Evidence for mRNA expression-mediated gene repression in human cells

Ina Hollerer[1], Juliet C. Barker[1,3*], Victoria Jorgensen[1*], Amy Tresenrider[1], Claire Dugast-Darzacq[2], Leon Y. Chan[1], Xavier Darzacq[2], Robert Tjian[2], Elçin Ünal[1§#] and Gloria A. Brar[1§#]

[1]Department of Molecular and Cell Biology, Barker Hall, University of California, Berkeley, CA, USA, 94720

[2]Department of Molecular and Cell Biology, Li Ka Shing Center, University of California, Berkeley, CA, USA, 94720

[3]Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA, 02139

*,§ These authors contributed equally to this work.

Correspondence: gabrar@berkeley.edu and elcin@berkeley.edu

Abstract:

We recently discovered a common mode of gene regulation in budding yeast, by which mRNA production represses protein expression. Whether this regulatory mechanism is conserved was unknown. Here we find that a similar mechanism regulates the human oncogene *MDM2*, which is transcribed from two promoters. Transcription from the distal *MDM2* promoter produces a poorly translated mRNA isoform and results in poor expression of the well-translated transcript produced from the proximal *MDM2* promoter, associated with the presence of repressive histone H3K36 trimethylation marks. Accordingly, down-regulation of transcription from the distal promoter actually *up-regulates* MDM2 protein levels. We conclude that this non-canonical transcript toggling mechanism, first defined in yeast, is conserved in human cells and propose that it might commonly be used to regulate mammalian gene expression.

Introduction:

Recently, we defined a form of gene regulation that challenges the broad assumption that mRNA production leads to increased protein production (Chen et al., 2017; Chia et al., 2017). We found that in budding yeast meiosis, the amount of protein for the conserved kinetochore protein Ndc80 is determined by an unexpected mechanism in which mRNA production from a more distal *NDC80* promoter inhibits Ndc80 protein synthesis through coordination of transcriptional and translational interference: the distal promoter-driven transcript cannot be efficiently translated into protein and its transcription interferes with the proximal *NDC80* promoter activity in *cis*. In this manner, a 5'-extended and poorly translated mRNA isoform represses the production of a canonical mRNA isoform and therefore inhibits Ndc80 protein production (Chen et al., 2017; Chia et al., 2017).

We showed that this integrated mode of regulation relies on three key features (Chen et al., 2017; Chia et al., 2017). First, a regulated switch between alternative promoters for the same gene leads to the usage of different transcription start sites (TSSs). Second, due to upstream open reading frame (uORF)-mediated translational repression, the distal promoter-generated transcript is inefficiently translated. Third, transcription from the distal promoter represses the expression of the canonical mRNA isoform through co-transcriptional histone modifications. As a net result of all three factors, the activation of the distal promoter for *NDC80* results in a *decrease* in protein production from this locus. We termed the distal promoter-generated and translationally silent transcript "*NDC80*^{LUTI}" for long undecoded transcript isoform, because, despite containing the

entire *NDC80* ORF, *NDC80*^{LUTI} is not efficiently translated into protein (Chen et al., 2017; Chia et al., 2017).

We also found that this mechanism is common in budding yeast cells, with 379 other genes showing protein levels that are regulated over time through meiotic differentiation by toggling between transcript isoforms, as a result of a LUTI mRNA-based mechanism (Cheng et al., 2018). While these studies were exclusively performed using budding yeast, two of the three hallmarks of LUTI mRNA-mediated gene repression are common in mammals. First, almost half of human genes show evidence of alternative promoter usage, resulting in transcripts that differ in their 5' leader (Wang et al., 2016). Second, transcripts with extended 5' leaders that contain uORFs result, in some cases, in a poorly translated transcript compared to isoforms with shorter 5' leaders (Floor and Doudna, 2016; Law et al., 2005). These features are observed in global studies, and they were also independently defined for several genes, including *TGFβ3*, *AXIN2*, and Mouse double-minute 2 homolog (*MDM2*), an oncogene and repressor of the tumor suppressor p53 (Arrick et al., 1994; Barak et al., 1994; Brown et al., 1999; Hughes and Brady, 2005). For example, the *MDM2* isoform produced from the distal P1 promoter contains a longer 5' leader than the one produced from the proximal P2 promoter (Figure 1B). This P1-driven *MDM2* isoform is poorly translated due to the presence of two uORFs in its extended 5' leader (Brown et al., 1999). Although it is well established that P2 can be activated by p53 (Barak et al., 1994; Honda et al., 1997; Wu et al., 1993), it is not known whether transcription from the P1 promoter regulates P2 activity. Our interpretation of previous studies suggested that the same type of transcript

toggling mechanism that we discovered for *NDC80* in yeast might control the human *MDM2* locus. Here, we report evidence that this is indeed the case and that LUTI-mediated regulation is conserved from yeast to human.

Results and discussion:

A key prediction, if *MDM2* is regulated by a LUTI mRNA mechanism, would be an inverse relationship between the presence of the two *MDM2* transcript isoforms, such that reduction in transcription from P1 should lead to increased transcription from P2. If instead, the two transcript isoforms vary independently, then no LUTI-based mechanism would apply. We first examined the relative abundance of the *MDM2* isoforms in human Embryonic Stem Cells (hESCs) using reverse transcription coupled with isoform-specific quantitative polymerase chain reaction (RT-qPCR) and found expression of both the distal P1 promoter-derived transcript and the proximal P2 promoter-derived *MDM2* transcript, hereon referred to as $MDM2^{PROX}$ (Figure 1C). Analysis of *MDM2* expression in data from a recent global study confirmed this result and further showed that $MDM2^{PROX}$ was enriched in polysomes in human embryonic stem cells (hESCs), while the P1-derived transcript, hereon referred to as $MDM2^{LUTI}$ (for reasons established below), was not [Figure 1D and (Blair et al., 2017)]. This translational signature is consistent with published data from several cell lines (Brown et al., 1999; Landers et al., 1997).

Because we observed LUTI-based regulation to be common during differentiation in yeast, as marked by temporally-regulated transcript toggling, we examined samples

taken during the process of hESC differentiation for evidence of this type of signature for *MDM2*. We found that, indeed, as neuronal differentiation progressed (Blair et al., 2017), a switch in transcript isoform expression from $MDM2^{PROX}$ to $MDM2^{LUTI}$ could be seen (Figure 1C). This inverse correlation between $MDM2^{PROX}$ *and* $MDM2^{LUTI}$ expression was most evident between hESCs and neuronal precursors (NPCs) (Figure 1C). We also observed an anti-correlation between $MDM2^{PROX}$ *and* $MDM2^{LUTI}$ expression as hESCs differentiated into an endodermal fate, as determined by endoderm-specific markers (Figure 1E, S1). Furthermore, using a siRNA-knockdown-validated antibody, we detected a clear decrease in MDM2 protein expression as hESCs differentiated into NPCs, which correlated with the timing of increased $MDM2^{LUTI}$ expression (compare Figure 1C and S2). This inverse pattern of proximal and distal promoter usage seen during hESC differentiation fits a LUTI-like model, and also suggests that production of the distal $MDM2^{LUTI}$ transcript might be capable of repressing the production of MDM2 protein through interference with transcription from the proximal promoter.

In order to directly test whether $MDM2^{LUTI}$ production causes repression of the P2 promoter activity, we inhibited transcription from P1 by using CRISPRi (Gilbert et al., 2013). To this end, we first examined MCF-7 breast cancer cells stably encoding the catalytically dead dCas9. Expressing each of four different single guide RNAs (sgRNAs) targeting the P1 promoter region led to a modest but significant increase, of up to 2-fold, in $MDM2^{PROX}$ transcript levels, which was associated with the reduction of transcription from P1 (Figure 2A, S3). This result was notable, given that the maximal knockdown of

P1 activity was only 40% relative to control cells in these lines (Figure 2A). We tried to enhance the P1 transcriptional knockdown by using CRISPRi in MCF-7 cells that carry a version of dCas9 fused to the Krüppel-associated box (KRAB) transcriptional repression domain (Gilbert et al., 2013). However, targeting of dCas9-KRAB to the P1 promoter led to repression of both the P1 and P2 promoters (Figure S4). This finding is consistent with the long-range effect of the KRAB domain up to 1Kb (Gilbert et al., 2014), beyond the 845 bp distance between the P1 and P2 regulated transcription start sites. Therefore, we performed all subsequent experiments using cell lines that stably expressed dCas9 without the KRAB domain, as this first-generation version of CRISPRi allowed us to achieve promoter-specific repression.

We further probed the relationship between P1 and P2 by knockdown of the gene encoding p53 (*TP53*) in MCF-7 cells using CRISPRi. Given that p53 is a well-characterized transcriptional activator for P2, it was not surprising that *TP53* knockdown resulted in a significant reduction (43%) of the P2-derived $MDM2^{PROX}$ transcript (Figure 2B, left panel). However, additional CRISPRi knockdown of $MDM2^{LUTI}$ still resulted in the transcriptional activation of P2, as evidenced by the 2- to 3-fold increase in $MDM2^{PROX}$ levels in this background compared to the *TP53* knockdown alone (Figure 2B, right panel; S5). The observation that the repression of $MDM2^{LUTI}$ leads to an increase in $MDM2^{PROX}$ expression, even in cells with reduced p53 levels, suggests that transcription from P1 actively represses P2 activity and that relief of this repression alone can lead to increased expression of $MDM2^{PROX}$ independent of p53.

To achieve higher transcriptional knock-downs we performed similar experiments in K562, a *TP53*$^{-/-}$ myeloid leukemia cell line that routinely shows robust CRISPRi-based repression (Gilbert et al., 2014). Inhibition of *MDM2*^{LUTI} transcription in these cells resulted in a dramatic increase (up to 10-fold) in *MDM2*^{PROX} expression (Figure 3A). We achieved a range of *MDM2*^{LUTI} knockdown efficiencies in this cell line and found that the degree of P1 downregulation generally correlated with the degree of P2 activation, as predicted by the LUTI mRNA-based gene regulatory mechanism (Figure 3A). Importantly, P1 inhibition and the resulting increase in P2 transcription resulted in an increase in MDM2 protein levels (Figure 3B), suggesting a functional importance to the transcript toggling that we observe and supporting the notion that P1-driven expression of *MDM2*^{LUTI} results in the downregulation of MDM2 protein levels.

H3K36me3 is a co-transcriptionally-laid modification that is seen in regions distal to active promoters (Bannister et al., 2005; Li et al., 2003; Mikkelsen et al., 2007; Xiao et al., 2003). In budding yeast, H3K36me3 is associated with a decrease in spurious transcription initiation from within bodies of transcribed genes (Carrozza et al., 2005; Keogh et al., 2005; Kim et al., 2016; Li et al., 2003; Xiao et al., 2003) and it plays a crucial role in the downregulation of the canonical *NDC80* transcript in meiotic prophase as a result of transcription from the distal promoter that produces the LUTI mRNA (Chia et al., 2017). Down-regulation of *MDM2*^{LUTI} expression resulted in more than 3-fold decrease in H3K36me3 signal at the P2 promoter (Figure 3C, S6), whereas the H3K36me3 signal remained high within the *MDM2* gene body, likely due to increased *MDM2*^{PROX} expression under these conditions. These data are consistent with a

mechanism whereby *MDM2*^{LUTI} expression represses transcription from the P2 promoter through co-transcriptional histone modifications, and provide further support for a model in which the two *MDM2* promoters are controlled by the same mechanism defined for *NDC80* in yeast. Based on these findings, we propose that LUTI-dependent regulation of gene expression is conserved from yeast to human.

Contrary to traditional gene regulatory models, mRNA and protein abundances show poor correlations over developmental programs in genome-scale yeast and vertebrate studies (Cheng et al., 2018; Peshkin et al., 2015). In yeast we have recently shown that hundreds of such cases can be explained by LUTI mRNA-based regulation, whereby developmentally-regulated transcript toggling, driven by timed induction of the regulatory transcription factors, is responsible for distal and proximal promoter usage and driving final protein output levels (Cheng et al., 2018). The natural toggling between *MDM2* isoforms during differentiation shown here suggests that broad use of this mechanism for developmental modulation of gene expression may be conserved.

In summary, we report here that the LUTI mRNA based mechanism, defined for *NDC80* in yeast, is conserved in humans. Based on the ubiquitous use of alternative promoters and uORF translation in humans (Floor and Doudna, 2016; Ingolia et al., 2011; Tresenrider and Unal, 2017; Wang et al., 2016), this may be the first of many such cases to be identified. Canonical models to explain the prevalence of mammalian alternative promoter usage suggest that one promoter might serve as a “back-up” or that the use of two promoters could simply allow activation by different transcription

factors that are present in different cell types (Davuluri et al., 2008). However, in the case of *MDM2*, we argue that its two promoters are fundamentally different in function. The P1 promoter produces a poorly translated $MDM2^{LUTI}$ transcript and the production of $MDM2^{LUTI}$ from this promoter interferes with P2 activity in *cis*, reducing the transcription of the well-translated $MDM2^{PROX}$ isoform. Therefore, P1-driven $MDM2^{LUTI}$ mRNA production serves to downregulate MDM2 protein expression (Figure 4).

MDM2 protein levels are elevated in a variety of cancers (reviewed in (Rayburn et al., 2005) and this elevation has been attributed in some cases to an increase in translation of the pool of *MDM2* transcripts, based on increased transcription from the P2 (Brown et al., 1999; Capoulade et al., 1998; Landers et al., 1997). Much research has focused on identifying alternate transcription factors that can activate P2—as it is clear that transcription can occur from this promoter in the absence of p53—and several have been found (Phelps et al., 2003; Zhang et al., 2012), but relatively little is known about P1 regulation. Our study argues that MDM2 expression levels can be modulated by changes in activation of P1 alone, suggesting a promising new general area for the development of gene regulatory tools that modulate P1 activity, and the activity of other yet-to-be-identified LUTI mRNA promoters, as a means to fine tune gene expression.

Acknowledgements:

We are grateful to Howard Chang and Sueng Woo Cho for providing dCas9 and dCas9-KRAB MCF-7 cell lines; Luke Gilbert, Max Horlbeck, and Jonathan Weissman for providing K562 dCas9 cell lines and for extensive advice; Helen Bateup, John Blair and

Dirk Hockemeyer for providing hESC differentiation samples; Stephen Floor for providing advice, data prior to publication and help with data re-analysis and Christiane Brune for technical support. We thank Barbara Meyer, Jasper Rine, Michael Rape, Folkert van Werven, Jingxun Chen and Eric Sawyer for their valuable feedback on the manuscript; Celeste Riepe for the suggestion of the MDM2 (SMP14) antibody; and members of the Tjian and Darzacq labs for tissue culture assistance and advice. This work was supported by funds from the Pew Charitable Trusts (00027344), Damon Runyon Cancer Research Foundation (35-15), National Institute of Health (DP2 AG055946-01) and Glenn Foundation for Medical Research to E.Ü.; funds from the Pew Charitable Trusts (00029624), the Alfred P. Sloan Foundation (FG-2016-6229), and the National Institute of Health (DP2 GM-119138) to G.B.; funds from the California Institute of Regenerative Medicine (CIRM, LA1-08013) to X.D., funds from the Howard Hughes Medical Institute (003061) to R.T. and funds from the Shurl and Kay Curci Foundation to L.Y.C.

Author contributions: IH, AT, XD, RT, LYC, EÜ, and GAB designed experimental strategies and methods and performed data analysis; IH, JCB, VJ, AT, CDD, and LYC executed experiments; IH, EÜ, and GAB wrote the manuscript.

References:

Arrick, B.A., Grendell, R.L., and Griffin, L.A. (1994). Enhanced translational efficiency of a novel transforming growth factor beta 3 mRNA in human breast cancer cells. Mol Cell Biol *14*, 619-628.

Bannister, A.J., Schneider, R., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2005). Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. J Biol Chem *280*, 17732-17736.

Barak, Y., Gottlieb, E., Juven-Gershon, T., and Oren, M. (1994). Regulation of mdm2 expression by p53: alternative promoters produce transcripts with nonidentical translation potential. Genes Dev *8*, 1739-1749.

Blair, J.D., Hockemeyer, D., Doudna, J.A., Bateup, H.S., and Floor, S.N. (2017). Widespread Translational Remodeling during Human Neuronal Differentiation. Cell Rep *21*, 2005-2016.

Brown, C.Y., Mize, G.J., Pineda, M., George, D.L., and Morris, D.R. (1999). Role of two upstream open reading frames in the translational control of oncogene mdm2. Oncogene *18*, 5631-5637.

Capoulade, C., Bressac-de Paillerets, B., Lefrere, I., Ronsin, M., Feunteun, J., Tursz, T., and Wiels, J. (1998). Overexpression of MDM2, due to enhanced translation, results in inactivation of wild-type p53 in Burkitt's lymphoma cells. Oncogene *16*, 1603-1610.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P.*, et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Chen, J., Tresenrider, A., Chia, M., McSwiggen, D.T., Spedale, G., Jorgensen, V., Liao, H., van Werven, F.J., and Unal, E. (2017). Kinetochore inactivation by expression of a repressive mRNA. Elife *6*.

Cheng, Z., Otto, G.M., Powers, E.N., Keskin, A., Mertins, P., Carr, S., Jovanovic, M., and Brar, G.A. (2018). Pervasive, coordinated protein level changes driven by transcript isoform switching during meiosis Cell *In press*.

Chia, M., Tresenrider, A., Chen, J., Spedale, G., Jorgensen, V., Unal, E., and van Werven, F.J. (2017). Transcription of a 5' extended mRNA isoform directs dynamic chromatin changes and interference of a downstream promoter. Elife *6*.

Davuluri, R.V., Suzuki, Y., Sugano, S., Plass, C., and Huang, T.H. (2008). The functional consequences of alternative promoter use in mammalian genomes. Trends Genet *24*, 167-177.

Floor, S.N., and Doudna, J.A. (2016). Tunable protein synthesis by transcript isoforms in human cells. Elife *5*.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C.*, et al.* (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A.*, et al.* (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442-451.

Honda, R., Tanaka, H., and Yasuda, H. (1997). Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. FEBS Lett *420*, 25-27.

Hughes, T.A., and Brady, H.J. (2005). Cross-talk between pRb/E2F and Wnt/beta-catenin pathways: E2F1 induces axin2 leading to repression of Wnt signalling and to increased cell death. Exp Cell Res *303*, 32-46.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of Mammalian proteomes. Cell *147*, 789-802.

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J.*, et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

Kim, J.H., Lee, B.B., Oh, Y.M., Zhu, C., Steinmetz, L.M., Lee, Y., Kim, W.K., Lee, S.B., Buratowski, S., and Kim, T. (2016). Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. Nat Commun *7*, 13534.

Landers, J.E., Cassel, S.L., and George, D.L. (1997). Translational enhancement of mdm2 oncogene expression in human tumor cells containing a stabilized wild-type p53 protein. Cancer Res *57*, 3562-3568.

Law, G.L., Bickel, K.S., MacKay, V.L., and Morris, D.R. (2005). The undertranslated transcriptome reveals widespread translational silencing by alternative 5' transcript leaders. Genome Biol *6*, R111.

Li, B., Howe, L., Anderson, S., Yates, J.R., 3rd, and Workman, J.L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. J Biol Chem *278*, 8897-8903.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Peshkin, L., Wühr, M., Pearl, E., Haas, W., Freeman, R.M., Gerhart, J.C., Klein, A.M., Horb, M., Gygi, S.P., and Kirschner, M.W. (2015). On the Relationship of Protein and mRNA Dynamics in Vertebrate Embryonic Development. Dev Cell *35*, 383-394.

Phelps, M., Darley, M., Primrose, J.N., and Blaydes, J.P. (2003). p53-independent activation of the hdm2-P2 promoter through multiple transcription factor response elements results in elevated hdm2 expression in estrogen receptor alpha-positive breast cancer cells. Cancer Res *63*, 2616-2623.

Rayburn, E., Zhang, R., He, J., and Wang, H. (2005). MDM2 and human malignancies: expression, clinical pathology, prognostic markers, and implications for chemotherapy. Curr Cancer Drug Targets *5*, 27-41.

Tresenrider, A., and Unal, E. (2017). One-two punch mechanism of gene repression: a fresh perspective on gene regulation. Curr Genet.

Wang, X., Hou, J., Quedenau, C., and Chen, W. (2016). Pervasive isoform-specific translational regulation via alternative transcription start sites in mammals. Molecular Systems Biology *12*, 875.

Wu, X., Bayle, J.H., Olson, D., and Levine, A.J. (1993). The p53-mdm-2 autoregulatory feedback loop. Genes Dev *7*, 1126-1132.

Xiao, T., Hall, H., Kizer, K.O., Shibata, Y., Hall, M.C., Borchers, C.H., and Strahl, B.D. (2003). Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev *17*, 654-663.

Zhang, X., Zhang, Z., Cheng, J., Li, M., Wang, W., Xu, W., Wang, H., and Zhang, R. (2012). Transcription factor NFAT1 activates the mdm2 oncogene independent of p53. J Biol Chem *287*, 30468-30476.

Figure legends

Figure 1. P1 and P2-driven *MDM2* transcript isoform toggling can be seen during human embryonic stem cell differentiation.

Model for LUTI mRNA expression-mediated gene repression. Top panel: LUTI mRNA production causes an increase in the co-transcriptional H3K36me3 marks at the proximal gene promoter and transcriptional repression of the canonical mRNA isoform. Because LUTI mRNA is not well translated due to uORFs in its extended 5' leader and because the well-translated canonical mRNA is repressed, the net effect of LUTI mRNA production is the downregulation of protein synthesis from the LUTI target gene locus. Bottom panel: In the absence of LUTI expression, transcription from the canonical gene promoter occurs, leading to protein production. **B.** Illustration of the *MDM2* gene structure**.** *MDM2* is transcribed from two different transcription start sites (TSS1 and TSS2) regulated by two different promoters (P1 and P2). Transcription from the distal TSS1 produces a 5'-extended, uORF-containing transcript, which is poorly translated.

Hereafter, the P1 promoter-driven transcript isoform is referred to as *MDM2*^{LUTI}, while the P2-driven isoform, transcribed from the proximal TSS2 is referred to as *MDM2*^{PROX}. The arrows describe the location of the isoform-specific primers used for the RT-qPCR analyses in this figure and all the subsequent figures (blue arrows: *MDM2*^{LUTI} specific primers; yellow arrows: *MDM2*^{PROX} specific primers). **C.** Upper panel: RT-qPCR data showing the fold differences of *MDM2*^{LUTI} *and MDM2*^{PROX} transcripts relative to *GAPDH* expression levels x1000 in human embryonic stem cells (hESCs). Error bars represent the range measured for two biological replicates. Lower panel: RT-qPCR data sowing the relative expression of *MDM2*^{LUTI} *and MDM2*^{PROX} transcripts in human embryonic stem cells (hESCs), neural progenitor cells (NPCs), Day 14 and Day 50 neurons (D14 and D50). Data were normalized relative to *MDM2*^{LUTI} or *MDM2*^{PROX} transcript abundance in hESCs. Error bars refer to the range measured for two biological replicates. **D.** Translation status of *MDM2*^{LUTI} and *MDM2*^{PROX}. Data from (Blair et al., 2017) were re-analyzed for the portion of the isoform-specific junction reads corresponding to the *MDM2*^{LUTI} or *MDM2*^{PROX} found in monosomes, low polysome fractions, and high polysome fractions. Relative abundance is calculated by dividing the number of isoform-specific junction reads for a given fraction by the total number of isoform-specific junction reads across all fractions for each transcript isoform. **E.** RT-qPCR data showing the changes in the expression of *MDM2*^{PROX} and *MDM2*^{LUTI} in hESCs differentiating into endoderm. D1-D4 refers to days after transfer of the hESCs to endoderm differentiation medium. Data were normalized relative to *MDM2*^{LUTI} or *MDM2*^{PROX} transcript abundance in hESCs.

Figure 2. Downregulation of *MDM2*^{LUTI} leads to an increase in the expression of *MDM2*^{PROX} in MCF-7 cells, independent of p53 expression.

A. RT-qPCR data displaying the changes of *MDM2*^{LUTI} and *MDM2*^{PROX} mRNA expression in MCF-7-dCas9 stable cells. The transcription of *MDM2*^{LUTI} was inhibited by CRISPRi using four different sgRNAs (#1-4). Data were normalized to *GAPDH*, and the fold change relative to the expression of *MDM2*^{LUTI} and *MDM2*^{PROX} in the cells transfected with an empty vector was calculated. Data points represent the mean of at least 3 independent biological replicates. Error bars represent standard error of the mean (SEM). Two-tailed Student's *t*-test was used to calculate the *P*-values in this figure and all of the subsequent figures. **B.** RT-qPCR data showing the change in the expression level of *TP53*, *MDM2*^{LUTI} and *MDM2*^{PROX} in MCF-7-dCas9 cells after CRISPRi-mediated *TP53* knockdown (left) or CRISPRi-mediated p53- and *MDM2*^{LUTI}-double knockdown (right, sgRNA #1 through #4), relative to the cells transfected with an empty vector. Data were normalized to *GAPDH.* Data points represent the mean of four biological replicates. Error bars represent SEM.

Figure 3. Downregulation of *MDM2*^{LUTI} reduces repression-associated histone marks at the P2 promoter and up-regulates MDM2 protein expression.

A. RT-qPCR data displaying the changes in *MDM2*^{LUTI} and *MDM2*^{PROX} expression levels in a stable K562-dCas9 cell line in which the transcription of *MDM2*^{LUTI} had been inhibited by CRISPRi using four different sgRNAs (#5, #6, #1, and #7). Data were normalized to *GAPDH*, and the fold change relative to cells transfected with the empty vector was calculated. Data points represent the mean of at least 3 independent

biological replicates. Error bars represent SEM. **B.** Top panel: Western Blot for MDM2 protein in the K562-dCas9 cells treated with empty vector (EV) or two different sgRNAs (#6 and #5) to inhibit *MDM2*^{LUTI} expression. β-actin was measured as a loading control. The ratio of MDM2 and β-actin in each sample was taken and then normalized to the EV value. Rep I, biological replicate 1; rep II, biological replicate 2. Bottom panel: Matched mRNA samples were analyzed for relative expression of *MDM2*^{LUTI} and *MDM2*^{PROX}. Data were normalized to *GAPDH* and the fold change relative to cells transfected with the empty vector was calculated **C.** Chromatin immunoprecipitation (ChIP) data displaying H3K36 trimethylation (H3K36me3) enrichment around the proximal TSS (TSS2) in K562-dCas9 cells after CRISPRi-mediated inhibition of *MDM2*^{LUTI} expression. Location of the four different primer pairs (A, B, C, and D) are shown in the schematic above the graph. Data points represent the mean of 4 independent biological replicates. Error bars represent SEM. n.s. = not significant

Figure 4. **Model of the LUTI mRNA based mechanism of the *MDM2* gene.**

The *MDM2* gene has two promoters, P1 and P2. *MDM2*^{PROX} is regulated by P2 whereas *MDM2*^{LUTI} is regulated by P1. In comparison to *MDM2*^{PROX}, *MDM2*^{LUTI} is poorly translated because of the existence of two upstream open reading frames (uORFs) in its extended 5'-leader. Top panel: When P1 promoter is active ("ON"), *MDM2*^{LUTI} transcription establishes H3K36 trimethylation at the downstream P2 promoter and causes repression of P2 ("OFF"). As a result, *MDM2*^{LUTI} becomes the predominant transcript product from the *MDM2* locus. Bottom panel: When P1 promoter is "OFF", transcriptional repression of the downstream P2 promoter is relieved, culminating in the

expression of *MDM2*^{PROX}. *MDM2*^{PROX} is efficiently translated, resulting in higher MDM2 protein levels.

Supplemental figure legends:

Figure S1. Validation of endodermal differentiation of human embryonic stem cells. RT-qPCR data showing the changes in expression of hESC- specific genes (*NANOG*, *SOX2,* and *OCT4*) and endoderm-specific genes (*SOX17* and *CXCR4)* during endodermal differentiation (Endo D1 through D4). Values normalized to the expression in hESCs.

Figure S2. Changes in MDM2 protein levels during neuronal differentiation of human embryonic stem cells. Left panel: Western Blot for MDM2 protein in the K562-dCas9 cells treated with *siMDM2* or a non-targeting siRNA to validate the specificity of the anti-MDM2 (SMP14) antibody. β-actin was used as a loading control. Right panel: Western blot for the neuronal differentiation samples analyzed in Figure 1C. * represents a lower, cross-reacting band of unknown origin in hESCs.

Figure S3. Information about the location of the single guide RNAs (sgRNAs) used for the *MDM2* locus in this study. Binding sites of the sgRNAs (red lines) used for the CRISPRi-mediated knockdown of *MDM2*^{LUTI} within the *MDM2* gene.

Figure S4. Both *MDM2*^{PROX} and *MDM2*^{LUTI} levels are reduced upon CRISPRi targeting of *MDM2*^{LUTI} transcription start site in MCF-7-dCas9-KRAB cell lines. RT-

qPCR data showing the changes in *MDM2*^{PROX} and *MDM2*^{LUTI} expression upon CRISPRi-mediated knockdown of *MDM2*^{LUTI} in MCF-7-dCas9 (solid bars) and MCF-7-dCas9-KRAB cell lines (checkered bars), using two different sgRNAs (#1 and #3). Data were normalized to *GAPDH* and the fold change relative to cells transfected with the empty vector was calculated.

Figure S5. *MDM2*^{PROX} can be upregulated upon *MDM2*^{LUTI} downregulation, even under conditions with low p53 levels. RT-qPCR data showing the expression of *MDM2*^{PROX} and *MDM2*^{LUTI} in the MCF-7-dCas9 cells treated with different sgRNAs (#1-4) targeting *MDM2*^{LUTI} in the presence (+) or absence (-) of a sgRNA targeting *TP53*. Data were normalized to *GAPDH* and the fold change relative to cells transfected with the empty vector was calculated. Data points represent the mean of at least 3 independent biological replicates. Error bars represent SEM.

Figure S6. Quality assessment for the H3K36me3 ChIP. qPCR analysis of chromatin immunoprecipitation performed with IgG or anti-H3K36me3. Same primers sets were used as in Figure 3C. Note that the H3K36me3 data are the same as shown in Figure 3C. Data points represent the mean of 4 independent biological replicates. Error bars represent SEM.

Materials and methods

Cell lines

MCF-7-dCas9 and -dCas9-KRAB cells were cultivated at 37°C with 5% CO_2 in high glucose Dulbecco's modified Eagle media (GlutaMax, Gibco) supplemented with 10% FBS and 1% penicillin-streptomycin. K562-dCas9 cell lines were cultivated at 37°C with 5% CO_2 in RPMI1640 media (Gibco) supplemented with 10% FBS and 10mM HEPES. hESCs (WIBR3 NIH#0079) were maintained in culture as described in (Lengner et al., 2010). The differentiation into definitive endoderm was performed using the STEMdiff™ Definitive Endoderm Kit (Stem Cell Technologies) following the manufacturer's instructions. MCF-7-dCas9 and -dCas9-KRAB cells were kindly provided by Howard Y. Chang (Stanford University). K562-dCas9 cells were kindly provided by Jonathan Weissman (University of California, San Francisco). Cell lines were authenticated by STR profiling and were tested to be negative for mycoplasma (MycoAlert™ Mycoplasma Detection Kit, Lonza).

RNA isolation, cDNA synthesis and quantitative polymerase chain reaction

Total RNA from hESCs differentiating into definitive endoderm and hESCs differentiating into neurons, MCF-7-dCas9, MCF-7-dCas9-KRAB and K562-dCas9 cells was isolated using Trizol (Life Technologies) according to the manufacturer's instructions. Equal amounts of RNA were primed with random hexamers and reverse transcribed using SuperScript II Reverse Transcriptase (Thermo Fisher) according to the manufacturer's instructions. RNA levels were quantified using SYBR Green/Rox (ThermoFisher) and the StepOnePlus Real-time PCR system (ThermoFisher).

Samples for total RNA isolation from hESCs differentiating into neurons (hESC, NPC, neurons D14, neurons D50) were a gift from Helen Bateup [University of California, Berkeley; (Blair et al., 2017)]

CRISPRi knockdowns

sgRNAs targeting *MDM2*^{LUTI} were designed and cloned into the lentiviral pU6-sgRNA EF1Alpha-puro-T2A-BFP vector. Lentivirus was packaged by co-transfecting sgRNA-expression plasmids and the packaging vectors pCMV-dR8.91 and pMD2.G into 293T cells using the TransIT®-LT1 Transfection Reagent (Mirus). Cells were treated with ViralBoost (Alstem) to allow for efficient lentivirus production and lentivirus was harvested 72 hours post-transfection. CRISPRi-directed gene knockdown was achieved by transducing MCF-7-dCas9, -dCAS9-KRAB and K562-dCas9 cell lines with sgRNA-containing lentivirus in the presence of 8µg/ml polybrene (Millipore Sigma). Successfully transduced cells were puromycin-selected (ThermoFisher; 1.3µg/ml for MCF-7 and 3µg/ml for K562 cells) and harvested 7 days post-infection.

The pU6-sgRNA EF1Alpha-puro-T2A-BFP vector was a gift from Jonathan Weissman (Addgene plasmid # 60955).

siRNA-mediated knockdowns

K562-dCas9 cells were transfected with 60pmol of siRNAs against MDM2 (SMARTpool: siGENOME MDM2 siRNA, Dharmacon) or 60pmol of non-targeting siRNAs (siGENOME Non-Targeting siRNA Pool #1, Dharmacon) using Lipofectamine®

RNAiMAX (Life Technologies) according to the manufacturer's instructions. Cells were harvested 72 hours post-transfection.

Western Blot analysis

Proteins were isolated by lysing cell pellets in RIPA buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with cOmplete™ Protease Inhibitor (Millipore Sigma). 10-20ug protein extracts were resuspended in reducing SDS sample buffer, fractionated by SDS-PAGE and transferred onto a nitrocellulose membrane. Washed membranes were blocked in Odyssey® PBS Blocking Buffer (LI-COR Biosciences) for 1.5h at room temperature and incubated with antibodies against MDM2 (SMP14, sc-965, Santa Cruz; 1:200) at 4°C overnight or β-actin (A2228, Millipore Sigma, 1:40,000) at room temperature for 1 hour. Membranes were washed and incubated in IRDye 800CW anti-donkey anti-mouse (MDM2) or IRDye 680RD anti-goat anti-mouse (β-actin) secondary antibodies (LI-COR Biosciences) for 1h (MDM2) or 30 minutes (β-actin) at room temperature and proteins were visualized using the Odyssey® CLx system (LI-COR Biosciences).

All antibodies were diluted in Odyssey® Blocking Buffer (PBS) (LI-COR Biosciences) with (secondary antibodies) or without (primary antibodies) 0.01% TWEEN. Total protein from hESCs differentiating into neurons (hESC, NPC, neurons D14, neurons D50) was a gift from Helen Bateup [University of California, Berkeley; (Blair et al., Cell Reports, 2017)].

H3K36me3 Chromatin immunoprecipitation (ChIP)

K562-dCas9 cells (5 x 15cm^2 plates per sample) were treated with 1% formaldehyde (16%, methanol free, Ultra Pure, Polysciences) for 10 minutes at room temperature to crosslink DNA and protein. The crosslinking reaction was stopped by adding 0.125M PBS-glycine and cells were harvested by centrifugation. Cells were subsequently resuspended in ice-cold PBS containing 0.25mM PMSF and 10ug/ml aprotinin (Millipore) and pelleted by centrifugation. Chromatin immunoprecipitation of these pellets was performed as previously described (Testa et al., 2005) with minor modifications. Chromatin was sonicated 50 x 30 seconds ON/30 seconds OFF with a Bioruptor® Pico (Diagenode) to obtain fragment sizes of ~200 bp. The sheared samples were incubated in RIPA buffer II (10 mM Tris-Cl, pH 8.0, 1 mM EDTA, pH 8.0, 0.5 mM EGTA, 1% Triton X-100, 0.1% SDS, 0.1% Na-deoxycholate, 140 mM NaCl) containing protease inhibitors and PMSF, with Dynabeads® Protein A

(Invitrogen) for 2 h at 4 °C on rotation. After removal of Dynabeads® Protein A, precleared lysates were immunoprecipitated overnight with 4 ug of rabbit anti-mouse IgG (Ab46540, Abcam) and anti-Histone H3 tri methyl lysine 36 (Ab9050, Abcam). Immunoprecipitates were recovered by incubation for 2 h at 4 °C with previously blocked Protein A Dynabeads in RIPA buffer II (1 µg/µl bovine serum albumin, protease inhibitors, and PMSF). Reverse crosslinked input DNA and immunoprecipitated DNA fragments were amplified with SYBR Green/Rox (ThermoFisher) and quantified with StepOnePlus Real-time PCR system (ThermoFisher).

Table 1. Primers used in this study

Target gene	Primer	5'-3' sequence
MDM2^{prox}	*MDM2*^{PROX} forward	GTGGCGATTGGAGGGTAGAC
	MDM2^{PROX} reverse	TTGTGCACCAACAGACTTTA
MDM2^{LUTI}	*MDM2*^{LUTI} forward	AAACTGGGGAGTCTTGAGGG
	MDM2^{LUTI} reverse	CAGACATGTTGGTATTGCACAT
GAPDH	*GAPDH* forward	AATCCCATCACCATCTTCCA
	GAPDH reverse	TGGACTCCACGACGTACTCA
NANOG	*NANOG forward*	CCAACATCCTGAACCTCAGCTAC
	NANOG reverse	GCCTTCTGCGTCACACCATT
SOX2	*SOX2 forward*	CACACTGCCCCTCTCACACAT
	SOX2 reverse	CATTTCCCTCGTTTTTCTTTGAA
OCT4	*OCT4 forward*	TCGAGAACCGAGTGAGAGGC
	OCT4 reverse	CACACTCGGACCACATCCTTC
CXCR4	*CXCR4 forward*	AGTGAGGCAGATGACAGATA
	CXCR4 reverse	ACAATACCAGGCAGGATAAG
SOX17	*SOX17 forward*	GCCGAGTTGAGCAAGATG
	SOX17 reverse	GGCCGGTACTTGTAGTTG
MDM2 (ChIP)	*MDM2* A forward	GAGTGGAATGATCCCCGAGG
	MDM2 A reverse	GGTTTTCGCGCTTGGAGTC
	MDM2 B forward	CAGACACGTTCCGAAACTGC
	MDM2 B reverse	CCAATCGCCACTGAACACAG

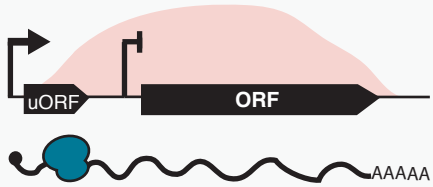
MDM2 C forward	CACAGATTCCAGCTTCGGAAC
MDM2 C reverse	GCCATGCTACAATTGAGGTATACG
MDM2 D forward	TGGCCAGTATATTATGACTAAACGA
MDM2 D reverse	CACGCCAAACAAATCTCCTA

Hollerer_Figure 1

A

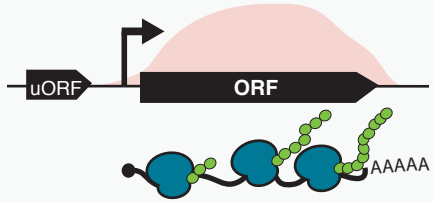
LUTI mRNA production is ON

H3K36me3 at the proximal gene promoter
Canonical mRNA transcription is OFF
ORF translation is OFF



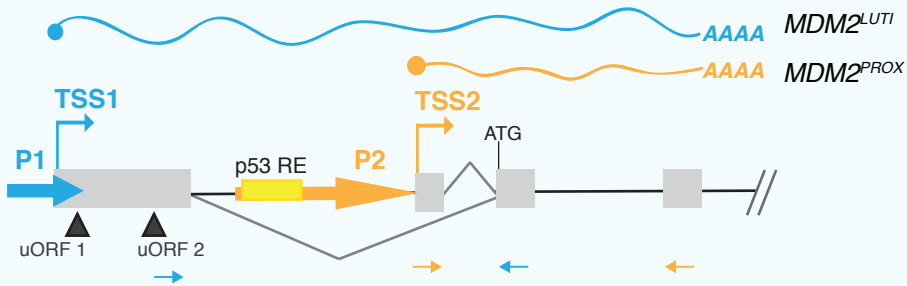
LUTI mRNA production is OFF

H3K36me3 in the gene body
Canonical mRNA transcription is ON
ORF translation is ON

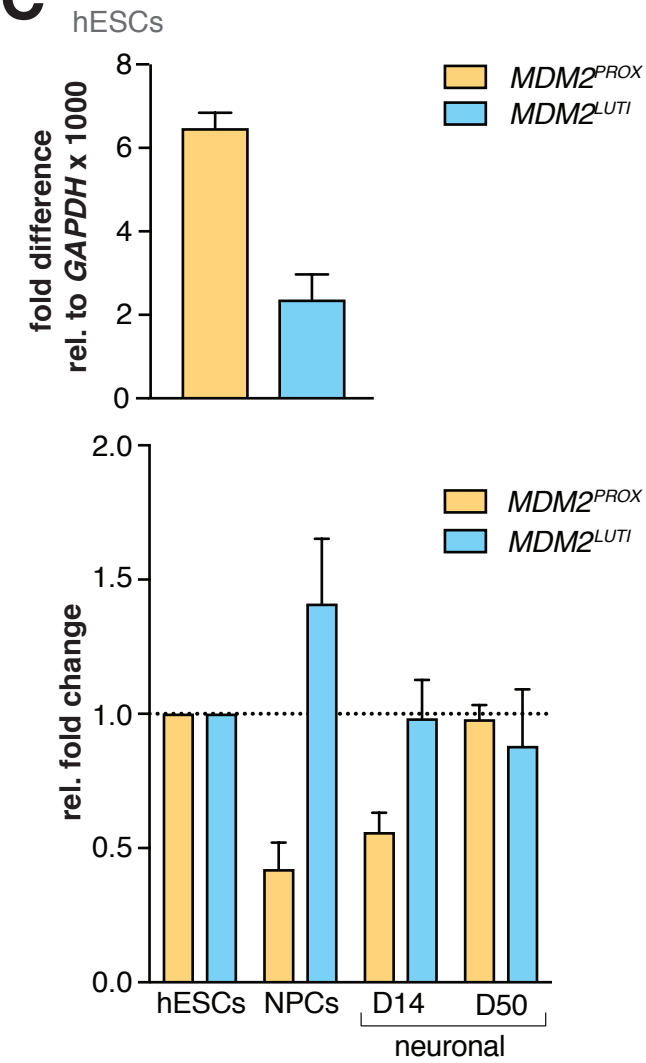


B

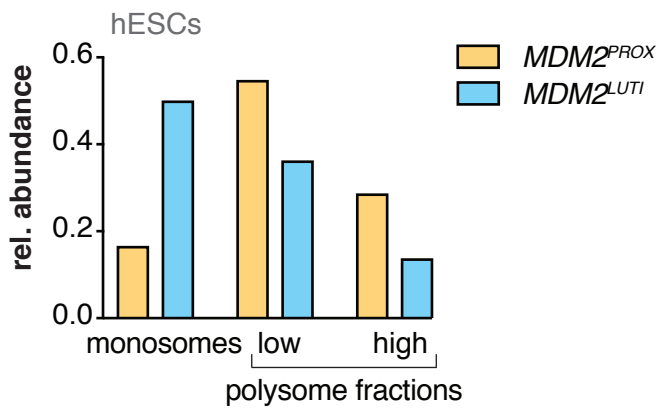
***MDM2* locus**



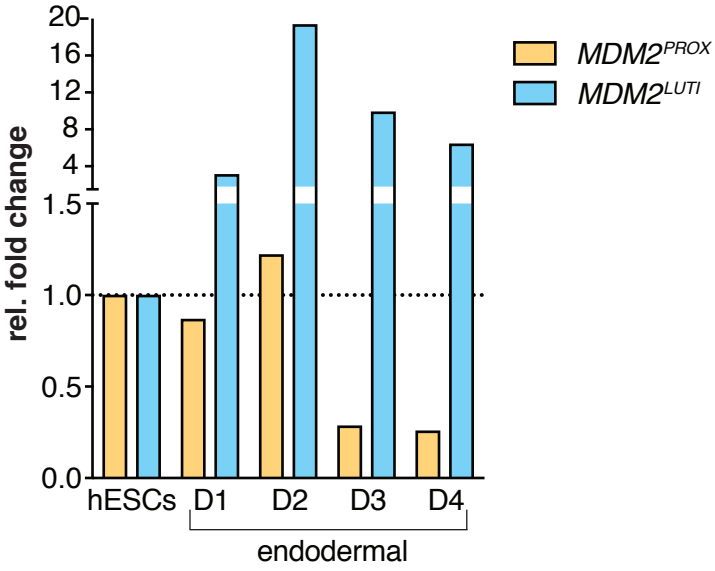
C

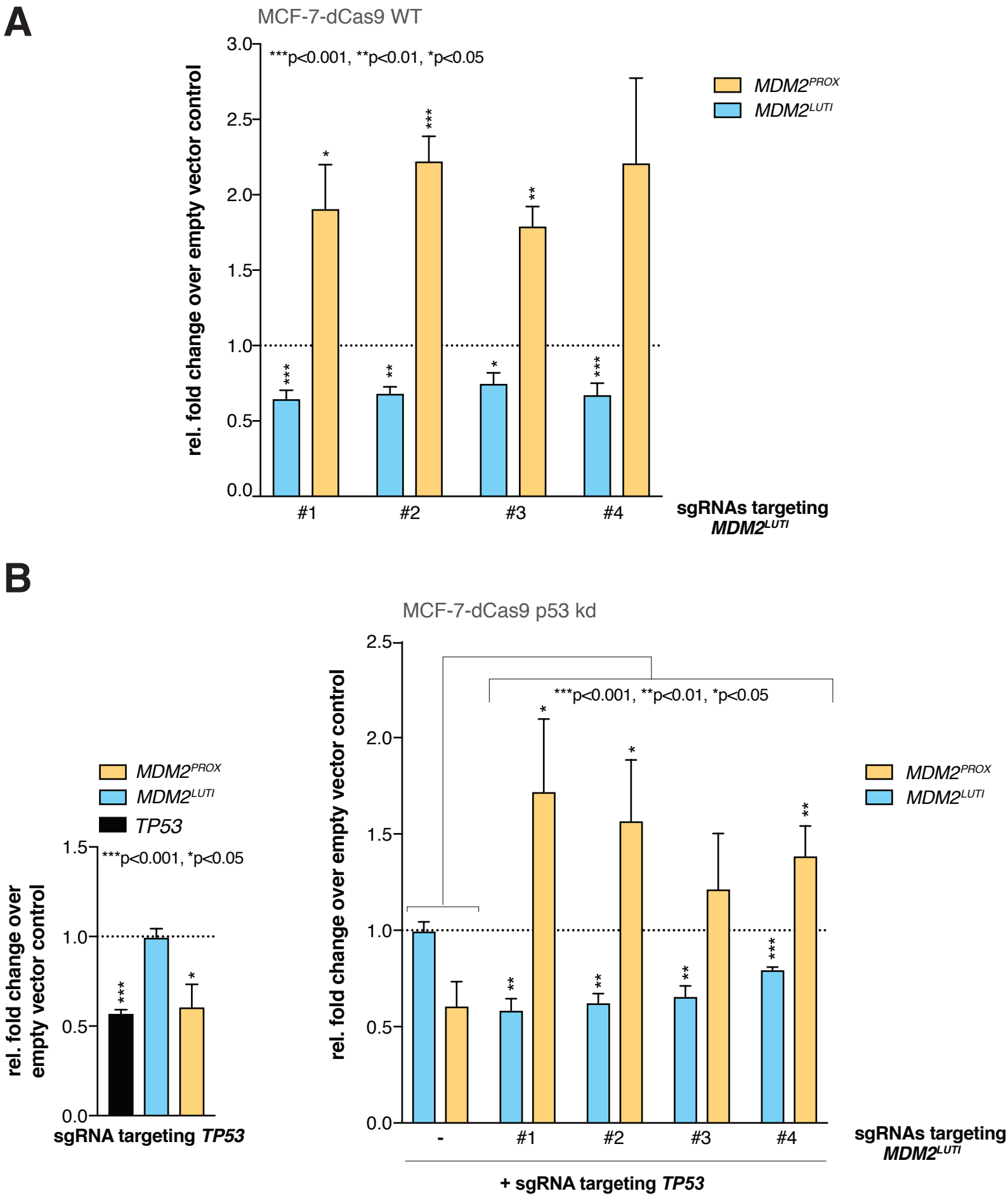


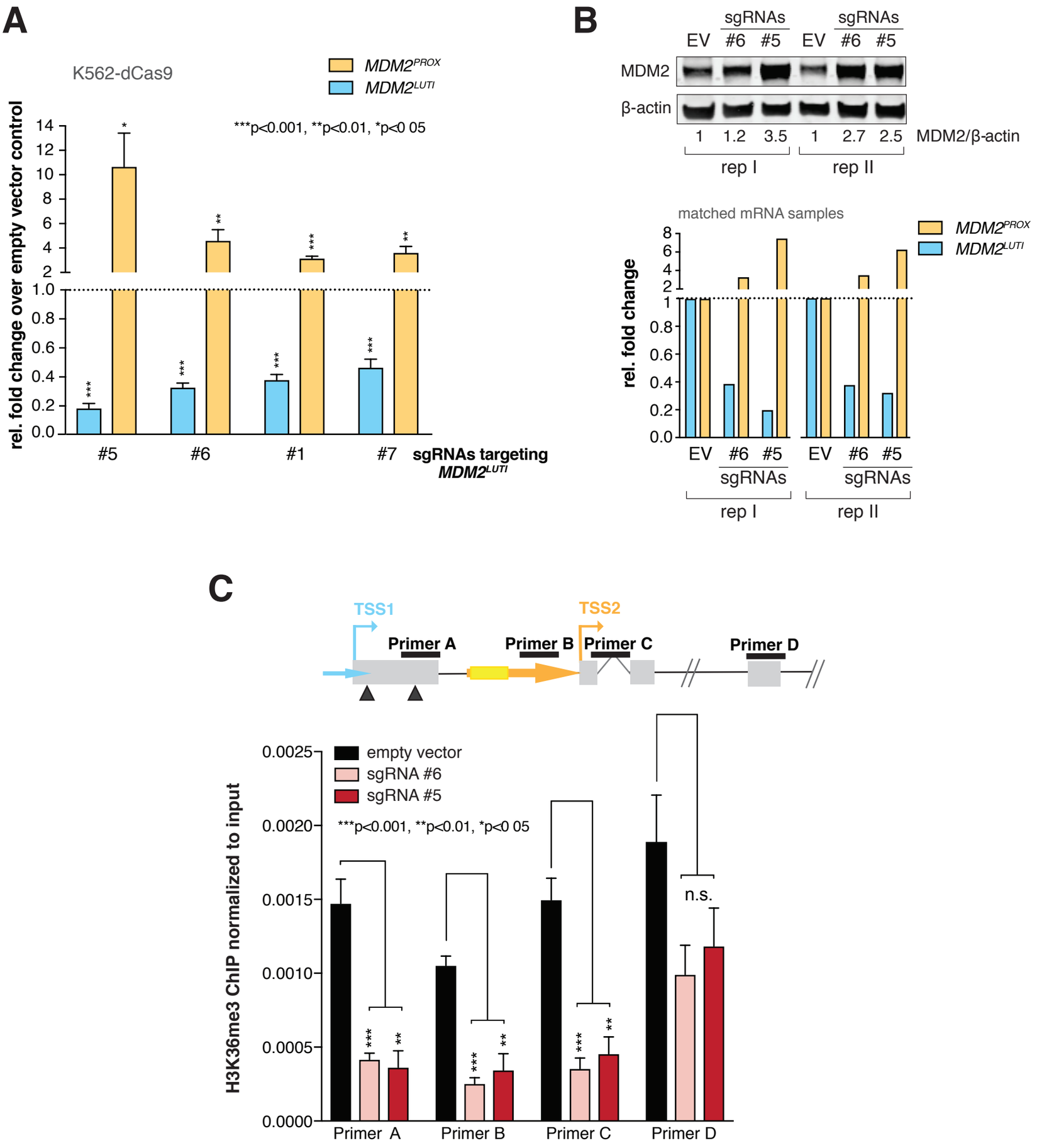
D

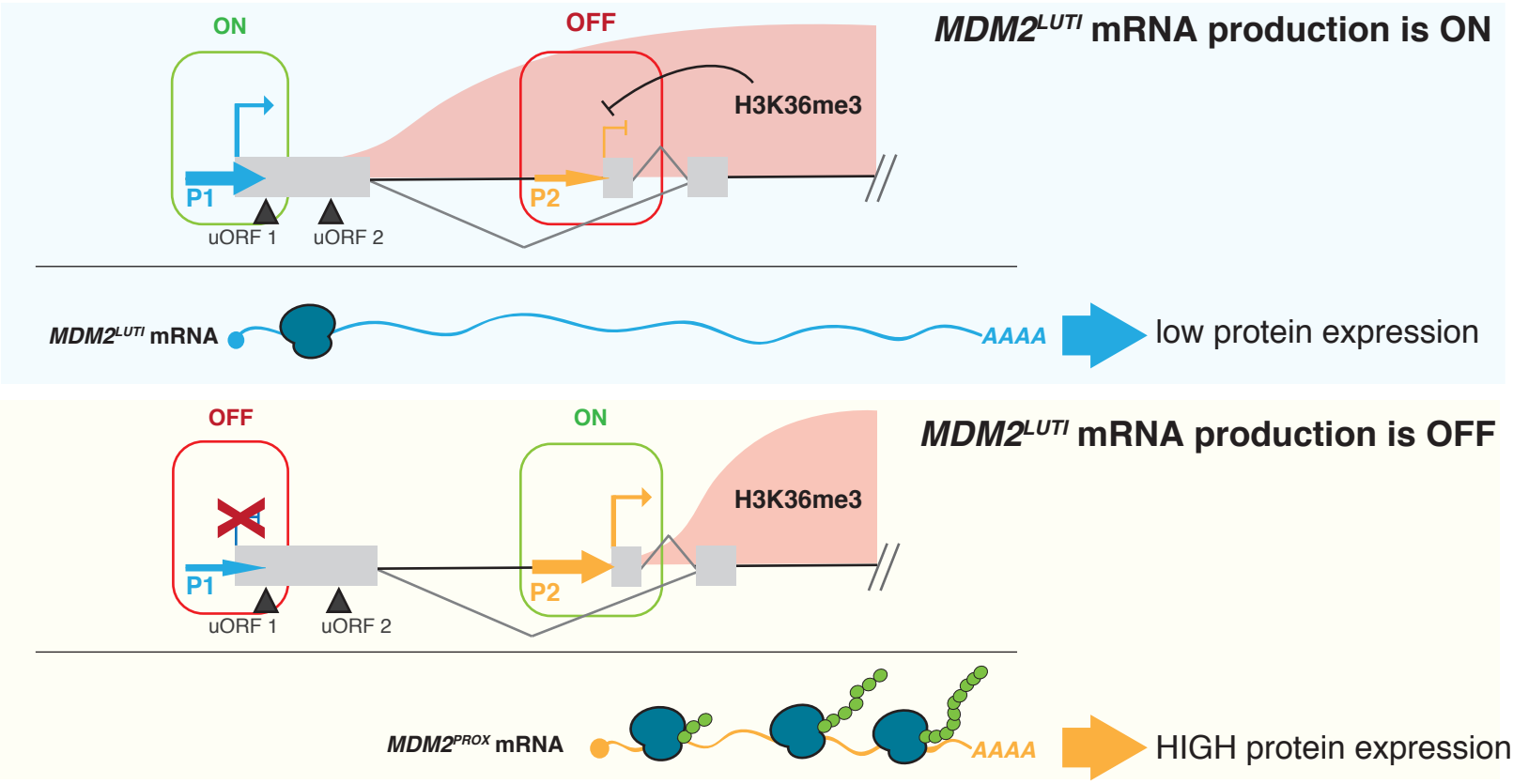


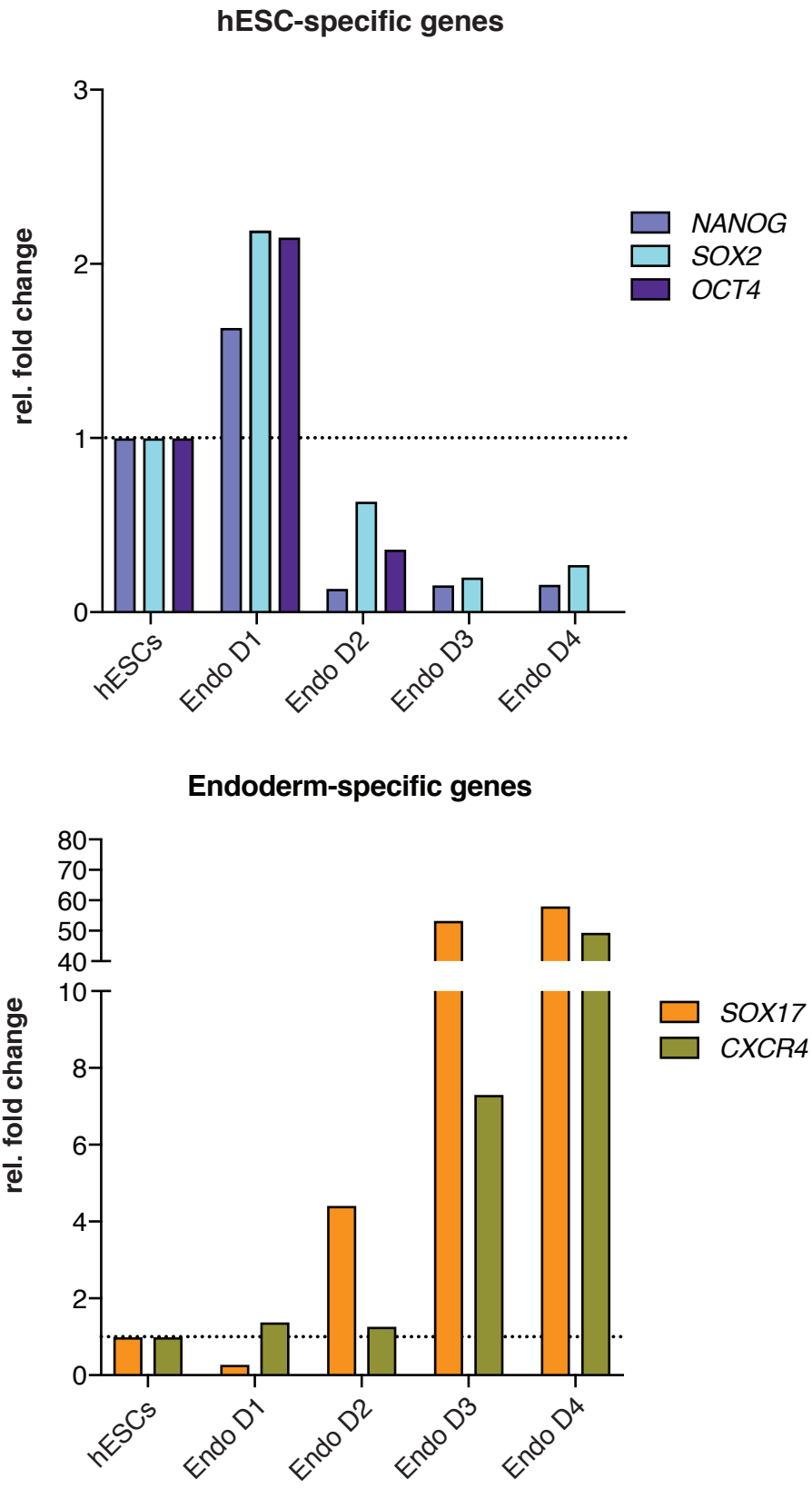
E

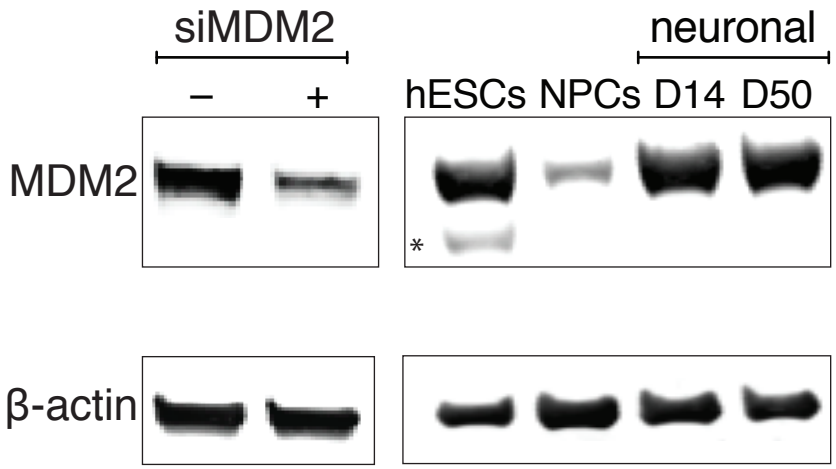


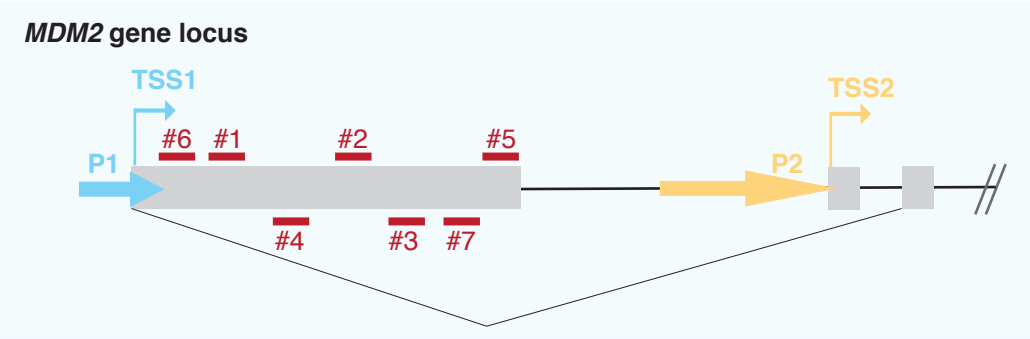

A
MCF-7-dCas9 WT
***p<0.001, **p<0.01, *p<0.05
rel. fold change over empty vector control
3.0
2.5
2.0
1.5
1.0
0.5
0.0
MDM2PROX
MDM2LUTI
#1
#2
#3
#4
sgRNAs targeting MDM2LUTI
B
MCF-7-dCas9 p53 kd
MDM2PROX
MDM2LUTI
TP53
***p<0.001, *p<0.05
rel. fold change over empty vector control
sgRNA targeting TP53
***p<0.001, **p<0.01, *p<0.05
rel. fold change over empty vector control
-
#1
#2
#3
#4
sgRNAs targeting MDM2LUTI
+ sgRNA targeting TP53

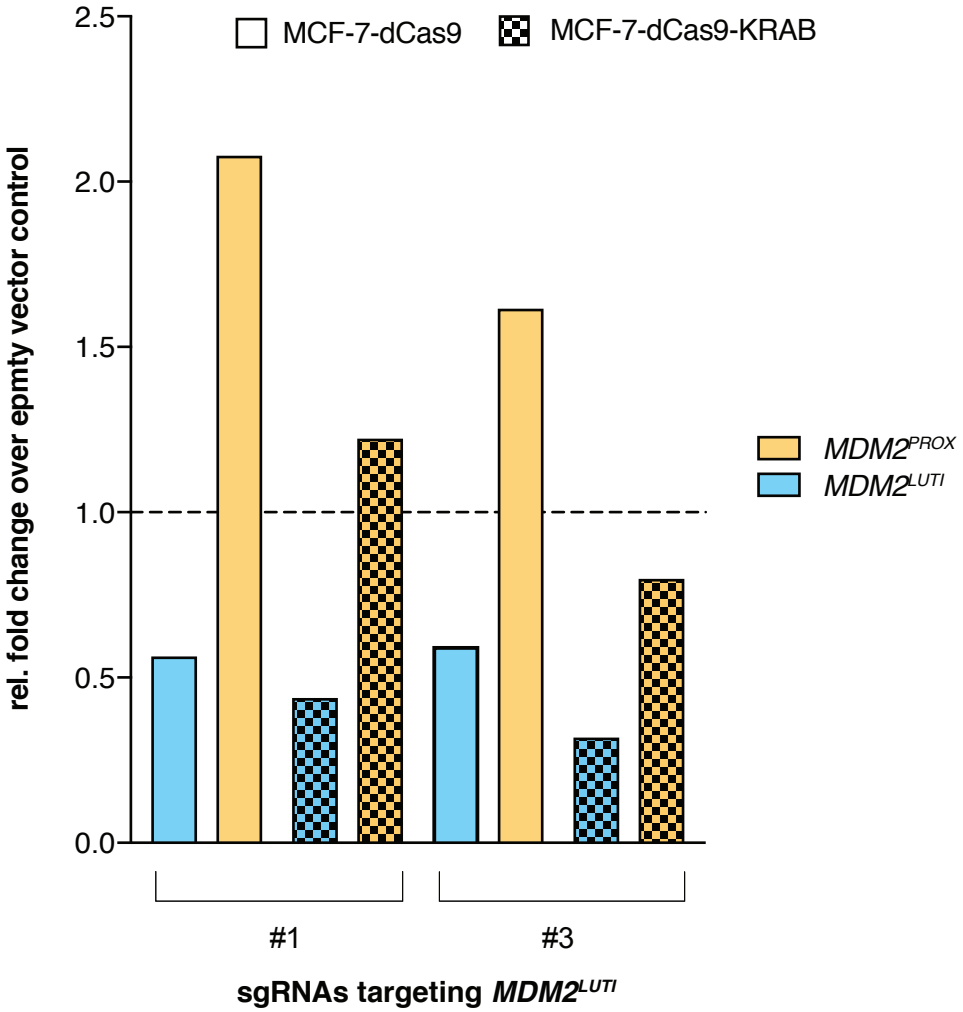

A
K562-dCas9
MDM2PROX
MDM2LUTI
***p<0.001, **p<0.01, *p<0 05
rel. fold change over empty vector control
14
12
10
8
6
4
2
1.0
0.8
0.6
0.4
0.2
0.0
#5
#6
#1
#7
sgRNAs targeting MDM2LUTI
B
sgRNAs
sgRNAs
EV
#6
#5
EV
#6
#5
MDM2
β-actin
1
1.2
3.5
1
2.7
2.5
MDM2/β-actin
rep I
rep II
matched mRNA samples
MDM2PROX
MDM2LUTI
rel. fold change
8
6
4
2
1
0.8
0.6
0.4
0.2
0.0
EV
#6
#5
EV
#6
#5
sgRNAs
sgRNAs
rep I
rep II
C
TSS1
TSS2
Primer A
Primer B
Primer C
Primer D
empty vector
sgRNA #6
sgRNA #5
***p<0.001, **p<0.01, *p<0 05
H3K36me3 ChIP normalized to input
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
n.s.
Primer A
Primer B
Primer C
Primer D

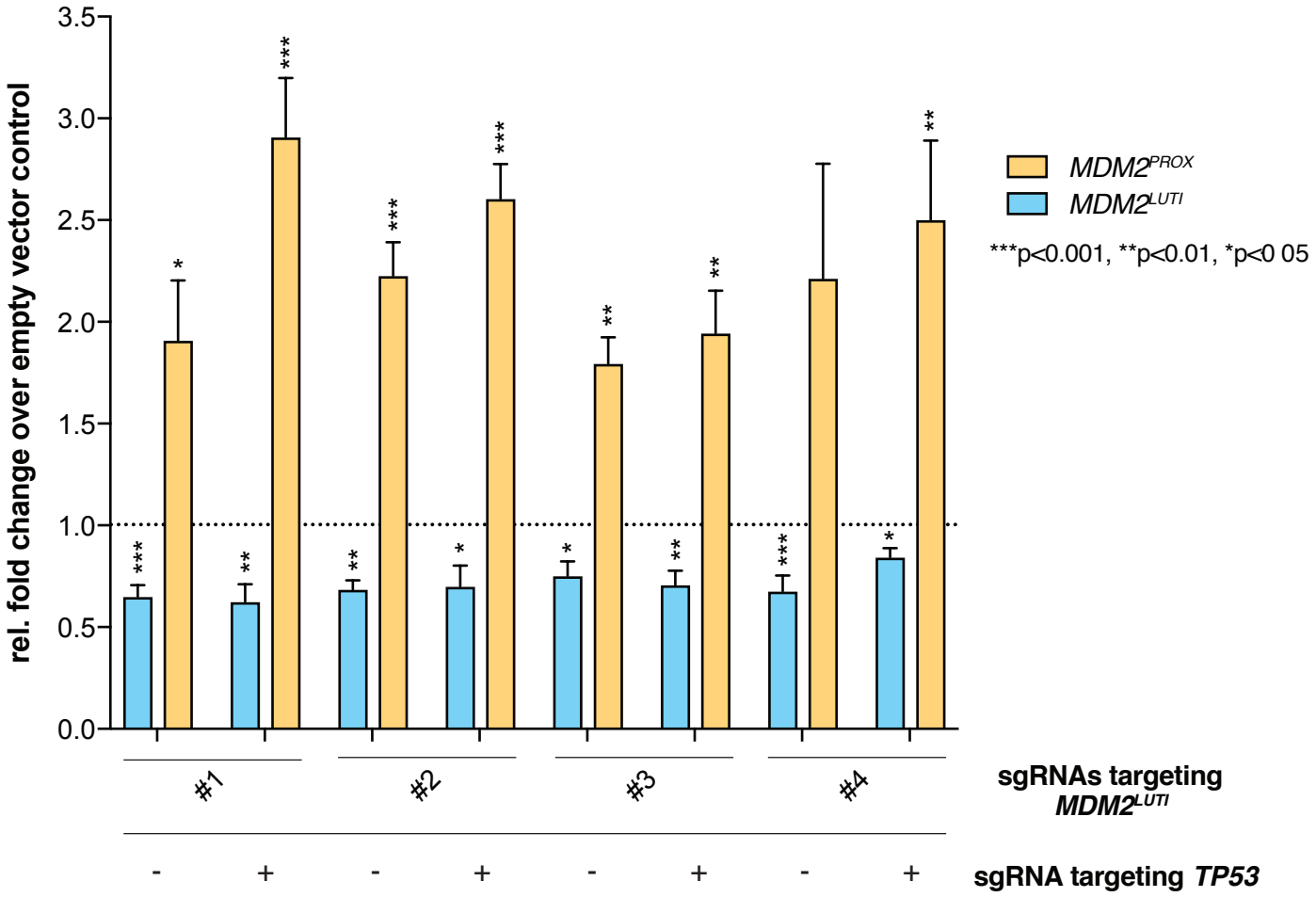

ON
OFF
H3K36me3
P1
P2
uORF 1
uORF 2
MDM2LUTI mRNA
AAAA
low protein expression
MDM2LUTI mRNA production is ON
OFF
ON
H3K36me3
P1
P2
uORF 1
uORF 2
MDM2PROX mRNA
AAAA
HIGH protein expression
MDM2LUTI mRNA production is OFF

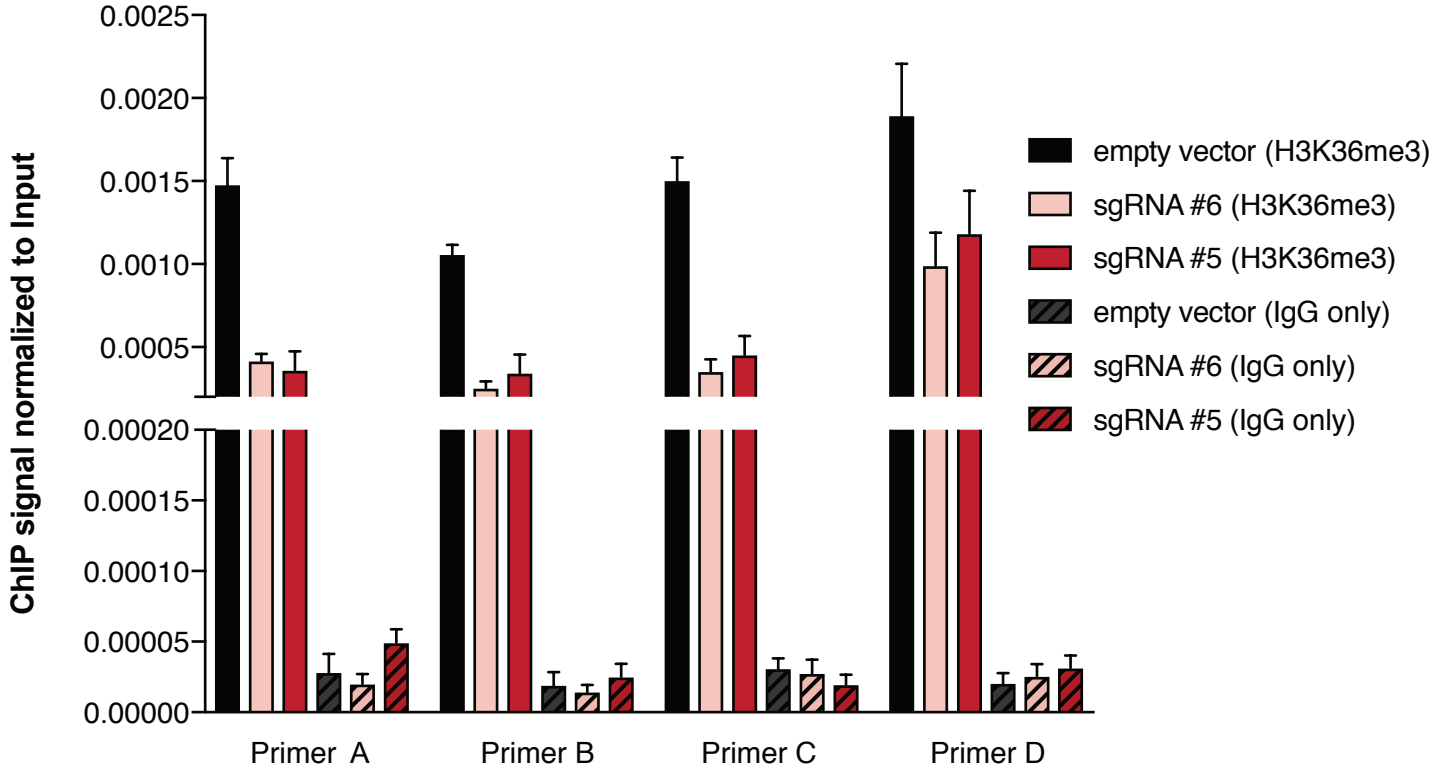

hESC-specific genes
rel. fold change
3
2
1
0
NANOG
SOX2
OCT4
hESCs
Endo D1
Endo D2
Endo D3
Endo D4
Endoderm-specific genes
rel. fold change
80
70
60
50
40
10
8
6
4
2
0
SOX17
CXCR4
hESCs
Endo D1
Endo D2
Endo D3
Endo D4


siMDM2
neuronal
–
+
hESCs
NPCs
D14
D50
MDM2
*
β-actin


MDM2 gene locus
TSS1
TSS2
P1
P2
#6
#1
#2
#5
#4
#3
#7


MCF-7-dCas9
MCF-7-dCas9-KRAB
rel. fold change over epmty vector control
2.5
2.0
1.5
1.0
0.5
0.0
MDM2PROX
MDM2LUTI
#1
#3
sgRNAs targeting MDM2LUTI


rel. fold change over empty vector control
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
MDM2PROX
MDM2LUTI
***p<0.001, **p<0.01, *p<0 05
#1
#2
#3
#4
sgRNAs targeting MDM2LUTI
- + - + - + - +
sgRNA targeting TP53


ChIP signal normalized to Input
0.0025
0.0020
0.0015
0.0010
0.0005
0.00020
0.00015
0.00010
0.00005
0.00000
Primer A
Primer B
Primer C
Primer D
empty vector (H3K36me3)
sgRNA #6 (H3K36me3)
sgRNA #5 (H3K36me3)
empty vector (IgG only)
sgRNA #6 (IgG only)
sgRNA #5 (IgG only)