Evidence for an Integrated Gene Repression Mechanism based on mRNA Isoform Switching 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 transcription from the proximal promoter produces a well-translated transcript. Remarkably, we find that down-regulation of transcription from the distal promoter markedly *up-regulates* expression from the proximal promoter and results in the loss of histone H3K36 trimethylation marks. Moreover, we observe transcript toggling between the two different *MDM2* isoforms as a natural part of two distinct human embryonic stem cell differentiation programs. We conclude that the integrated gene repression mechanism recently identified in yeast is conserved in human cells.

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 translation through integration 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 (Figure 1A) (CHEN *et al.* 2017; CHIA *et al.* 2017). First, a developmentally 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. When all three factors act together, and only then, the activation of *NDC80* transcription can result in a *decrease* in translation from this locus. We termed

the distal promoter-generated 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 further found that this mechanism is common in budding yeast cells, with 379 other genes showing protein levels that are regulated in this manner over time through meiotic differentiation (CHENG *et al.* 2018). While these studies were exclusively performed using budding yeast, two of the three hallmarks of LUTI-based gene repression are common in mammals, suggesting possible conservation of this integrated mechanism. First, almost half of human genes show evidence of alternative promoter usage, resulting in transcript isoforms 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 (LAW *et al.* 2005; FLOOR AND DOUDNA 2016). 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-derived *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 (WU *et al.* 1993; BARAK *et al.* 1994; HONDA *et al.* 1997), it is not known whether transcription from the P1 promoter regulates P2

activity. Evidence of transcriptional interference from nearby transcription has been previously established in human cells (reviewed in (SHEARWIN *et al.* 2005; PALMER *et al.* 2011), but these cases have not involved production of ORF-encoding mRNAs, and most importantly, the link between this type of transcriptional control and differentially translated transcripts has not been explored. Our interpretation of previous studies suggested that the same type of mechanism that we discovered in yeast might control the human *MDM2* locus. Here, we report evidence that this is indeed the case and that this integrated mode of LUTI-based regulation is conserved from yeast to human.

RESULTS AND DISCUSSION

A key prediction, if *MDM2* were regulated by a LUTI-based 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. To directly test this prediction, 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 the P2-derived $MDM2^{PROX}$ transcript levels, which was associated with the reduction of transcription from P1 (Figure 2A and Figure S1). 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 S2). 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, 43% reduction of the P2-derived $MDM2^{PROX}$ transcript (Figure 2B, left panel). However, additional CRISPRi knockdown of the P1-derived transcript, hereon referred to as $MDM2^{LUTI}$, 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; Figure S3). The observation that $MDM2^{LUTI}$ repression 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 test whether *MDM2*^{LUTI} based repression occurs in a different cell type, 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, importantly, found that the degree of P1 downregulation generally correlated with the degree of P2 activation (Figure 3A), as predicted by the LUTI-based gene regulatory mechanism (CHIA *et al.* 2017).

H3K36me3 is a co-transcriptionally established modification that occurs in regions downstream of active promoters (XIAO *et al.* 2003; BANNISTER *et al.* 2005; MIKKELSEN *et al.* 2007). In budding yeast, H3K36me3 is associated with a decrease in spurious transcription initiation from within the body of transcribed genes (LI *et al.* 2003; XIAO *et al.* 2003; CARROZZA *et al.* 2005; KEOGH *et al.* 2005; KIM *et al.* 2016). H3K36me3 is also implicated in noncoding RNA transcription-dependent repression of gene promoters in budding yeast (VAN WERVEN *et al.* 2012). Importantly, H3K36me3 mark is enriched at the proximal *NDC80* promoter as a result of *NDC80*^{LUTI} transcription and is involved in the transcriptional downregulation of the canonical *NDC80* mRNA isoform (CHIA *et al.* 2017). In mammalian cells, H3K36me3 has been implicated in the repression of spurious intragenic transcription, alternative splicing and DNA repair (DHAYALAN *et al.* 2010; XIE *et al.* 2011; WAGNER AND CARPENTER 2012; BAUBEC *et al.* 2015; SUZUKI *et al.* 2017). However its involvement in promoter repression has been less clear. Interestingly, in zebrafish, H3K36me3 mark is present in the promoter regions of

quiescent genes that are developmentally regulated during spermatogenesis (WU *et al.* 2011).

We found that down-regulation of *MDM2*^{LUTI} expression resulted in more than 3-fold decrease in the H3K36me3 signal over the P2 promoter (Figure 3B, Figure S4). In contrast, H3K36me3 signal remained high within the *MDM2* gene body, likely due to increased *MDM2*^{PROX} transcription 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 proximal *MDM2* promoter is controlled by the same mechanism defined in yeast. Based on these findings, we propose that the LUTI-based gene regulation is conserved from yeast to human.

In budding yeast, developmentally controlled switching between the LUTI and canonical mRNA isoforms occurs pervasively during meiotic differentiation (CHENG *et al.* 2018). We tested whether such transcript toggling naturally occurs in human cells by using two different human Embryonic Stem Cell (hESC) differentiation models. In human hESCs, both *MDM2* transcript isoforms were expressed (Figure 4A). When these cells were induced to undergo neuronal differentiation, a switch in transcript isoform expression from *MDM2*^{PROX} to *MDM2*^{LUTI} was evident (Figure 4B). This inverse correlation between *MDM2*^{PROX} *and MDM2*^{LUTI} transcript levels was most pronounced between hESCs and neuronal precursors (NPCs) (Figure 4B). 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 4C, Figure S5). This inverse pattern of proximal and distal promoter usage seen during hESC differentiation suggests that the LUTI-based mechanism regulates *MDM2* expression during normal cellular differentiation.

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

In summary, we report here that the LUTI-based mechanism, initially defined for *NDC80* in yeast, is conserved in humans. Based on the ubiquitous use of alternative promoters and uORF translation in humans (INGOLIA *et al.* 2011; FLOOR AND DOUDNA 2016; WANG *et al.* 2016; TRESENRIDER AND UNAL 2018), 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* expression (Figure 5).

MDM2 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 (LANDERS *et al.* 1997; CAPOULADE *et al.* 1998; BROWN *et al.* 1999). 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.

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]

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.

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) or 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).

Data availability

All the reagents generated in this study are available upon request.

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 technical help; Helen Bateup, John Blair and Dirk Hockemeyer for providing hESC differentiation samples; Stephen Floor for providing advice and sharing data prior to publication 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; 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.

FIGURE LEGENDS

Figure 1. Illustrations of LUTI-based gene repression and *MDM2* gene locus

A. Model for LUTI-based 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 translation from the LUTI target gene locus. Bottom panel: In the absence of LUTI expression, transcription from the canonical gene promoter occurs, leading to translation. **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). p53 RE refers to the location of p53 response element within the P2 promoter.

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.

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.** 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 Arrowheads indicate the location of uORFs in exon 1 and yellow box indicates the p53 response element. Data points represent the mean of 4 independent biological replicates. Error bars represent SEM. n.s. = not significant

Figure 4. P1 and P2-driven *MDM2* transcript isoform toggling can be seen during human embryonic stem cell differentiation.

A. RT-qPCR data showing the fold differences of $MDM2^{LUTI}$ *and* $MDM2^{PROX}$ transcript levels relative to *GAPDH* expression levels x1000 in human embryonic stem cells (hESCs). Error bars represent the range measured for two biological replicates. **B.** 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. **C.** 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 5. **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 translation.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. 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 S2. 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 S3. *MDM2*^{PROX} can be up-regulated 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 (with our without sgRNA targeting *TP53*) was calculated. Data points represent the mean of at least 3 independent biological replicates. Error bars represent SEM.

Figure S4. 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.

Figure S5. 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.

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

REFERENCES

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

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

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

4. Baubec, T., D. F. Colombo, C. Wirbelauer, J. Schmidt, L. Burger *et al.*, 2015 Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature 520**:** 243-247.

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

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

7. Capoulade, C., B. Bressac-de Paillerets, I. Lefrere, M. Ronsin, J. Feunteun *et al.*, 1998 Overexpression of MDM2, due to enhanced translation, results in inactivation of wild-type p53 in Burkitt's lymphoma cells. Oncogene 16**:** 1603-1610.

8. Carrozza, M. J., B. Li, L. Florens, T. Suganuma, S. K. Swanson *et al.*, 2005 Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 123**:** 581-592.

9. Chen, J., A. Tresenrider, M. Chia, D. T. McSwiggen, G. Spedale *et al.*, 2017 Kinetochore inactivation by expression of a repressive mRNA. Elife 6.

10. Cheng, Z., G. M. Otto, E. N. Powers, A. Keskin, P. Mertins *et al.*, 2018 Pervasive, Coordinated Protein-Level Changes Driven by Transcript Isoform Switching during Meiosis. Cell 172**:** 910-923 e916.

11. Chia, M., A. Tresenrider, J. Chen, G. Spedale, V. Jorgensen *et al.*, 2017 Transcription of a 5' extended mRNA isoform directs dynamic chromatin changes and interference of a downstream promoter. Elife 6.

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

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

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

15. Gilbert, L. A., M. A. Horlbeck, B. Adamson, J. E. Villalta, Y. Chen *et al.*, 2014 Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell 159**:** 647-661.

16. Gilbert, L. A., M. H. Larson, L. Morsut, Z. Liu, G. A. Brar *et al.*, 2013 CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154**:** 442-451.

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

18. Hughes, T. A., and H. J. Brady, 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.

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

20. Keogh, M. C., S. K. Kurdistani, S. A. Morris, S. H. Ahn, V. Podolny *et al.*, 2005 Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell 123**:** 593-605.

21. Kim, J. H., B. B. Lee, Y. M. Oh, C. Zhu, L. M. Steinmetz *et al.*, 2016 Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. Nat Commun 7**:** 13534.

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

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

24. Lengner, C. J., A. A. Gimelbrant, J. A. Erwin, A. W. Cheng, M. G. Guenther *et al.*, 2010 Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell 141**:** 872-883.

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

26. Mikkelsen, T. S., M. Ku, D. B. Jaffe, B. Issac, E. Lieberman *et al.*, 2007 Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448**:** 553-560.

27. Palmer, A. C., J. B. Egan and K. E. Shearwin, 2011 Transcriptional interference by RNA polymerase pausing and dislodgement of transcription factors. Transcription 2**:** 9-14.

28. Peshkin, L., M. Wuhr, E. Pearl, W. Haas, R. M. Freeman, Jr. *et al.*, 2015 On the Relationship of Protein and mRNA Dynamics in Vertebrate Embryonic Development. Dev Cell 35**:** 383-394.

29. Phelps, M., M. Darley, J. N. Primrose and J. P. Blaydes, 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.

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

31. Shearwin, K. E., B. P. Callen and J. B. Egan, 2005 Transcriptional interference--a crash course. Trends Genet 21**:** 339-345.

32. Suzuki, S., Y. Murakami and S. Takahata, 2017 H3K36 methylation state and associated silencing mechanisms. Transcription 8**:** 26-31.

33. Testa, A., G. Donati, P. Yan, F. Romani, T. H. Huang *et al.*, 2005 Chromatin immunoprecipitation (ChIP) on chip experiments uncover a widespread distribution of NF-Y binding CCAAT sites outside of core promoters. J Biol Chem 280**:** 13606-13615.

34. Tresenrider, A., and E. Unal, 2018 One-two punch mechanism of gene repression: a fresh perspective on gene regulation. Curr Genet 64**:** 581-588.

35. van Werven, F. J., G. Neuert, N. Hendrick, A. Lardenois, S. Buratowski *et al.*, 2012 Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. Cell 150**:** 1170-1181.

36. Wagner, E. J., and P. B. Carpenter, 2012 Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol 13**:** 115-126.

37. Wang, X., J. Hou, C. Quedenau and W. Chen, 2016 Pervasive isoform-specific translational regulation via alternative transcription start sites in mammals. Mol Syst Biol 12**:** 875.

38. Wu, S. F., H. Zhang and B. R. Cairns, 2011 Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. Genome Res 21**:** 578-589.

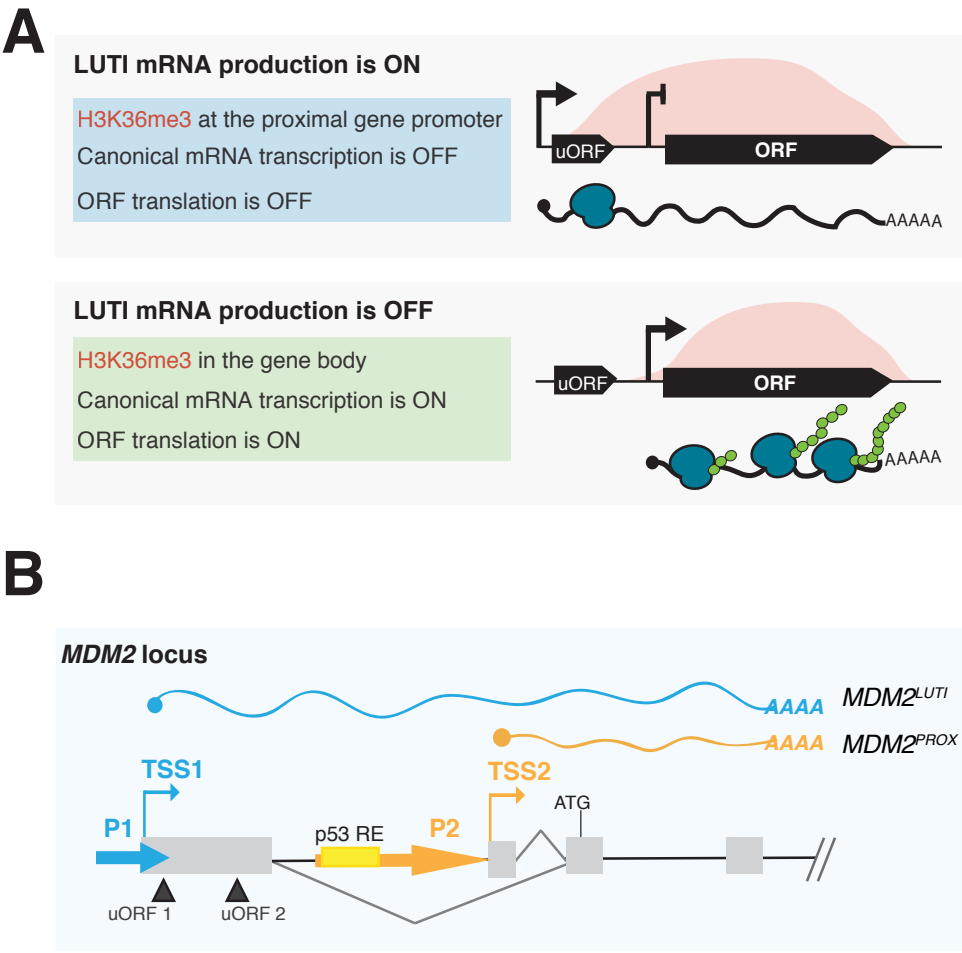
39. Wu, X., J. H. Bayle, D. Olson and A. J. Levine, 1993 The p53-mdm-2 autoregulatory feedback loop. Genes Dev 7**:** 1126-1132.

40. Xiao, T., H. Hall, K. O. Kizer, Y. Shibata, M. C. Hall *et al.*, 2003 Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev 17**:** 654-663.

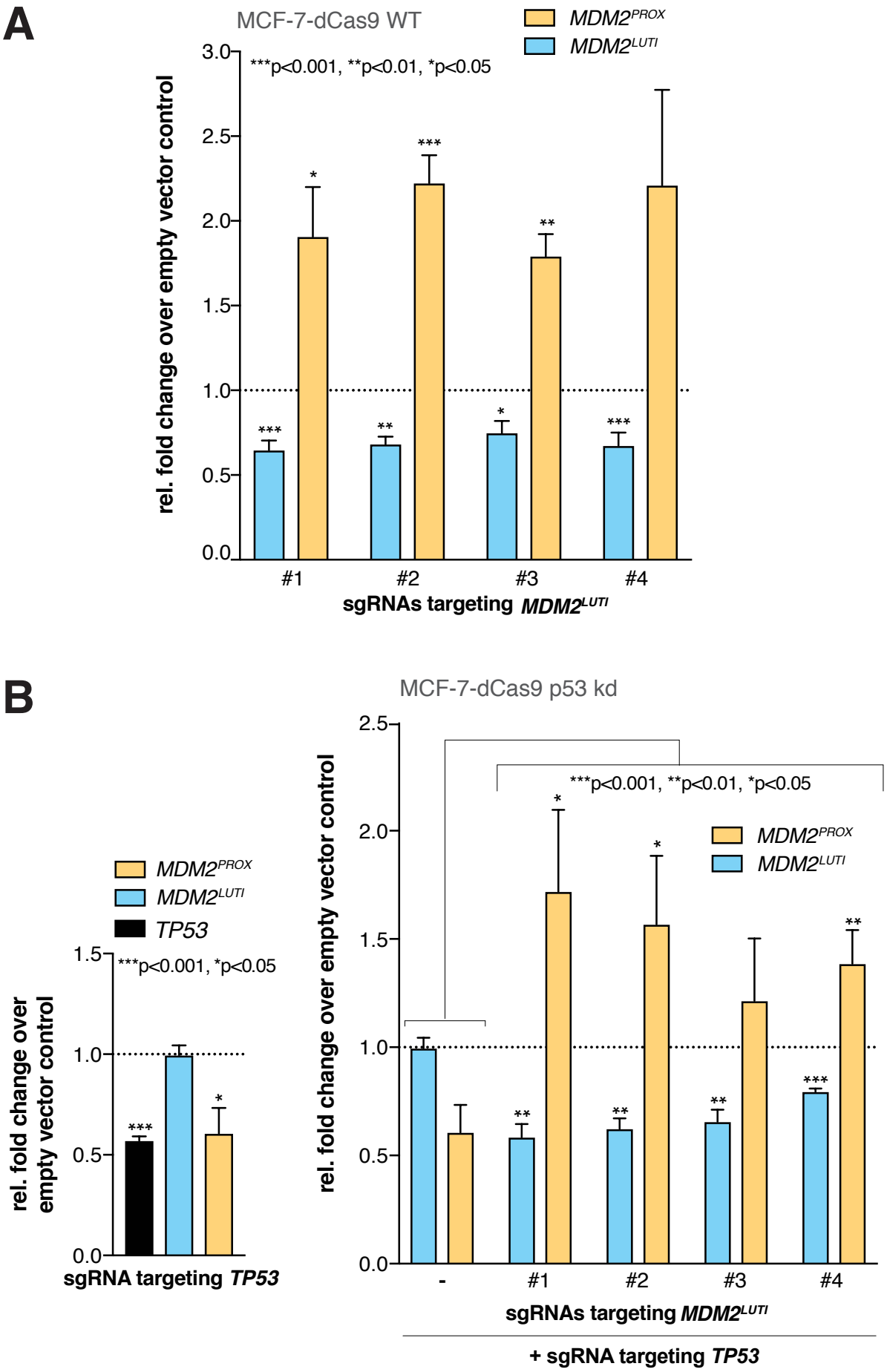
41. Xie, L., C. Pelz, W. Wang, A. Bashar, O. Varlamova *et al.*, 2011 KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. EMBO J 30**:** 1473-1484.

42. Zhang, X., Z. Zhang, J. Cheng, M. Li, W. Wang *et al.*, 2012 Transcription factor NFAT1 activates the mdm2 oncogene independent of p53. J Biol Chem 287**:** 30468-30476.

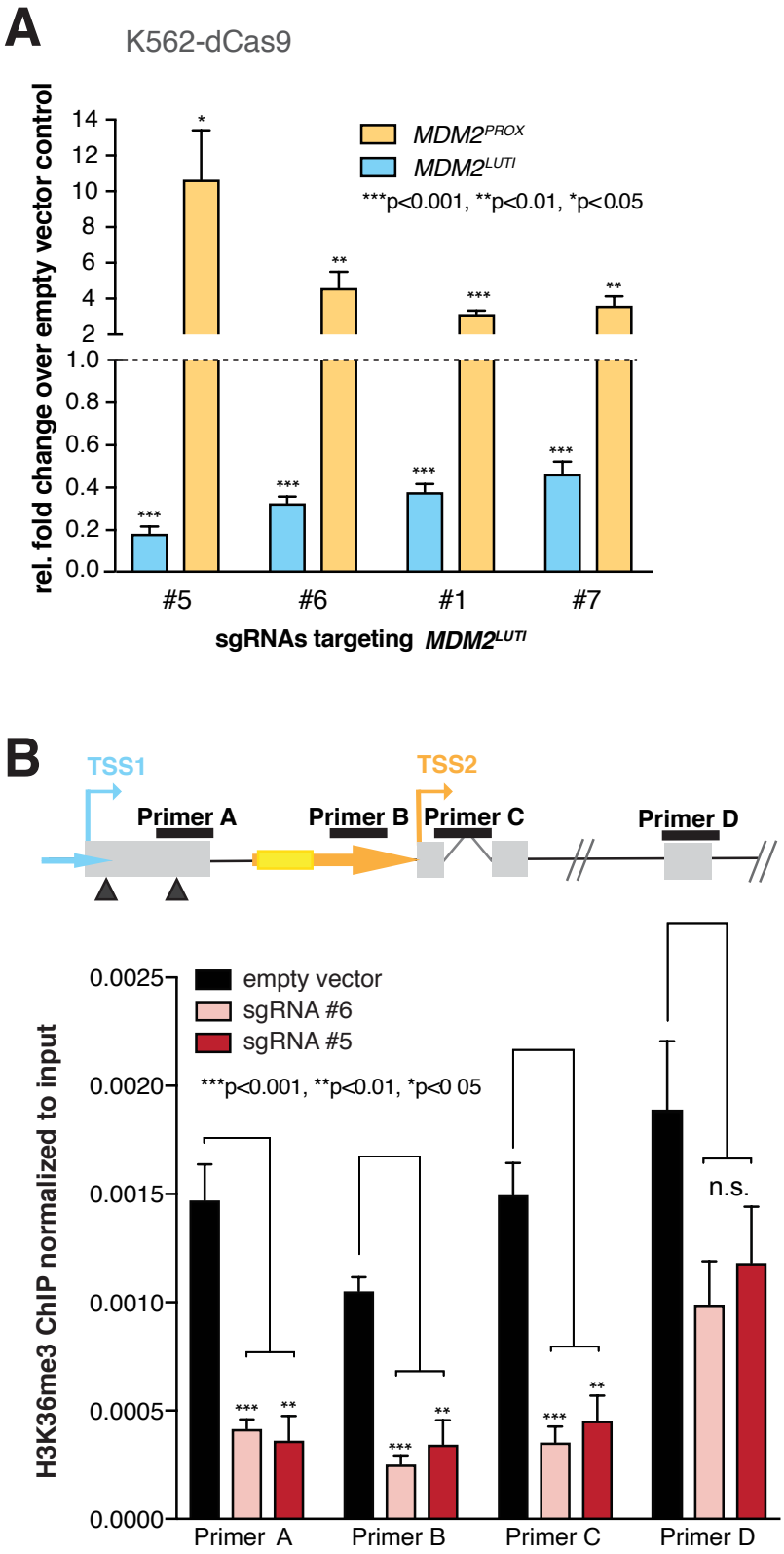
Hollerer et al. Figure 1



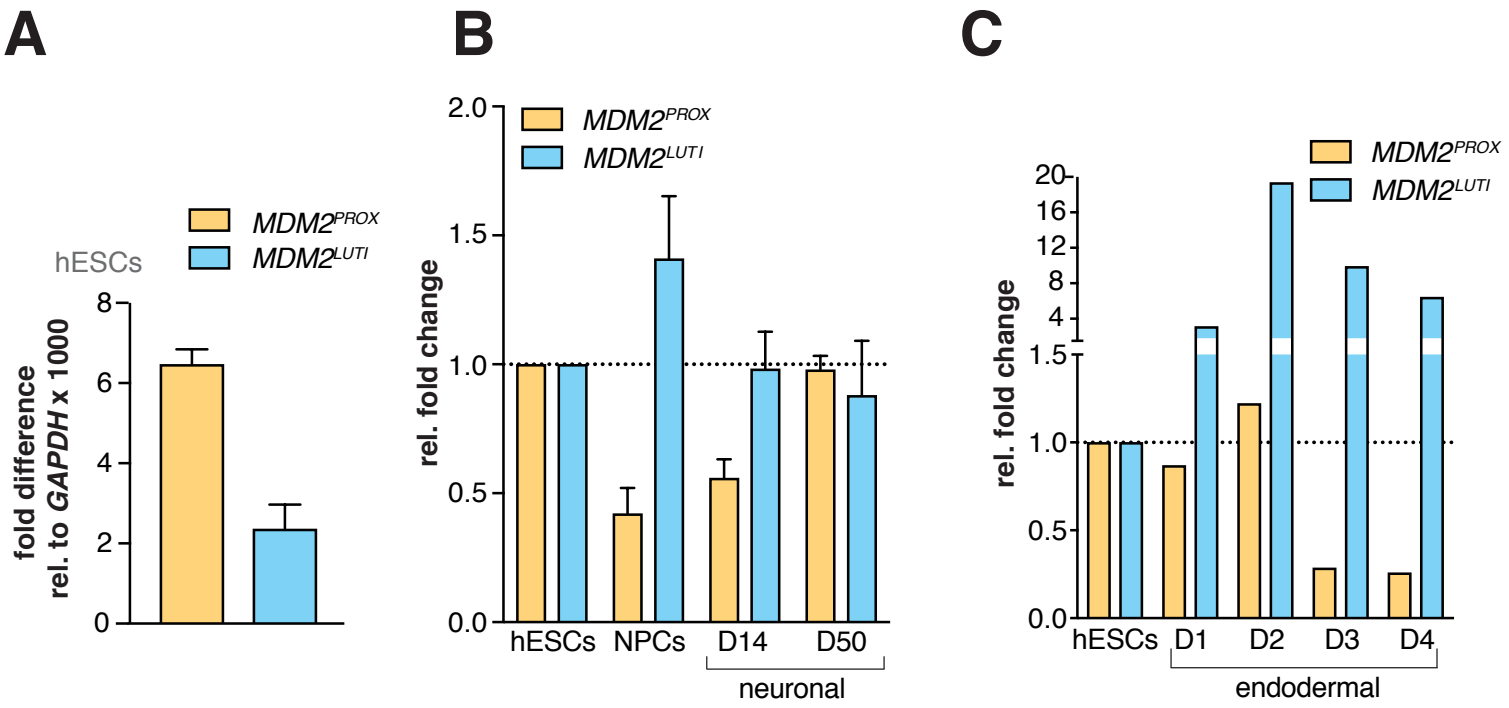
Hollerer et al. Figure 2



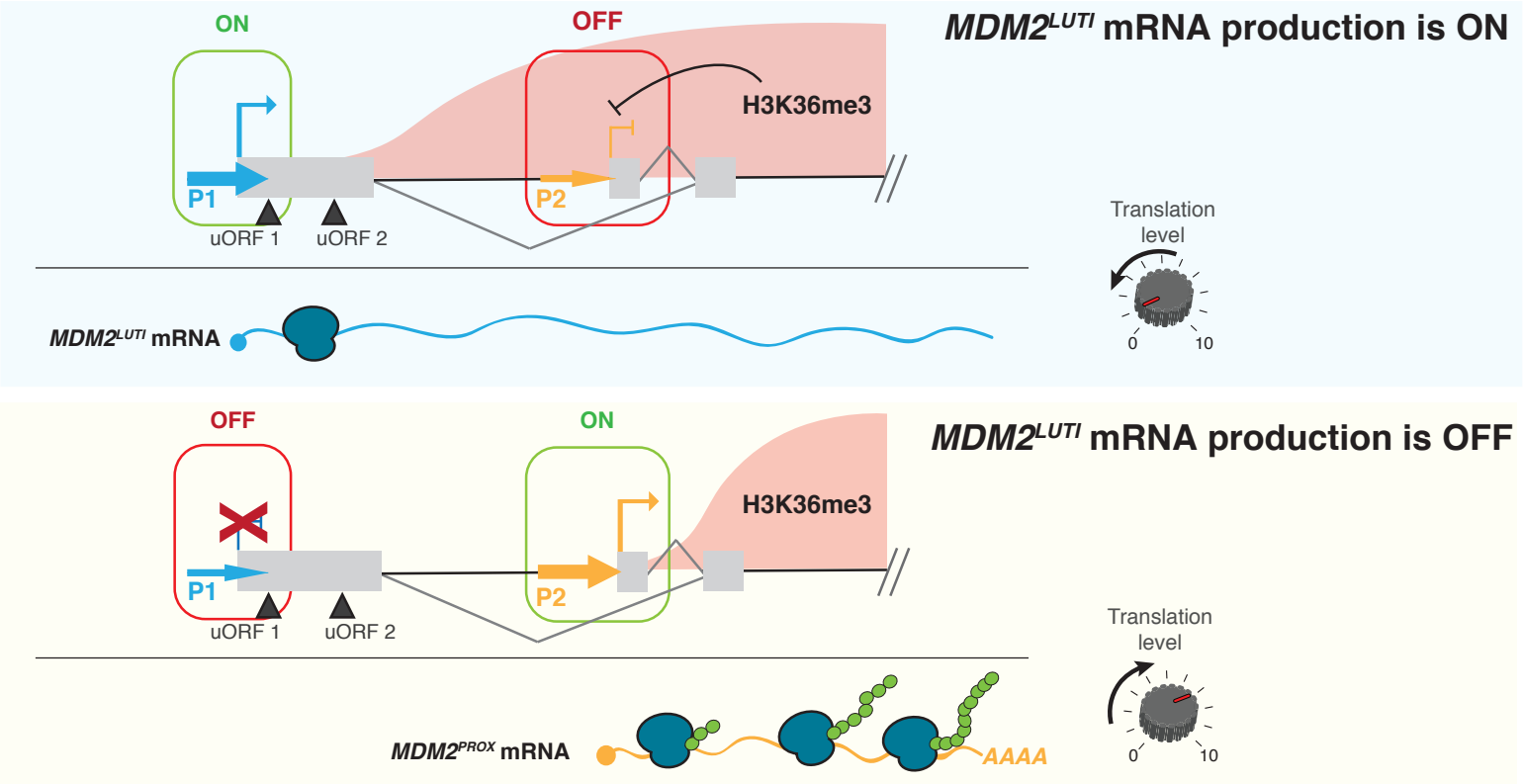
Hollerer et al. Figure 3



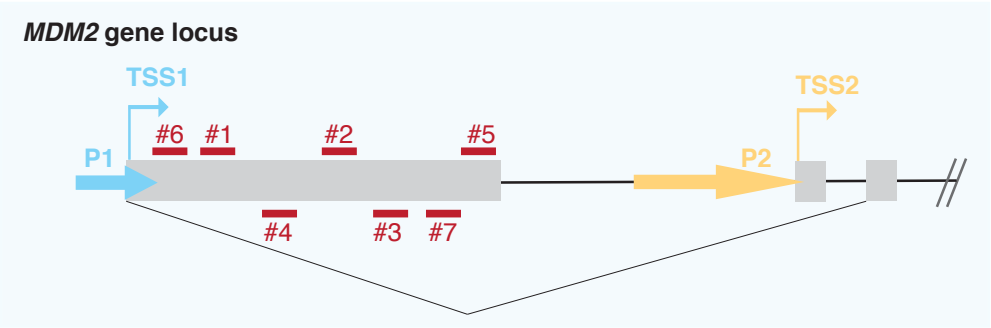
Hollerer et al. Figure 4



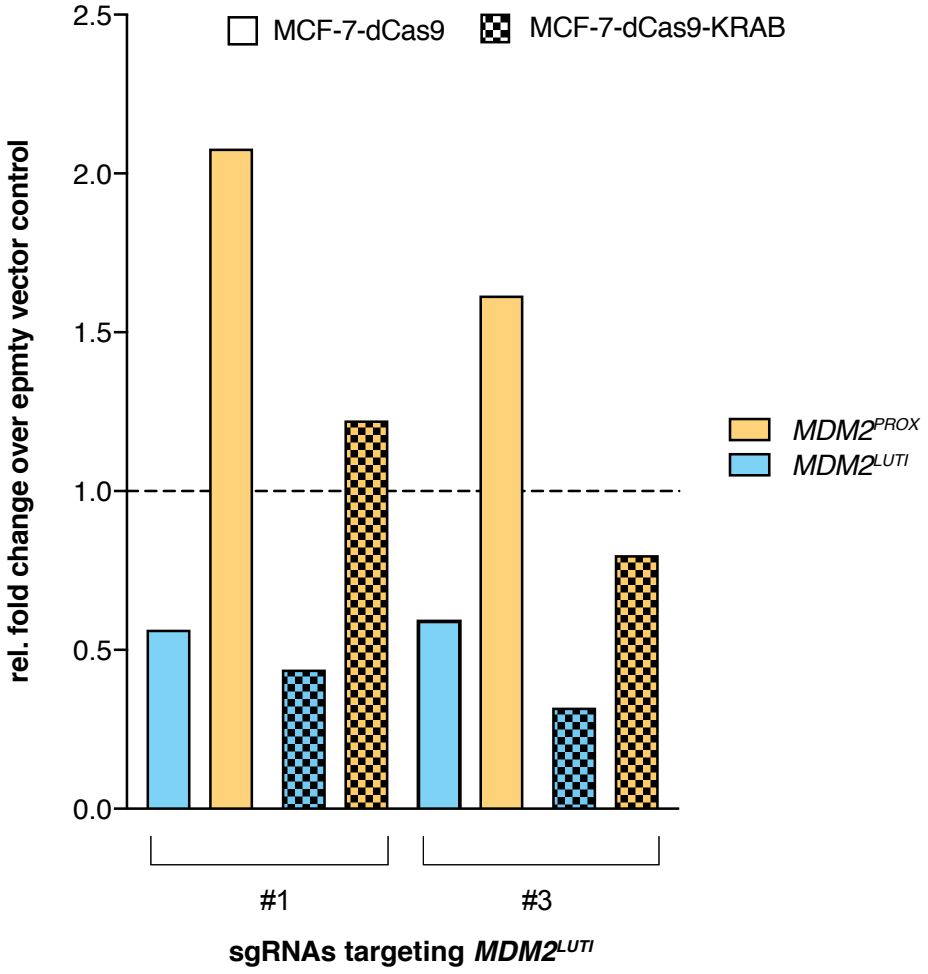
Hollerer et al. Figure 5



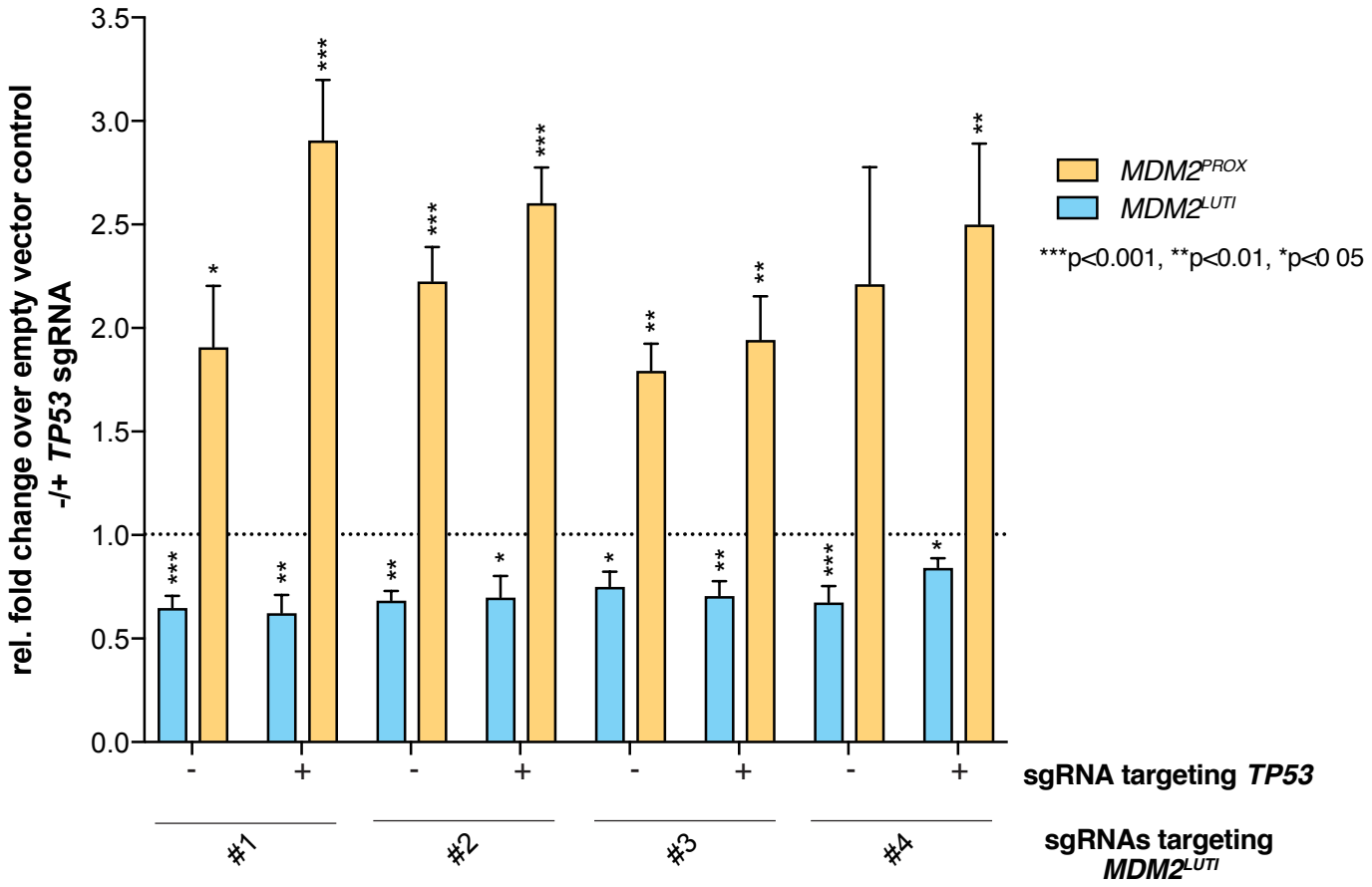
Hollerer et al. Figure S1



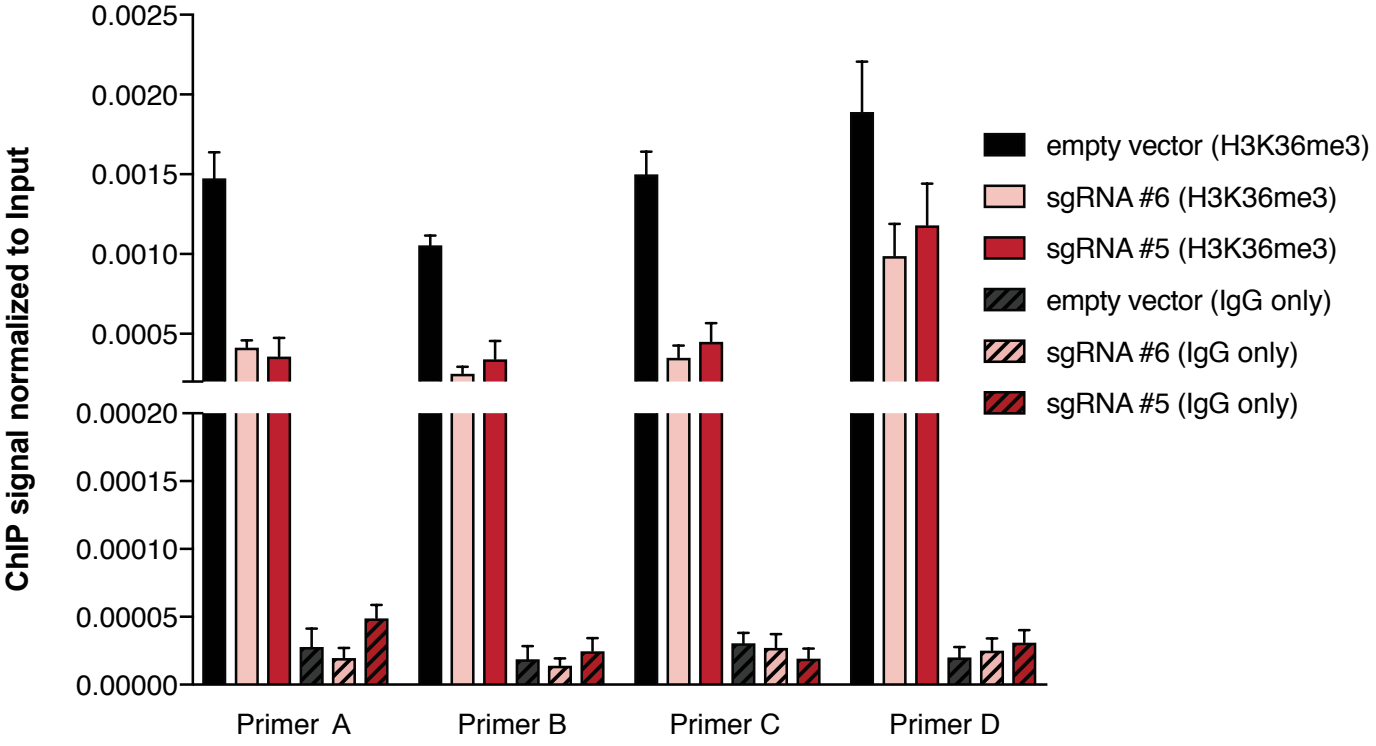
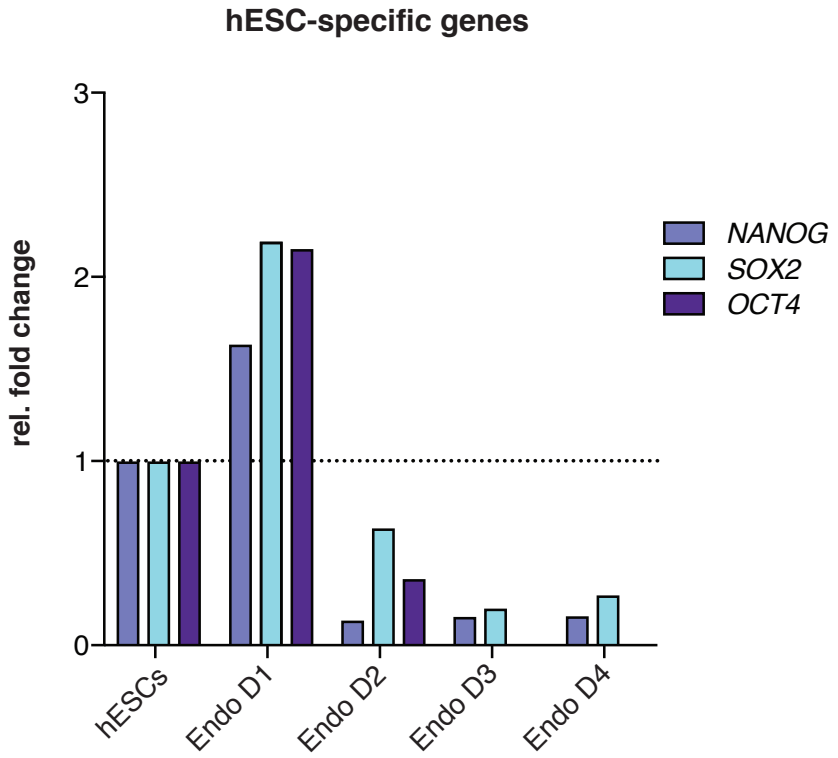
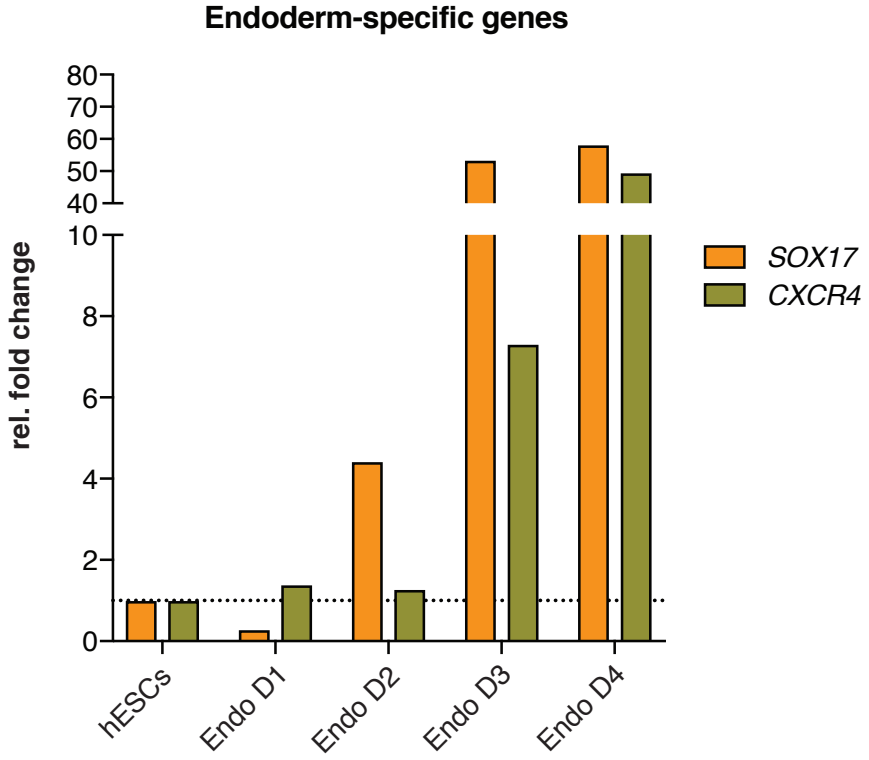
Hollerer et al. Figure S2



Hollerer et al. Figure S3



Hollerer et al. Figure S4



Hollerer et al. Figure S5