanoma and breast cancer cell lines

KDM5B is up-regulated in many cancers and is a luminal lineage-driving oncogene in breast cancer^{41,53}. However, KDM5B acts apparently as a tumor suppressor in basal breast cancer cell lines^{29,33}. Recent research suggested the expression of a KDM5B isoform (RBP2-H1), which includes the alternative exon-6, correlates with the malignant transformation of melanocytes⁴⁹. Therefore, we started studying the inclusion frequency of this alternative exon-6 in KDM5B transcripts in breast cancer cell lines compared to melanoma cell lines. We performed a computational analysis of transcriptomic data from a set of 49 melanoma cell lines and a set of 56 breast cancer cell lines, retrieved from the Cancer Cell Line Encyclopedia (CCLE) database. Although the exon-6 read counts are relatively lower compared to other side exons, indicating this exon is generally skipped, the analysis revealed a similar frequency of exon-6 inclusion in melanoma and breast cancer cell lines. As shown in the Sashimi plot (Fig.4), about 12% of transcripts in breast cancer cell lines and 13% of transcripts in melanoma cell lines include the exon-6. This analysis suggests the existence of KDM5B variants with exon-6 also in breast cancer cells, as already observed in melanoma cells⁵⁰.

To detect splicing variants of KDM5B in breast cancer cells, we performed RT-PCR experiments using a primer pair (NM primers in Table S1) to amplify the region between exon-4 and exon-7. In that way, PCR fragments of different lengths are generated depending on the inclusion of exon-5 and exon-6 (Fig.S1, panel A). We compared the pattern of KDM5B variants in luminal and basal breast cancer cell lines to other reference and normal-control cell lines, respectively SK-MEL-28 melanoma cell line and Epstein-Barr virus (EBV) immortalized lymphocytes and peripheral blood mononuclear cells (PBMC). RT-PCR results confirm the existence of different KDM5B splicing variants and suggest that a similar expression pattern characterizes both the melanoma cell line and the three analyzed breast cancer cell lines (Fig.S1, panel B). KDM5B transcripts including exon-6 were detected in all the

analyzed cancer cell lines and appeared barely detectable in human PMBC or EBV-immortalized lymphocytes.



Alternative transcript variants from GFF annotation

Fig.4 Sashimi plot. Sashimi plot showing the average level of exon-6 inclusion in KDM5B transcripts in 49 melanoma cell lines (purple) and 56 breast cancer cell lines (red). The General Feature Format (GFF) annotation of the alternative exon-6 splicing variants is shown below the sashimi plot.

3.2. Although the KDM5B transcripts level is significantly higher in MCF7 compared to MDA-MB-231, the fraction of exon-6 including transcripts is similar in the two cell lines

Previous research demonstrated KDM5B transcripts level was lower in MDA-MB-231 triple-negative breast cancer cells, compared to luminal breast cancer cells MCF7^{20,29}. However, only the total level of KDM5B transcripts was considered. Thus, we quantified through RT-qPCR the level of KDM5B transcripts with exon-6 by comparing it with the total KDM5B transcripts level. Using a primer pair spanning in the region of exon-8 and exon-9 (ALL KDM5B primers in Table S1), we determined the total KDM5B transcripts level. Using a primer pair with a forward primer on exon-5 and a reverse primer specific for exon-6 (KDM5B Exon-6 primer pair in Table S1), we quantified the subgroup of KDM5B transcripts with exon-6. In line with Klein et al. (2014)²⁹, we found a significantly increased level of total KDM5B mRNAs in MCF7 versus MDA-MB-231 (Fig.5A), supporting the idea of KDM5B as a driver oncogene in the expression programs of luminal cells^{15,41}. On the contrary, no significant difference was found in the level of KDM5B transcripts with exon-6 between these two cell lines (Fig.5B), and the percentage of the KDM5B transcripts including the exon-6 was approximately 4-9%. A large-scale bioinformatic analysis on publicly available data from the Cancer Cell Line Encyclopedia project (CCLE, Broad Institute) and The Cancer Genome Atlas (TCGA) confirmed the level of KDM5B transcripts which include the exon-6 is relatively low (<10%) and constant in all the analyzed cancer cell lines and patients' samples with few exceptions (Fig.S2).

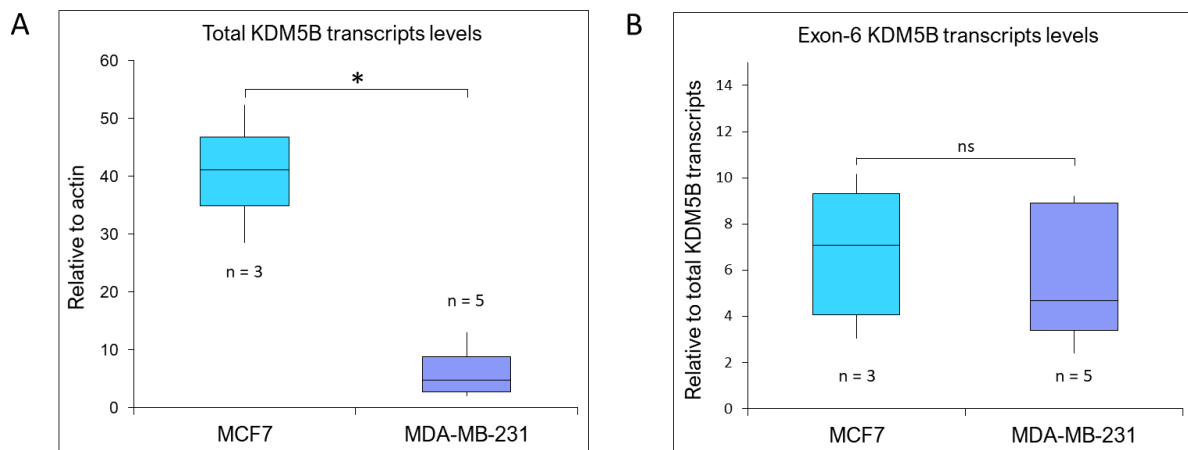


Fig.5 RT-qPCR KDM5B transcript levels in MCF7 and MDA-MB-231. A. The total KDM5B transcripts level in MCF7 is significantly higher compared to MDA-MB-231. **B.** The fraction of KDM5B transcripts including the exon-6 is similar in MCF7 and MDA-MB-231 cells. ns: $p > 0.05$, * $p \leq 0.05$

3.3. Western blot analyses suggest that a predicted N-terminal truncated KDM5B isoform including exon-6 is expressed in breast cancer cell lines

A comparable level of KDM5B transcripts with exon-6 was detected in MCF7 and MDA-MB-231, despite the considerably different levels of total KDM5B transcripts. However, transcription and translation don't have a linear and simple relationship and the existence of a transcript doesn't guarantee the expression of the corresponding protein product⁵⁴. In light of that, we performed an explorative Western blot in MCF7 cells, using two different primary antibodies previously described by Kuźbicki et al. (2016)⁴⁹. The first antibody (Ab1) allows the detection of all the possible KDM5B isoforms; the second one (Ab2) detects the KDM5B isoforms which include the exon-6 encoded residues; indeed, Ab1 binds the C-terminal portion of KDM5B that is included in all the isoforms, whereas Ab2 is specific for the residues encoded by exon-6 (Fig.3B). Interestingly, using the primary antibody against the C-terminal region of KDM5B isoforms (Ab1) we found two main bands, with an apparent molecular weight of approximately 175kDa (corresponding to KDM5B-PLU-1) and 162kDa, whereas no band was detected around 180kDa (corresponding to RBP2-H1 isoform with exon-6, the isoform previously reported in melanoma^{50,55}). Using instead the antibody specific for the exon-6 region (Ab2) we found the same band at 162kDa; once again no band was

detected around 180kDa (Fig.6A), as expected for RBP2-H1 (Fig.6B).

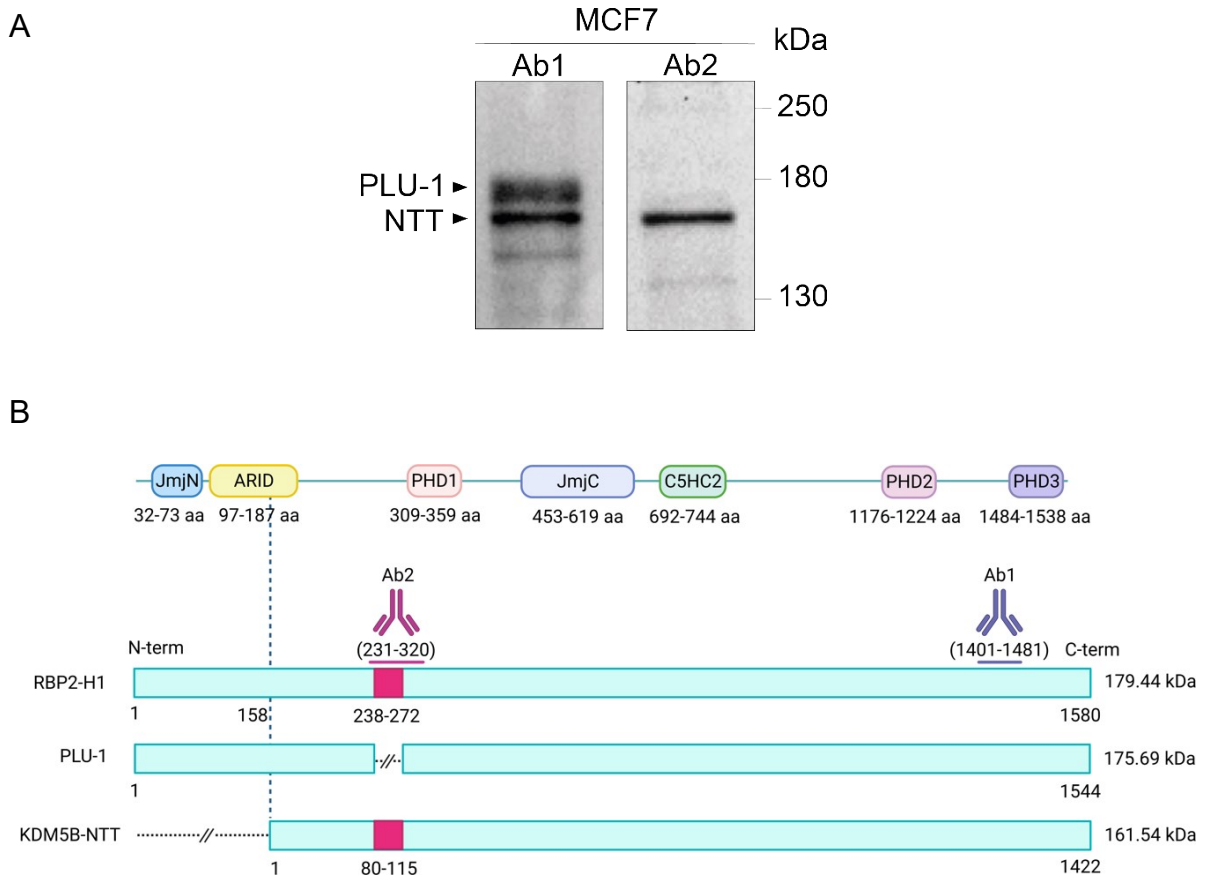


Fig.6 KDM5B isoforms in MCF7 cells. A. WB of KDM5B isoforms in MCF7 cells confirm the expression of PLU-1 (175kDa) and NTT (162kDa). On the left, the detection using Ab1 against the C-terminal region shows both the isoforms; on the right, the blot done after stripping, using Ab2, specific for the residues encoded by the exon-6, shows only the band at 162kDa. **B.** Schematic of KDM5B protein domains and comparison of RBP2-H1, PLU-1 and KDM5B-NTT isoforms. The band detected at 162kDa using both the antibodies (Ab1 and Ab2) corresponds to the predicted N-terminal truncated isoform of KDM5B including the alternative exon-6 encoded residues.

Both the antibodies recognized the same band at 162kDa, thus the corresponding KDM5B isoform includes both the exon-6 encoded residues and the C-terminal region but seems to be smaller than KDM5B-PLU1 (175kDa), despite the insertion of additional residues. According to the NCBI Reference Sequence database, the protein we detected at 162kDa may correspond to a predicted KDM5B isoform that is N-terminal truncated and with the exon-6 encoded residues (Fig.6B), (hereafter called KDM5B-NTT). To verify if the expression of this putative isoform interested other cancer cells, we extended the explorative western blot analysis to other cell lines, such as the luminal T47D and the triple-negative MDA-MB-231 breast cancer

cell lines, and the SK-MEL-28 human melanoma cell line, compared to the testis, as a normal adult tissue reference control. We found that all tumor cell lines express both KDM5B isoforms (PLU-1 and NTT), with the MDA-MB-231 cell line which seems predominantly express KDM5B-NTT, that was instead barely detected in testis (Fig.7, upper panel, Ab1). Once again, hybridization using Ab2 after membrane stripping proved that the exon-6 encoded residues are included in the smaller isoform (Fig.7, lower panel, Ab2).

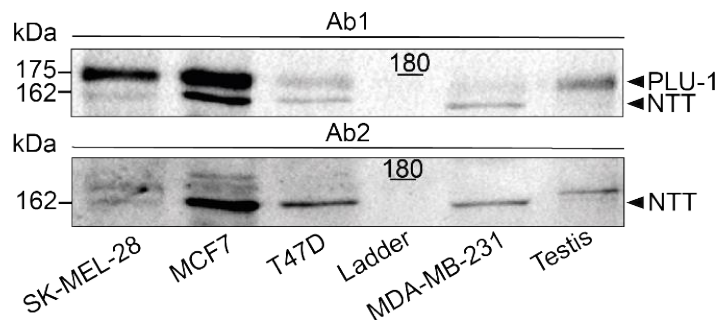


Fig.7 Western blot KDM5B isoforms. Preliminary investigation of KDM5B isoforms in different human cancer cell lines, compared to the testis, the only normal adult tissue showing a high expression level of KDM5B. Ab2 (lower panel) was used after the stripping of the membrane hybridized with Ab1 (upper panel). All tumor cell lines express both KDM5B isoforms (PLU-1 and NTT), with the MDA-MB-231 cell line which seems predominantly express KDM5B-NTT, that was instead barely detected in testis.

3.4. 5'-RLM-RACE demonstrates the presence of a new downstream transcriptional start site, compatible with the expression of a truncated and catalytically inactive isoform of KDM5B

A recent analysis of full-length mRNAs in the breast cancer MCF-7 cell line suggests that transcription start site and splicing pattern may be coordinated: it is indeed becoming increasingly clear how splicing and transcription influence each other⁵⁶⁻⁵⁸. According to the NCBI Reference Sequence database, the truncated KDM5B isoform including exon-6 might be generated by two predicted transcripts of KDM5B, namely X1 and X2. It is particularly noteworthy these predicted variants differ from the PLU-1 mRNA for the transcriptional start site and for the inclusion of the exon-6 (Fig.3A). The translation of these predicted transcript variants does not start from the first exon, but from the fourth, producing the predicted isoform N-terminal truncated

(KDM5B-NTT in Fig.6B). X1 and X2 transcripts differ from each other only at the 5'-untranslated region (5'UTR), therefore we used the RNA Ligase Mediated Rapid Amplification of 5' cDNA Ends (5'-RLM-RACE) method in MCF7 cells to verify the existence of a downstream transcriptional start site (TSS) which could generate transcripts able to code for the predicted protein isoform. This 5'RACE method is designed to amplify cDNA only from full-length capped mRNA: indeed, Calf Intestine Alkaline Phosphatase (CIP) removes free 5'-phosphates from molecules such as ribosomal RNAs, fragmented mRNAs, tRNAs, and contaminating genomic DNA; the Tobacco Acid Pyro-phosphatase (TAP) removes the cap structure from full-length mRNA, leaving a 5'-monophosphate; after de-capping reaction, an RNA Adapter Oligonucleotide is ligated to the 5'-end of the mRNAs population using T4 RNA Ligase, excluding the dephosphorylated RNA molecules which lack the 5'-phosphate necessary for ligation; after a random-primed reverse transcription reaction, two nested PCR amplifies the 5' end of a specific transcript. In the nested PCRs the forward primers are always complementary to the 5'-anchor region, while the reverse primers are specific for the target transcripts (see Table S1 for primer sequences). Since western blot results showed two predominant bands (around 175kDa, corresponding to the well-known canonical PLU-1 isoform and around 162kDa, corresponding to the predicted isoform including the exon-6), we decided to focus our investigation on transcript variants with exon-6. Therefore, in the first nested PCR we used a reverse primer on exon-6 (Reverse Ex-6) and in the second nested PCRs we used reverse primers (either Reverse 1 or Reverse 2) as described in Fig.8A. The final PCR products produced using the two different reverse primers (Reverse 1 for PCR1 and Reverse 2 for PCR2) are shown in Fig.8B. The sequencing of both PCR products revealed a new transcriptional start site (TSS), specifically 369 nucleotides before the second translation start site. We found a perfect match with the sequence of KDM5B (see complete alignment in Supplementary File S1), surprisingly starting at exon-1 (Fig.8C-D) instead of intron-1, as expected for the KDM5B predicted transcripts. Nevertheless, this new TSS skips the first ATG in exon-1 and is compatible with the expression of KDM5B-NTT isoform translated by the second ATG in the exon-4 (Fig.8C). By integrating the 5'RACE results and the Western blot data, we found a new KDM5B transcript variant that correlates with the expression of the N-terminal truncated isoform of KDM5B (KDM5B-NTT) which includes the insertion of 36 residues encoded by the exon-6 after the Glu237 of PLU-1 isoform

(Fig.6B). Supporting this finding, a putative active promoter between the first and the second ATG was found using a bioinformatic tool based on ChIP-Seq data of transcription factors occupancy and histone modifications (Fig.S3)^{33,41}.

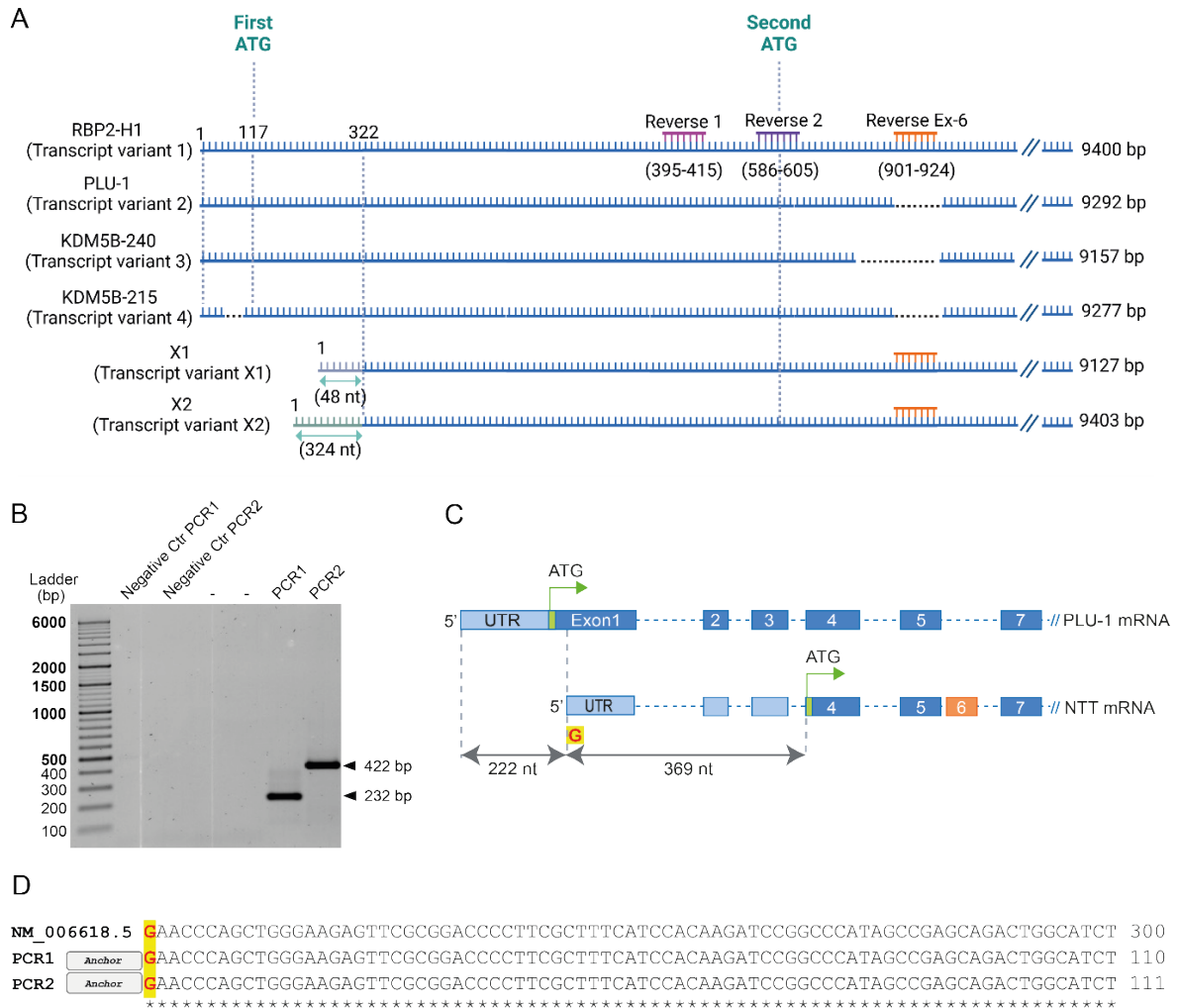


Fig.8 5'RLM-RACE of KDM5B transcripts. A. Schematic representation of reverse primers used for the nested PCRs of 5'RLM-RACE. Primer-binding nucleotides are indicated underneath the RBP2-H1 transcript. In the first nested PCR, a forward primer (not shown) complementary to the 5'-anchor region was used together with the reverse primer on the exon-6 region (Reverse Ex-6). In the second nested PCRs, an inner forward primer (not shown) complementary to the 5'-anchor region was used together with either Reverse 1 (PCR1) or Reverse 2 (PCR2). **B.** PCR products of PCR1 and PCR2 are shown. After purification, both PCR products were sequenced. **C.** Schematic of the 5' end of the KDM5B-NTT transcript compared to the 5' end of the KDM5B-PLU1 transcript. We found a new transcriptional start site 222 nucleotides downstream of the canonical one. The position of the translation start sites (ATG) in exon-1 and exon-4 are shown. **D.** Partial alignment sequences of the two PCR products, using the PLU-1 sequence (NM_006618.5) as a reference. For a complete alignment see Supplemental File 1.

3.5. MCF7 and MDA-MB-231 breast cancer cell lines show a different relative amount of KDM5B-NTT with respect to the PLU-1 isoform

The preliminary western blot analyses (Fig.7) suggested differences in KDM5B-PLU-1 and KDM5B-NTT expression in different breast cancer cell lines. As previously demonstrated^{32,35}, the catalytic core formation of KDM5B requires the tight association between the JmjN domain and jmjC domain, together with seven residues of the ARID domain (residues 94–100)³⁵. The demethylase function of KDM5B is noted to regulate transcription in luminal breast cancer cell lines, acting as a driver oncogene⁴¹. KDM5B negatively regulates the levels of many tumour suppressor genes, such as BRCA1 and Cav-1¹⁷. This demethylase is indeed a therapeutic target and many small anti-cancer molecules targeting KDM5B were identified so far⁵⁹. KDM5B inhibitors are especially used as a therapeutic treatment in luminal breast cancer²². In this regard, the KDM5B-NTT isoform is interesting because doesn't include the whole JmjN domain and part of the ARID domain (Fig.6B), resulting in a catalytically inactive protein. Therefore, we questioned the potential relevance of the expression of this catalytic domain-lacking protein in different breast cancer subtypes.

Analyzing all the molecular features of currently available breast cancer cell lines, Dai et al. (2017)⁶⁰ found a correspondence with breast cancer tumours. Luminal tumors have a better prognosis and luminal cell lines are less aggressive, while triple-negative tumors have poor prognoses and basal breast cancer cell lines are more aggressive (Fig.S4). To test the expression of these two isoforms in different subtypes of breast cancer, we analyzed the relative abundance of NTT/PLU-1 in breast cancer cell lines classified as luminal, HER2+, triple negative A (TNA) and triple negative B (TNB). Fig.9A shows that all the breast cancer cell lines express KDM5B-NTT, detected using both Ab1 and Ab2, and KDM5B-PLU1, without a clear correlation with the cancer subtype classification (Fig.9B). However, KDM5B-NTT is significantly higher expresses in MDA-MB-231 (0.57 average) than in MCF7 (0.40 average) (Fig.9B; Fig.10; Fig.S5). The significant increased expression of KDM5B-NTT isoform compared to PLU-1 in MDA-MB-231 versus MCF7 cells (Fig.10) is interesting because of the different functional roles of KDM5B in these cell lines²⁹. In MDA-MB-231 KDM5B inhibits cell growth and metastatic and angiogenic capability³³, whereas in MCF7 it enhances cell proliferation^{15,41}.

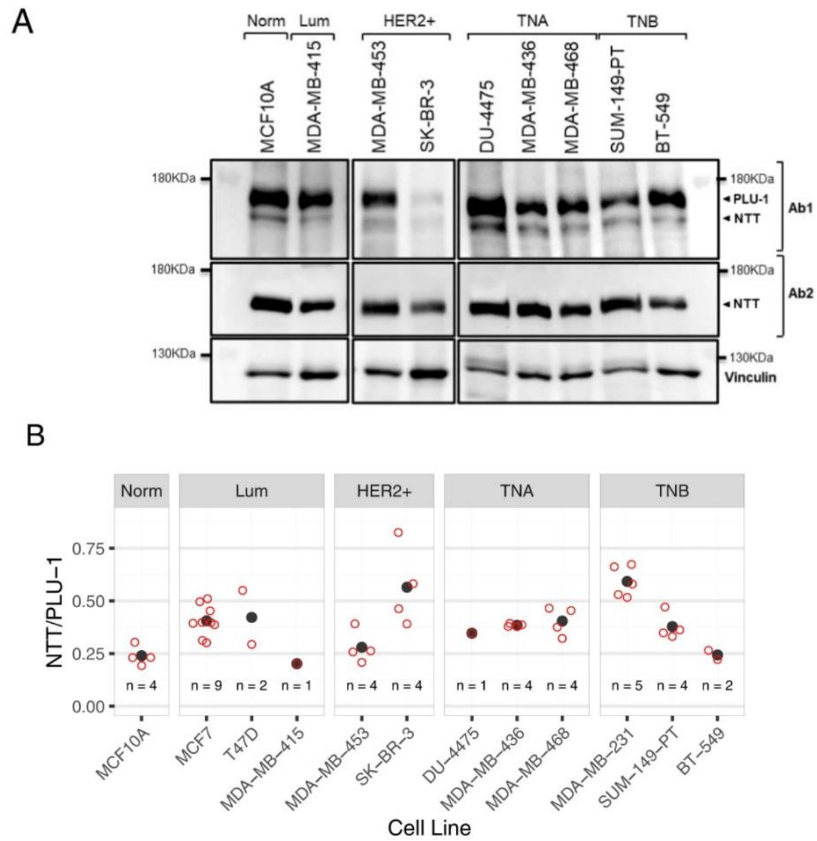


Fig.9 KDM5B isoforms expression in different breast cancer subtypes. A. In all the analyzed breast cancer cell lines we found the expression of KDM5B-PLU-1 and KDM5B-NTT isoforms. (MCF7 and MDA-MB-231 blots are shown if Fig.S5) **B.** No clear association between breast cancer subtypes and relative KDM5B-NTT expression was highlighted. The relative NTT abundance was calculated as NTT/PLU1 ratio of densitometric values of the hybridization with the anti-C-terminal primary antibody (Ab1).

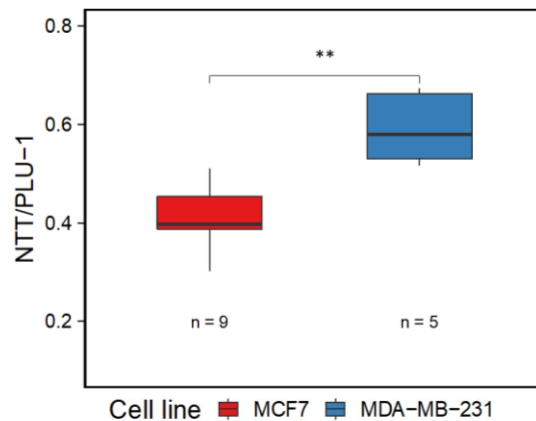


Fig.10 Relative expression of KDM5B-NTT in MCF7 and MDA-MB-231 cells. Direct comparison with the data reported in Fig.S5. The NTT/PLU-1 ratio is higher in MDA-MB231 versus MCF7, around 50-60% and 30-40% respectively. ** $p \leq 0.01$

Although the total KDM5B transcripts level was significantly higher in MCF7 compared with MDA-MB-231 cells (Fig.5A), we found a similar percentage fraction of KDM5B transcripts with exon-6 over the total pool of KDM5B transcripts in these two cell lines (Fig.5B). Moreover, the KDM5B-NTT expression quantified by western blot analyses showed a significant higher relative abundance in MDA-MB-231 cells (Fig.10). Therefore, we raised two questions:

- 1) how a subgroup representing less than 10% of the total transcripts (Fig.5B) can produce a much higher protein product (around 40% in MCF7 and 57% in MDA-MB-231) (Fig.10);
- 2) why the relative abundance of the KDM5B-NTT isoform is different in the two cell lines if the transcripts that generate this isoform represent a similar fraction of the total transcripts in both cell lines (Fig.5B and Fig.10);

We reasoned that the different relative amounts of KDM5B-PLU-1 and KDM5B-NTT isoforms may be not related to transcriptional regulation, but rather to post-translational regulation.

3.6. Cycloheximide treatment highlighted the greater instability of PLU-1 compared to the KDM5B-NTT isoform

No correspondence between the relative expression of KDM5B-NTT (compared to KDM5B-PLU-1) and the level of exon-6 including transcripts was found. Often the individual half-life of proteins is the major factor that regulates the relationship between mRNA molecules and protein expression levels⁵⁴, therefore we decided to investigate a possible regulation at the protein stability level of the two isoforms. The rates of protein synthesis and protein degradation determine cellular protein abundance and many factors can influence the lifetime of a protein, such as intrinsic protein stability, post-translational modifications (like phosphorylation and ubiquitination), subcellular localization and the first amino-terminal residue (N-end rule)^{54,61}. We analyzed the protein stability of KDM5B-PLU-1 and KDM5B-NTT using a cycloheximide (CHX) assay to block the protein synthesis⁶². Upon the CHX treatment, we used the Ab1 (anti-C-term antibody) to compare the stability of the two isoforms in a time course of

12 hours (0h - 12h, every two hours). The results of the time course of CHX treatment show clearly that the NTT isoform is much more stable than the PLU-1 isoform in both the analyzed breast cancer cell lines (Fig.11). As shown in Fig.11C-D, at 12h upon CHX treatment, in both cell lines KDM5B-PLU-1 level is less than 20%, whereas KDM5B-NTT rises more than 60%, compared to control (DMSO-treated cells). KDM5B-NTT is very stable for at least 12 hours, as confirmed using the antibody Ab2 (Fig.11B). Hence, the observed discrepancy between the expression of KDM5B-NTT at the protein level and the low level of exon-6 including transcripts may be explained by considering that the NTT isoform is highly stable. Moreover, the rate of PLU-1 degradation is faster in MDA-MB-231, with a half-life of around two hours (Fig.11D and Fig.12), rather than in MCF7 cells, with a half-life of about four hours (Fig.11C and Fig.12). Overall, these results show that the different relative amount NTT/PLU-1 in MCF7 and MDA-MB-231 depends on different protein stability and turn-over rate of these two isoforms.

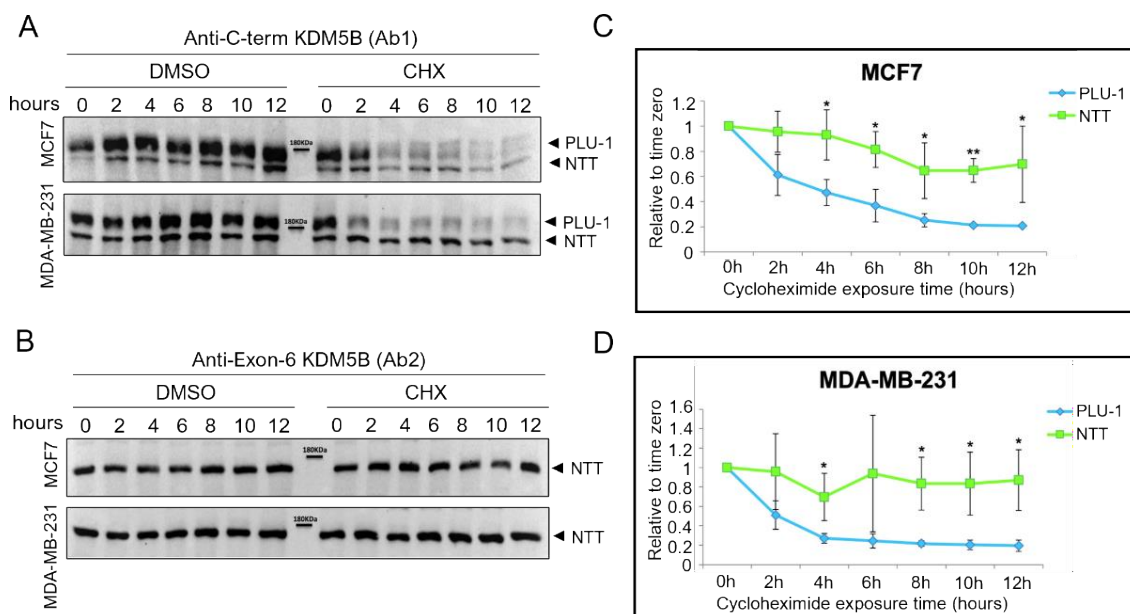


Fig.11. Protein stability of KDM5B isoforms. **A.** After CHX treatment, the KDM5B-NTT results more stable than KDM5B-PLU-1 in both MCF7 and MDA-MB-231, as shown in the blots using the anti-C-terminal KDM5B. **B.** The hybridization with Ab2 shows that KDM5B-NTT is highly stable in both cell lines for at least 12 hours. **C.** Degradation kinetics of KDM5B isoforms in a time course in MCF7 cells; (n = 3 for each time point). **D.** Degradation kinetics of KDM5B isoforms in a time course in MDA-MB-231 cells; (n = 3 for each time point). ns: p > 0.05, *p < 0.05, **p < 0.01, ***p < 0.001; Statistics by unpaired Welch's t-test.

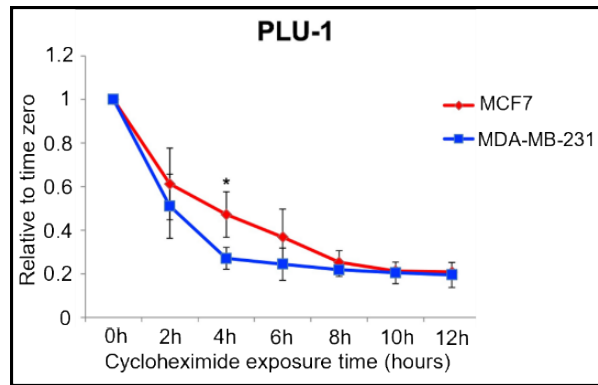


Fig.12. Direct comparison of PLU-1 degradation kinetics in MCF7 and MDA-MB-132 cells. By comparing the degradation curves of PLU-1 isoform in the two analyzed breast cancer cell lines, it was found a different rate of degradation, that is significantly faster in MDA-MB-231, where the mean half lifetime of the PLU-1 is of about 2h, versus MCF7, where instead it is around 4h; (n = 3 for each time point). ns: $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Statistics by unpaired Welch's t-test.

3.7. The proteasomal degradation activity regulates PLU-1 turn-over

To verify whether the proteasomal degradation was the cause of PLU-1 turn-over, we performed experiments with the 26S proteasome inhibitor MG132⁶³. This inhibitor considerably counteracted protein degradation following CHX treatment, demonstrating that the proteasomal activity regulates PLU-1 stability in both cell lines (Fig.13).

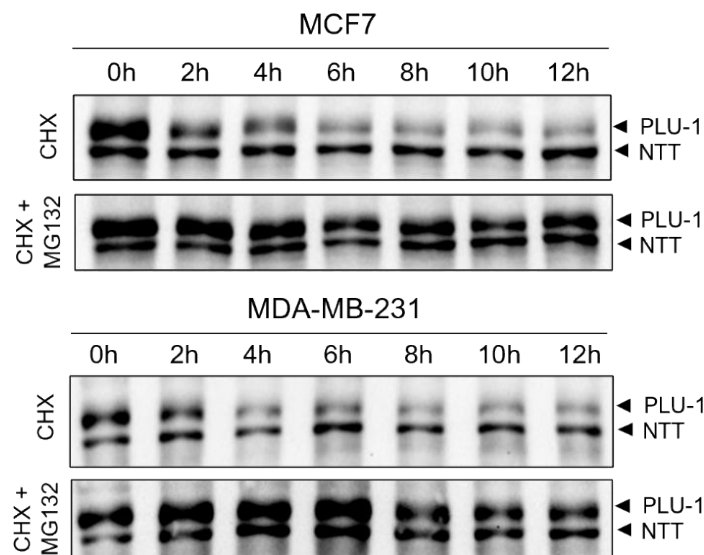


Fig.13. Proteasomal regulation of KDM5B-PLU-1. The cycloheximide (CHX) treatment highlights the PLU-1 instability, which is instead rescued by MG132 treatment, demonstrating that the proteasomal activity regulates the turn-over of PLU-1 isoform in both cell lines.

Indeed, previous observations suggested that the triple-negative breast cancer cells MDA-MB-231 might be particularly addicted to the 26S proteasome level⁶⁴. The greater proteasomal activity in MDA-MB-231, combined with the lower PLU-1 expression in MDA-MB-231 compared to MCF7, and with the higher stability of NTT compared to PLU-1 may explain the significantly higher relative abundance ratio NTT/PLU-1 in MDA-MB-231 cells. For time being, we don't know why PLU-1 and NTT isoforms have so remarkably different stability. However, our hypothesis is that PLU-1 should have something that is missing in NTT and that specifically induces its proteasomal degradation. Reasoning on the possibility of determinants in the first 158 residues at the N-terminal of PLU-1, that are missing in the shorter NTT isoform, we found a *bona fide* degron protein motif (Fig.14 and Fig.S6). Moreover, Bueno et al. (2013)⁶⁵ showed that the proteasomal-dependent degradation of KDM5B is mediated by the E3 SUMO-dependent ubiquitin ligase RNF4. RNF4 targets KDM5B to proteasomal degradation upon the SUMOylation at Lysine residues 242 and 278 (K242 and K278) by the SUMO E3 ligase human polycomb protein 2 (hPC2) (also commonly referred to as CBX4)⁶⁵. Strikingly, both K242 and K278 map to a disordered region of PLU-1, whereas in the predicted 3D protein structure of KDM5B-NTT the first one lysine (corresponding to K120 in NTT) appears less accessible as shown in Fig.14. To assess if the N-terminal residues which are missing in KDM5B-NTT play specifically a role in targeting PLU-1 to proteasomal degradation targeted experiments are needed. Therefore, we are currently working to the cloning of the first 158 residues of the N-terminal of PLU-1 at the N-term of the reporter gene GFP. Later new experiment with cycloheximide and MG132 will be performed using as a control the GFP without the N-terminal residues of PLU-1.

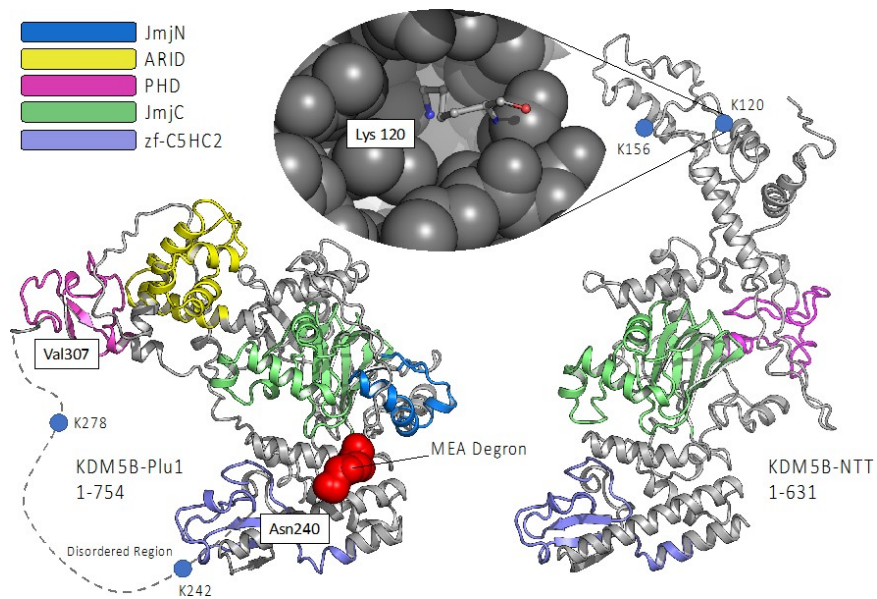


Fig.14. AlphaFold2 models of KDM5B-Plu1 (1-754; left) and the corresponding N-terminal region of KDM5B-NTT (1-631; right). The positions of the JmjN, ARID, PHD, JmjC and zf-C5HC2 domains were obtained by PFAM, mapped on the corresponding 3D-structures and colored according to the legend shown. The region 740-307, comprising the ubiquitylated residues K242 and K278 (blue spheres) of KDM5B-Plu1, is predicted as intrinsically disordered and shown as dashed lines. The corresponding residues of NTT, K120 and K156, are predicted on a folded domain. In particular K120 is predicted as buried and not accessible for ubiquitinylation. Sequence analysis with the ELM server indicates that Plu-1 has an N-terminal Degron which is missing in NTT ("MEA"). These structural differences may help partly explain the increased stability of NTT over Plu-1.

Part 2 – Investigating the putative role of KDM5B-NTT isoform in breast cancer

The results shown so far demonstrate that a new KDM5B isoform including exon-6, namely NTT, is expressed in breast cancer cells. KDM5B-NTT is a truncated and catalytically inactive isoform generated by an alternative transcript that includes the exon-6 and is translated from a downstream ATG. KDM5B-NTT is almost constantly expressed in different cancer cell lines and patients' samples. No correlation between breast cancer subtypes and the expression of KDM5B-NTT was highlighted, however, the relative ratio NTT/PLU-1 resulted significantly higher in MDA-MB-231 cells rather than MCF7 cells. The relative abundance of NTT over PLU-1 was higher in MDA-MB-231 than MCF7 despite the similar level of KDM5B-NTT transcripts compared to the total level of KDM5B transcripts in the two cell lines. We observed a different protein stability of PLU-1 and NTT isoforms, indicating that KDM5B-NTT is mainly regulated at the protein level and is much more stable than PLU-1, which is targeted by proteasomal degradation. This suggests that KDM5B-NTT can accumulate in cells. KDM5B-NTT is an interesting isoform because differs from PLU-1 for the presence of 36 additional residues and for the lack of 158 residues at the N-terminal, as shown in the sequence alignment in Fig.S6. KDM5B-NTT is a loss-of-function isoform because is a catalytic JmjN domain-lacking protein; however, at the same time, the NTT isoform may become a gain-of-function for the presence of additional residues not included in the canonical PLU-1 isoform (Fig.6B). To investigate the putative function of KDM5B-NTT in breast cancer cells we decided to use two complementary approaches: downregulate specifically KDM5B-NTT in MDA-MB-231 cells, where the NTT/PLU-1 ratio was greater than in MCF7 cells; on the opposite, over-express NTT in MCF7 cells, where the NTT/PLU-1 ratio was lower compared to MDA-MB-231 cells.

3.8. RNA interference and transient over-expression as tools to investigate the function of KDM5B-NTT in breast cancer cells

To downregulate the KDM5B-NTT isoform in MDA-MB-231 cells we designed a customized siRNA specific for exon-6 containing transcripts (Fig.S7). A small fluorescent oligo RNA was used as a transfection indicator (Fig.S8). Performing RNAi experiments at different time points and using different final concentrations of siRNA, we, unfortunately, observed that the interference was reducing not only the level of NTT protein isoform (Fig.15) but also the level of PLU-1 (Fig.16).

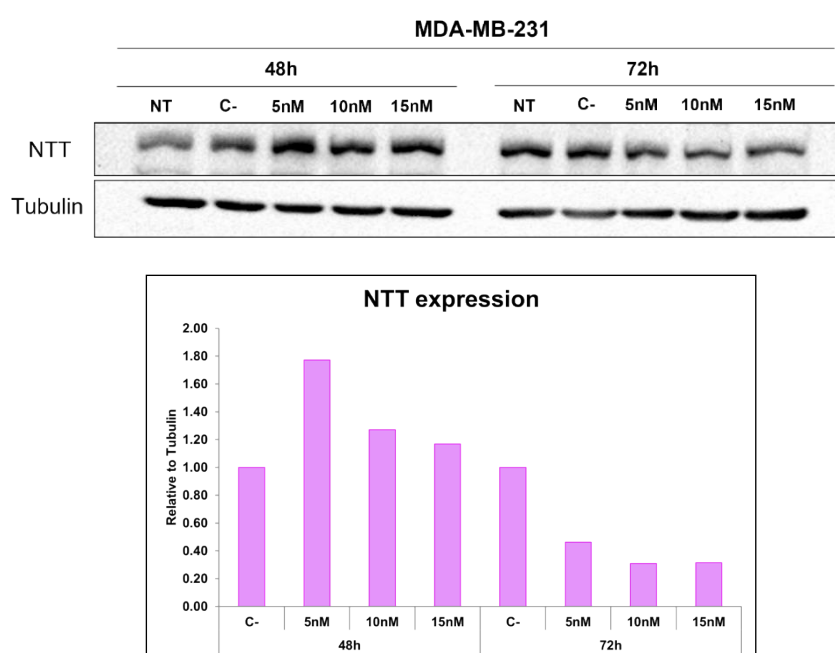


Fig.15. RNAi for the downregulation of KDM5B-NTT in MDA-MB-231 cells. Western blot of preliminary RNAi experiments showed that the siRNA reduced the level of NTT (detected using the Ab2 specific for the exon-6 encoded residues), especially at 72h at the final concentrations of 10nM and 15nM, as shown in the quantification at the bottom panel (direct comparison with the blot in the upper panel); the NTT expression level relative to tubulin of each final siRNA concentration (5nM, 10nM and 15nM) was compared to the negative control condition (C-). N=1 for each different condition.

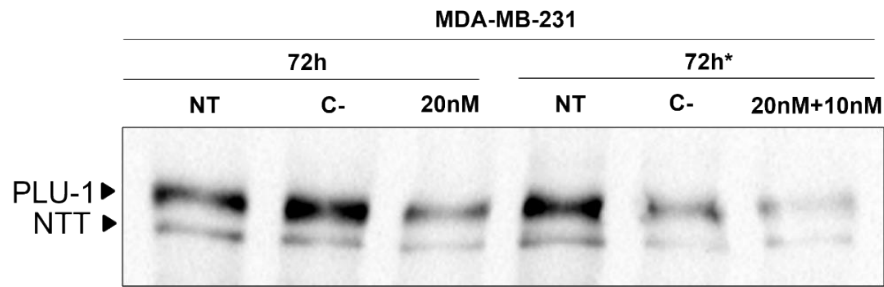


Fig.16. The siRNA used for the depletion of KDM5B-NTT affected PLU-1. Using the Ab1 against the C-terminal region we observed that the siRNA targeting the exon-6 region was also reducing the PLU-1 levels, as evident in the time point 72h* (where * indicates a double pulse of siRNA transfection; the first pulse was done 24h after the seeding of cells and the second pulse 24h after the first pulse)

RT-qPCR confirmed that the interference severely reduces the exon-6 containing transcripts, but it affects the total mRNA level more than you would expect if the interference was limited to it (Fig.17). This effect may be due to the nuclear activity of our siRNA on mRNA precursor (pre-mRNA). Indeed, RNA interference can occur also in the nucleus of mammalian cells⁶⁶. In light of these preliminary results, we decided to pause this path, focusing instead on the over-expression strategy. We're working to make the downregulation more specific for NTT, designing two new siRNAs spanning the junctions between exon-5 and exon-6 and between exon-6 and exon-7, to act exclusively on already spliced transcripts.

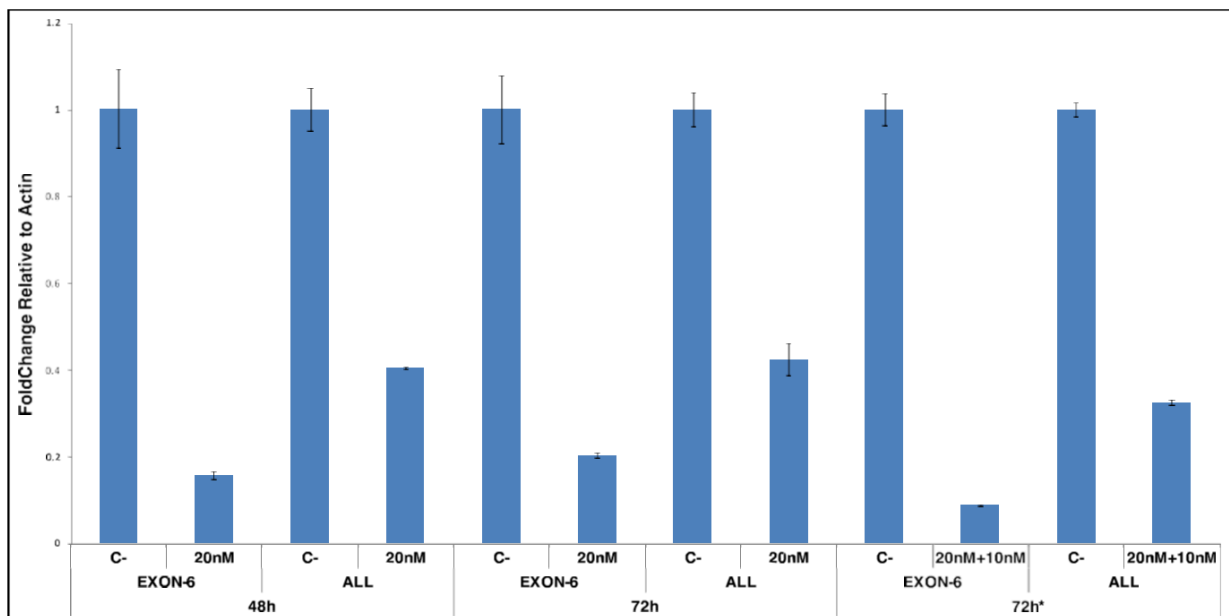


Fig.17. RT-qPCR analyses showed an effect on total KDM5B mRNA level larger than expected if the interference was limited to exon-containing transcripts. Upon three different conditions of RNAi, the reduction of the level of transcripts with the exon-6 (EXON-6 at 48h, 72h, and 72h*) was appreciable compared to the negative control (C-); however, the reduction level of the total KDM5B transcripts (ALL) was much more than expected in each RNAi condition considering that the percentage of spliced KDM5B mRNA with exon-6 is around 4-9% of the total KDM5B transcripts (see Fig.5B).

We continued the investigation of the potential role of NTT through the over-expression of NTT in both MCF7 and MDA-MB-231. Both cell lines were transiently transfected to over-express KDM5B-NTT. The transfection efficiency was determined in each experiment using as a positive transfection control the over-expression of GFP as shown in Fig.S9. NTT isoform was efficiently over-expressed in both cell lines, as visible by the increased signal of the 162kDa band at 24 or 48 hours after transfection (Fig.18 and Fig.S10). At 24h after transfection, KDM5B-NTT increased by 2-3 fold in MCF7 cells and 4-8 fold in MDA-MB-231 cells compared to the empty control vector. In the transfected cells, especially at 24h post-transfection, NTT became more prevalent than PLU-1, being approximately 2- or 5-fold more abundant than PLU-1 in MCF7 and MDA-MB-231 cells, respectively (Fig.S10).

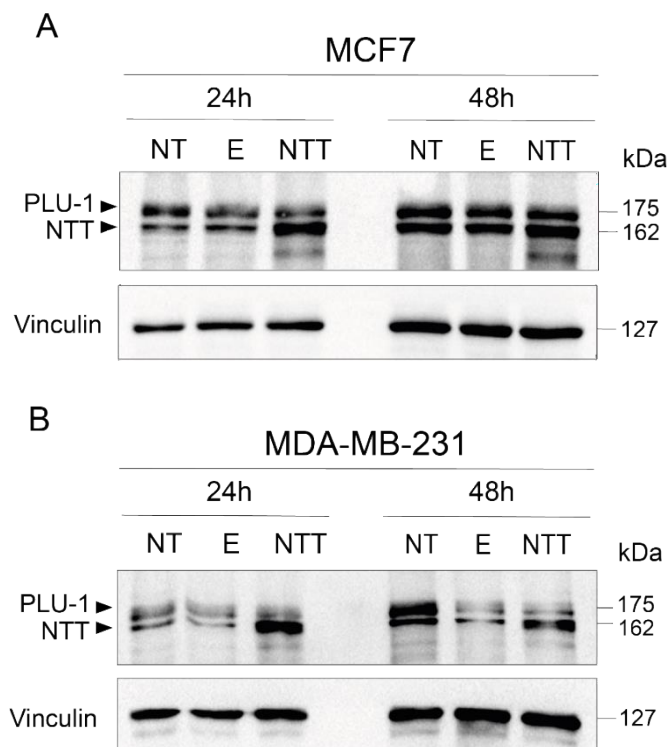


Fig.18. Over-expression of KDM5B-NTT. A. The over-expression of KDM5B-NTT (NTT) at 24h and 48h after transient transfection in MCF7 cells, NT indicates not transfected cells; E indicated transfection using the empty vector as a negative control. **B.** The over-expression of KDM5B-NTT (NTT) at 24h and 48h after transient transfection in MDA-MB-231 cells, NT indicates not transfected cells; E indicated transfection using the empty vector as a negative control.

Although catalytically inactive, KDM5B-NTT retains all the nuclear localization signals reported for PLU-1 (Fig.S11), therefore it should also maintain the ability to enter the nucleus and associate to chromatin by the PHD domains^{29,33-35}. By analyzing the localization of PLU-1 and KDM5B-NTT in the nuclear and cytoplasmic fractions of MDA-MB-231 cells (Fig.19), we confirmed that both isoforms are highly enriched in the nuclear fraction. It is thus possible that KDM5B-NTT may interfere with the nuclear functions of KDM5B-PLU-1. For that reason, we tested the effect of KDM5B-NTT over-expression on histone methylation.

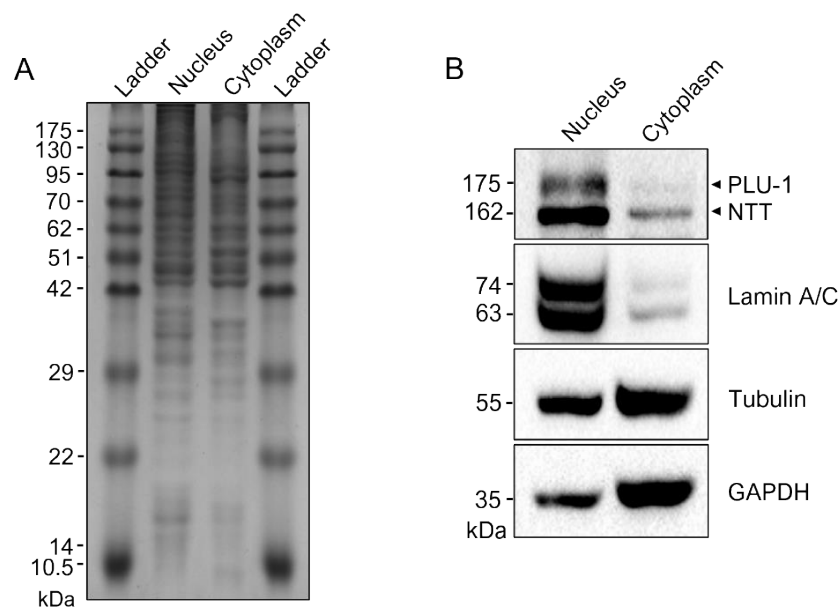


Fig.19. Nuclear localization of KDM5B-NTT. **A.** Coomassie blue after fractionation nucleus/cytoplasm in MDA-MB-231 cells and SDS-PAGE in 12% acrylamide gel. The same amount of nuclear and cytoplasmic lysates was loaded to verify the localization of KDM5B-NTT. **B.** Both NTT and PLU-1 are enriched in the nucleus. Lamin A/C, Tubulin and GAPDH were used as references for the nuclear and cytoplasmic proteins. The SDS-PAGE for the KDM5B isoforms was performed in a 5% gel (37.5:1 Acrylamide - Bis-Acrylamide solution), while the SDS-PAGE to detect all the other targets was done using a 12% gel (29:1 Acrylamide - Bis-Acrylamide solution).

3.9. KDM5B-NTT affects the global H3K4me3 level

After NTT over-expression we performed a mass-spec analysis to quantify the H3K4 mono-, di- and tri-methylation level (Fig.20) at 24h upon transfection when the level of KDM5B-NTT was generally higher (Fig.18 and S10). The basal H3K4 di- and tri-methylation level in not transfected cells (NT) was higher in MDA-MB-231 cells than in MCF7 cells (NT condition in Fig.20A). At 24h after NTT over-expression, we found a significant increase of H3K4me3 level in MCF7 cells (Fig.20C), that reached the level

quantified in MDA-MB-231 cells, where no significant changes in trimethylation of H3K4 were detected. In both cell lines, no significant change levels of H3K4me2 or H3K4me1 were observed (Fig.20A-B). The increased H3K4me3 level in MCF7 cells after KDM5B-NTT overexpression may be explained by a competition effect exerted by the catalytically inactive isoform on PLU-1 and eventually other KDM5 histone demethylase family members. In MDA-MB-231 cells no changes in H3K4me3 level were observed, probably because of the higher basal level of H3K4 tri-methylation, likely due to lower PLU1 abundance.

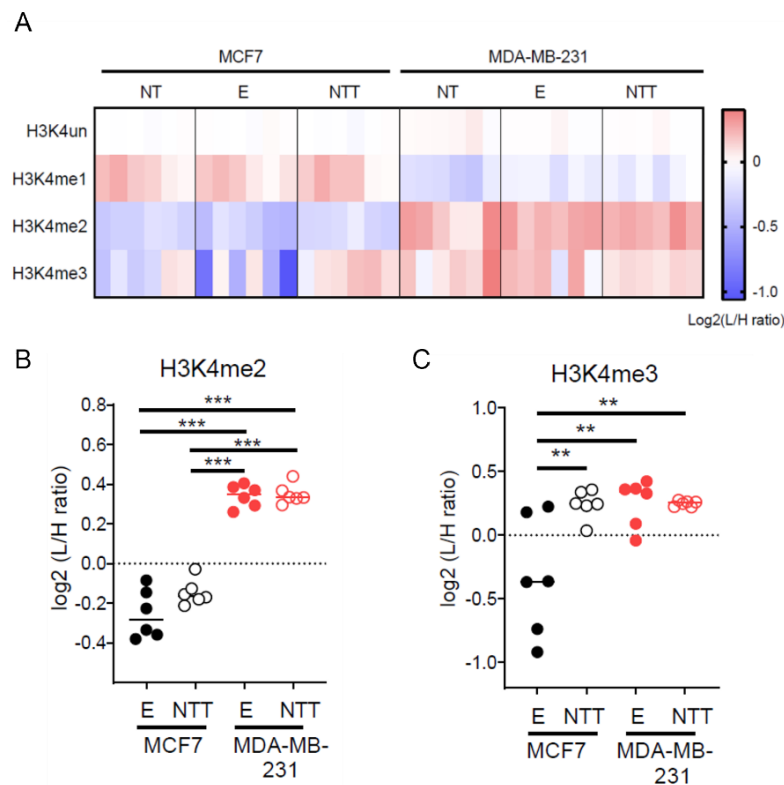


Fig.20. MS-based analysis of H3K4 methylation levels in MCF7 and MDA-MB-231 cells upon KDM5B-NTT over-expression. **A.** Heatmap display of the log₂ L/H ratios (where L is sample and H is internal standard) obtained for H3K4 mono-, di- and tri-methylation in MCF7 and MDA-MB-231 cells in three different conditions: not transfected (NT); transfected with empty vector (E); over-expressing KDM5B-NTT isoform (NTT). Data were normalized over the average value across all the samples. Six different biological replicates were considered. **B.** H3K4me2 and H3K4me3 levels (expressed as L/H ratios) in MCF7 and MDA-MB-231 cells overexpressing KDM5B-NTT, compared to the negative control (empty vector; E). Statistical analysis was performed by repeated measures ANOVA, followed by Tukey's multiple comparison test. **p < 0.005, ***p < 0.001.

3.10. KDM5B-NTT regulates gene expression

To evaluate the possible role of KDM5B-NTT in the general transcription profile, we performed RNA-Seq analysis after KDM5B-NTT transient over-expression in MCF7 and MDA-MB-231 cells. RNA-Seq was performed on total RNA purified from four different batches of MCF7 and MDA-MB-231 cells 24 hours after NTT over-expression. The empty vector transfection was used as a negative control. After batch effect correction, principal component analysis (PCA) showed a clear clustering of the KDM5B-NTT transfected samples compared to empty vector in both cell lines (Fig.S12). Data analysis of RNA-Seq showed that 123 transcripts were significantly ($FDR < 0.1$) regulated in MCF7 cells and only 47 in MDA-MB-231 cells. Moreover, the modulated genes in MCF7 cells were mostly induced (91), and the majority (66) were induced at least 2-fold (Fig.21A). On the opposite, in MDA-MB-231, only 7 genes were significantly induced and 8 repressed at least 2-fold (Fig.21B). Gene set enrichment analysis (GSEA) revealed that KDM5B-NTT over-expression modulated the genes belonging to the interferon-alpha and -gamma response and inflammation (Fig.22A) in MCF7 cells, but not in MDA-MB-231 cells. These genes play a role in tumour progression and regression and, as previously reported⁶⁷, they can be induced in colon cancer cells by epigenetic drug treatment, leading to a decrease in H3K9 methylation and, at the same time, to an increase in H3K4 tri-methylation⁶⁸. To evaluate whether the gene induction in MCF7 may be due to epigenetic de-repression by NTT, we used a comparison with the effect of two epigenetic drugs: PDCA, a selective KDM5 inhibitor⁶⁹, and MTA, a general DNA and histone methyltransferase inhibitor⁷⁰, which has been previously shown to reactivate epigenetically silenced promoters. We noted that a cluster of KDM5B-NTT-derepressed genes was also significantly induced by one or both treatments and/or previously shown to be bound by KDM5B or by KDM1A (or LSD1), a known interactor of KDM5B, in ChIP experiments (Fig.20B)^{33,41,53}. These results suggested that these are good candidates for a direct competitive de-repression by KDM5B-NTT.

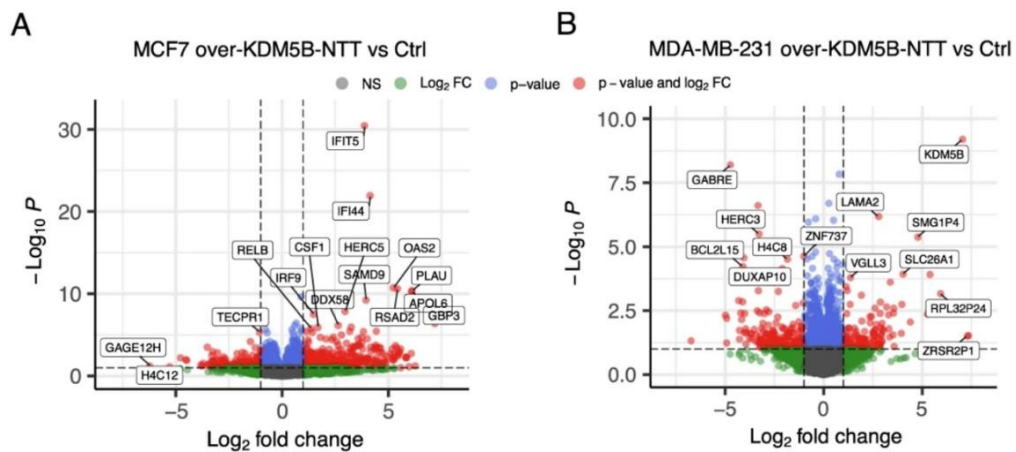


Fig.21. KDM5B-NTT over-expression regulates gene expression. **A.** Volcano plot showing the differentially expressed genes (DEGs) in MCF7 cells upon KDM5B-NTT over-expression. Red dots indicate DEGs associated with $p < 0.1$ and $\log_2(\text{FC}) < -1$ or $\log_2(\text{FC}) > 1$; blue dots indicate genes with $p < 0.1$ and $-1 < \log_2(\text{FC}) < 1$; green dots indicate genes with $\log_2(\text{FC}) < -1$ or $\log_2(\text{FC}) > 1$, and $p > 0.1$. **B.** Volcano plot showing the differentially expressed genes (DEGs) in MDA-MB-231 cells upon KDM5B-NTT over-expression. Red dots indicate DEGs associated with $p < 0.1$ and $\log_2(\text{FC}) < -1$ or $\log_2(\text{FC}) > 1$; blue dots indicate genes with $p < 0.1$ and $-1 < \log_2(\text{FC}) < 1$; green dots indicate genes with $\log_2(\text{FC}) < -1$ or $\log_2(\text{FC}) > 1$, and $p > 0.1$.

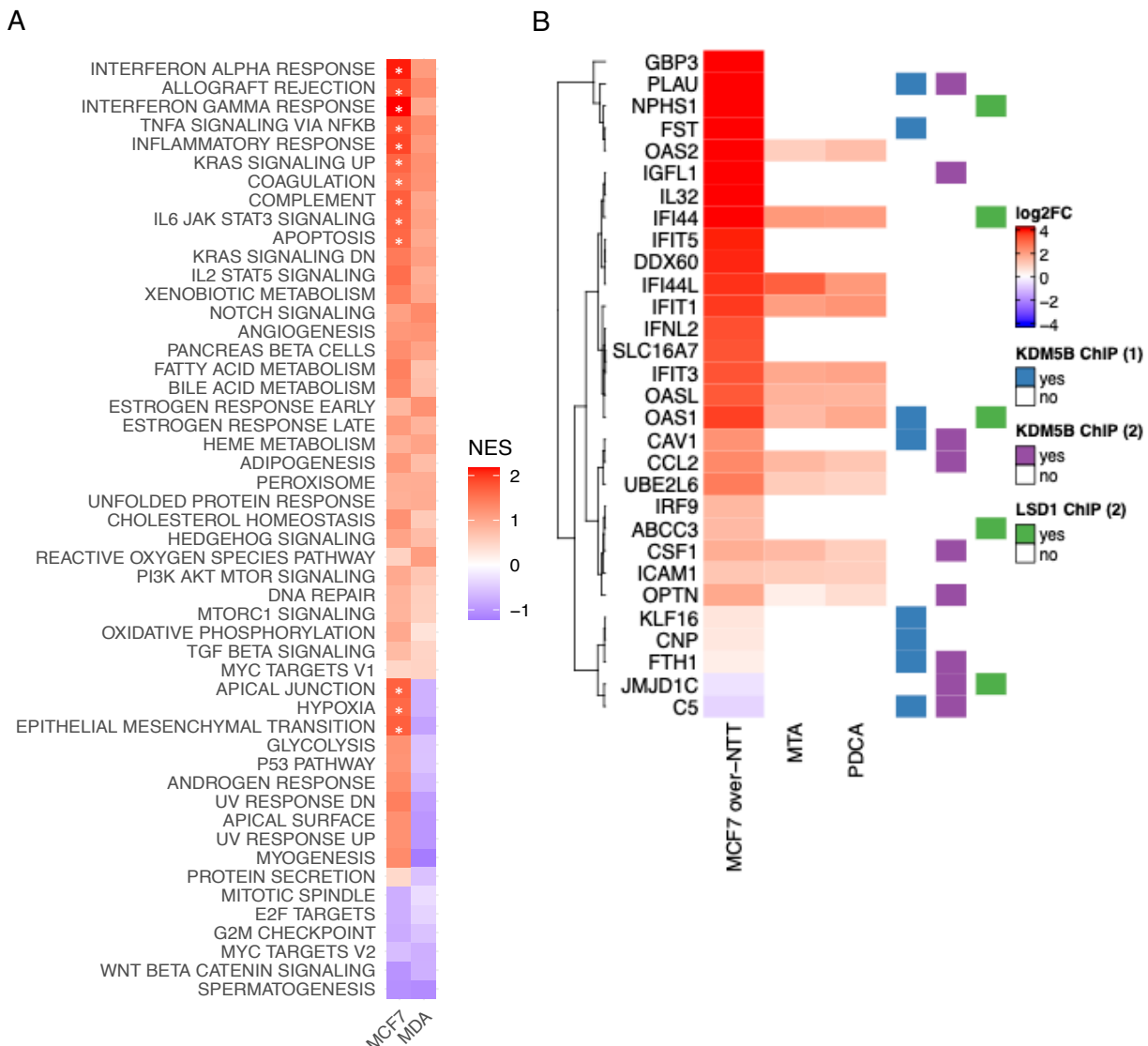


Fig.22. Effects of KDM5B-NTT over-expression on the transcriptome. A. Heatmap showing in color scale the GSEA Normalized Enriched Score (NES) of gene expression profile resulting from over-expression of KDM5B-NTT in MCF7 and MDA-MB-231 cells in comparison with control cell samples. (White*=FDR<0.01). **B.** Cluster of KDM5B-NTT-deregulated genes, which are also significantly induced by one or both epigenetic drug (PDCA and MTA) and/or previously shown to be bound by KDM5B in different ChIP experiments^{33,41,53}.

3.11. The putative regulatory action of KDM5B-NTT

Recently, Jamshidi et al. (2021)⁷¹ described a strain of mice expressing a KDM5B splicing variant that lacks the ARID domain and five amino acids in the JmjN domain. Interestingly, unlike KDM5B knock-out mice which display an embryonic lethal phenotype, these mice were viable and fertile although the protein was catalytically inactive. Overall, these data strengthen many experimental observations regarding the biological effects of KDM5 demethylases, independently of their catalytic activity^{24,25,36,72,73}. Our hypothesis is that KDM5B-NTT may be recruited to genomic sites through the PDH domains^{29-31,34,35} and keep some non-catalytic regulatory functions and/or act by competing with the catalytic function of the PLU-1 isoform. The over-expression of KDM5B-NTT positively correlates with a global increase of H3K4me3 and a modulation of genes, which were mostly induced, in MCF7 cells suggesting that NTT may have a functional role although being catalytically inactive. On the contrary, no significant changes in H3K4 tri-methylation level or gene expression were observed in the MDA-MB-231 cells, probably because the basal relative abundance of NTT is higher than in the MCF7 cells. These observations suggest a possible dominant-negative role of NTT in lowering PLU-1 activity in the basal cell lines where PLU-1 functions as a transcriptional repressor for genes involved in inflammatory responses, cell proliferation, adhesion, and migration^{29,33}. It is plausible that KDM5B-NTT could interfere with PLU-1 at some target genes, by competing with its repressive activity. Supporting this hypothesis, many of the induced genes upon KDM5B-NTT over-expression in MCF7 cells were previously shown to be regulated by the direct action of PLU-1 in MCF7 or other breast cancer cell lines^{15,33,41}. Strikingly, KDM5B-NTT over-expression increased the level of Caveolin-1, a tumour suppressor gene known to be directly repressed by PLU-1 through the lowering of H3K4me3 at its promoter region^{17,53}. The putative competing function of NTT in the regulation of H3K4me3 levels and in the following transcriptional regulation is described in Fig.23. However, it needs to be verified whether the truncated isoform KDM5B-NTT can directly associate with chromatin affecting the PLU-1 recruitment on

specific common targets. Moreover, it is not granted the causality relationship between the increase of H3K4me3 and the observed gene inductions. Therefore, to ascertain a direct correlation between transcriptomic changes and local increase of H3K4 trimethylation upon NTT over-expression, we performed a Cleavage Under Targets & Release Using Nuclease (CUT&RUN) assay (CST, #86652) to define the spatial and temporal distribution of the H3K4me3 mark in the genome of MCF7 cells. However, data analysis is still ongoing, and we can't rule out indirect effects on the significant observed transcriptomic changes. An alternative hypothesis is that the functions of KDM5B-NTT depend on its ability to locally recruit transcription and RNA-processing factors at specific genomic sites, independently of its demethylase activity. In this scenario, the putative role played by the additional residues encoded by the exon-6 is still unknown. To identify interactors of KDM5B-NTT, a total of four pull-down experiments using the anti-flag resin after the over-expression of the NTT and PLU-1 isoforms (Fig.S13) have been done for the mass-spec analysis that is ongoing. In fact, it will be interesting to compare the NTT and PLU-1 interactomes to better understand if NTT could also regulate the PLU-1 functions by sequestering common interactors. Whether NTT may play a role in migration and invasiveness remains to be investigated. Preliminary results showed that NTT over-expression in MCF7 does not determine significant changes in proliferation or cell cycle dynamics (Fig.S14), leaving still many open questions about the functions of this new truncated and catalytically inactive isoform in breast cancer cells.

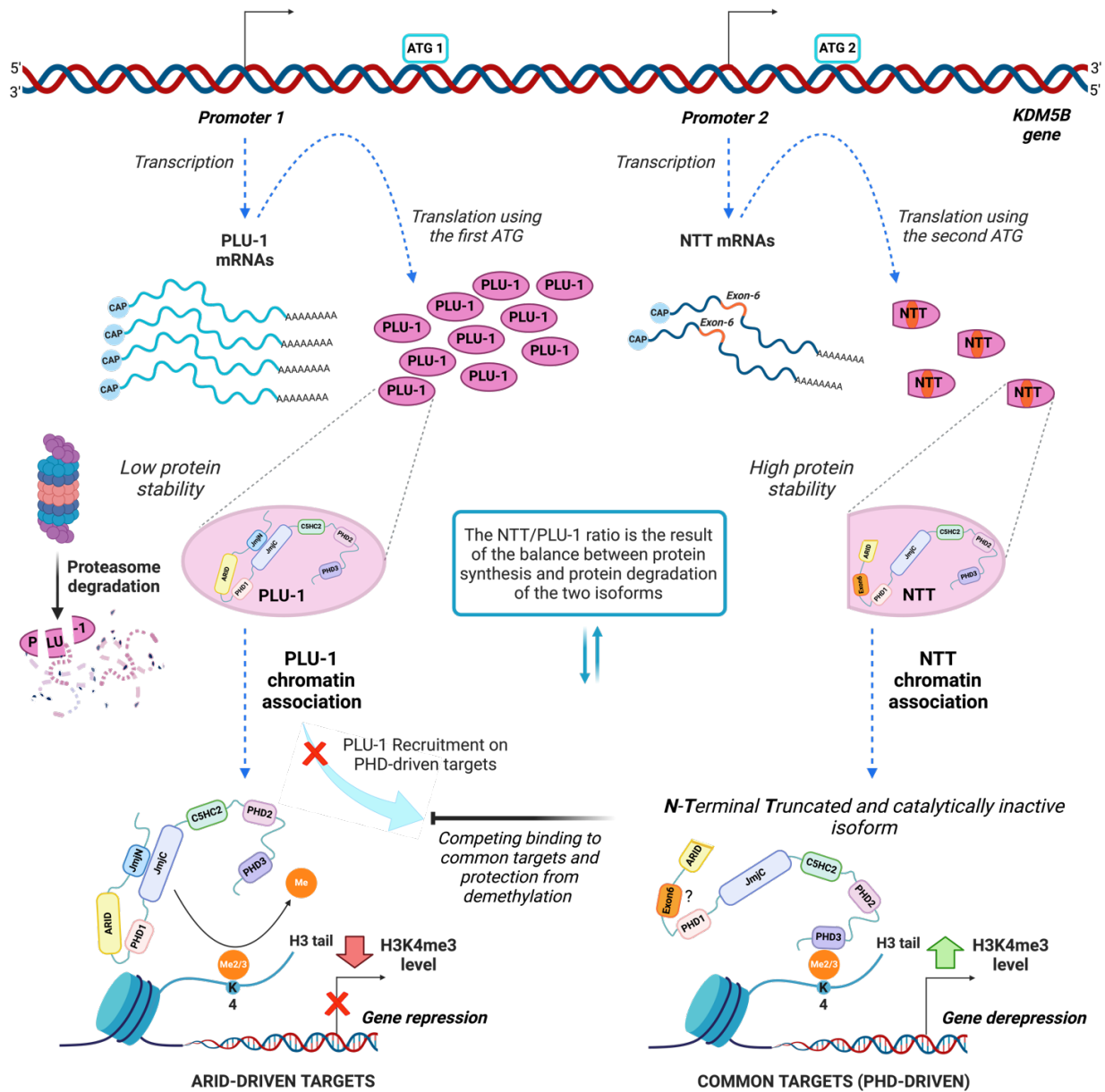


Fig.23. Diagram illustrating the proposed mechanism for the regulatory action of KDM5B-NTT. Transcription from Promoter 2 produces NTT mRNAs which include the exon-6; NTT mRNAs generate the expression of the KDM5B-NTT protein isoform. A second ATG (ATG2) is used for translation of NTT isoform, resulting in a KDM5B N-Terminal Truncated protein isoform including 36 more residues encoded by the exon-6 (in orange). Even though the NTT mRNAs are around 4-9% of the total KDM5B transcriptional events, the NTT protein isoform accumulates in cells because of its higher stability compared to PLU-1, which is targeted by the proteasome degradation system. NTT chromatin association can be driven by the PHD3 domain, which preferentially recognizes the trimethylated H3K4. This association may compete with the PLU-1 recruitment on common targets, preventing their demethylation and leading to an increase of H3K4me3 levels with a following gene induction. Created with BioRender.com.

4. CONCLUSIONS AND FUTURE PERSPECTIVES

Understanding the biological relevance of different KDM5B isoforms in tumour progression might be essential for the identification of new combined therapeutic technologies in cancers where KDM5B targeting plays a key role, such as in breast cancer. In this study we found a new KDM5B isoform, N-terminal truncated and catalytically inactive, which include additional residues of still unknown function. On the putative role of this isoform in breast cancer progression many questions are still open. Indeed, the relative abundance of KDM5B-NTT isoform compared to PLU-1 in our preliminary analysis didn't show a clear correlation with the molecular breast cancer subtype classification. NTT appears quite constantly expressed in all the analyzed cell lines and the variability in the relative ratio NTT/PLU-1 is likely being mainly affected by the balance of PLU-1 synthesis and degradation, tightly controlled by the proteasome activity. Considering that the level of proteasome activity may be different in each cell line and may be also influenced by the state of the cells, many variables can affect the outcome, adding higher levels of complexity. To fully understand the relevance of KDM5B-NTT in the fine-tuning regulation of KDM5B-PLU-1 activity, this heterogeneity should be considered.

Although the NTT biological functions remain still elusive, this study highlights an interest aspect regarding the faster proteasomal degradation of PLU-1 compared to NTT, suggesting a role of the N-terminal portion of PLU-1, missing in the truncated isoform, in regulating the protein turnover. However, the residues involved in proteasomal targeting are still unknown and worth further investigation. A more complete understanding of the KDM5B-PLU-1 degradation could open the possibility of tuning the levels of the two isoforms by a combination of epigenetic drugs and proteasomal inhibitors.

So far, KDM5B was known to act as an oncogene in luminal breast cancer cell lines, supporting high proliferation rates^{15,17,41}, and as a tumour suppressor in triple-negative cell lines inhibiting angiogenesis, migration and metastatic potential^{29,33}. The higher NTT/PLU-1 ratio observed in MDA-MB-231 than in MCF7 might contribute to the contraposing effects of KDM5B reported in these two cell lines. However, further investigations about the mechanistic details are needed.

It will be fascinating to assess the KDM5B-NTT level or the ratio between the two isoforms in different breast cancer subtypes, in both pre-clinical and clinical setups. Additionally, understanding cellular and tissue response to changes in this ratio could open new therapeutic approaches or biomarker studies.

Moreover, considering that the inhibition of the demethylase activity of KDM5B helps to overcome radio resistance in cancer by changes in H3K4 methylation at DNA double strand breaks^{18,22}, it will be interesting to understand the contribution of NTT isoform and/or of any alteration of the NTT/PLU-1 ratio upon DNA damage.

Finally, it is still unknown, and deserving future understanding, if other non-catalytical isoforms of histone demethylases exist and whether they play a relevant role in cancer epigenomics.

5. MATERIAL AND METHODS

Cell cultures

Human breast cancer MCF7, T47D and MDA-MB-231 cells were a gift from Dr Lucia Gabriele (Istituto Superiore di Sanità, Rome, Italy), who purchased them from ATCC and maintained them in culture. Cells were cultured at 37°C in a 5% CO₂ atmosphere for, at most, 10-12 passages, then replaced by cells from early passages, kept in liquid nitrogen since collection. The DU4475 cells and other T47D cells were obtained from the MRC-PPU, University of Dundee (Scotland, UK). The MCF10A, MDA-MB-415, MDA-MB-453, SKBR3, MDA-MB-436, MDA-MB-468, SUM-149-PT, and BT-549 were obtained from the European Institute of Oncology (IEO, Milan, Italy). Mycoplasma contamination was checked monthly (and before every freezing and after each thawing) using a Mycoplasma PCR Detection Kit (Abm). MCF7 and MDA-MB-231 cells were grown in high-glucose DMEM medium plus FBS 10%, 1 mmol L⁻¹ L-glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin (Thermo Fisher Scientific, Waltham, MA, USA). T47D cells were grown in RPMI medium plus FBS 10%, 1 mmol L⁻¹ L-glutamine, 1% insulin, 100 U/mL penicillin and 100 µg/mL streptomycin (Thermo Fisher Scientific, Waltham, MA, USA). DU4475 cells were grown in RPMI 1640 medium (Gibco, UK) added with 10% FBS (Gibco, UK), 1 mmol/L L-glutamine (Gibco, UK), 100 U/mL penicillin (Gibco, UK), 100 µg/mL streptomycin (Gibco, UK) and 1 mM sodium pyruvate (Gibco, UK). MCF10A were cultured in Dulbecco's modified Eagle's medium (DMEM) (Euroclone) and Ham's F12 (Gibco, UK) medium (1:1 v/v) with stable glutamine, 5% Horse Serum, 0.5 mg/mL Hydrocortisone, and 100 ng/mL Cholera toxin, with the addition of 10 µg/mL EGF at each passage. SKBR3, MDA-MB-415, and MDA-MB-453 were grown in DMEM with stable glutamine supplemented with 10% North American FBS (HyClone). BT549 were grown in DMEM with stable glutamine supplemented with 10% South American FBS (Microgem). ZR-751 were grown in RPMI 1640 (Euroclone) supplemented with 10% South American FBS, 2 mM L-Glutamine, 1 mM Sodium Pyruvate, and 10 mM HEPES. MDA-MB-468 and MDA-MB-436 were cultured in DMEM and Ham's F12 medium (1:1 v/v) with stable glutamine and supplemented with 10% South American FBS. SUM-149-PT were grown in Ham's F12 media supplemented with 5% South American FBS, 2 mM L-Glutamine, 5 µg/mL insulin, 1 µg/mL Hydrocortisone, and 10 mM HEPES. All media were supplemented

with penicillin (100 µg/mL), and streptomycin (100 mg/mL). Cells were cultured at 37 °C in a 5% CO₂ humidified atmosphere and passaged at 80% confluence. SUM-149-PT cell line was cultured at 37 °C in a 10% CO₂ humidified atmosphere. Epstein–Barr virus-immortalized lymphocytes and peripheral blood mononuclear cells (PBMC) were kindly provided by Dr. Alessandra Fragale (Istituto Superiore di Sanità, Rome, Italy). Human testicular total protein lysates were purchased from Takara Bio (Shiga, US).

Protein extraction and quantification

The Procedure for lysis of adherent cells in RIPA Buffer with protease inhibitor cocktail (Complete, EDTA- free Protease Inhibitor Cocktail, Roche, Basel, Switzerland) was used for MCF7, T47D, MDA-MB-231 breast cancer cell lines according to producer's protocol guidelines (Sigma-Aldrich); the procedure for lysis of suspension cultured cells in RIPA Buffer with protease inhibitor cocktail (Complete, EDTA-free Protease Inhibitor Cocktail, Roche, Basel, Switzerland) was used for the frozen cellular pellets, according to producer's protocol guidelines (Sigma-Aldrich). Cell lysates were stored at –80°C. Proteins were quantified using the Pierce Detergent Compatible Bradford Assay Kit (ThermoFisher Scientific, US). The Human testis protein medley (catalog no. 635309 Lot number: 1602008A) was used as normal control adult tissue. Nuclear and cytoplasmic fractions of protein lysates of MDA-MB-231 cells were obtained using the NE-PER Nuclear and Cytoplasmic Extraction Reagents (ThermoFisher, cat.no.78833).

Western blot

Western blot analysis was used to evaluate the amount of KDM5B isoforms. For each sample, 30 µg of RIPA extracted total proteins were run on an 5% denaturing PAGE to resolve both PLU-1 (179kDa) and NTT (162kDa). The running was in TGS buffer (Biorad) at 80 V until the marker separation and at 100 V for two hours after the marker separation. Transfer efficiency was verified via ponceau staining Ponceau (Sigma, Cat#P7767). Filters were hybridized with α – C term KDM5B (Rabbit) 1:7000, (Atlas Antibodies HPA0217179) at 4°C over- night. The Restore Western blot Stripping Buffer (Termo Scientific #21059) was used according to the manufacturer guidelines

to restore the filters and the stripping efficiency was ever tested with a ChemiDoc XRS using the Clarity Western ECL Substrate Peroxide solution/Enhancer solution (BIO-RAD #1705061). After stripping, the filters were hybridized with α – esone 6 KDM5B (Mouse) 1:1.000, BIO- RAD (MCA4340Z). Goat anti-rabbit HRP conjugate (1:10 000, Thermo Fisher Scientific Cat#31460, RRID: AB_228341) and anti-mouse (1:10000, Cell Signaling Cat#7076) were used as a secondary antibodies. Anti-vinculin (Rabbit) 1:1000 in 5% BSA-TBST (ab129002, Abcam) and anti-Tubulin (Mouse) 1:4000 in 5% MILK-TBST (T5168-100UL, Sigma) were used as protein loading controls. For quantitative analysis, western blots were analyzed with a ChemiDoc XRS+ Imaging System (Bio-Rad, Hercules, CA, USA). We used 5 μ L of Thermo Scientific Spectra Multicolor High Range Protein Ladder (Cat#26625) as marker. The molecular weights were calculated with MW analysis tool of the Image Lab 3.0 software. The signal for NTT isoform was normalized to PLU-1 signal detected on the same blot (hybridized with α – C term KDM5B) to evaluate the percentage fraction of NTT over PLU-1.

RNA extraction, RT-PCR and RT-qPCR

Total RNA extraction from MCF7, T47D and MDA-MB-231 cells was performed using RNeasy mini kit (Qiagen, US) or miRNeasy Mini kit with DNase on column digestion (Qiagen, US) according to the manufacturer's protocol. The RNA quality was tested by electrophoresis gel 1% TBE 0.5x after the RNA extraction. Retro-transcription of mRNAs was performed using sensiFAST cDNA synthesis kit, according to producer's protocol guidelines (Bioline). PCR amplification was performed using DeamTaq Polymerase according to producer's protocol guidelines (Thermo Fisher Scientific), as described in the table below. RT-qPCR was performed using a SensiFAST SYBR Hi-ROX Kit ((Meridian, US). PrimerBlast and Primer3 software were used for primer design. The primers used for RT-PCR and RT-qPCR are listed in Supplementary Table S1. RT-PCR was run as follows: 95 °C for 3 min, 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min for 35 cycles, with a final elongation at 72°C for 7 min. RT-qPCR was run in a two-step cycle at 95°C for 2 min, then 95 °C for 5 s, and 60 °C for 30 s for 40 cycles. Relative quantification of gene expression was conducted using the Applied Biosystems Step One Plus PCR System and data analysis was performed using the comparative $\Delta\Delta$ Ct method. Actin was used as an endogenous reference gene. For the electrophoresis of the RT-PCR products we used 2%

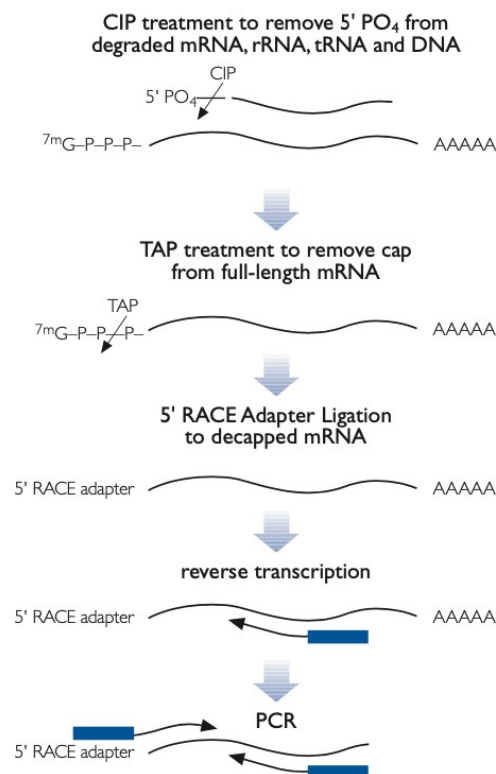
agarose gel TBE 1x, at 100 V; as DNA Ladder markers we used 5 μ L of Thermo Scientific O'GeneRuler DNA Ladder Mix, ready to use (SM1173) or 5 μ L of Thermo Scientific O'RangeRuler 100+500 bp DNA Ladder, ready-to-use (SM0653).

5'RLM-RACE

The 5'RACE assay was performed using the FirstChoice RLM-RACE Kit (ThermoFisher Scientific, US). RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) was designed to amplify cDNA only from full-length capped mRNA. Briefly, total RNA from MCF7 cells was treated with CIP phosphatase to dephosphorylate uncapped RNAs and then de-capped using the TAP enzyme. After the de-capping reaction, an RNA adapter oligonucleotide was ligated to the residual phosphorylated 5'-end from the capped-mRNAs using T4 RNA ligase (see schematic below, from datasheet of FirstChoice RLM-RACE Kit, AM1700, Invitrogen). RNA quality was checked after each enzymatic reaction, before RT. After a random-primed RT reaction, allowing the production of a cDNA population encompassing the 5'- initiation sites fused to the anchor, two nested PCR amplified the 5'-end of a specific transcript. As a negative control, total dephosphorylated RNA, untreated with TAP de-capping enzyme (negative TAP-control), was used in the PCR reactions. Both nested PCRs gave rise to a dominant product, which was purified and sequenced, whereas non-specific products were obtained using a Minus-TAP control reaction as a template (data not shown). Platinum Taq DNA Polymerase, High Fidelity (ThermoFisher Scientific, US) was used for PCR reactions. The primer sequences used for the 5'RACE PCR reactions are listed in Supplementary Table1. In the first 5'RACE PCR, a Reverse Ex-6 primer and the 5'RACE Outer Fw primer were used; in the second nested 5'RACE PCR, the downstream 5'RACE Inner Fw primer was used together with the reverse primer 1 (PCR1) or the reverse primer 2 (PCR2); both Fw primers were complementary to the 5'-anchor sequence and provided with the kit. For the electrophoresis of the 5'RACE-PCR products we used 1,5% agarose gel TBE 1x, at 100 V; as DNA Ladder markers we used 5 μ L of Thermo Scientific O'RangeRuler 100+500 bp DNA Ladder, ready-to-use (SM0653). The images were acquired using a Typhoon 9300 scanning instrument. The purified PCR products were analyzed by Sanger sequencing. Here the table of the PCR setting and the general workflow (schematic workflow from the Invitrogen kit datasheet):

Step	Cycles	Temp	Time
Initial denaturation	1	94°C	2 min
Denaturation	35	94°C	15 sec
Annealing		56°C	30 sec
Extension		68°C	1 min 40 sec
Hold	1	4°C	indefinitely

5' RLM-RACE



Cycloheximide assay and proteasome inhibitor

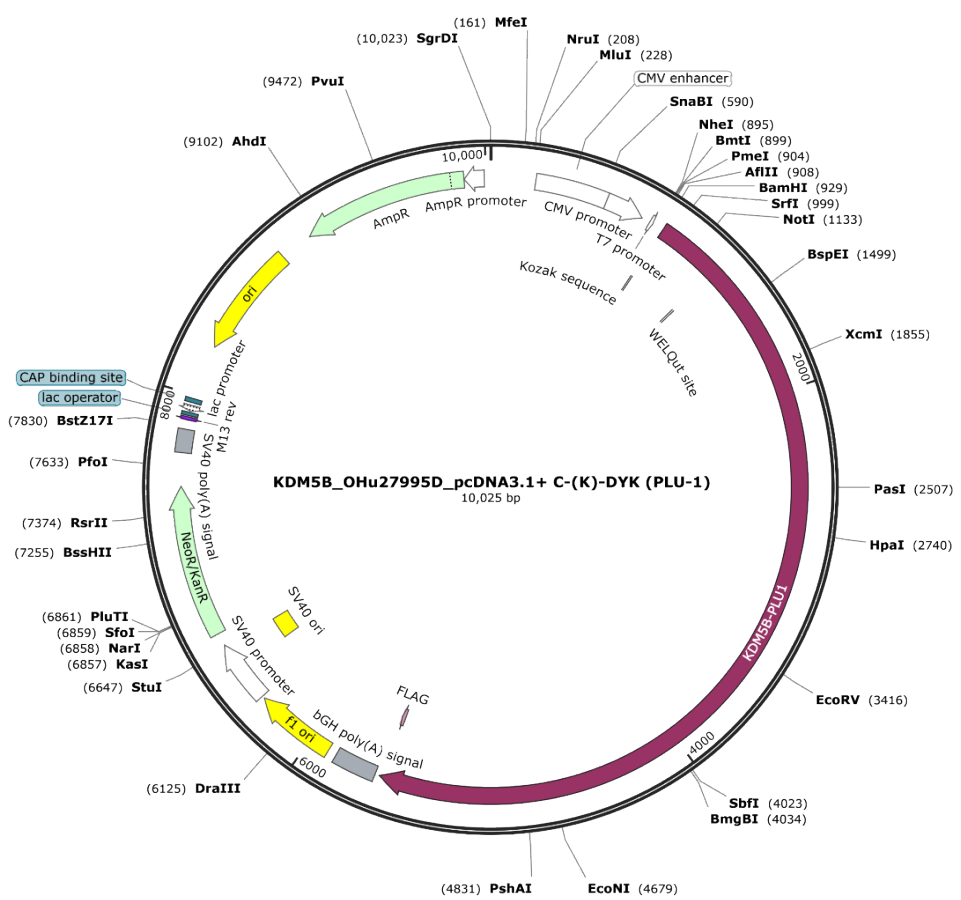
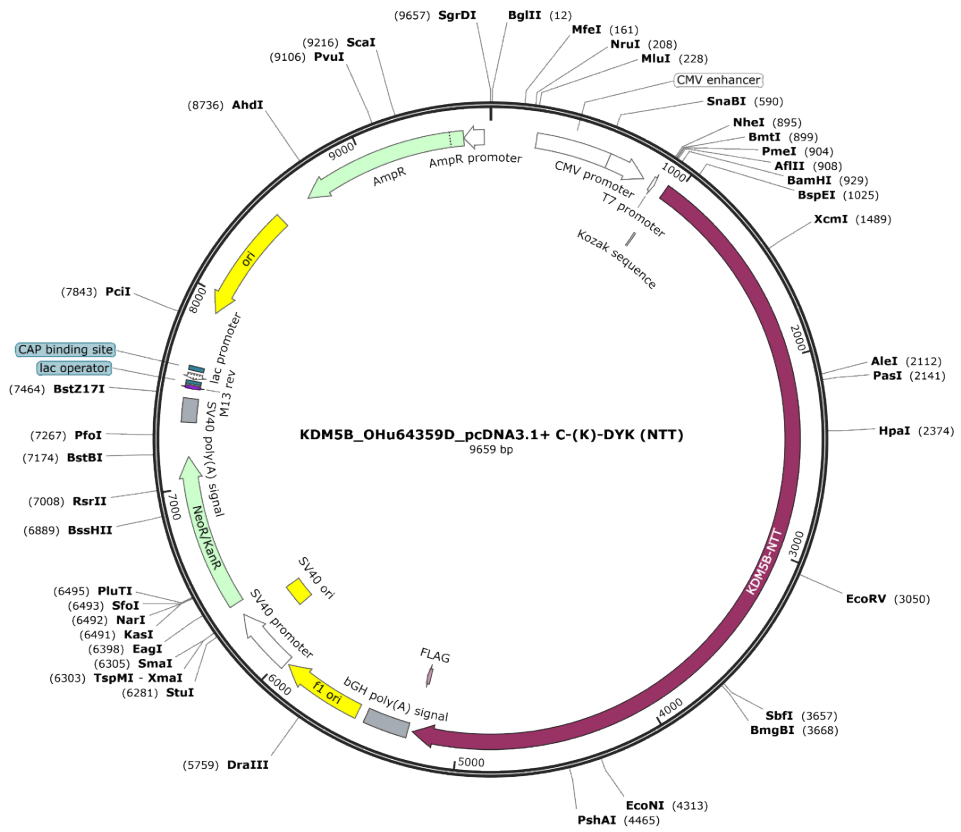
MCF7 and MDA-MB-231 cells were seeded on 6-well plates at a density of 30x10³ cells/cm² and 25x10³ cells/cm² respectively. After 24h from cell seeding, the cells were treated with cycloheximide (CHX, 50 µg/ml) (Sigma, C-6255, Lot 122H1110), using DMSO (Sigma, S-D2438-5X10ML) as a mock negative control, or CHX (50 µg/ml) and (R)-MG132 (10µM) (Sigma-Aldrich, M8699-1MG); RIPA cell lysates were prepared every two hours (from 0 hours to 12 hours).

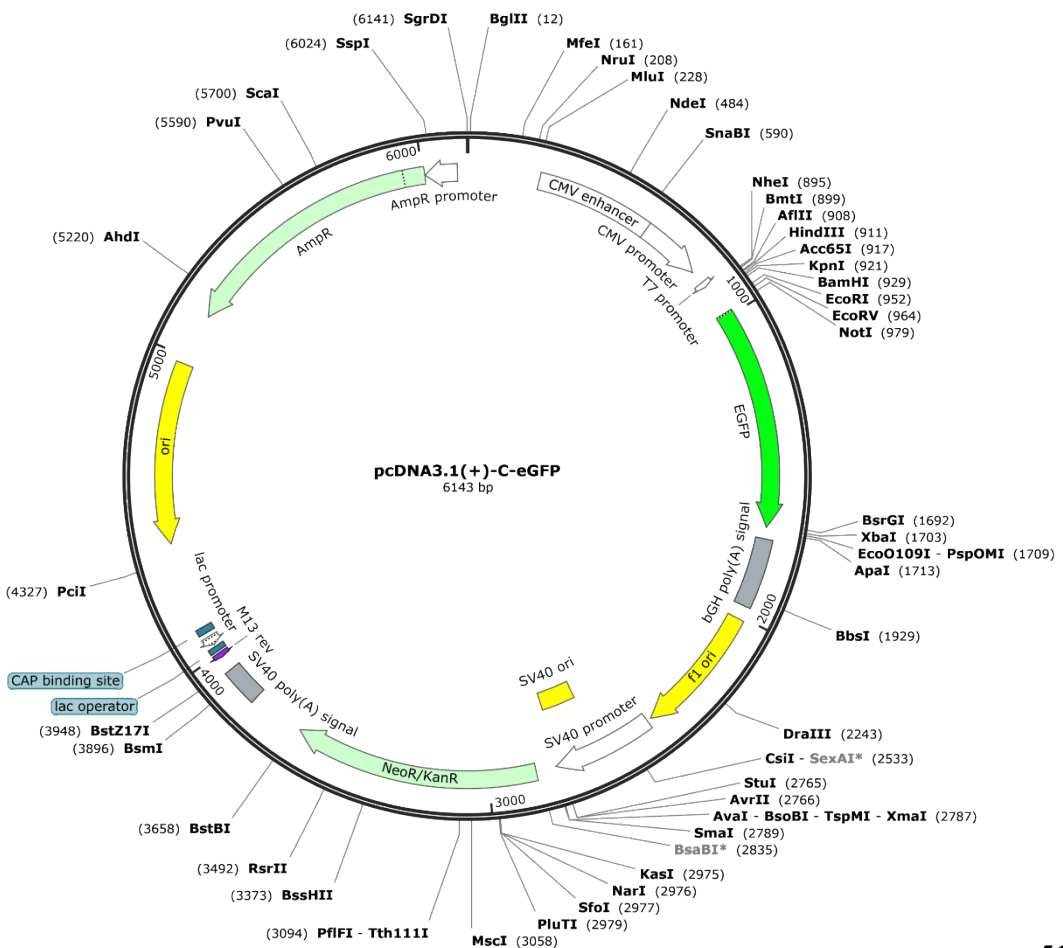
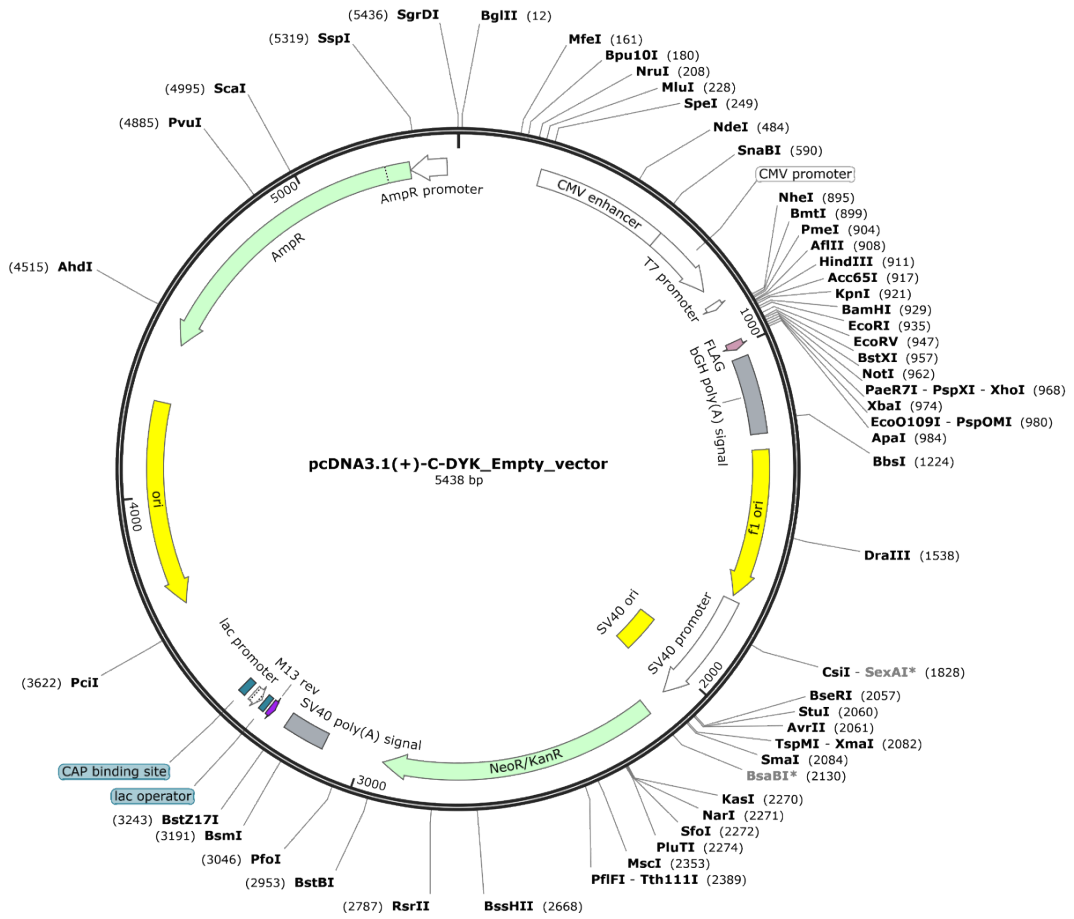
Cell Transfection

MCF7 and MDA-MB-231 cells were cultured in high-glucose DMEM with 1% L-Glut and 10% FBS without P/S at 37°C in 5% CO₂. 30 x10³ cells/cm² and MCF7 cells and 25 x10³ cells/cm² of MDA-MB-231 cells were seeded the day before transfection. Transfection using LIPO3000 (ThermoScientific, US) and Opti-MEM I Reduced Serum Medium (Gibco, ThermoFisher Scientific, US) was performed in 60-mm plates, using 9 µg of NTT overexpression plasmid or empty plasmids according to the manufacturer's protocol, and cell lysates were prepared 24 and 48 hours after plasmid transfection for subsequent analysis. For MDA-MB-231 transfection 25'000 cells/cm² were seeded in 6-well plate and transfected with RNAiMAX (Invitrogen, cat.no.13778150) to deliver exon-6 custom siRNA (5nM, 10nM, 15nM, 20nM or double transfection pulse 20nM+10nM) into cells following the manufacturer's recommended protocol. Cell lysates were prepared at 48 and 72 hours after siRNA transfections for subsequent analysis. The fluorescent oligonucleotide duplex siGLO green Transfection Indicator was used as a positive control of siRNA transfection (Darmachon Reagents, cat.no. D-001630-01-05).

Plasmids

KDM5B-NTT isoform over-expression plasmid (KDM5B_OHu64359D_pcDNA3.1+/C-(K)-DYK, Cat. OHu64359D), KDM5B-PLU-1 isoform over-expression plasmid (KDM5B_OHu27995D_pcDNA3.1+/C-(K)-DYK), empty vector (pcDNA3.1(+)-C-DYK) as a negative control, and GFP plasmid (pcDNA3.1(+)-C-eGFP) as transfection control were purchased from GenScript (US) and amplified in DH5α competent cells (ThermoFisher Scientific, US). The heat-shock was used for the transformation of competent cells, which were selected in LB-agar plates with Ampicillin (100µg/mL). Plasmids were extracted using the Qiagen plasmid MIDI kit (Qiagen, US). The detailed maps of all the used plasmids are shown below.





Protein structure analysis

The crystal structure of the catalytic domain of PLU-1 in complex with N-oxalylglycine was retrieved from Protein Data Bank (www.rcsb.org; PDB: 5A1F). KDM5B-NTT (residues 1-631) and the first 25 residues that were missing in Plu1 were *ab initio* modeled using the official release of the neural network implemented in AlphaFold2 (version 2.0.1). PFAM was used to map the positions of the JmjN, ARID, PHD, JmjC and zf-C5HC2 domains on the three-dimensional structures. The ELM server was used to identify potential degradation signals on the sequences of PLU1 and KDM5B-NTT Protein sequences and structures analysis was carried out using PyMod 3. See Supplementary references for details.

Mass spectrometry (MS) analysis

Histones were enriched from 1×10^6 breast cancer cells (MDA-MB-231 and MCF7) in 6 biological replicates for each condition (NT, E, and KDM5B-NTT), as previously described. Prior to digestion, approximately 3 μg of histone octamer were mixed with an equal amount of heavy-isotope labelled histones, which were used as an internal standard for quantification, and were separated on a 17% SDS-PAGE gel. Histone bands were excised, chemically acylated with propionic anhydride and in-gel digested with trypsin, followed by peptide N-terminal derivatization with phenyl isocyanate (PIC). Peptide mixtures were separated by reversed-phase liquid chromatography (RP-LC) on an EASY-Spray column (ThermoFisher Scientific, US), 25-cm long (inner diameter 75 μm , PepMap C18, 2 μm particles), which was connected online to a Q Exactive Plus instrument (Thermo Fisher Scientific) through an EASY-Spray™ Ion Source (ThermoFisher Scientific, US).

The differentially modified peptides containing H3K4 were quantified manually using QualBrowser version 2.0.7 (ThermoFisher Scientific, US), as previously described. For each histone modified peptide, the light (sample) and heavy (internal standard) percentage relative abundances (%RAs) were estimated by dividing the area under the curve (AUC) of each modified peptide for the sum of the areas corresponding to all the observed forms of that peptide. Light/Heavy (L/H) ratios of %RA were then calculated (Table S2). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (11) via the PRIDE partner repository with the

dataset identifier PXD033337. Data display and statistical analysis were carried out using GraphPad Prism 9.3.1 (GraphPad). Changes in single histone modifications among groups were analyzed by one-way ANOVA, followed by Tukey's multiple comparisons test performed on log₂ transformed L/H ratios. See Supplementary references for details.

Computational analysis of breast cancer and melanoma cell lines

Gene expression data was obtained for the CCLE database. For this research, datasets of RNA-Seq from 49 melanoma cell lines and 56 breast cancer cell lines were used. Computational analysis was performed using R version 3.4.3. in collaboration with the CINECA Institute. The reads were aligned to the human genome (GRCh38) using STAR (Spliced Transcripts Alignment to a Reference) (v2.6.1c)⁴⁴ and the rMATS tool v4.0.2 (turbo) was used to evaluate the frequency of inclusion of the exon-6 in the alternative splice variants of KDM5B. See Supplementary references for details.

RNA-Sequencing

Data for estimation of exon-1 and exon-6 containing transcripts of KDM5B are taken from the CCLE repository for tumour cell lines related data and from TCGA BRCA project for patients affected by breast cancer related data.

For RNA-Sequencing (RNA-Seq), total RNA was extracted from purified from cell pellets of control and KDM5B-NTT-overexpressing cultures using Qiagen RNeasy kits. RNA-Seq libraries from total RNA (1 µg) from each sample were prepared using the Illumina TruSeq Stranded mRNA Kit (Illumina, San Diego, CA) according to manufacturer's instructions. The amplified fragmented cDNA of 300 bp in size were sequenced in paired-end mode using a NovaSeq6000 (Illumina) sequencer with a read length of 150 bp. Processing raw data for both format conversion and de-multiplexing were performed by Bcl2Fastq version 2.20 of the Illumina pipeline. Sequence reads quality was evaluated using *FastQC* (version 0.11.8, Babraham Institute Cambridge, UK) tool then adapter sequences were masked with *Cutadapt* version 1.11 from raw fastq data using the following parameters: *--anywhere --overlap5 --times 2 --minimum-length 35 --mask-adapter*.

Then reads were mapped to the mouse Ensembl GRCh38 build reference genome using *HISAT2* version 2.2.0 using Gene annotations according to the Ensembl database. Gene-level quantification was performed using *salmon* version 1.8.0. See Supplementary references for details.

Differential gene expression analysis

To identify differentially expressed genes (DEGs) data was filtered to remove from the analysis the genes having < 10 counts per million across all replicates for each comparison. DEGs were assessed with a comparison of MCF7 and MDA-MB-231 cell lines with constitutive expression of KDM5B-NTT over control cell lines, using a Wald's test with a false discovery rate threshold of (FDR) < 0.1. The data normalization and differential analysis for gene expression were performed using Bioconductor R package *DESeq2* version 1.34 accounting for the presence of batch effects. The figures were obtained using the R environment with base plot functions and from package *ggplot2* version 3.3.6. Volcano plots were created using Bioconductor R package *EnhancedVolcano* version 1.12.0. See Supplementary references for details.

Gene set enrichment analysis (GSEA)

Broad Institute's Gene Set Enrichment Analysis (GSEA) method was used to assess the enrichment of the ranked list of DEGs of the KDM5B-NTT unique gene expression signatures of MCF7 and MDA-MB-231 samples versus the curated "Hallmark" gene set collections from the BROAD molecular signature database (MSigDb version 7.4.1). The human version of the MSigDb "Hallmark" gene set collections were obtained from the *msigdb* R package version 7.4.1.

This ES of GSEA is calculated by walking down the ranked list of genes, increasing a running-sum statistic when a gene is in the gene set and decreasing it when it is not. A normalized enrichment score (NES) is also calculated by GSEA in which differences in pathway size (i.e., gene set size) are considered, allowing for comparisons between pathways within the analysis. The enrichments were considered significant when FDR < 0.01. See Supplementary references for details.

Transcriptome of MCF7 cells treated with epigenetic drugs

MCF7 cells were treated with 60 μ M 5'-deoxy-5'-methylthioadenosine (MTA) or with 300 μ M 2,4-pyridine-dicarboxylic acid (PDCA) and 0.6% DMSO for negative controls. After 24h cells were harvested, and total RNA purified. The transcriptomes of MTA or PDCA treated cells were compared with DMSO treated cells.

Cell Proliferation Assay

Cell proliferation assay was performed with Cell Counting Kit-8 (Sigma-Aldrich), according to the manufacturer's instructions. MCF7 cells were seeded in a 96-well plate at a density of 10^4 cells/well the day before transfection. The transfection was performed with KDM5B-NTT over-expression or empty vectors as described in the 'Cell Transfection' paragraph. At indicated time of cells growth a volume of 10 μ L of CCK-8 solution was added to each well and plate incubated for 3 hours at 37°C in a 5% CO₂ atmosphere. The absorbance at 450 nm was read using a Clariostar microplate reader (BMG Labtech).

Flow-cytometry

Flow-cytometry analysis of DNA content was performed using an EPICS xl flow-cytometer (Beckman-Coulter). At the indicated times from transfection, MCF7 cells were trypsinized, pelleted, washed with PBS and, finally, resuspended in PBS containing 1% Triton- X-100 and 40 μ g/mL propidium iodide (Sigma-Aldrich). Samples were incubated for 20 min at 37°C and then analyzed, acquiring 10 000 events for each sample. Acquired data were analyzed using the WinMDI software by Joe Trotter, available at <http://facs.scripps.edu>.

Co-Immunoprecipitation

For Co-IP experiments MCF7 cells were cultured in high-glucose DMEM with 1% L-Glut (Gibco), 10% FBS (Gibco), 1% NEAA (Gibco), 1% Sodium Pyruvate (Gibco) without P/S at 37°C in 5% CO₂. The day before transfection, 50×10^3 cells/cm² of MCF7 cells were seeded in 150mm dishes (using the automated cell counter, DeNovix CellDrop FL Fluorescence Cell Counting).

Transfection using PEI (Sigma) and Opti-MEM I Reduced Serum Medium (Gibco) was performed with 72 μg of NTT-flag or PLU-1-flag overexpression plasmids or empty-flag plasmid using a ratio 1:3 of DNA:PEI. The over-expression of the GFP plasmid was used as a positive control of transfection. Cell lysates were prepared 24 after plasmid transfection for subsequent Co-IP experiments. Cells were washed twice using cold PBS, then harvested and lysed using the cell scraper and 800 μL of cell lysis buffer containing 50 mM Tris-HCl pH 7.5, 250 mM NaCl, 0.5% Triton X-100, 5% glycerol and proteases and phosphatases inhibitors (Roche). The lysates were quantified using Bradford Coomassie reagent (Pierce). Anti-FLAG M2-conjugated agarose beads (Sigma) were washed three times using 700 μL of cold PBS and four times using 700 μL of the complete cell-lysis buffer. Between each wash, the beads were centrifuged for 1 min at 4°C, 2500 g, and kept on ice for 2 min before removing the supernatant washing buffer. 20 μL of 50% slurry (10 μL of resin bed volume for each IP) were incubated with 1mg of each lysate (up to a final volume of 500 μL) overnight at 4°C. The beads were then washed 6 times with 800 μL cold cell-lysis buffer, and the bound proteins were eluted in 20 μL of 10% SDS buffer. The samples were cleaned using Spin-X columns (0.45 μm) (Costar – Corning) by centrifugation for 1 min at RT, 8000 g. The eluted samples were analyzed using Western blotting or processed for the ongoing mass-spec analysis using the S-Trap micro-columns (ProtiFi) according to the manufacturer's protocol.

Statistical analysis

Data were analyzed using R version 4.2, Prism (version 6.0; GraphPad Software Inc.) and Microsoft Excel version 16. Statistical analysis was conducted using unpaired two-sided Welch's *t*-test when the variances of experimental groups were not similar, or one-way ANOVA for normal distributed values, as indicated in the figure legends. *p*-value of less than 0.05 was considered statistically significant. Error bars reported in graphs as S.D. Sample size (*n*) values are indicated on the corresponding graphs or in the legend. All replicates in every experiment were collected independently.

Key Resources table.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-KDM5B	Atlas Antibodies	Cat#HPA027179
Mouse monoclonal anti-Human JARID1B	BIORAD	Cat#MCA430Z
Goat anti-Rabbit IgG (H+L) Secondary Antibody, HRP	Invitrogen	Cat#31460
Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP	Invitrogen	Cat# 31430
Chemicals, Peptides, and Recombinant Proteins		
Dulbecco Modified Eagle Medium (DMEM) high glucose w/o sodium pyruvate and w/o L-glutamine)	Corning	Cat#15-013-CV
RPMI 1640 without L-glutamine	Corning	Cat#15-040-CV
Penicillin-Streptomycin, with 10,000 units penicillin and 10 mg streptomycin per mL in 0.9% NaCl, sterile- filtered, BioReagent, suitable for cell culture	Sigma-Aldrich	Cat#S-P0781-100ML
L-Glutamine solution, 200 mM, solution, sterile-filtered, BioXtra, suitable for cell culture	Sigma-Aldrich	Cat#S-G7513-100ML
Fetal Bovine Serum, South American sourced	Corning	Cat#35-079-CV
PBS w/o calcium and w/o magnesium	Corning	Cat#21-040-CV
Trypsin/EDTA	Sigma-Aldrich	Cat#T4299-100ML
Incuwater-Clean 100 ml	PanReac AppliChem	Cat#A5219,0100
Aquabator-CleanT (100X)	PanReac AppliChem	Cat#A9390,0250
RIPA buffer	Sigma-Aldrich	Cat#R0278-50ML
Protease inhibitoComplete(TM), Mini, EDTA-free Protease Inhibitor Cocktail, Tablets	Roche	Cat#S-4693159001
Ponceau	Sigma-Aldrich	Cat#P7767
Bovine Serum Albumin	Sigma-Aldrich	Cat#A3069
2-β-mercaptoethanol	Sigma-Aldrich	Cat#63689
2-Propanol, BioReagent, for molecular biology, >=99.5%	Sigma-Aldrich	Cat#S-I9516-500ML
Clarity Western ECL Substrate	BIORAD	Cat#170-5061
Tris-Glycine buffer	BIORAD	Cat#1610734
Tris-Glycine-SDS buffer	BIORAD	Cat#161-0771
Sodium Dodecyl Sulfate solution 10%	Sigma-Aldrich	Cat#71736
Resolving Gel Buffer for PAGE	BIORAD	Cat#161-0798
Stacking Gel Buffer for PAGE	BIORAD	Cat#161-0799
Acrylamide/Bis Solution, 37.5:1	SERVA	Cat#10688.01
Ammonium persulfate	BIORAD	Cat#161-0700
N,N,N',N'-Tetramethylethylenediamine	Sigma-Aldrich	Cat# T9281
Spectra Multicolor High Range Protein Ladder (40 to 300kDa) 2 x 250 µl	Thermo Scientific	Cat#26625
Restore™ Western Blot Stripping Buffer	Thermo Fisher Scientific	Cat#21059
DAPI	Sigma-Aldrich	Cat#D9542-1MG
VECTASHIELD® Antifade Mounting Medium	Maravai LifeScience	Cat#H-1000
Platinum Taq High Fidelity	LIFE TECHNOLOGIES	Cat#11304011
DreamTaq DNA Polymerase	LIFE TECHNOLOGIES	Cat#11304011
UltraPure™ DNase/RNase-Free Distilled Water	Invitrogen	Cat#10977035
RNase AWAY, decontamination reagent for RNase	Sigma-Aldrich	Cat#S-83931-250ML
O'Gene Ruler 100bp+500bp, Ready to use	LIFE TECHNOLOGIES	Cat#SM0653
RNase-Free DNase Set (50)	Qiagen	Cat#79254
Poly-L-lysine solution, 0.01%, sterile-filtered, BioReagent, suitable for cell culture (50ml)	Sigma-Aldrich	Cat#P4707-50ML
Critical Commercial Assays		
RNeasy mini kit (50)	Qiagen	Cat#74104
SensiFAST cDNA synthesys kit	Bioline	Cat#BIO-65053:50
RLM-RACE KIT	LIFE TECHNOLOGIES	Cat#AM1700
High Pure PCR Product Purification Kit	Roche	Cat#11 732 668 001
PCR Mycoplasma Detection Kit 100 rxn	ABM	Cat#G238
Biological samples		
Human testis protein medley	Takara Bio	Cat#635309 Lot number: 162008A

Experimental Models: Cell Lines		
MCF7	ATCC	HTB-22
T47D	ATCC	HTB-133
MDA-MB-231	ATCC	HTB-26
SK-MEL-28	ATCC	HTB-72TM
Databases		
The Cancer Genome Atlas (TCGA) Database	Broad Institute	https://www.cancer.gov/tcga
Cancer Cell Line Encyclopedia (CCLE) Database	Broad Institute	https://portals.broadinstitute.org/ccle
UniProt database	UniProt Consortium	https://www.uniprot.org/
Reference Sequence (RefSeq) Database	NCBI	https://www.ncbi.nlm.nih.gov/refseq/
Software and Algorithms		
STAR (Spliced Transcripts Alignment to a Reference)(v2.6.1c)	Dobin et al. 2013	http://code.google.com/p/rna-star/ .
rMATS v4.0.2 (turbo)	Shen et al. 2014	maseq-mats.sourceforge.net/

6. REFERENCES

1. Luger, K., Dechassa, M. L. & Tremethick, D. J. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol* **13**, 436–447 (2012).
2. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381–395 (2011).
3. Dahl, J. A. *et al.* Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**, 548–552 (2016).
4. Jambhekar, A., Dhall, A. & Shi, Y. Roles and regulation of histone methylation in animal development. *Nat Rev Mol Cell Biol* **20**, 625–641 (2019).
5. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
6. Zhang, B. *et al.* Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* **537**, 553–557 (2016).
7. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp Mol Med* **49**, e324–e324 (2017).
8. Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet* **7**, 715–727 (2006).
9. Dimitrova, E., Turberfield, A. H. & Klose, R. J. Histone demethylases in chromatin biology and beyond. *EMBO Rep* **16**, 1620–1639 (2015).
10. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
11. Xie, L. *et al.* KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. *EMBO J* **30**, 1473–1484 (2011).
12. Santos-Rosa, H. *et al.* Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407–411 (2002).
13. Vermeulen, M. *et al.* Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. *Cell* **131**, 58–69 (2007).
14. Lauberth, S. M. *et al.* H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021–1036 (2013).
15. Yamane, K. *et al.* PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Mol Cell* **25**, 801–812 (2007).
16. Klose, R. J. *et al.* The retinoblastoma binding protein RBP2 is an H3K4 demethylase. *Cell* **128**, 889–900 (2007).

17. Scibetta, A. G. *et al.* Functional analysis of the transcription repressor PLU-1/JARID1B. *Mol Cell Biol* **27**, 7220–7235 (2007).
18. Li, X. *et al.* Histone demethylase KDM5B is a key regulator of genome stability. *Proc Natl Acad Sci U S A* **111**, 7096–7101 (2014).
19. Gong, F., Clouaire, T., Aguirrebengoa, M., Legube, G. & Miller, K. M. Histone demethylase KDM5A regulates the ZMYND8-NuRD chromatin remodeler to promote DNA repair. *J Cell Biol* **216**, 1959–1974 (2017).
20. Mocavini, I. *et al.* JARID1B expression and its function in DNA damage repair are tightly regulated by miRNAs in breast cancer. *Cancer Sci* **110**, 1232–1243 (2019).
21. Di Nisio, E., Lupo, G., Licursi, V. & Negri, R. The Role of Histone Lysine Methylation in the Response of Mammalian Cells to Ionizing Radiation. *Front Genet* **12**, 639602 (2021).
22. Pippa, S. *et al.* Small Molecule Inhibitors of KDM5 Histone Demethylases Increase the Radiosensitivity of Breast Cancer Cells Overexpressing JARID1B. *Molecules* **24**, (2019).
23. Gong, F. & Miller, K. M. Histone methylation and the DNA damage response. *Mutat Res* **780**, 37–47 (2019).
24. Ren, Z. *et al.* A PRC2-Kdm5b axis sustains tumorigenicity of acute myeloid leukemia. *Proc Natl Acad Sci U S A* **119**, e2122940119 (2022).
25. Zhang, S.-M. *et al.* KDM5B promotes immune evasion by recruiting SETDB1 to silence retroelements. *Nature* **598**, 682–687 (2021).
26. Outchkourov, N. S. *et al.* Balancing of histone H3K4 methylation states by the Kdm5c/SMCX histone demethylase modulates promoter and enhancer function. *Cell Rep* **3**, 1071–1079 (2013).
27. Pavlenko, E., Ruengeler, T., Engel, P. & Poepsel, S. Functions and Interactions of Mammalian KDM5 Demethylases. *Front Genet* **13**, 906662 (2022).
28. Tu, S. *et al.* The ARID domain of the H3K4 demethylase RBP2 binds to a DNA CCGCC motif. *Nat Struct Mol Biol* **15**, 419–421 (2008).
29. Klein, B. J. *et al.* The histone-H3K4-specific demethylase KDM5B binds to its substrate and product through distinct PHD fingers. *Cell Rep* **6**, 325–335 (2014).
30. Torres, I. O. *et al.* Histone demethylase KDM5A is regulated by its reader domain through a positive-feedback mechanism. *Nat Commun* **6**, 6204 (2015).
31. Zhang, Y. *et al.* The PHD1 finger of KDM5B recognizes unmodified H3K4 during the demethylation of histone H3K4me_{2/3} by KDM5B. *Protein Cell* **5**, 837–850 (2014).
32. Horton, J. R. *et al.* Characterization of a Linked Jumonji Domain of the KDM5/JARID1 Family of Histone H3 Lysine 4 Demethylases. *Journal of Biological*

- Chemistry* **291**, 2631–2646 (2016).
33. Li, Q. *et al.* Binding of the JmjC demethylase JARID1B to LSD1/NuRD suppresses angiogenesis and metastasis in breast cancer cells by repressing chemokine CCL14. *Cancer Res* **71**, 6899–6908 (2011).
 34. Wang, G. G. *et al.* Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. *Nature* **459**, 847–851 (2009).
 35. Johansson, C. *et al.* Structural analysis of human KDM5B guides histone demethylase inhibitor development. *Nat Chem Biol* **12**, 539–545 (2016).
 36. Zhang, Y., Liang, J. & Li, Q. Coordinated Regulation of Retinoic Acid Signaling Pathway by KDM5B and Polycomb Repressive Complex 2: Regulation of RA Pathway by KDM5B. *J. Cell. Biochem.* **115**, 1528–1538 (2014).
 37. Harmeyer, K. M., Faconpre, N. D., Herlyn, M. & Basu, D. JARID1 Histone Demethylases: Emerging Targets in Cancer. *Trends in Cancer* **3**, 713–725 (2017).
 38. Ohguchi, Y. & Ohguchi, H. Diverse Functions of KDM5 in Cancer: Transcriptional Repressor or Activator? *Cancers* **14**, 3270 (2022).
 39. Taylor-Papadimitriou, J. & Burchell, J. JARID1/KDM5 demethylases as cancer targets? *Expert Opin Ther Targets* **21**, 5–7 (2017).
 40. Pedersen, M. T. & Helin, K. Histone demethylases in development and disease. *Trends Cell Biol* **20**, 662–671 (2010).
 41. Yamamoto, S. *et al.* JARID1B is a luminal lineage-driving oncogene in breast cancer. *Cancer Cell* **25**, 762–777 (2014).
 42. Lu, P. J. *et al.* A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer. *J Biol Chem* **274**, 15633–15645 (1999).
 43. Xiang, Y. *et al.* JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. *Proc Natl Acad Sci U S A* **104**, 19226–19231 (2007).
 44. Hayami, S. *et al.* Overexpression of the JmjC histone demethylase KDM5B in human carcinogenesis: involvement in the proliferation of cancer cells through the E2F/RB pathway. *Mol Cancer* **9**, 59 (2010).
 45. Roesch, A. *et al.* A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* **141**, 583–594 (2010).
 46. Barrett, A. *et al.* PLU-1 nuclear protein, which is upregulated in breast cancer, shows restricted expression in normal human adult tissues: a new cancer/testis antigen? *Int J Cancer* **101**, 581–588 (2002).
 47. Madsen, B. *et al.* PLU-1, a transcriptional repressor and putative testis-cancer antigen, has a specific expression and localisation pattern during meiosis. *Chromosoma* **112**, 124–132 (2003).

48. Roesch, A. *et al.* Re-expression of the retinoblastoma-binding protein 2-homolog 1 reveals tumor-suppressive functions in highly metastatic melanoma cells. *J Invest Dermatol* **126**, 1850–1859 (2006).
49. Kuźbicki, Ł., Lange, D., Strączyńska-Niemiec, A. & Chwirot, B. W. Altered Splicing of JARID1B in Development of Human Cutaneous Melanoma? *Applied Immunohistochemistry & Molecular Morphology* **24**, 188–192 (2016).
50. Kuźbicki, Ł., Lange, D., Stanek-Widera, A. & Chwirot, B. W. Prognostic significance of RBP2-H1 variant of JARID1B in melanoma. *BMC Cancer* **17**, 854 (2017).
51. Vogt, T. *et al.* Deficiency of a novel retinoblastoma binding protein 2-homolog is a consistent feature of sporadic human melanoma skin cancer. *Lab Invest* **79**, 1615–1627 (1999).
52. Kashuba, V. *et al.* Isolation and chromosomal localization of a new human retinoblastoma binding protein 2 homologue 1a (RBBP2H1A). *Eur J Hum Genet* **8**, 407–413 (2000).
53. Yamane, K. *et al.* PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Mol Cell* **25**, 801–812 (2007).
54. Maier, T., Güell, M. & Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett* **583**, 3966–3973 (2009).
55. Roesch, A. *et al.* Retinoblastoma-binding protein 2-homolog 1: a retinoblastoma-binding protein downregulated in malignant melanomas. *Mod Pathol* **18**, 1249–1257 (2005).
56. Anvar, S. Y. *et al.* Full-length mRNA sequencing uncovers a widespread coupling between transcription initiation and mRNA processing. *Genome Biol* **19**, 46 (2018).
57. Braunschweig, U., Gueroussov, S., Plocik, A. M., Graveley, B. R. & Blencowe, B. J. Dynamic integration of splicing within gene regulatory pathways. *Cell* **152**, 1252–1269 (2013).
58. Fiszbein, A., Krick, K. S., Begg, B. E. & Burge, C. B. Exon-Mediated Activation of Transcription Starts. *Cell* **179**, 1551-1565.e17 (2019).
59. Zheng, Y.-C. *et al.* Lysine demethylase 5B (KDM5B): A potential anti-cancer drug target. *European Journal of Medicinal Chemistry* **161**, 131–140 (2019).
60. Dai, X., Cheng, H., Bai, Z. & Li, J. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. *J Cancer* **8**, 3131–3141 (2017).
61. Varshavsky, A. The N-end rule. *Cell* **69**, 725–735 (1992).
62. Schneider-Poetsch, T. *et al.* Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol* **6**, 209–217 (2010).
63. Goldberg, A. L. Development of proteasome inhibitors as research tools and cancer drugs. *Journal of Cell Biology* **199**, 583–588 (2012).

64. Tsvetkov, P. *et al.* Oncogenic addiction to high 26S proteasome level. *Cell Death Dis* **9**, 773 (2018).
65. Bueno, M. T. D. & Richard, S. SUMOylation negatively modulates target gene occupancy of the KDM5B, a histone lysine demethylase. *Epigenetics* **8**, 1162–1175 (2013).
66. Gagnon, K. T., Li, L., Chu, Y., Janowski, B. A. & Corey, D. R. RNAi factors are present and active in human cell nuclei. *Cell Rep* **6**, 211–221 (2014).
67. Jorgovanovic, D., Song, M., Wang, L. & Zhang, Y. Roles of IFN- γ in tumor progression and regression: a review. *Biomark Res* **8**, 49 (2020).
68. Fragale, A. *et al.* Antitumor Effects of Epidrug/IFN α Combination Driven by Modulated Gene Signatures in Both Colorectal Cancer and Dendritic Cells. *Cancer Immunol Res* **5**, 604–616 (2017).
69. Kristensen, L. H. *et al.* Studies of H3K4me3 demethylation by KDM5B/Jarid1B/PLU1 reveals strong substrate recognition in vitro and identifies 2,4-pyridine-dicarboxylic acid as an in vitro and in cell inhibitor. *FEBSJ* **279**, 1905–1914 (2012).
70. Williams-Ashman, H. G., Seidenfeld, J. & Galletti, P. Trends in the biochemical pharmacology of 5'-deoxy-5'-methylthioadenosine. *Biochem Pharmacol* **31**, 277–288 (1982).
71. Jamshidi, S. *et al.* KDM5B protein expressed in viable and fertile Δ ARID mice exhibit no demethylase activity. *Int J Oncol* **59**, 96 (2021).
72. DiTacchio, L. *et al.* Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock. *Science* **333**, 1881–1885 (2011).
73. Gaillard, S. *et al.* KDM5A and KDM5B histone-demethylases contribute to HU-induced replication stress response and tolerance. *Biol Open* **10**, bio057729 (2021).

7. SUMMPLEMENTAL FILES

FileS1.

CLUSTAL O(1.2.4) multiple sequence alignment

PCR1_oligoRev	-----	0
PCR2_oligoRev	-----	0
PCR1_oligoFw	-----	0
NM_006618.5	GTACAACTCGGACTTGCTGTTGCTCGAGCCGCTCTGCACGGGTCTCGGACCGAGCGGAG	60
PCR2_oligoFw	-----	0
PCR1_oligoRev	-----	0
PCR2_oligoRev	-----	0
PCR1_oligoFw	-----	0
NM_006618.5	CTCGCAGCCTCGGTCCCGGAGCCACCTTCGCCTCGCCCTTGCCAGCCTGCGGTG ATGG	120
PCR2_oligoFw	-----	0
PCR1_oligoRev	-----	0
PCR2_oligoRev	-----	0
PCR1_oligoFw	-----	0
NM_006618.5	AGGCGGCCACCACACTGCACCCAGGCCCGCCCGCGCTGCCCTCGGGGCCCCGGGCC	180
PCR2_oligoFw	-----	0
PCR1_oligoRev	-----T CGCGGATCCGAACACTGCGTTTGCTGGCTTC GAACCCAGCTGGGAAGAGT	51
PCR2_oligoRev	----- CGCGGATCCGAACACTGCGTTTGCTGGCTTC GAACCCAGCTGGGAAGAGT	50
PCR1_oligoFw	-----GGCTGACAGCTGGGAGA	17
NM_006618.5	CGCTGGGGCAGTTCTCGCTCCACCCGAGTGCCTCGGTCTTCGAACCCAGCTGGGAAGAGT	240
PCR2_oligoFw	-----CCCTTCGATCAGCTGTTCCAGAGTC	25
PCR1_oligoRev	TCGCGGACCCCTTCGCTTTCATCCACAAGATCCGGCCCATAGCCGAGCAGACTGGCATCT	111
PCR2_oligoRev	TCGCGGACCCCTTCGCTTTCATCCACAAGATCCGGCCCATAGCCGAGCAGACTGGCATCT	110
PCR1_oligoFw	GTCGCGGACCCCTTCGCTTTCATCCACAAGATCCGGCCCATAGCCGAGCAGACTGGCATCT	77
NM_006618.5	TCGCGGACCCCTTCGCTTTCATCCACAAGATCCGGCCCATAGCCGAGCAGACTGGCATCT	300
PCR2_oligoFw	GCGAGATGCTTCGACTTCTCATCCACAAGATCCGGCCCATAGCCGAGCAGACTGGCATCT	85
	* * * * * *	
PCR1_oligoRev	GTAAGGTGCGGCCCGCCCGGATTGGCAGCCACCATTTCATGTGATGTTGAAAACATCAT	171
PCR2_oligoRev	GTAAGGTGCGGCCCGCCCGGATTGGCAGCCACCATTTCATGTGATGTTGATAAACTTC	170
PCR1_oligoFw	GTAAGGTGCGGCCCGCCCGGATTGGCAGCCACCATTTCATGTGATGTTGATAAACTTC	137
NM_006618.5	GTAAGGTGCGGCCCGCCCGGATTGGCAGCCACCATTTCATGTGATGTTGATAAACTTC	360
PCR2_oligoFw	GTAAGGTGCGGCCCGCCCGGATTGGCAGCCACCATTTCATGTGATGTTGATAAACTTC	145
	* * * * * *	
PCR1_oligoRev	TACGCCAAGACACAGCCCTATAG-----	194
PCR2_oligoRev	ATTTTACGCCACGTATCCAGAGACTGAATGAATTGGAGGCCAAACTCGTGTAATAATTGA	230
PCR1_oligoFw	ATTTTACGCCACGTATCCAGAGACTGAATGAATTGGAGGCCAAACTCGTGTAATAA----	193
NM_006618.5	ATTTTACGCCACGTATCCAGAGACTGAATGAATTGGAGGCCAAACTCGTGTAATAATTGA	420
PCR2_oligoFw	ATTTTACGCCACGTATCCAGAGACTGAATGAATTGGAGGCCAAACTCGTGTAATAATTGA	205
	* * * * *	
PCR1_oligoRev	-----	194
PCR2_oligoRev	ATTTCTTGGACCAGATTGCAAAGTACTGGGAGTTACAGGGAAGTACTCTGAAAATTCCAC	290
PCR1_oligoFw	-----	193
NM_006618.5	ATTTCTTGGACCAGATTGCAAAGTACTGGGAGTTACAGGGAAGTACTCTGAAAATTCCAC	480
PCR2_oligoFw	ATTTCTTGGACCAGATTGCAAAGTACTGGGAGTTACAGGGAAGTACTCTGAAAATTCCAC	265
PCR1_oligoRev	-----	194
PCR2_oligoRev	ATGTGGAGAGGAAGATCTTGGACTTATTTTCAGCTTAATAAGTTAGTTGCAGAAGCAAGGT	350
PCR1_oligoFw	-----	193
NM_006618.5	ATGTGGAGAGGAAGATCTTGGACTTATTTTCAGCTTAATAAGTTAGTTGCAGAAGCAAGGT	540
PCR2_oligoFw	ATGTGGAGAGGAAGATCTTGGACTTATTTTCAGCTTAATAAGTTAGTTGCAGAAGCAAGGT	325
PCR1_oligoRev	-----	194
PCR2_oligoRev	GGATTTGTCAGTTGTTGCAAGGATATAGTAGCCTTCTAAAG-----	391
PCR1_oligoFw	-----	193
NM_006618.5	GATTTGCAG-TTGTTTGCAGGATAGAAAATGGACCAAAAATGCTACCAAGATGGGGTTT	599
PCR2_oligoFw	GATTTGCAG-TTGTTTGCAGGATAGAAAATGGACCAAAAATGCTACCAAGATGGGGTTT	384

PCR1_oligoRev	-----	194
PCR2_oligoRev	-----	391
PCR1_oligoFw	-----	193
NM_006618.5	GCTCCTGGCAAAGCAGTGGGCTCACATATCAGAGGGCATTATGAACGAATTCTCAACCCC	659
PCR2_oligoFw	GCTCCTGA-----	392

FileS1. Sequence alignment of KDM5B-PLU-1 mRNA (NM_006618.5) and the PCR products obtained upon the 5'RLM-RACE. Both PCR products (PCR1 and PCR2) of 5'RACE were sequenced using the reverse and forward used in the relative nested PCR reactions. The first ATG in PLU-1 transcript is highlighted in heavenly; the second ATG used for the translation of the NTT transcript is highlighted in green. Anchor primer sequence and transcriptional start site of NTT transcript are highlighted respectively in grey and yellow. Direct comparison with Figure 8D.

8. SUPPLEMENTAL TABLES AND FIGURES

Table S1. Primer sequences

RT-PCR Primers	Primer Sequences
NM Fw	5'-CCAAGATGGGGTTTGCTCCT-3'
NM Rv	5'-CAGACATACAGGTCCACAGCA-3'
RT-qPCR Primers	Primer Sequences
Actin Fw	5'-TCGTGCGTGACATTAAGGAG-3'
Actin Rv	5'-AGGAAGGAAGGCTGGAAGAG-3'
ALL KDM5B Fw	5'-TGCTTGATCCCACCTCTCCA-3'
ALL KDM5B Rv	5'-AACGCATCTGCCATTTCCCC-3'
KDM5B Exon-6 Fw	5'-CGAGCAAAACGCATGAGAGCA-3'
KDM5B Exon-6 Rv	5'-AGGCAGAAGAATTGCTGGAATCTA-3'
5'RLM-RACE Primers	Primer Sequences
5'RACE Outer Fw primer	5'-GCTGATGGCGATGAATGAACACTG-3'
5'RACE Inner Fw primer	5'-CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG-3'
Reverse Ex-6	5'-AGGCAGAAGAATTGCTGGAATCTA-3'
Reverse 1	5'-TTTACACGAGTTTGGGCCTCC-3'
Reverse 2	5'-AGGAGCAAACCCCATCTTGG-3'

Table S2. MS-based quantification of H3K4 methylations (L/H ratios)

MCF7	NT						E						NTT					
	R#1	R#2	R#3	R#4	R#5	R#6	R#1	R#2	R#3	R#4	R#5	R#6	R#1	R#2	R#3	R#4	R#5	R#6
H3K4un	0.99	0.99	0.99	0.98	0.99	1.00	1.00	0.99	0.99	0.98	1.01	1.00	0.99	0.99	0.99	0.96	0.99	1.00
H3K4me1	1.16	1.21	1.13	1.11	1.04	1.02	1.14	1.17	1.14	1.06	1.02	1.07	1.13	1.21	1.15	1.15	1.02	1.02
H3K4me2	0.84	0.87	0.87	0.93	0.92	0.88	0.79	0.94	0.90	0.85	0.78	0.77	0.89	0.90	0.92	0.98	0.88	0.86
H3K4me3	0.85	0.98	0.88	0.93	1.18	1.16	0.60	1.13	0.78	1.17	0.77	0.53	1.03	1.17	1.19	1.26	1.28	1.19

MDA-MB-231	NT						E						NTT					
	R#1	R#2	R#3	R#4	R#5	R#6	R#1	R#2	R#3	R#4	R#5	R#6	R#1	R#2	R#3	R#4	R#5	R#6
H3K4un	1.00	1.01	1.01	1.01	1.04	0.98	1.00	1.00	1.00	1.02	0.98	0.98	1.00	1.00	1.00	1.01	0.99	0.99
H3K4me1	0.88	0.87	0.85	0.80	0.78	0.92	0.95	0.96	0.96	0.87	0.96	0.92	0.94	0.90	0.97	0.86	0.95	1.00
H3K4me2	1.31	1.29	1.20	1.12	1.12	1.38	1.32	1.20	1.26	1.22	1.29	1.31	1.26	1.29	1.26	1.23	1.36	1.26
H3K4me3	1.26	1.05	1.16	1.24	1.20	1.46	1.28	1.25	1.29	0.97	1.34	1.06	1.17	1.20	1.19	1.17	1.21	1.20

Fig.S1

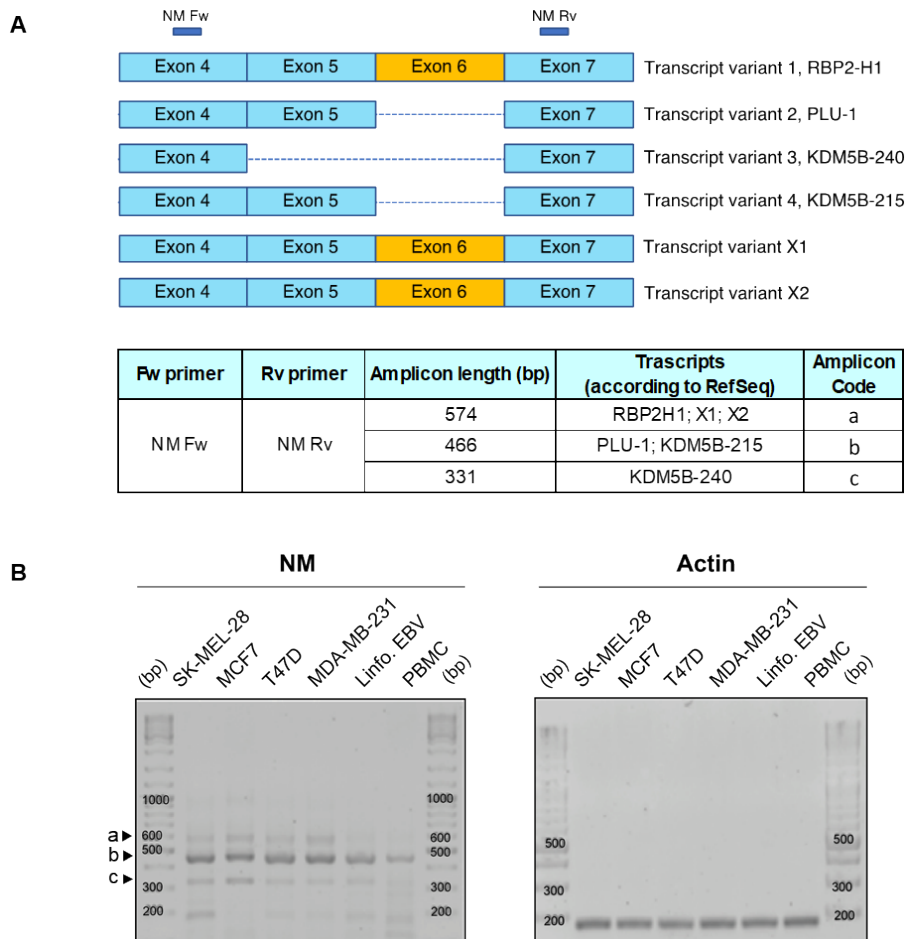


Figure S1. KDM5B splicing variants. A. Schematic representation of NM primer pair used in the RT-PCR shown in panel B. The table under the scheme reports the amplicon length (bp) corresponding to the respective KDM5B splicing variants (Transcripts according to RefSeq in the table). Depending on the inclusion of exon-5 and exon-6, three different PCR products can be generated (a, b and c are the codes for these PCR products). **B.** The PCR using NM primer pair (left panel) produced all the expected amplicons (a, b and c), demonstrating the existence of different KDM5B splicing variants in breast cancer cell lines with a similar expression pattern to melanoma ones. In SK-MEL-28, MCF7, T47D, and MDA-MB-231 there are KDM5B transcripts that include the exon-6 (574 bp, product code = a) and

which are barely detected in the normal-control cell lines, such as Epstein-Barr virus (EBV) immortalized lymphocytes (Linfo. EBV) and peripheral blood mononuclear cells (PBMC). Actin was used as a internal control (panel right).

Fig.S2

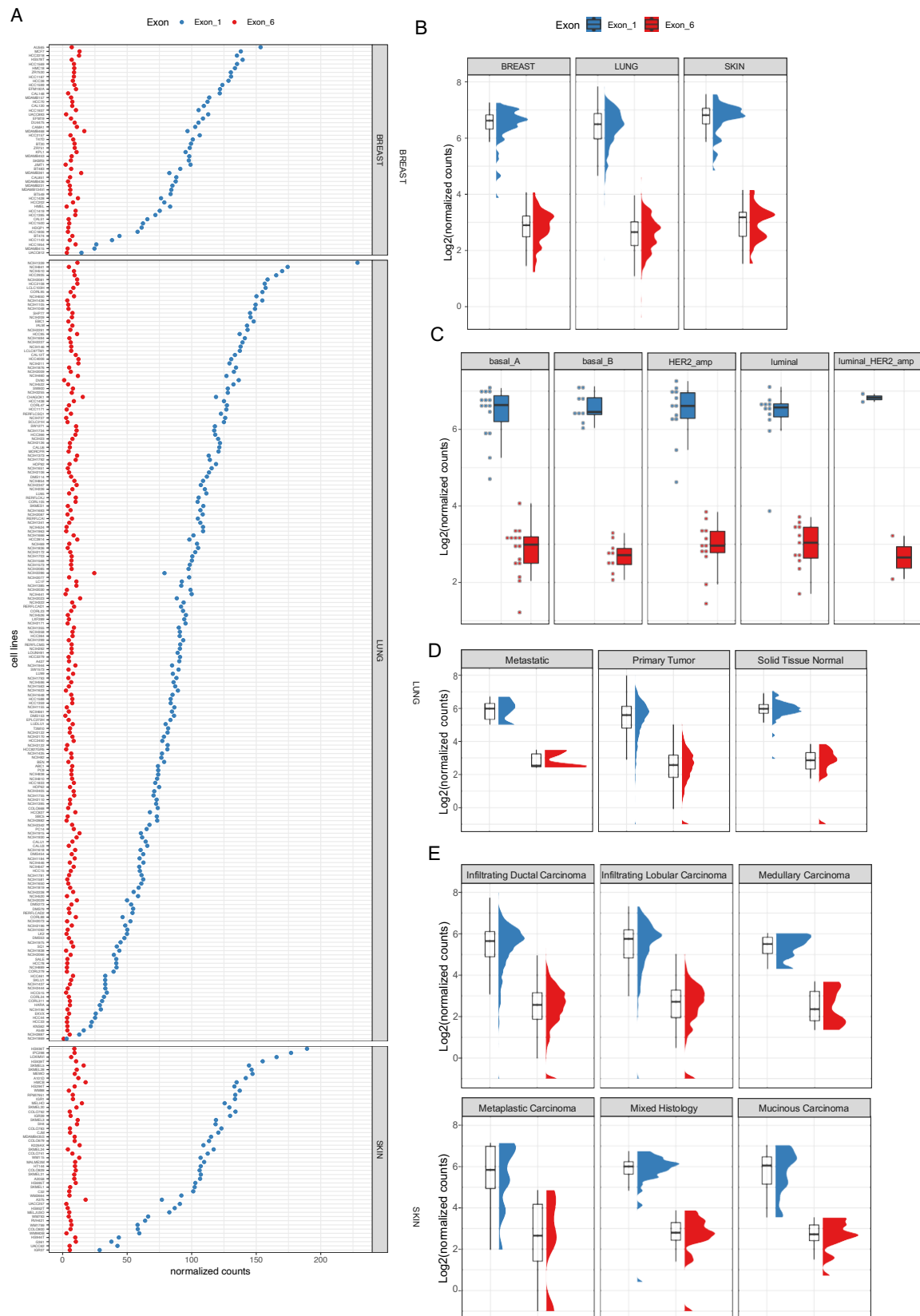


Fig.S2. Large scale data analysis on publicly available data on KDM5B transcripts expression. A, Dot plots show the expression levels quantified by RNA-Seq of the KDM5B exon-1 and exon-6 for all the cell lines of Breast, Lung and Skin analyzed by the Cancer Cell Line Encyclopedia project (CCLE, Broad Institute, <https://sites.broadinstitute.org/ccle/datasets>). Data are reported in normalized count levels. B, Summaries of the data showed in (A) represented with boxplots and density curves. C, Boxplots showing the expression levels of the KDM5B exon-1 and exon-6 containing transcripts of Breast cancer cell lines of (A) grouped by tumor subtypes. D, Boxplots showing the expression levels of the KDM5B exon-1 and exon-6 containing transcripts in Breast Cancer samples of patients collected in TCGA project. The data are from 7 metastatic samples, 1127 primary tumors and 112 normal breast tissues samples. E, Summaries of the data showed in (D) stratified by histological breast cancer subtypes.

Fig.S3

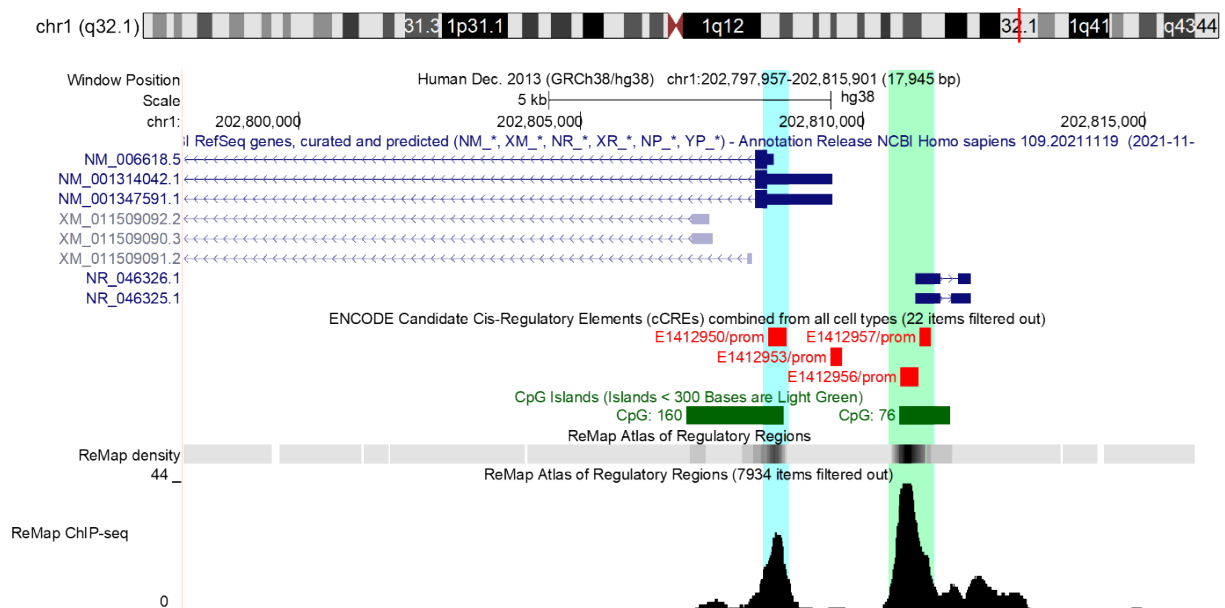


Fig.S3. Promoter region of KDM5B gene using the UCSC genome browser. The figure shows the TSS region and the 5' end of KDM5B transcripts with curated (NM) and predicted (XM) RefSeq gene annotation (the last version of RefSeq available on UCSC genome browser is the 109.20211119 of 2021-11-23), followed by track with the ENCODE Registry of candidate cis-Regulatory Elements (cCREs) in the human genome. The red boxes of this track represent the cCREs with promoter-like signatures (cCRE-PLS) provided by the [SCREEN](https://screen.cancer.gov/) (Search Candidate cis-Regulatory Elements) web tool. The green boxes represent prediction of two CpG islands. The last track shows the [ReMap Atlas](https://remap.broadinstitute.org/) of regulatory regions, which consists of a large-scale integrative analysis of all Public ChIP-seq data for transcriptional regulators from GEO, ArrayExpress, and ENCODE. In particular here there are two ReMap subtracks with the first one showing the whole ChIP-seq dataset in "dense" mode while the second one showing the ChIP-seq dataset filtered for "breast" keyword in the biotype field of the dataset in "full" mode, respectively. These data clearly supports the existence of two promoter regions which are highlighted in light green and light blue. The figure was created with UCSC genome browser.

Fig.S4

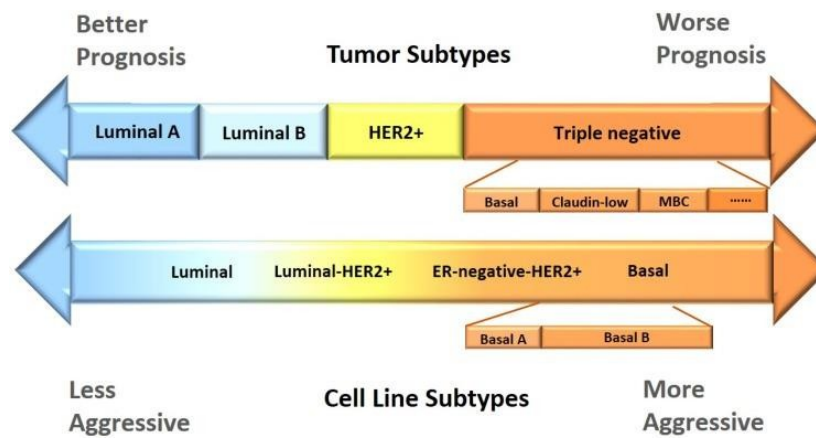


Fig.S4. Breast cancer cell lines as a model for breast tumour subtypes. The features of breast cancer cell lines recapitulate the features of breast tumour subtypes. Depending on the status of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor 2 receptor (HER2), breast cancers are classified as Luminal (A and B), HER2 positive and triple negative. MCF7 are classified as luminal A (LA) and MDA-MB-231 are classified as triple negative B (TNB). From Dai et al., 2017⁶⁰.

Fig.S5

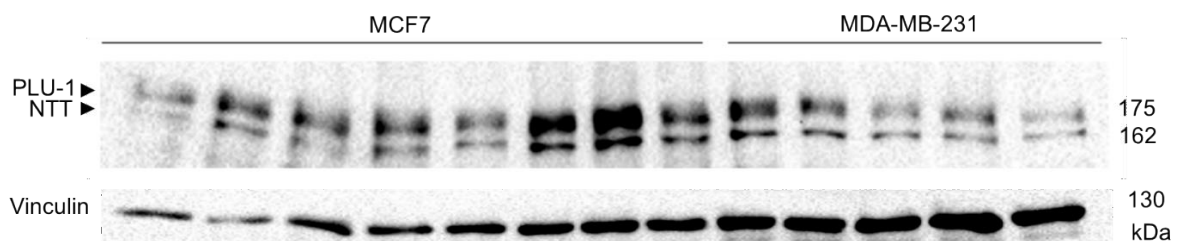


Fig.S5. Expression of KDM5B protein isoforms and transcripts in MCF7 and MDA-MB-231 cells. Western blot using the anti-C-terminal Ab showing the KDM5B-PLU-1 (175KDa) and the KDM5B-NTT (162KDa) bands in several cultures of MCF7 and MDA-MB-231.

```

RBP2H1 1 MEAATTIHLPGPRPALPLGGGFLGFEFLPPPECPVFEPSWEEFADFFAFHKIRPIAEGTGI CKVRPPDQWPFACDVKLHFTPTRIORNELEEAQTRVKNLNFIDQIAKWELOGSTLIKIPHVVERKILDLFQNLKVAEBEGGFVAVCKDRKWKIATMGGFAPGKAVGSHIRGHE 176
PLU-1 1 MEAATTIHLPGPRPALPLGGGFLGFEFLPPPECPVFEPSWEEFADFFAFHKIRPIAEGTGI CKVRPPDQWPFACDVKLHFTPTRIORNELEEAQTRVKNLNFIDQIAKWELOGSTLIKIPHVVERKILDLFQNLKVAEBEGGFVAVCKDRKWKIATMGGFAPGKAVGSHIRGHE 176
NTT 1 -----MGFAPGKAVGSHIRGHE 18

RBP2H1 177 RIINPNYLFISGDSIRCLQKPNLITDIDKQEKYKPHDIPOQSVQSPSETCPPARAKRMRABEPOSIAVLVPLECSGAILAHCNLRLLDSSNSASASANNIKIEPEETTEARTHLNRRMGCPTEKENEKEMKSSIKQEPIERKDYIVENEKEPKRSRKRATNAVLDLVVCLLG 352
PLU-1 177 RIINPNYLFISGDSIRCLQKPNLITDIDKQEKYKPHDIPOQSVQSPSETCPPARAKRMRABEPOSIAVLVPLECSGAILAHCNLRLLDSSNSASASANNIKIEPEETTEARTHLNRRMGCPTEKENEKEMKSSIKQEPIERKDYIVENEKEPKRSRKRATNAVLDLVVCLLG 316
NTT 177 RIINPNYLFISGDSIRCLQKPNLITDIDKQEKYKPHDIPOQSVQSPSETCPARRAKRMRABEPOSIAVLVPLECSGAILAHCNLRLLDSSNSASASANNIKIEPEETTEARTHLNRRMGCPTEKENEKEMKSSIKQEPIERKDYIVENEKEPKRSRKRATNAVLDLVVCLLG 194

RBP2H1 353 SGNDEDRLLLCGGDDSYHTFCLIPPLHDPKGDWRCPKLAQECSPQEAAGFEQAAARDYILRTFGEMADAFKSDYFNNPVHVPTLVEKEFWRIVSTIEEDVTVVEYGADIASKEFGSGFVDRGDKIKLSPREEEYLDSDGNNLNPNVPMQSVLAHITADICGMKLPMLVYVGM 528
PLU-1 317 SGNDEDRLLLCGGDDSYHTFCLIPPLHDPKGDWRCPKLAQECSPQEAAGFEQAAARDYILRTFGEMADAFKSDYFNNPVHVPTLVEKEFWRIVSTIEEDVTVVEYGADIASKEFGSGFVDRGDKIKLSPREEEYLDSDGNNLNPNVPMQSVLAHITADICGMKLPMLVYVGM 492
NTT 195 SGNDEDRLLLCGGDDSYHTFCLIPPLHDPKGDWRCPKLAQECSPQEAAGFEQAAARDYILRTFGEMADAFKSDYFNNPVHVPTLVEKEFWRIVSTIEEDVTVVEYGADIASKEFGSGFVDRGDKIKLSPREEEYLDSDGNNLNPNVPMQSVLAHITADICGMKLPMLVYVGM 370

RBP2H1 529 FSSFCWHIEDHWSYSINYLHWGEPKTYWVGYAAEQLENVWKKLAPLQVSDPDLHLQVLTIMNPNTLMTHEVYVYRTNOCAGFEVITPPRAYHSGFNQGNFAEAVNFCVDMWLPGRQCVEHYRLLHRYCVFSDHEMICMASADYLDVVVASTVQKDMAIMIDEKALRET 704
PLU-1 493 FSSFCWHIEDHWSYSINYLHWGEPKTYWVGYAAEQLENVWKKLAPLQVSDPDLHLQVLTIMNPNTLMTHEVYVYRTNOCAGFEVITPPRAYHSGFNQGNFAEAVNFCVDMWLPGRQCVEHYRLLHRYCVFSDHEMICMASADYLDVVVASTVQKDMAIMIDEKALRET 668
NTT 371 FSSFCWHIEDHWSYSINYLHWGEPKTYWVGYAAEQLENVWKKLAPLQVSDPDLHLQVLTIMNPNTLMTHEVYVYRTNOCAGFEVITPPRAYHSGFNQGNFAEAVNFCVDMWLPGRQCVEHYRLLHRYCVFSDHEMICMASADYLDVVVASTVQKDMAIMIDEKALRET 546

RBP2H1 705 VRKLGVIDSERMDFELLPDDERQCCKTKTCFMSAISCSCKPGLLVCLHHVKELCSCPPYKYKRYRYTDDLYPMNNAIKLRAESYNEWALNVALEAKINRKKSLIVSFKALIESEMKKFPDNDLRLHRLVTDQAEKASVAQQLNGKRQTRVRSGGKSONLTVNELRQ 880
PLU-1 669 VRKLGVIDSERMDFELLPDDERQCCKTKTCFMSAISCSCKPGLLVCLHHVKELCSCPPYKYKRYRYTDDLYPMNNAIKLRAESYNEWALNVALEAKINRKKSLIVSFKALIESEMKKFPDNDLRLHRLVTDQAEKASVAQQLNGKRQTRVRSGGKSONLTVNELRQ 844
NTT 547 VRKLGVIDSERMDFELLPDDERQCCKTKTCFMSAISCSCKPGLLVCLHHVKELCSCPPYKYKRYRYTDDLYPMNNAIKLRAESYNEWALNVALEAKINRKKSLIVSFKALIESEMKKFPDNDLRLHRLVTDQAEKASVAQQLNGKRQTRVRSGGKSONLTVNELRQ 722

RBP2H1 861 FVTQLYALPCVLSQTPLLKDLNLRVEDFQQHSQKLLSEETPSAAELQDLLDVSEFFDELQVLAEMRIRLEQARWLEEVQQAQLDPSLTLDDMRRLIDILGVGLAPYSAVEKAMARLQELLTVSEHDDKAKSLLKARPHRSINSLATAVKIEEIPAYLPNGAALKDSVQARAW 1056
PLU-1 845 FVTQLYALPCVLSQTPLLKDLNLRVEDFQQHSQKLLSEETPSAAELQDLLDVSEFFDELQVLAEMRIRLEQARWLEEVQQAQLDPSLTLDDMRRLIDILGVGLAPYSAVEKAMARLQELLTVSEHDDKAKSLLKARPHRSINSLATAVKIEEIPAYLPNGAALKDSVQARAW 1020
NTT 723 FVTQLYALPCVLSQTPLLKDLNLRVEDFQQHSQKLLSEETPSAAELQDLLDVSEFFDELQVLAEMRIRLEQARWLEEVQQAQLDPSLTLDDMRRLIDILGVGLAPYSAVEKAMARLQELLTVSEHDDKAKSLLKARPHRSINSLATAVKIEEIPAYLPNGAALKDSVQARAW 898

RBP2H1 1057 LQDVEGLQAGRVVPLDTLIEIVTRGRSIPVHNSLPRLETIVAEVQAWKECAVNTFLETNSPYSLLLEVLCPKCDI GLLGLKRRQKRLKEPLPNGKXKTKLESLSLIERALTESKETA SAMATLGEARLREMEALQSLRANEGKLLSPLQVDIKICLQKAPAPMIQCELCR 1232
PLU-1 1021 LQDVEGLQAGRVVPLDTLIEIVTRGRSIPVHNSLPRLETIVAEVQAWKECAVNTFLETNSPYSLLLEVLCPKCDI GLLGLKRRQKRLKEPLPNGKXKTKLESLSLIERALTESKETA SAMATLGEARLREMEALQSLRANEGKLLSPLQVDIKICLQKAPAPMIQCELCR 1196
NTT 899 LQDVEGLQAGRVVPLDTLIEIVTRGRSIPVHNSLPRLETIVAEVQAWKECAVNTFLETNSPYSLLLEVLCPKCDI GLLGLKRRQKRLKEPLPNGKXKTKLESLSLIERALTESKETA SAMATLGEARLREMEALQSLRANEGKLLSPLQVDIKICLQKAPAPMIQCELCR 1074

RBP2H1 1233 DAFHTSCVAVPSISQGLRIWLCPHCRSRSEKPLEKILPILASLQRIVRIPGSDALRYMTERVNWQRAQQLISSGNLKFVQDRVSGGLLSRWQASQVSDTNKVSQPFQTSLSLDDNDNRITSYLHSPFTGRSCIPLHGVSPEVNELLMEALQVLSLFEIQELYQTLLA 1408
PLU-1 1197 DAFHTSCVAVPSISQGLRIWLCPHCRSRSEKPLEKILPILASLQRIVRIPGSDALRYMTERVNWQRAQQLISSGNLKFVQDRVSGGLLSRWQASQVSDTNKVSQPFQTSLSLDDNDNRITSYLHSPFTGRSCIPLHGVSPEVNELLMEALQVLSLFEIQELYQTLLA 1372
NTT 1075 DAFHTSCVAVPSISQGLRIWLCPHCRSRSEKPLEKILPILASLQRIVRIPGSDALRYMTERVNWQRAQQLISSGNLKFVQDRVSGGLLSRWQASQVSDTNKVSQPFQTSLSLDDNDNRITSYLHSPFTGRSCIPLHGVSPEVNELLMEALQVLSLFEIQELYQTLLA 1250

RBP2H1 1409 KESPAQQTDRSSFPVPSSEKNDCCGRKDGINSLEKLRKRLEREGLSSEWERVKKMRTPKKKIKLISHPKDMNFKLERERSVELVRSAAETHSLSPDTSYSEQDESEDAICPAVSCIQPEGDEVDMVQDGGSCNQWFHQVCGVSPSEMAEKEDIYICVRCITVKDAPSRK 1580
PLU-1 1373 KESPAQQTDRSSFPVPSSEKNDCCGRKDGINSLEKLRKRLEREGLSSEWERVKKMRTPKKKIKLISHPKDMNFKLERERSVELVRSAAETHSLSPDTSYSEQDESEDAICPAVSCIQPEGDEVDMVQDGGSCNQWFHQVCGVSPSEMAEKEDIYICVRCITVKDAPSRK 1544
NTT 1251 KESPAQQTDRSSFPVPSSEKNDCCGRKDGINSLEKLRKRLEREGLSSEWERVKKMRTPKKKIKLISHPKDMNFKLERERSVELVRSAAETHSLSPDTSYSEQDESEDAICPAVSCIQPEGDEVDMVQDGGSCNQWFHQVCGVSPSEMAEKEDIYICVRCITVKDAPSRK 1422

```

Fig.S6.

Fig.S6. Sequence alignment of KDM5B isoforms. RBP2-H1, PLU-1 and NTT are compared. The residues encoded by exon-6 region (pink rectangle), the N-terminal region (yellow rectangle) missing in NTT and the putative N-terminal degron motif (orange) are highlighted.

Fig.S7.

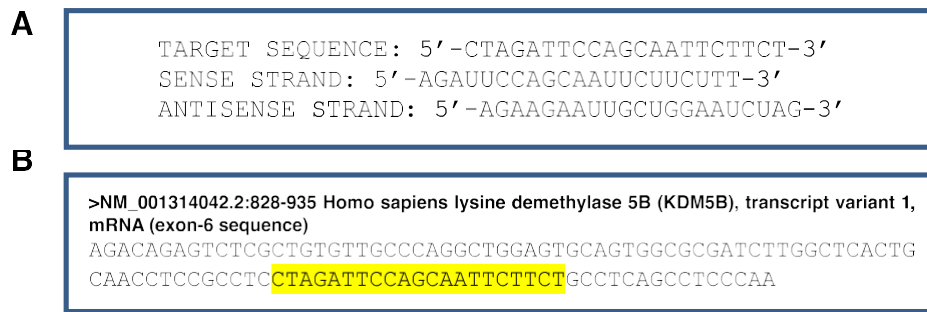


Fig.S7. Custom siRNA targeting exon-6 of KDM5B. **A.** The target sequence (exon-6 region of KDM5B transcripts) for the custom siRNA and the siRNA sequence are shown. **B.** The exon-6 sequence of KDM5B-RBP2-H1 transcript variant) is reported according to RefSeq (NCBI). In yellow, the region targeted from the custom siRNA, that correspond to a region in the middle of the exon-6 sequence.

Fig.S8

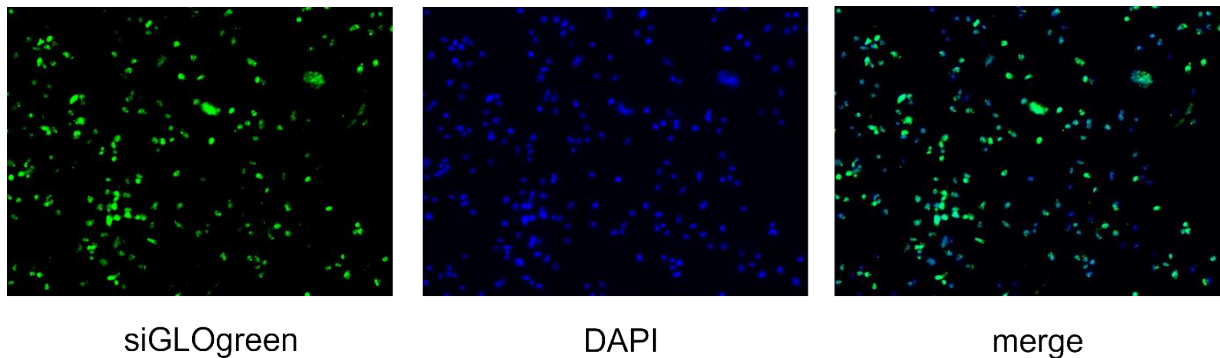


Fig.S8. RNAi transfection efficiency. A fluorescent oligo RNA (siGLOgreen) was used to check the transfection efficiency of MDA-MB-231 cells.

Fig.S9

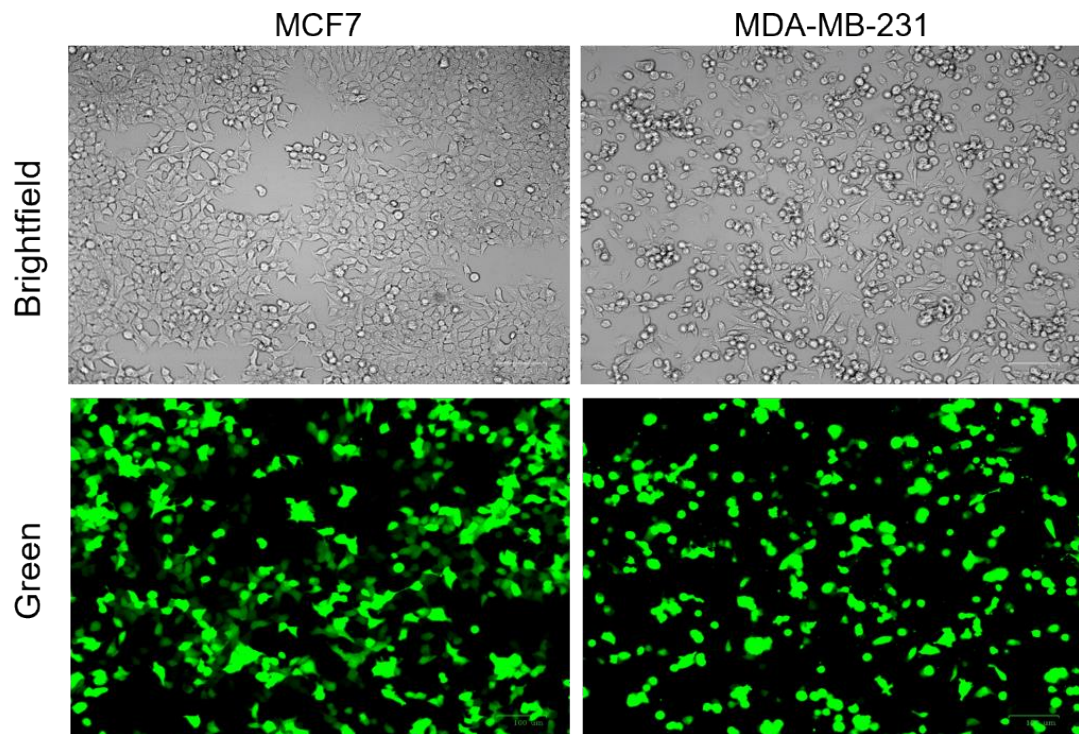


Fig.S9. Post-transfection analysis of cells. Green fluorescence and brightfield images demonstrate about 60% and 50% transfection efficiency in MCF7 and MDA-MB-231.

Fig.S10

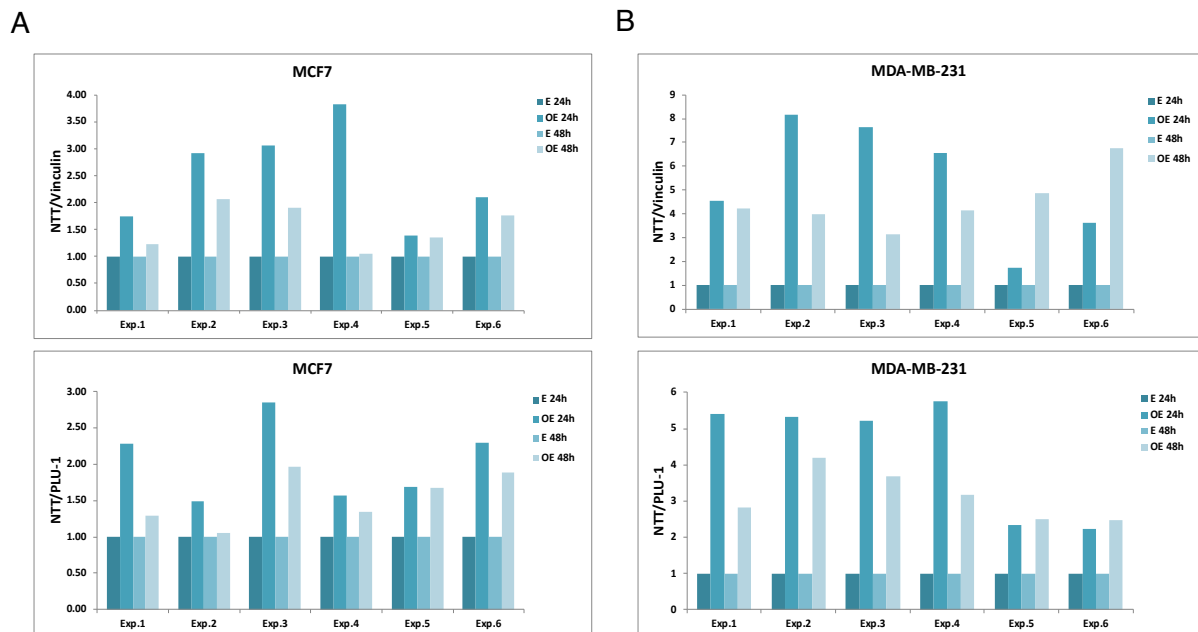


Fig.S10. Quantification of NTT overexpression at 24h and 48h upon transfection. A. The overexpression of NTT compared to the control condition (E, Empty vector) was higher at 24 hours (OE 24h) rather than at 48 hours (OE 48h) from transfection in MCF7 cells; the quantification using vinculin

as reference (upper panel) and the relative amount of NTT over PLU-1 (lower panel) are shown. **B.** The NTT protein was efficiently expressed also in MDA-MB-231, both at 24h and 48h after transfection; over-expression quantification relative to Empty control conditions is shown both using vinculin as a calibrator (upper panel) and using the relative amount of NTT over PLU-1 (lower panel).

Fig.S11

PLU-1_nuclear localization signals: **I**, aa 227–234; **II**, aa 295–302; **III**, aa 1103–1120; **IV**, aa 1400–1417; **V**, aa 1434–1440. From Lu et al. (1999)

>(NP_006609.3) Lysine-specific demethylase 5B; PLU-1 isoform		>Lysine-specific demethylase 5B; NTT isoform	
MEAAATTLHPGPRPALPLGGPGFLGFLPPECEPVFPEFSWEEFADPPFAFTHKIRPIAEQTG	60	MGFAPGKAVGSHTRGHYERTLNFPYNIPLSGDSLRLCLQKPNLTDTDKKEYKPHDTPQRQS	60
ICKYRFPFPDQWPPFACDVKLHFTPIQRLNLEAQTRVKNLFLDQIAKYWELQGSTLKI	120	VQPSETCP PARRAKRM RAER QSLAVLPRLECSGAILAHCNLRLLDSSNSASASQ AMNIK	120
PHVERKILDLFLQKLVAEEGGFAVVCCKDRKWKIATKMGFAPGKAVGSHIRGHYERILN	180	IEPEETTEARTHNLRRMGCPTEKCEKEMKSSIKQEPIERKDYIVENEK PKKRSRSK	180
PYNLFLSGDSLRLCLQKPNLTDTDKKEYKPHDIPQRQSVQFSETCP PARRAKRM RAEAMN	240	ATNAVDLYVCLCGSNDDEDRLLCDGCDSDYHTFCLIPPLHDVPGDWRCPKCLAQES	240
IKIPEEETTEARTHNLRRMGCPTEKCEKEMKSSIKQEPIERKDYIVENEK PKKRSRS	300	KPQEAQFQEAARDYTLRTFGEMADAFKSDYFNMPVHMVPTLVEKEFWRLVSTIEEDVT	300
KKATNAVDLYVCLCGSNDDEDRLLCDGCDSDYHTFCLIPPLHDVPGDWRCPKCLAQ	360	VEYGADIASKEFGSGFPVRDQGIKLSPEEEYLDGWNLNMPVMEQSVLAHITADICGM	360
CSKPOEAFQEAARDYTLRTFGEMADAFKSDYFNMPVHMVPTLVEKEFWRLVSTIEED	420	KLPLVYGMCFSSFCWHLEDHWSYINYLHWGEFTKTYVYVGGYAAEQLENVMKKLAPLFL	420
VITYEGADIASKEFGSGFPVRDQGIKLSPEEEYLDGWNLNMPVMEQSVLAHITADIC	480	VSQPDLLHQLVTIMNPNTLMTHEVPPYRTNQCAGEFVITFPFRAYHSGFNQGFNAEAVNF	480
GMKLPWLVYGMCFSSFCWHLEDHWSYINYLHWGEFTKTYVYVGGYAAEQLENVMKKLAPLFL	540	CTVDWLPGRQCVHEYRLHRYCVFSDHEMICKMASKADVLVAVASTVQKDMAIMIEDE	540
LFVSPQDLLHQLVTIMNPNTLMTHEVPPYRTNQCAGEFVITFPFRAYHSGFNQGFNAEAVNF	600	KALRETVRKLGVIDSERMDFELLPDDEQCVCKCTTCFMSAISCSCPKCLLVCLHHVKEL	600
NFCTVDWLPGRQCVHEYRLHRYCVFSDHEMICKMASKADVLVAVASTVQKDMAIMIEDE	660	CSCPPYKLYRYTLDDLYPMNALKLRASYNWALNVNEALEAKINKKSLVSPKAL	660
DEKALRETVRKLGVIDSERMDFELLPDDEQCVCKCTTCFMSAISCSCPKCLLVCLHHVKEL	720	IEESEMKKFPDNDLLRHLRLVTDQAEKCAVAQQLNGKQRQTRYSRGGKSNQNLTVNEL	720
ELCSCPPYKLYRYTLDDLYPMNALKLRASYNWALNVNEALEAKINKKSLVSPKAL	780	RQFVTQLYALPCVLSQTPLLKDLLNRVEDFQHSQKLLSEETPSAAELQDLDLVSEFEDV	780
ALIEESEMKKFPDNDLLRHLRLVTDQAEKCAVAQQLNGKQRQTRYSRGGKSNQNLTVNEL	840	ELPQLAEMRIREQARWLEEVQACLDPSLLTDDMRRLIDLGVGLAPYSAVEKAMARLQ	840
ELRQFVTQLYALPCVLSQTPLLKDLLNRVEDFQHSQKLLSEETPSAAELQDLDLVSEFEDV	900	ELLTVSEHWDKAKSLLKARPRHSLNSLATAVKEIEEPAYLFGAALKDSVQQRARDWLQ	900
DVELPQLAEMRIREQARWLEEVQACLDPSLLTDDMRRLIDLGVGLAPYSAVEKAMARLQ	960	DVEGLQAGGRVPLDITLIELVTRGRSIPVHLNLSLPRLETVAEVOAMKECAVNTFLTENS	960
LQELLTVSEHWDKAKSLLKARPRHSLNSLATAVKEIEEPAYLFGAALKDSVQQRARDWLQ	1020	PYSLVLELCPHCDIGLLGLK RQKRLKLEPLNGKKS TKLESLSLERALTESKETASAM	1020
LQDVEGLQAGGRVPLDITLIELVTRGRSIPVHLNLSLPRLETVAEVOAMKECAVNTFLTENS	1080	ATLGEARLRREMEALQSLRLANEGKLLSPLQDVIDIKICLQKAPAAPMIQCELCRDAFHTS	1080
NSPYSLELVCPRCDIGLLGLK RQKRLKLEPLNGKKS TKLESLSLERALTESKETASAM	1140	CVAVPISIQGLRIVLCPHRRSEKPPLEKILPLLASLQIRIVRLEPEGDALRYMERTVNW	1140
AMATLGEARLRREMEALQSLRLANEGKLLSPLQDVIDIKICLQKAPAAPMIQCELCRDAFHTS	1200	QHRAQQLSSGNLKFVQDRVGSGLYSRWQASAGQVSDTNKVSQPPGTTFSLSLDDWDR	1200
TSCVAVPISIQGLRIVLCPHRRSEKPPLEKILPLLASLQIRIVRLEPEGDALRYMERTVNW	1260	TSYLHSPFSTGRSCLPIHGVSPVNEELMEAQQLQVSLPITQELYQTLAKPSAQQDTR	1260
NWQHRAQQLSSGNLKFVQDRVGSGLYSRWQASAGQVSDTNKVSQPPGTTFSLSLDDWDR	1320	SSPVRPSSSEKNDCCRGK RDGINSLEKRLKRRLER GLSSERWERVKMRT PKKKIKL SH	1320
NRTSYLHSPFSTGRSCLPIHGVSPVNEELMEAQQLQVSLPITQELYQTLAKPSAQQDTR	1380	SHFQDMNFKLERERSYELVRSATHSLSFSDTSYSEQEDSEDAICPAVSCIQPEGDEV	1380
DRSSPVRPSSSEKNDCCRGK RDGINSLEKRLKRRLER GLSSERWERVKMRT PKKKIKL SH	1440	VQCDGSCNQWHEQVCGVSPSEMAEKEDYICVROTVKDAERSK	1422
SHFQDMNFKLERERSYELVRSATHSLSFSDTSYSEQEDSEDAICPAVSCIQPEGDEV	1500		
DWVQCDGSCNQWHEQVCGVSPSEMAEKEDYICVROTVKDAERSK	1544		

Fig.S11. Nuclear Localization Signals (NLSs) of PLU-1 according to Lu et al. 1999⁴². All the NLSs of PLU-1 (sequence on the left) are also in the NTT isoform (sequence on the right) suggesting that it can localize into the nucleus.

Fig.S12

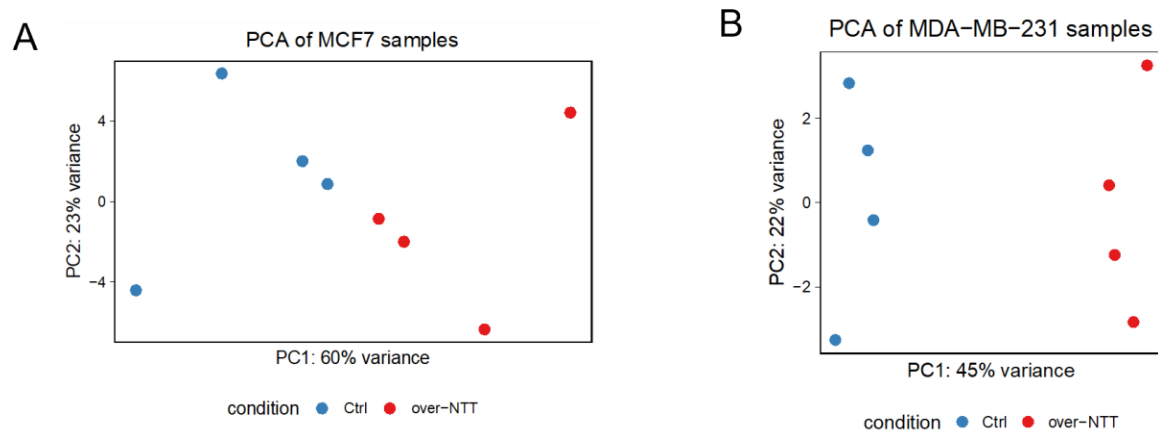


Fig.S12. PCA analysis upon NTT overexpression. The replicates (N=4 for each cell line) over-expressing NTT cluster differently from control replicates and most of the variance observed (explained by PC1) can be associated to the NTT over-expression, both in MCF7 cells (panel A) and MDA-MB-231 (panel B).

Fig.S13

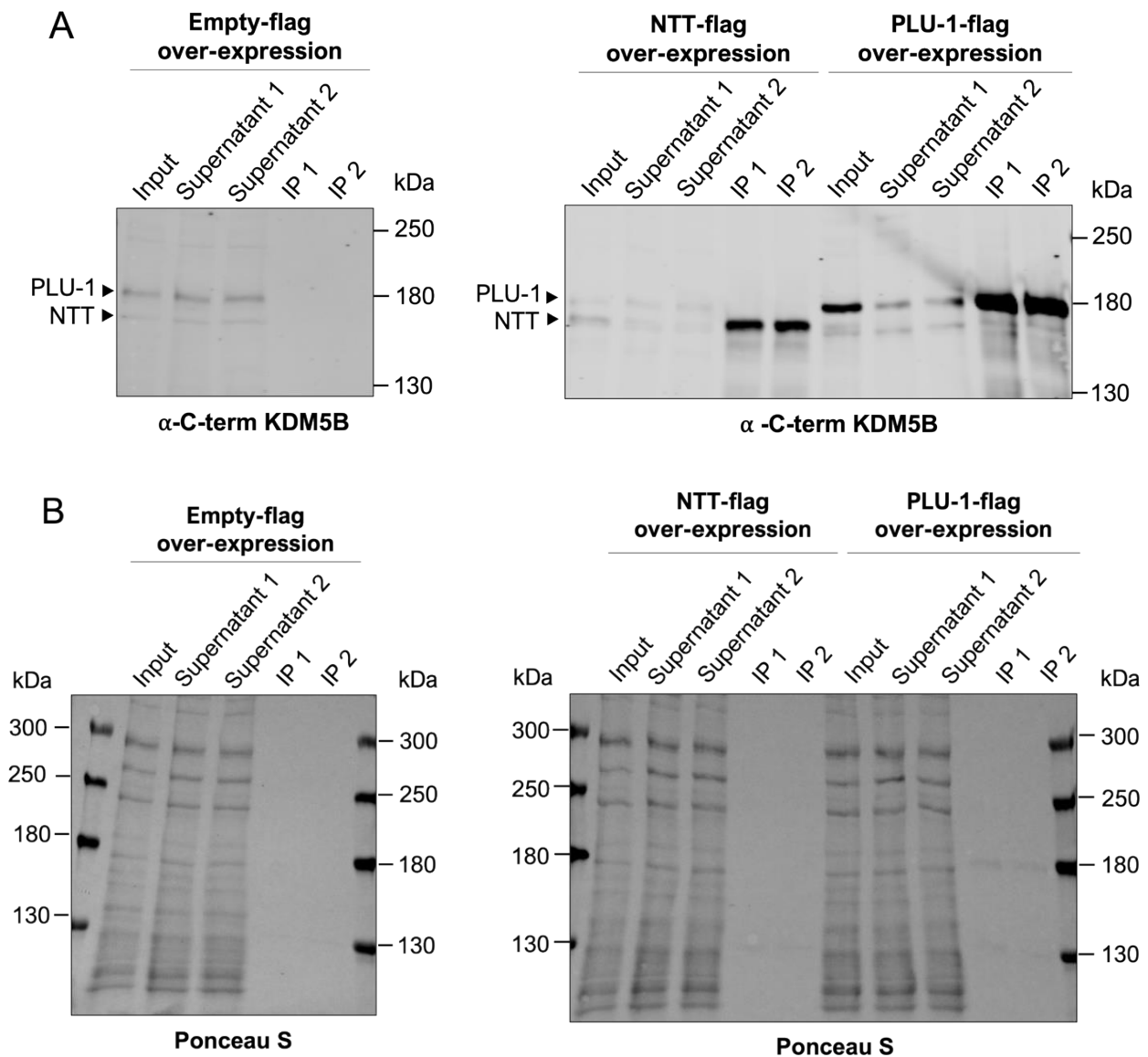


Figure S13. Co-IP experiments in MCF7 cells. A. Using the anti-flag resin, the two KDM5B isoforms (NTT and PLU-1) were immunoprecipitated to study the molecular interactors through mass-spectrometry. Here it is shown the blot upon two IP experiments done for the setting up of the Co-IP protocol. In each IP1, 20 μ L of 50% slurry (10 μ L of resin bed volume) was used; in each IP2 (40 μ L of 50% slurry (20 μ L of resin bed volume) was used. The blots using the anti-C-terminal antibody (Ab1, see schematic in Fig.6B) are shown. **B.** Ponceau staining of the western blots shown in panel A.

Fig.S14

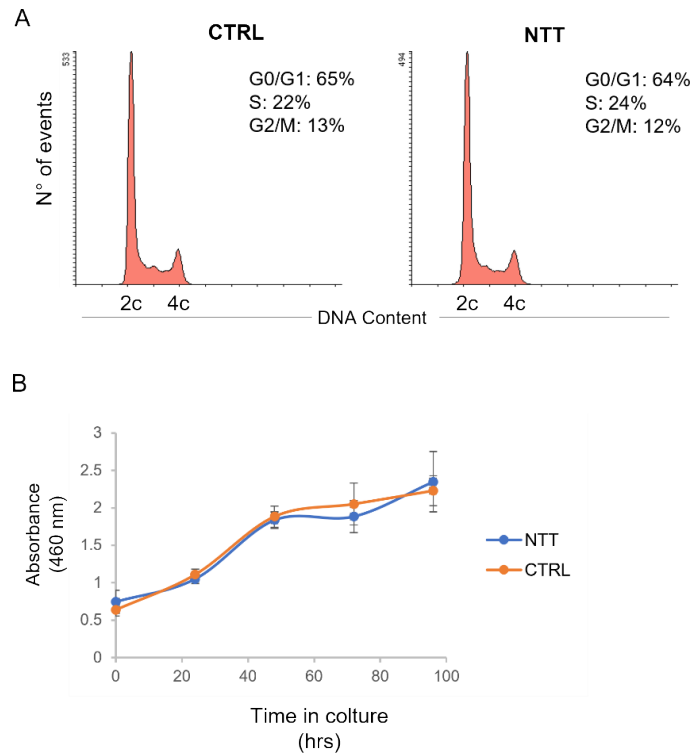


Figure S14. NTT over-expression in MCF7 cells does not affect cell cycle. **A.** Examples of cytofluorimetric patterns obtained from control cells (CTRL) and PLU-1-NTT over-expressing cells (NTT). Cells were transfected with empty and PLU-1-NTT vectors, respectively. Flow-cytometry analysis of DNA content was done at 24 hours from transfection. Cell cycle phase distribution is reported as an inset in the figures. **B.** NTT over-expression in MCF7 cells does not affect cell proliferation. MCF7 cells were transfected with NTT over-expression vector or empty control vector. Cell proliferation were assayed by the WST-8 colorimetric assay. No significant differences were observed between control and NTT overexpressing cells. Each point represents the average of 8 replicates, vertical bars indicate standard deviation values.

9. SUPPLEMENTAL REFERENCES

- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Res.* 2000;28:235–42.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596:583–9.
- Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* 2021;49:D412–9.
- Puntervoll P, Linding R, Gemünd C, Chabanis-Davidson S, Mattingsdal M, Cameron S, et al. ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res.* 2003;31:3625–30.
- Janson G, Paiardini A. PyMod 3: a complete suite for structural bioinformatics in PyMOL. *Bioinformatics.* 2021;37:1471–2.
- Noberini R, Restellini C, Savoia EO, Bonaldi T. Enrichment of histones from patient samples for mass spectrometry-based analysis of post-translational modifications. *Methods.* 2020;184:19–28.
- Noberini R, Bonaldi T. A Super-SILAC Strategy for the Accurate and Multiplexed Profiling of Histone Posttranslational Modifications. *Methods Enzymol.* 2017;586:311–32.
- Noberini R, Savoia EO, Brandini S, Greco F, Marra F, Bertalot G, et al. Spatial epi-proteomics enabled by histone post-translational modification analysis from low-abundance clinical samples. *Clin Epigenetics.* 2021;13:145.
- Noberini R, Osti D, Miccolo C, Richichi C, Lupia M, Corleone G, et al. Extensive and systematic rewiring of histone post-translational modifications in cancer model systems. *Nucleic Acids Res.* 2018;46:3817–32.
- Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol.* 2014;32:223–6.
- Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 2017;14, 417–419.

- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5, R80.
- Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* 2015;12, 115–121.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014;15, 550.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545-50.

ACKNOWLEDGMENTS

Vorrei ringraziare di cuore tutte le persone che mi hanno supportato in questo stupendo cammino. Ringrazio in particolar modo il prof. Rodolfo, mentore lungimirante e ottimista, sempre pronto a incoraggiare. Grazie per avermi dato la possibilità di vivere un'esperienza arricchente fin dall'inizio di questo percorso. A tal proposito ringrazio molto anche Virginia, che mi ha ospitata nel suo laboratorio di Dundee per un anno, periodo in cui ho avuto modo di incontrare persone che mi hanno fatta crescere non solo sotto l'aspetto professionale, ma anche personale. Grazie anche a tutti i miei collaboratori e colleghi di laboratorio. Un grazie particolare a Cecilia, Valerio, Carlo e Silvia che sono sempre stati dei capi saldi. Ringrazio dal profondo la mia famiglia e i miei amici che credono sempre in me e nelle mie capacità, a volte più di quanto non faccia io stessa.

APPENDIX








Publications produced during the Ph.D. by the candidate and related to the thesis topic:

- Di Nisio, E., Licursi, V., Mannironi, C., Buglioni, V., Paiardini, A., Robusti, G., et al. (2023). A truncated and catalytically inactive isoform of KDM5B histone demethylase accumulates in breast cancer cells and regulates H3K4 tri-methylation and gene expression. *Cancer Gene Therapy*. doi: [10.1038/s41417-022-00584-w](https://doi.org/10.1038/s41417-022-00584-w).

Cancer Gene Therapy www.nature.com/cgt

ARTICLE OPEN Check for updates

A truncated and catalytically inactive isoform of KDM5B histone demethylase accumulates in breast cancer cells and regulates H3K4 tri-methylation and gene expression

Elena Di Nisio ^{1,2}, Valerio Licursi ³, Cecilia Mannironi ³, Valentina Buglioni ¹, Alessandro Paiardini ⁴, Giulia Robusti ⁵, Roberta Noberini⁵, Tiziana Bonaldi^{5,6} and Rodolfo Negri ^{1,3,8}✉

© The Author(s) 2023

KDM5B histone demethylase is overexpressed in many cancers and plays an ambivalent role in oncogenesis, depending on the specific context. This ambivalence could be explained by the expression of KDM5B protein isoforms with diverse functional roles, which could be present at different levels in various cancer cell lines. We show here that one of these isoforms, namely KDM5B-NTT, accumulates in breast cancer cell lines due to remarkable protein stability relative to the canonical PLU-1 isoform, which shows a much faster turnover. This isoform is the truncated and catalytically inactive product of an mRNA with a transcription start site downstream of the PLU-1 isoform, and the consequent usage of an alternative ATG for translation initiation. It also differs from the PLU-1 transcript in the inclusion of an additional exon (exon-6), previously attributed to other putative isoforms. Overexpression of this isoform in MCF7 cells leads to an increase in bulk H3K4 methylation and induces derepression of a gene cluster, including the tumor suppressor *Cav1* and several genes involved in the interferon- α and - γ response. We discuss the relevance of this finding considering the hypothesis that KDM5B may possess regulatory roles independent of its catalytic activity.

Cancer Gene Therapy; <https://doi.org/10.1038/s41417-022-00584-w>

INTRODUCTION


Histone methylation is a dynamic modification that is tightly controlled during cellular differentiation by the coordinated key action of histone methyltransferases and demethylases [1]. Histone lysine demethylases (KDMs) act as both readers and erasers of this modification. Specifically, there are two enzymatic classes of histone demethylases in human: the FAD-dependent amine oxidases, belonging to the lysine-specific histone demethylase (LSD) family, also known as the KDM1 subfamily, and the Fe(II) and 2-oxoglutarate-dependent oxygenases with a conserved JmjC catalytic domain, belonging to the Jumoni histone demethylases (JHDMs) family, also known as the KDM2-KDM7 subfamilies [2, 3]. KDMs show a high level of substrate specificity, with each acting on specific lysine residues with different degrees of methylation. Histone methylation may lead to different transcriptional outcomes, such as gene activation or repression, depending on the residue involved, the degree of methylation, and the location of the mark in the genome [4, 5]. Among the KDMs, KDM5 (or JARID1) enzymes act on H3K4me2/me3, playing an important role in transcriptional regulation, especially during development and differentiation. KDM5 enzymes can act as epigenetic repressors by removing the transcriptional activating mark H3K4me3 at gene promoter regions, but they can also act as transcriptional activators by removing H3K4me3 from the body of actively transcribed genes to safeguard transcriptional elongation by repressing spurious intragenic transcription [6]. Moreover, they can contribute to transcriptional activation by converting H3K4me3 to H3K4me1 at enhancer regions, as H3K4me1 modification combined with acetylated H3K27 is predictive of active enhancers. KDM5 proteins are also involved in DNA double-strand break repair [7, 8] and chemoresistance [9–11]. The mammalian KDM5 subfamily consists of KDM5A (also known as JARID1A or RBP2), KDM5B (also known as JARID1B or PLU-1), KDM5C (also known as JARID1C or SMCX), and KDM5D (also known as JARID1D or SMCY). The deregulated expression of KDM5 enzymes has been extensively documented in many cancer types and contributes significantly to tumor initiation and progression [12]. Specifically, KDM5B is a master regulator of H3K4-methylome in stem cells, development, and cancer [13]. It was initially identified as a markedly upregulated gene in breast cancer [14] and was later found to be overexpressed in prostate [15], bladder, and lung cancers [16], and in stem-like subpopulations in melanomas [17]. In normal adult tissues, KDM5B shows a highly restricted expression, except in the testis, where it may contribute to transcriptional control during spermatogenesis, suggesting that it could belong to the class of testis/cancer antigens [18]. In breast

¹Department of Biology and Biotechnologies “C. Darwin”, Sapienza University of Rome, via dei Sardi 70, 00185 Rome, Italy. ²MRC Protein Phosphorylation and Ubiquitylation Unit, School of Life Sciences, University of Dundee, Sir James Black Centre, Dow Street, DD1 5EH Dundee, Scotland, UK. ³Institute of Molecular Biology and Pathology (IBPM), National Research Council (CNR) of Italy, Via degli Apuli 4, 00185 Rome, Italy. ⁴Department of Biochemical Sciences, Sapienza University of Rome, p.le Aldo Moro 5, 00185 Rome, Italy. ⁵Department of Experimental Oncology, IEO, European Institute of Oncology IRCCS, Via Adamello 16, 20139 Milan, Italy. ⁶Department of Oncology and Hematology-Oncology, University of Milan, Milan 20122, Italy. ✉email: rodolfo.negri@uniroma1.it


Received: 25 July 2022 Revised: 20 December 2022 Accepted: 21 December 2022
Published online: 26 January 2023

SPRINGER NATURE

- Di Nisio, E., Lupo, G., Licursi, V., and Negri, R. (2021). The Role of Histone Lysine Methylation in the Response of Mammalian Cells to Ionizing Radiation. *Front Genet* 12, 639602. doi: 10.3389/fgene.2021.639602.



REVIEW
published: 30 March 2021
doi: 10.3389/fgene.2021.639602



The Role of Histone Lysine Methylation in the Response of Mammalian Cells to Ionizing Radiation

Elena Di Nisio¹, Giuseppe Lupo¹, Valerio Licursi^{1} and Rodolfo Negri^{1,2*}*

¹ Department of Biology and Biotechnology Charles Darwin, Sapienza University of Rome, Rome, Italy, ² Institute of Molecular Biology and Pathology, National Research Council (IBPM-CNR), Rome, Italy

OPEN ACCESS

Edited by:
Ki Moon Seong,
Korea Institute of Radiological
and Medical Sciences, South Korea

Reviewed by:
Abhijit Shukla,
Memorial Sloan Kettering Cancer
Center, United States
Joo Mi Yi,
College of Medicine, Inje University,
South Korea

***Correspondence:**
Valerio Licursi
valerio.licursi@uniroma1.it
Rodolfo Negri
rodolfo.negri@uniroma1.it

Specialty section:
This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 09 December 2020
Accepted: 11 March 2021
Published: 30 March 2021

Citation:
Di Nisio E, Lupo G, Licursi V and
Negri R (2021) The Role of Histone
Lysine Methylation in the Response
of Mammalian Cells to Ionizing
Radiation. *Front. Genet.* 12:639602.
doi: 10.3389/fgene.2021.639602

Eukaryotic genomes are wrapped around nucleosomes and organized into different levels of chromatin structure. Chromatin organization has a crucial role in regulating all cellular processes involving DNA-protein interactions, such as DNA transcription, replication, recombination and repair. Histone post-translational modifications (HPTMs) have a prominent role in chromatin regulation, acting as a sophisticated molecular code, which is interpreted by HPTM-specific effectors. Here, we review the role of histone lysine methylation changes in regulating the response to radiation-induced genotoxic damage in mammalian cells. We also discuss the role of histone methyltransferases (HMTs) and histone demethylases (HDMs) and the effects of the modulation of their expression and/or the pharmacological inhibition of their activity on the radio-sensitivity of different cell lines. Finally, we provide a bioinformatic analysis of published datasets showing how the mRNA levels of known HMTs and HDMs are modulated in different cell lines by exposure to different irradiation conditions.

Keywords: DNA damage, ionizing radiation, DNA repair, HPTMs, histone methylation

INTRODUCTION

The first and basic level of chromatin organization consists in the wrapping of genomic DNA around histone octamers forming the nucleosomes (Luger et al., 2012). Nucleosomes mediate the interactions of the genomic DNA with all the effectors involved in fundamental biological processes such as transcription, replication, recombination, damage response and repair (Groth et al., 2007; McGinty and Tan, 2015). Histone tails, which protrude from the nucleosome core, are the target of a plethora of post-translational chemical modifications, such as: acetylation, methylation, phosphorylation, ubiquitylation, sumoylation and ADP-ribosylation (Campos and Reinberg, 2009; Bannister and Kouzarides, 2011). These Histone Post-Translational Modifications (HPTMs) form a sophisticated code of signals, known as histone code, which regulates the interactions of the genome with very important cellular effectors involved in several biological processes. The HPTMs

Abbreviations: ATM, Ataxia-telangiectasia mutated; DDR, DNA damage response; DivA, DSB Inducible via AsiSI; DSBs, Double-strand breaks; HDMs, Histone demethylases; HMTs, Histone methyltransferases; HPTMs, Histone post-translational modifications; HR, Homologous recombination; KDMs, Lysine demethylases; KMTs, Lysine methyltransferases; L3MBTL1, Lethal-3-malignant brain tumor-like protein-1; LEDGE, Lens epithelium-derived growth factor; LET, Linear energy transfer; MMS, Methyl methanesulfonate; NHEJ, Non-homologous end joining; SSBs, Single-strand breaks.

Frontiers in Genetics | www.frontiersin.org
1
March 2021 | Volume 12 | Article 639602