Long noncoding RNAs (lncRNAs) are emerging as regulators of fundamental biological processes. Here we report on the characterization of an intergenic lncRNA expressed in epithelial tissues which we termed EPR (Epithelial cell Program Regulator). EPR is rapidly downregulated by TGF-β and its sustained expression largely reshapes the transcriptome, favors the acquisition of epithelial traits, and reduces cell proliferation in cultured mammary gland cells as well as in an animal model of orthotopic transplantation. EPR generates a small peptide that localizes at epithelial cell junctions but the RNA molecule per se accounts for the vast majority of EPR-induced gene expression changes. Mechanistically, EPR interacts with chromatin and regulates Cdkn1a gene expression by affecting both its transcription and mRNA decay through its association with SMAD3 and the mRNA decay-promoting factor KHSRP, respectively. We propose that EPR enables epithelial cells to control proliferation by modulating waves of gene expression in response to TGF-β.
Human transcriptome analysis has revealed the existence of a surprisingly high number of noncoding RNAs that have been classified in multiple families based on their size and biogenesis. Long noncoding RNAs (lncRNAs) are defined as transcripts longer than 200 nucleotides transcribed by RNA polymerase II and commonly originated from intergenic regions. LncRNAs can be capped, spliced, and polyadenylated and usually show limited protein coding potential (refs. 1, 2, and literature cited therein).

LncRNAs are emerging as a fundamental aspect of biology due to their ability to reprogram gene expression and influence distinct cellular functions including cell fate determination, cell cycle progression, apoptosis, and aging12,13. Their expression is usually tissue restricted, developmentally regulated, and can change under specific pathological conditions. Many lncRNAs influence hallmarks of cancer such as uncontrolled proliferation, evasion of cell death, as well as metastasis formation and it has been suggested that lncRNAs can act as oncogenes or tumor suppressors—either directly or indirectly—by interfering with different pathways3,4. From a mechanistic point of view, lncRNAs may influence the function of transcriptional complexes, modulate chromatin structures, serve as scaffolds to form ribonucleoprotein (RNP) complexes or as decoys for proteins and micro-RNAs (miRNAs)3,5. Thus, lncRNA-mediated control of gene expression may take place at transcriptional and/or posttranscriptional levels5–9.

Recently, lncRNAs have been described as important components of the transforming growth factor β (TGF-β) signaling pathway10,11. TGF-β belongs to a large family of structurally related cytokines that regulate growth, survival, differentiation, and migration of many cell types including mammalian gland epithelial cells (ref. 12, for a review). TGF-β activates membrane kinase receptors and induces phosphorylation of cell-specific SMAD proteins that, in complex with the common SMAD4, transduce extracellular stimuli into gene expression changes due to its ability to interact with several molecular partners14. KHSRP is able to posttranscriptionally regulate gene expression by promoting decay of unstable mRNAs, favoring maturation of select miRNAs from precursors, and controlling alternative splicing events14. Recently, we reported that KHSRP affects the alternative splicing of a cohort of pre-mRNAs that encode regulators of cell migration of many cell types including mammary gland epithelial cells (ref. 12, for a review). TGF-β activates membrane kinase receptors and induces phosphorylation of cell-specific SMAD proteins that, in complex with the common SMAD4, accumulate into the nucleus to regulate gene expression at different levels (ref. 13, for a recent review).

In our previous studies, we showed that the multifunctional RNA-binding protein KHSRP acts as a regulatory hub that conveys extracellular stimuli into gene expression changes due to its ability to interact with several molecular partners14. KHSRP is able to posttranscriptionally regulate gene expression by promoting decay of unstable mRNAs, favoring maturation of select miRNAs from precursors, and controlling alternative splicing events14. Recently, we reported that KHSRP affects the alternative splicing of a cohort of pre-mRNAs that encode regulators of cell adhesion and motility—such as CD44 and FGFR2—favoring their epithelial type exon usage and that miRNA-mediated KHSRP silencing is required for TGF-β-induced epithelial-to-mesenchymal transition (EMT) in immortalized NMuMG mammary gland cells15. Further, we found that Resveratrol—a natural polyphenolic compound endowed with anti-inflammatory, anti-proliferative, as well as proapoptotic activities—prevents TGF-β-dependent KHSRP downregulation. Thus shifting Cdk4 and Fgfr2 pre-mRNA alternative splicing from the mesenchymal-specific to the epithelial-specific isoforms16. Our previous observation that the lncRNA H19 interacts with KHSRP and affects its mRNA decay-promoting function17 prompted us to identify additional KHSRP/lncRNAs interactions endowed with regulatory potential.

Here we describe a previously uncharacterized mammalian lncRNA expressed in epithelial tissues that we termed EPR (after Epithelial Program Regulator). EPR came to our attention due to its ability to interact with KHSRP and to counteract TGF-β-induced EMT. EPR contains an open reading frame (ORF) that is translated into a small peptide localized at epithelial cell junctions. However, we found that EPR regulates the expression of a large set of target transcripts independently of the peptide biogenesis. Our studies have revealed that EPR interacts with chromatin, regulates Cdkn1a gene expression by affecting both its transcription and mRNA decay, and controls cell proliferation in both immortalized and transformed mammary gland cells as well as in a mouse model of orthotopic transplantation.

**Results**

**Identification of EPR, an epithelial cell-enriched lncRNA.** This study was initiated in an attempt to identify lncRNAs which are able to interact with KHSRP and whose expression is regulated by TGF-β in immortalized murine mammary gland NMuMG cells. To this end, we leveraged RNA-seq analysis performed in untreated or TGF-β-treated NMuMG cells. TGF-β treatment significantly reduced or increased the levels of 110 and 194 lncRNAs, respectively ([log2 fold changes] > 2.0, p < 0.01 (Student’s t test); Supplementary Table 1a) while RIP-seq analysis showed that TGF-β modulates the interaction of KHSRP with 67 lncRNAs ([log2 fold changes] > 2.0, p < 0.01 (Student’s t test); Supplementary Table 1b), among a set of lncRNA candidates of potential interest in EMT, we focused on the previously uncharacterized BC030870 (ENSMUSG00000074300, located on mouse chromosome 8 and transcribed in reverse orientation) that we renamed EPR (highlighted in yellow in Supplementary Table 1a and 1b). RIP analysis followed by quantitative RT-PCR (qRT-PCR) as well as bandshift analysis confirmed that EPR directly interacts with KHSRP (Supplementary Fig. 1a, b). TGF-β induced a small increase in EPR levels followed by rapid downregulation (Fig. 1a) that accounts for the reduced interaction between KHSRP and EPR upon a 6-h treatment (Supplementary Table 1b). TGF-β-dependent modulation of EPR expression requires TGF-β type 1 receptor signaling as shown by the ability of SB431542 (a selective inhibitor of ALK5, 4, and 7) to abrogate the effect of the cytokine on EPR expression (Supplementary Fig. 1c). SMAD complexes are major effectors of TGF-β-dependent transcriptional regulation13 and our ChIP-qPCR showed that SMAD3 interacts with EPR promoter in a TGF-β-modulated way (Supplementary Fig. 1d, upper panel). Positive (Serpinel) and negative (Mett9) controls for ChIP experiments are provided in Supplementary Fig. 1d (lower panel) and Supplementary Fig. 1e, respectively. Our data are consistent with the hypothesis that SMAD3 interacts with a corepressor complex on EPR promoter region to modulate its transcription19. De novo protein synthesis is not required for TGF-β-induced downregulation of EPR expression as revealed by the use of cycloheximide (Supplementary Fig. 1f, upper panel; Zeb2 (also known as Sip1) represents the control for cycloheximide activity20).

EPR is expressed during embryonic development (Supplementary Fig. 1g) and in epithelial tissues of adult mice with a prevalence in the gastrointestinal tract, lung, kidney and mammary gland (Fig. 1b). EPR is polyadenylated and spliced (Supplementary Fig. 1h) and it is almost equally distributed in the cytoplasm, nucleoplasm, and chromatin of NMuMG cells (Fig. 1c; see Supplementary Fig. 2a for an immunoblot-based validation of cell fractionation). LINCO1207 (a.k.a. SMIM31, located on chromosome 4 and transcribed in forward orientation; hereafter indicated as h.EPR) is the human ortholog of EPR and displays superimposable epithelial tissue-enriched expression (as evaluated through the Human BodyMap 2.0 data from Illumina; Supplementary Fig. 2b). Bioinformatics analysis performed on RNA-Seq data derived from different subpopulations of normal breast cells isolated by FACS analysis from reduction mammary plasty specimens21 revealed that h.EPR is expressed exclusively in differentiated luminal cells of the mammary gland (Fig. 1d).
In order to investigate the potential role of EPR in TGF-β-induced EMT, we decided to counteract TGF-β-dependent EPR downregulation by stably overexpressing the lncRNA in NMuMG cells (overexpression was 3- to 12-fold compared to the respective mock cells (empty vector-transfected), in different transfectant pools). EPR overexpression prevented TGF-β-induced downregulation of epithelial factors (Cdh1, Ocln) and induction of mesenchymal markers (Fn1, Fstl1, Zeb2, Adam12) as well as morphological changes (Fig. 1e, Supplementary Fig. 2c). Strikingly, we observed that EPR overexpression affects the levels of...
epithelial and mesenchymal markers (Fig. 1f, Supplementary Fig. 2d), and induced a cobblestone-like cell morphology in untreated cells (Fig. 1g). Further, EPR overexpression significantly limited the migratory potential of NMuMG cells (Supplementary Fig. 2e). Conversely, transient silencing of EPR downregulated the mRNA levels of epithelial markers, enhanced the levels of mesenchymal markers (Fig. 1h), and rescued the gene expression changes induced by stable overexpression of the IncRNA (Supplementary Fig. 2f). Interestingly, bioinformatics analysis of RNA-Seq data derived from human normal breast samples revealed a statistically significant positive correlation between the expression of h.EPR and epithelial markers such as CDH1 and OCLN and a negative correlation with mesenchymal markers such as VIM and SNAI1 (Supplementary Fig. 2g). This observation is in agreement with the evidence that EPR expression is mutually exclusive with the expression of the EMT factor Cdhl2 as revealed by bioinformatics analysis of datasets derived from single-cell RNA-Seq analysis performed in mice (Supplementary Fig. 2h).

In conclusion, the name EPR that we assigned to IncRNA BC030870 (after Epithelial Program Regulator) is consistent with its enriched expression in epithelial cells and with the upregulation of epithelial markers and downregulation of mesenchymal markers induced by its overexpression.

EPR encodes a small polypeptide. A few recent reports show that certain IncRNAs contain short ORFs that can be translated into peptides endowed with regulatory functions22-25. The analysis of EPR sequence revealed the presence of a 213 nucleotide-long ORF potentially encoding a 71-amino acid polypeptide that, interestingly, corresponds to the IncRNA region that displays the highest identity with the human ortholog (Fig. 2a). The putative polypeptide sequence is well conserved among mammalian species and in silico methods identified a conserved α-helical transmembrane domain while a predicted second α-helix was found in the putative cytosolic domain (Fig. 2b). Importantly, polysome fractionation followed by qRT-PCR analysis revealed that EPR localizes to actively translating polysomes (Supplementary Fig. 3a).

To investigate whether EPR ORF is translated, we inserted a FLAG tag at its 3’ end and transiently transfected the resulting construct into HEK-293 cells (Fig. 2c, left). As shown in Fig. 2c (right), the ORF was translated into a short polypeptide of the expected molecular mass. To unambiguously prove the existence of the endogenous small EPR-encoded peptide (EPRp), the ORF was expressed in bacteria and the resulting peptide was purified and utilized as immunogen to generate a rabbit polyclonal antibody. Polyclonal anti-EPRp recognized a recombinant polypeptide transiently expressed in HEK-293 cells (Supplementary Fig. 3b) and, most importantly, a ~8 KDa polypeptide in mouse gastrointestinal tract organs and breast (Supplementary Fig. 3c). In keeping with EPR downregulation upon TGF-β treatment, the expression of the EPRp was downregulated in response to treatment with TGF-β for 24 h (Fig. 2d).

To investigate the molecular partners of EPRp, we performed immunoaffinity purification of proteins interacting with EPRp in NMuMG cells. Mass spectrometry (MS) analysis of coimmunoprecipitating proteins separated by SDS-PAGE (Fig. 2e) revealed an enrichment in junctional and cytoskeletal proteins (Supplementary Data 1). Coimmunoprecipitation experiments confirmed that EPRp interacts with the tight junction proteins TJP1 (ZO-1) and CGN (Gingulin), with the tight and adherens junction protein CGNL1 (Paracingulin) as well as with the actin-associated proteins CTNNT (Cortactin) and MYH9 (epithelial myosin-II) (Fig. 2f, Supplementary Fig. 3d).

EPR regulates gene expression in NMuMG cells. We set out to investigate the function(s) of EPR in NMuMG cells. First, in order to answer the question whether the phenotypic changes that we observed by overexpressing EPR were caused by the IncRNA per se, the peptide or both, we performed transcriptome-wide RNA-Seq analyses in mock cells as well as in NMuMG cells overexpressing either EPR or a point-mutant version unable to produce the peptide (EPRSTOPe, for details see above and Fig. 3a). Bioinformatics analyses of RNA-Seq data revealed a vast rearrangement of the transcriptome as a consequence of both EPR and EPRSTOPe overexpression (Supplementary Data 2). Gene ontology (GO) analysis of RNA-Seq results revealed the enrichment of terms related to epithelial morphogenesis, cell motility, cell migration, and epithelial cell proliferation among the top regulated categories. Representative examples of transcripts either upregulated or downregulated by both EPR and EPRSTOPe are shown in Fig. 3b, c. In keeping with the sequence conservation between EPR and h.EPR, the overexpression of the human IncRNA in murine NMuMG cells yielded gene expression changes superimposable to those obtained by overexpressing the murine IncRNA (Supplementary Fig. 4a).

Interestingly, overexpression of either EPR or EPRSTOPe caused largely overlapping gene expression changes when compared to mock cells (Fig. 3d, upper panel). When we directly compared gene expression changes induced by EPR or EPRSTOPe by applying stringent statistical criteria, we noticed that only a relatively small group of genes displayed expression changes dependent on the presence of EPRp (Fig. 3d, lower panel). The analysis of three independent NMuMG transfectant pools overexpressing EPRSTOPe, followed by qRT-PCR-based validation, allowed us to further restrict the number of transcripts whose levels are affected by the peptide per se (Fig. 3e, Supplementary Fig. 4b). These include transcripts encoding a calcium-dependent cell adhesion protein (Pcdh19), two ion transporters (Slc9a2, Slc39a4), a cytokine receptor (Fgfr2) as well as a modulator of membrane transport and actin dynamics (Anxa6). Further, analysis of an additional EPR mutant (referred to as EPRSTOPe in which the start codon has been mutagenized to a STOP codon, see below) confirmed the restricted number of gene expression changes that can be ascribed to the peptide translation (Supplementary Fig. 4c).
Fig. 2 A small peptide (EPRp) originates from EPR and displays junctional localization. a Schematic of the exon–intron structure (not in scale) of EPR, a putative ORF is dashed. Percentage of human/mouse identity in the putative ORF and in its flanking RNA regions is presented. b Alignment of the predicted mammalian amino acids sequence encoded by the putative ORF present in EPR and in its orthologs. The position of the predicted transmembrane α-helix is shown as a solid black bar while the position of an additional cytoplasmic α-helix is shown as a solid green bar. c Left, diagram of the FLAG-fusion construct used for transfection (EPRp-FLAG); right, immunoblot analysis of total cell extracts from either mock or EPRp-FLAG-transfected HEK-293 cells. d Immunoblot analysis of total cell extracts from NMuMG cells serum-starved and either treated with TGF-β (+) for 24 h or untreated (−); extracts from EPRp-FLAG-transfected HEK-293 cells (EPRp-FLAG) represent a positive control. e SDS-PAGE analysis of total cell extracts from either mock or EPRp-FLAG-overexpressing (EPRp-FLAG) NMuMG cells immuno-purified using anti-FLAG monoclonal antibody. A representative silver-stained gel is shown. Asterisks indicate the position of immunoglobulin heavy and light chains. f Coimmunoprecipitation of FLAG-tagged EPRp and distinct junctional proteins (as indicated) in total extracts from NMuMG cells stably transfected with EPRp-FLAG. g Immunofluorescence analysis of either mock or EPRpSTOP-FLAG- or EPRpSTOP-FLAG-stably transfected NMuMG cells cultured to confluence, to allow formation of cell–cell junctions. Arrows point to CGN and EPRp-FLAG junctional localization. Scale bar: 10 μm. For immunoblots, the indicated antibodies were used; the position of molecular mass markers is presented on the left and representative gels are shown. ACTB is also known as Actin Beta.
Fig. 3 EPR overexpression reshapes NMuMG cells transcriptome and reduces cell proliferation. a Schematic of the point mutation introduced in the second codon of the EPR ORF to generate EPRSTOPE (left panel). HEK-293 cells were transiently transfected with either EPR or EPRSTOPE and the translation of EPR was assessed by immunoblotting (right panel). b qRT-PCR analysis of the indicated transcripts in either mock, EPR- or EPRSTOPE-expressing NMuMG cells. c Immunoblot analysis of total cell extracts from either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. d Upper panel, TopTable Venn diagram of the limma one-factor contrast analysis showing the overlap between transcripts either upregulated (numbers in red) or downregulated (numbers in blue) in cells stably overexpressing either EPR or EPRSTOPE compared with mock cells. Only transcripts displaying |log2 fold expression difference| > 2.5 (p < 0.0001, Student’s t test) were included in further comparisons. Lower panel, Pie diagram showing the percentage of gene expression changes common to NMuMG cells overexpressing EPR and EPRSTOPE as well as the percentage of changes that can be specifically attributed to the EPR translation. e qRT-PCR analysis of the indicated transcripts in either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. f Immunoblot analysis of total cell extracts from either mock, EPR-, EPRSTOPE- or EPRSTOPM-overexpressing NMuMG cells. g qRT-PCR analysis of Cdkn1a mRNA levels in either control siRNA (siC) or siEPR-transfected NMuMG cells (upper panel). Immunoblot analysis of total cell extracts from either control siRNA (siC) or siEPR-transfected NMuMG cells (lower panel). h Proliferation analysis (Operetta CLS High-Content Analysis System) of either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. i Cell cycle distribution in cultures grown at similar density (near 90% confluence). Cell cycle was analyzed by flow cytometry after double staining with EdU and propidium iodide of either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. Bars plot the relative proportion of cells in G1, S, and G2 phases for each cell line. The values of qRT-PCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: *p < 0.01, **p < 0.001 (Student’s t test). For immunoblots, the indicated antibodies were used; the position of molecular mass markers is indicated on the left and representative gels are shown.
Considering the emerging evidence that some IncRNAs act locally (in cis) to regulate the expression of nearby genes, we investigated this possibility and RNA-Seq analysis revealed that the expression levels of genes proximal to EPR (Palld, Cpe, Sc4mol, Kllh2, Tnem192, Tnma16, Naf16, Nat2 and Psal, localized over 8 MB of chromosome 8) are unaffected by the almost complete EPR downregulation that occurs in NMuMG cells treated with TGF-β for 24 h (R.G. and G.B., unpublished observation).

Altogether, transcriptome-wide analyses showed a EPR-dependent wide rearrangement of the transcriptome in NMuMG cells with relatively restricted effects on gene expression ascribed to the peptide. Thus, we decided to focus our further studies on the EPR functions that are independent of the peptide biogenesis.

EPR regulates Cdkn1a gene expression and cell proliferation.

Among the GO terms significantly enriched by the overexpression of either EPR or EPRSTOPE, we identified the category Regulation of Epithelial Cell Proliferation. Indeed, both EPR and EPRSTOPE overexpression significantly affected the levels of a group of transcripts belonging to this category including the cyclin-dependent kinase inhibitor Cdkn1a (a.k.a. p21WAF1/Cip1) (Supplementary Fig. 4d). Immunoblots presented in Fig. 3f show that overexpression of either EPR or EPRSTOPE or EPRSTOPM strongly enhanced Cdkn1A levels. Conversely, EPR silencing strongly reduced Cdkn1A expression (Fig. 3g). As expected, Cdkn1A levels were enhanced by overexpression of the human ortholog of EPR (Supplementary Fig. 4e). Most important, we found that overexpression of either EPR or EPRSTOPE as well as of h.EPR strongly reduces cell proliferation rate in NMuMG cells (Fig. 3h, Supplementary Fig. 4f). Cell cycle analysis demonstrated a relevant increment of cells arrested in the G1 phase in the case of NMuMG cells transfected with either EPR or EPRSTOPE in comparison to mock cells (Fig. 3i). To exclude the possibility that gene expression changes that we observed (Fig. 3b) might be dependent on the EPR-induced G1 arrest, we sorted cells in the G1 phase and analyzed gene expression changes by qRT-PCR. Data presented in Supplementary Fig. 4g indicate that the expression changes induced in G1-enriched cells by overexpression of either EPR or EPRSTOPE are superimposable to those observed in the total cell population (Fig. 3b).

Together, these results provide evidence that modulation of EPR levels regulates Cdkn1a gene expression and affects cell proliferation in NMuMG cells. Given the role of Cdkn1A in promoting cell cycle arrest in response to many stimuli—including TGF-β26—we decided to focus our further mechanistic studies on the role of EPR in TGF-β-dependent regulation of Cdkn1a gene expression.

EPR regulates TGF-β-dependent Cdkn1a gene expression.

Analysis of newly synthesized transcripts revealed that overexpression of either EPR or EPRSTOPE strongly enhances Cdkn1a transcription (Fig. 4a) and the kinetic analysis of mRNA decay indicated that overexpression of either EPR or EPRSTOPE induces also a significant stabilization of Cdkn1a mRNA (Fig. 4b).

TGF-β signaling promotes tissue growth and morphogenesis during embryonic development while, as tissues mature, many cell types gain the ability to respond to TGF-β with growth arrest that is primarily due to imbalance of G1 events27. As similarly reported in other cell types28,29, treatment of NMuMG cells with TGF-β for 1 h caused a rapid induction of Cdkn1a gene expression that was followed by return to baseline levels after 6 h (Fig. 4c). The observation that Cdkn1a return to baseline levels matches EPR downregulation (Fig. 4c) and that EPR overexpression strongly enhances Cdkn1a levels, prompted us to hypothesize a role for EPR in the TGF-β-dependent modulation of Cdkn1a gene expression. Our ChIP-qPCR assays showed that TGF-β treatment for 1 h stimulates the binding of SMAD3 to Cdkn1a promoter that returns to basal levels after 6 h (Fig. 4d, see also ref. 28). TGF-β-dependent control of Cdkn1a mRNA decay was never investigated in detail but, considering that cells often achieve rapid changes of gene expression by integrating gene transcription control with regulated mRNA decay30,31, we addressed the possibility that TGF-β could affect Cdkn1a mRNA decay. Figure 4e showed that Cdkn1a mRNA stability is unaffected by 1 h of TGF-β treatment (upper panel) but is reduced when the treatment is prolonged up to 6 h (lower panel). Thus, the TGF-β-dependent rapid fluctuations of Cdkn1a expression depend on the regulation of both transcription and mRNA decay in NMuMG cells.

Our hypothesis that EPR plays a role in the regulation of TGF-β-dependent Cdkn1A expression was supported by the evidence that EPR silencing abrogated Cdkn1a mRNA induction upon TGF-β treatment for 1 h (Fig. 4f) while its overexpression enhances Cdkn1a levels and blunts its rapid modulation by TGF-β (Fig. 4g).

Together, our results indicate that EPR plays a dual role in TGF-β-dependent Cdkn1a gene expression control.

EPR affects both Cdkn1a gene transcription and mRNA decay.

We investigated the molecular mechanism(s) by which EPR regulates Cdkn1a gene transcription. The evidence of enhanced RNA-Pol II occupancy and reduced presence of the H3K27me3 repressive mark at the Cdkn1a promoter in EPR-overexpressing cells (Supplementary Fig. 5a) together with our finding that EPR is present in the chromatin fraction (see Fig. 1c) prompted us to explore the possibility that EPR affects Cdkn1a transcription through direct interaction with its promoter region. Chromatin isolation by RNA Purification (ChiRIP)-Seq experiments (P.B., G.B., E. Zapparoli et al., unpublished) as well as ChiRIP-qPCR experiments revealed the direct interaction of EPR with Cdkn1a promoter (Fig. 5a). The interaction of EPR with Cdkn1a promoter is not significantly affected by a 1 h TGF-β treatment (Supplementary Fig. 5b).

RIP experiments showed that SMAD3 interacts with EPR and the interaction is enhanced by treatment with TGF-β for 1 h (Supplementary Fig. 5c). In keeping with growing evidence suggesting that the interaction between IncRNAs and specific transcription factors can affect gene expression5,32, ChIP-qPCR experiments showed that EPR overexpression enhances SMAD3-Cdkn1a promoter association and abrogates its dismissal after 6 h of TGF-β treatment (Fig. 5b). These effects are reproduced by overexpression of EPRSTOPE (Fig. 5b). Cell treatment with SB431542 abrogated the TGF-β-dependent enhancement of SMAD3-Cdkn1a promoter association in mock as well as in NMuMG cells overexpressing either EPR or EPRSTOPE (Supplementary Fig. 5d). Notably, EPR overexpression favored SMAD3-Cdkn1a promoter interaction also in untreated cells and this was not modified by SB431542 treatment (Fig. 5b and Supplementary Fig. 5d). To explain the association of SMAD3 with Cdkn1a promoter in cells overexpressing EPR also in the absence of TGF-β treatment, we hypothesize that EPR overexpression favors the association of SMAD3 molecules present in NMuMG cell nuclei of untreated cells with Cdkn1a promoter. Indeed, it is known that, although SMAD proteins rapidly accumulates into nuclei upon TGF-β treatment33, a certain amount of SMAD3 is present in the nuclei of untreated cells (ref. 34; Supplementary Fig. 5e).

In keeping with results shown in Fig. 4b, we found that EPR overexpression prevents Cdkn1a mRNA destabilization induced by a treatment with TGF-β for 6 h (Fig. 5c). Our initial observation that EPR interacts with KHSRP, a factor able to
promote rapid decay of select labile mRNAs in many cell types, prompting us to explore whether KHSRP regulates Cdkn1a mRNA accumulation and prevented its rapid degradation (Supplementary Fig. 6a and Fig. 5d, upper panel) while transient KHSRP overexpression in NMuMG cells stably expressing EPR promoted Cdkn1a mRNA destabilization (Fig. 5d, lower panel). KHSRP is predominantly nuclear in NMuMG cells and we found that mature Cdkn1a mRNA is abundant in nuclear fractions of these cells (Supplementary Fig. 6b) where it

![Image](https://example.com/image1)

**Fig. 4** Dual regulation of Cdkn1a gene expression by EPR. a qRT-PCR analysis of either total (left) or nascent (right) Cdkn1a transcript in either mock, EPR- or EPRSTOP- overexpressing NMuMG cells. b Either mock, EPR- or EPRSTOP-overexpressing NMuMG cells were treated with 100 µM 5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) for different times (as indicated). Cdkn1a gene expression was analyzed by qRT-PCR. c qRT-PCR analysis of Cdkn1a (black line) and EPR (red line) in NMuMG cells serum-starved (2% FBS, 16 h) and either treated with TGF-β (10 ng ml⁻¹) for the indicated times or untreated (time 0). d Chromatin prepared from NMuMG cells serum-starved and either treated with TGF-β for either 1 h (top panel) or 6 h (bottom panel) or left untreated (control in both panels). Subsequently, cells were treated with 100 µM DRB for the indicated times and total RNA was isolated and analyzed by qRT-PCR to quantify Cdkn1a mRNA levels. Please note the slight, reproducible difference in the decay kinetic of Cdkn1a mRNA between nontransfected and mock-transfected NMuMG cells. 

![Image](https://example.com/image2)

f qRT-PCR analysis of Cdkn1a in NMuMG cells transiently transfected with either control siRNA (scC) or siRNA designed to silence EPR expression (siEPR), serum-starved (2% FBS, 16 h) and then either treated with TGF-β (10 ng ml⁻¹) for 1 h (-) or left untreated (--). g qRT-PCR analysis of Cdkn1a expression in either mock (blue line) or EPR overexpressing (red line) NMuMG cells serum-starved (2% FBS, 16 h) and either treated with TGF-β (10 ng ml⁻¹) for the indicated times or left untreated (time 0). Please note the logarithmic scale of the Y-axis. The values of both qRT-PCR and qPCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: *p < 0.01, **p < 0.001 (Student’s t test)
undergoes rapid decay and is stabilized by EPR overexpression as measured by two independent techniques (Supplementary Fig. 6c, d). These findings suggested that EPR might interfere with the ability of KHSRP to interact with Cdkn1a mRNA. RIP experiments presented in Fig. 5e show that KHSRP association with Cdkn1a mRNA was strongly reduced by EPR and EPRSTOP overexpression. Based on these results, we investigated whether EPR downregulation, that is physiologically obtained by TGF-β treatment, affects the interaction of KHSRP with Cdkn1a mRNA. RIP analyses indicated that KHSRP-EPR association is abrogated while KHSRP interaction with Cdkn1a mRNA is increased upon TGF-β treatment (6 h) (Fig. 5f).

Altogether, these results suggest that EPR controls Cdkn1a expression by sustaining its transcription and by impairing its mRNA decay in response to TGF-β.

**EPR overexpression reduces breast cancer cell proliferation.** We investigated EPR in murine breast cancer cell lines and
observed that its expression is severely reduced when compared with immortalized NMuMG cells (Supplementary Fig. 7a). Similarly, the expression of hEPR was below detection levels in highly aggressive human breast cancer cell lines (M.J.M. and F.N., unpublished observation). hEPR could be detected in about 75% of breast cancer primary samples (780/1043 cases from The Cancer Genome Atlas (TCGA) database36; Supplementary Fig. 7b) and, according to PAM50 molecular subtype classification, it was more expressed in Luminal A and Her2 tumors while it was almost absent in Basal-like tumors, the most frequent subtype of triple-negative breast cancers37 (Fig. 6a).

Based on these observations, we decided to express EPR in triple-negative mesenchymal-like breast cancer cells, such as murine 4T1 and human MDA-MB-231 cell lines, respectively (Supplementary Fig. 7c).

Overexpression of either EPR or EPRSTOPE in 4T1 cells resulted in a strong induction of Cdkn1a gene expression as well as in a significant reduction of clonogenic potential, cell proliferation, and anchorage-independent cell growth (Fig. 6b-e). EPR overexpression in 4T1 cells also downregulated the expression of mesenchymal factors such as Cdh2 and Adam12 (Supplementary Fig. 7d). Very similar results were observed by overexpressing either human or murine EPR in human MDA-MB-231 cells (Fig. 6f, g, Supplementary Fig. 7e–g).

Finally, to interrogate the activity of EPR on cell proliferation control in vivo, we orthotopically injected either mock or EPR-expressing 4T1 cells into syngenic BALB/c mice. In concordance with our observations in cultured cells, EPR expression resulted in a remarkable reduction of tumor volume after 10 days (Fig. 6h, left panel). A significant reduction of the tumor mass was still evident and statistically significant also at 2 weeks after the transplant when mice were sacrificed (Fig. 6h, right panel, Supplementary Fig. 7h).

Altogether, our results indicate that EPR overexpression modulates cell proliferation and epithelial/mesenchymal markers levels in breast cancer cells and restrains cell proliferation in transplanted mice.

Discussion

Here we report on the initial functional characterization of the long intergenic noncoding RNA EPR well conserved among mammalian species and expressed in select epithelial tissues including differentiated luminal cells of human breast. The levels of EPR are rapidly downregulated by TGF-β/SMAD signaling in immortalized mammary gland cells and its sustained expression largely reshapes the transcriptome by inducing epithelial traits while preventing the acquisition of mesenchymal markers upon TGF-β treatment. Remarkably, EPR overexpression enhances the levels of the cyclin-dependent kinase inhibitor CDKN1A and strongly reduces cell proliferation in both immortalized and transformed mammary gland cells as well as in transplanted mice.

EPR is almost equally distributed in chromatin, nucleoplasm and cytoplasm and the cytoplasmic component associates with polysomes where a small peptide (EPRp) is translated. EPRp interaction with a variety of cytoskeletal and junctional proteins accounts for its junctional localization. However, the analysis of the phenotype that we observed in cells overexpressing EPR mutants unable to originate the peptide clearly indicates that the vast majority of gene expression changes that we describe here are independent of the peptide biogenesis.

In this report, we investigated how the lncRNA molecule per se controls gene expression and we focused our studies on the EPR-dependent regulation of CDKN1A that functions as both a sensor and an effector of multiple antiproliferative signals and promotes cell cycle arrest in response to TGF-β26. In NMuMG cells, TGF-β induces an early wave of Cdkn1a expression due, in part, to an increased SMAD complex-dependent gene transcription while a prolonged treatment causes the return of Cdkn1a levels to the baseline. Our data suggest that Cdkn1a promoter-bound EPR recruits SMAD3 molecules—that accumulate into the nucleus upon TGF-β treatment for 1 h—to induce rapid gene transcription. In parallel, EPR interacts with KHSRP limiting its association with Cdkn1a mRNA and this results in the transcript stabilization. We propose that EPR downregulation upon 6 h of TGF-β treatment causes SMAD3 dismissal from Cdkn1a promoter that results in a return of Cdkn1a transcription to basal levels and, in parallel, enables KHSRP to destabilize the Cdkn1a transcript. Our data suggest that EPR-regulated molecular events shape the rapid wave of Cdkn1a expression in response to TGF-β. The evidence that CDKN1A is more abundant in cells overexpressing EPR than in mock cells (independently of any treatment with TGF-β) allows us to hypothesize that overexpressed EPR is able to recruit SMAD3 molecules already present in cell nuclei to the Cdkn1a promoter region and, possibly, to distal enhancers as well as to block KHSRP-induced Cdkn1 mRNA degradation.

Our data indicate that EPR couples Cdkn1a transcriptional regulation with mRNA decay control. Indeed, the integration of transcription and mRNA decay provides a kinetic boost to a series of processes that would be otherwise slower and less efficient. This report strengthen the idea that coupling transcription to mRNA decay enables cells to rapidly modulate waves of gene
expression in response to a variety of stimuli thus achieving optimal mRNA homeostasis.38,39

Since the decay of mature mRNAs is generally thought to occur in the cytoplasm, our finding that spliced and polyadenylated Cdkn1a mRNA is abundant in the nucleus of NMuMG cells, where it undergoes regulated decay, might be somehow surprising. However, the possibility that mature mRNAs accumulate in the nuclei of mammalian cells has been described and a previous report has suggested that Cdkn1a mRNA undergoes degradation by the nuclear exosome in mammalian cells.40

Further, a report about mature RNA degradation controlled by nuclear histone acetyltransferases and deacetylases provocatively goes degradation by the nuclear exosome in mammalian cells.40,41 A previous report has suggested that mature mRNAs accumulate in the nuclei of mammalian cells.40

Fig. 6 Antiproliferative effect of EPR when expressed in transformed mammary gland cells. a Box plot shows the expression of h.EPR in the TCGA Breast Cancer (BRCA) dataset annotated according to PAM50 molecular subtype classification. The number of samples in each subtype is presented. Asterisks mark significant values (Wilcoxon’s test; *p < 0.05, **p < 0.001). b qRT-PCR analysis (left panel) and immunoblot analysis (right panel) of Cdkn1a expression in either mock, EPR- or EPRSTOPE-overexpressing 4T1 cells. c Either mock, EPR- or EPRSTOPE-overexpressing 4T1 cells were seeded at low density and colony-formation assays were performed after 4 days. A representative plate is shown. The numbers are averages (±SEM) of four independent experiments performed in triplicate. Statistical significance: ***p < 0.000001 (Student’s t test). d Cell proliferation analysis of either mock, EPR- or EPRSTOPE-overexpressing 4T1 cells. e Cells were cultured in soft agar for 21 days and phase contrast micrographs were taken at ×10 magnification. The values are averages (±SEM) of four independent experiments performed in triplicate. Statistical significance: ***p < 0.000001 (Student’s t test). f qRT-PCR analysis of CDKN1A mRNA levels in transfected MDA-MB-231 cells. g Cell proliferation analysis of either mock or h.EPR-overexpressing MDA-MB-231 cells. h Tumor volume (n = 6 tumors/group) was measured by digital caliper assessment 10 days after injection of either mock or EPR-expressing 4T1 cells in BALB/c mice. Box plot analysis of tumor volume is shown (left). Box plots display summaries of the data distribution: the lower whisker is the minimum, the lower box edge is the first quartile, the middle line is the median, the upper box edge is the third quartile, and the upper whisker the maximum value. Data were analyzed in R version 3.4.3, using the Wilcoxon Unpaired Test as implemented in “stat_compare_means” in “ggplot 2.2.” Images of the tumors at the end of the experiment, 2 weeks after injection (right). The values of qRT-PCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: **p < 0.001 (Student’s t test).
suggested the possibility that mRNA degradation pathways might operate also in the nucleus.

The TGF-β cysteotic program in epithelial cells involves, among other molecular events, the induction of Cldn1 and Cldn2b but cancer cells utilize any opportunity to circumvent TGF-β ability to inhibit cell proliferation. Inactivating mutations in the TGF-β Type-II receptors and SMAD4 have been described in tumors even though cancer cells can lose the cystostatic responsiveness due to defects downstream to SMAD factors. We hypothesize that the absence of EPR/h.EPR, that occurs in breast cancer cell lines and in certain breast cancers, may contribute to the loss of TGF-β ability to restrain cell proliferation while may enable the cytokine to sustain their carcinogenic potential.

The notion that lncRNAs are devoid of coding potential has been recently challenged by reports demonstrating that translated short peptides are responsible for the biological functions of the respective lncRNA. However, Yu et al. reported that lnc-RAM, an lncRNA that per se enhances myogenic differentiation by interacting with MYOD, is the transcript encoding myogenin, a small peptide previously reported as a mediator of muscle performance through inhibition of the pump activity of SERCA. In this case, RNA and peptide functions cooperate in modulating muscle physiology. Intriguingly, under our model the lncRNA per se is responsible for most of the gene expression changes while the peptide might be responsible for specific functions related to its cytoskeletal/junctional localization. We hypothesize that the peptide could participate in multiprotein complexes serving as permeability barriers and/or it could be implicated in the establishment of apico-basal polarity as well as in the transduction of signals to the cell interior. Further studies will be needed to clarify the specific functions of the peptide but we can predict that EPR and the peptide might synergize in executing epithelial cell-specific programs.

EPR was discovered as a highly regulated lncRNA in NMuMG mammary gland cells and, by exploiting the TGF-β responsiveness of these cells, we were able to clarify some details of its molecular functions. How EPR influences normal physiology and disease in tissues undergoing repeated rounds of proliferation/differentiation, such as the gastrointestinal tract where it is highly expressed, will represent an important area of future research.

Methods

Cell lines. Marine immortalized NMuMG cells (ATCC, no. CRL-1636) were cultured in Dulbecco's Modified Eagle Medium (DMEM) plus 10% Fetal Bovine Serum (FBS) and 10 μg ml⁻¹ bovine insulin (Sigma-Aldrich), 4T1 mouse mammary gland cancer cells (obtained from ATCC, no. CRL-2539) were cultured in DMEM/F12 plus 10% FBS, human mammary gland adenocarcinoma cells MDA-MB-231 (obtained from DSMZ, Germany, through Dr. G. Fronza, authenticated by STR DNA profiling) were cultured in DMEM plus 5% FBS, and human HEK-293 cells (ATCC, no. CRL-1573) were cultured in DMEM plus 10% FBS. HEK-293 cells were used to verify transient expression of FLAG-tagged EPRs based on their highly efficient transfectability. NMuMG cells were maintained in DMEM supplemented with 2% for 16 h prior to the addition of 1×10⁹ mg ml⁻¹ human recombinant TGF-β1 purchased from R&D Systems. All cell lines were tested for mycoplasma contamination and resulted negative. SB431542 compound was purchased from Sigma-Aldrich, dissolved in Dimethyl sulfoxide (DMSO), and used at a 1 μM concentration. Cycloheximide, dissolved in DMSO, was purchased from Sigma-Aldrich and used at a 5 μM concentration.

Antibodies and immunobots. Anti-EPR polyclonal rabbit antibody was generated by injecting rabbits with recombinant purified EPR expressed in E. Coli using the pQE-EPR at Cambridge Research Biochemicals (Billingham, Cleveland, UK). Anti-CDH1 goat polyclonal antibody (sc-31020, used at 1:500 final dilution), anti-CDKN1A mouse monoclonal antibody (sc-6246, used at 1:200 final dilution), and anti-HDAC1 rabbit polyclonal antibody (sc-7872, used at 1:500 final dilution) were from Santa Cruz; anti-TIP1 rabbit polyclonal antibody (ab96587, used at 1:100 final dilution), anti-SMAD3 rabbit polyclonal antibody (ChIP grade ab28379), anti-GFP rabbit polyclonal antibody (ChIP grade ab90, used at 1:200 final dilution) were from Abcam; mouse monoclonal anti-FLAG (F1804, used at 1:500 final dilution), mouse monoclonal anti-TUBA (DM1, used at 1:1000 final dilution) and mouse monoclonal anti-ACTB (AC-74, used at 1:30,000 final dilution) were from Sigma-Aldrich. Mouse monoclonal anti-RNA Polymerase II (clone CTD4H8) and rabbit polyclonal antibody to H3K27me3 (CS200603) were from Millipore. Rabbit polyclonal anti-CGN serum (CS52, used at 1:5000 final dilution) against a purified recombinant 50 kDa C-terminal fragment of chicken cingulin as well as anti-angiostatin (sc-2133, used at 1:100) rabbit polyclonal antibody was purchased from Santa Cruz Biotechnology, Inc. Anti-EPR antibodies raised at the University of Geneva. Images of the uncropped and unprocessed scans of the most important Immunoblots are presented in Supplementary Fig. 8.

Plasmids. Plasmid EPR was obtained by inserting the sequence from nucleotide 1 to 1487 of murine BC038070 into pBlCEP-CMV-2 vector (Sigma-Aldrich); plasmid h.EPR was obtained by inserting the sequence from nucleotide 1 to 126 of LINCO1207 into pBlCEP-CMV-2 vector. Plasmids EPR-PSTOE and EPR-STOMP were obtained by Site-Directed Mutagenesis of plasmid EPR using the QuickChange II mutagenesis kit (Agilent Technologies) and the oligonucleotides 5’—CACCCGTAGTCCTAGCAGCTACATC—3’ and 5’—CAGCGTGATGTCCATAGCATTACTAC—3’, respectively. Plasmids EPR-FLAG and EPRSTOE-PAGE were generated by inserting the sequence from nucleotide 1 to 560 of murine BC038070 by PCR and flagged at its 3’ (either wild-type or mutated as above) into pIREShyg vector. Plasmids GMP- mouse cortactin (262722, CTTN-GFP) and CMV-GFP-human MYH9-GFP (from Addgene, Plasmid pQ5-EPR was obtained by inserting the sequence from nucleotide 345 to 560 of murine BC038070 into pQ-E30 vector (Qiagen).

For insertions obtained by RT-PCR, the Pfu DNA Polymerase (Promega) was used. The inserts cloned in all constructs were sequenced on both strands (BMR Genomics, Padova, Italy).

Cell transfections. NMuMG, 4T1, and MDA-MB-231 cells were transfected with Lipofectamine 2000 (ThermoFischer) while HEK-293 cells were transfected with Attractene transfection Reagent (Qiagen). NMuMG, 4T1, and MDA-MB-231 cells stably transfected with recombinant pBlCEP-CMV-2-based vectors were maintained in selective medium containing 800, 350, and 750 μg ml⁻¹ G418 (Sigma-Aldrich), respectively. NMuMG cells stably transfected with recombinant pIREShyg-based vectors were maintained in selective medium containing 600 μg ml⁻¹ Hygromycin B (Sigma-Aldrich). Specific mock cells were generated (for every cell line and every plasmid backbone) by transfecting the corresponding empty vector in each cell type. Mock cells were subjected to a selection procedure identical to the other transfectants. siRNAs utilized to knockdown murine EPR (5′—GAAACAGAAAAGUUCUAAG—3′) were purchased from Thermo Fisher. Stable KHSRP knockdown in NMuMG cells was obtained using previously described silencing sequences and pu3pere-Neo (Oligoengine) according to the manufacturer's instructions. The adenoaviral vectors pAdCMVnull (AdNull) and pAdKHSRP (full-length human KHSRP cDNA cloned into an Adenoviral-Type 5-based vectors) were maintained in selective medium containing 600 μg ml⁻¹ G418 (Sigma-Aldrich). A wound was introduced by scratching the confluent cell layer in each cell type. Cell mock were subjected to a selection procedure identical to the other transfectants. siRNAs utilized to knockdown murine EPR (5′—GAGCAAAGAGAAUGCUA—3′) were purchased from Thermo Fisher. Stable KHSRP knockdown in NMuMG cells was obtained using previously described silencing sequences and pu3pere-Neo (Oligoengine) according to the manufacturer's instructions. The adenoaviral vectors pAdCMVnull (AdNull) and pAdKHSRP (full-length human KHSRP cDNA cloned into an Adenoviral-Type 5-based vectors) were purchased from Vector Biolabs.

Scratch wound closure assay. Either mock or EPR-overexpressing NMuMG cells were cultured in six-wells plates up to confluence and pretreated for 2 h with 5 μg ml⁻¹ Mitomycin C (Sigma-Aldrich). A wound was scratched into monolayers and cells were cultured for up to 48 h in the presence of 5 μg ml⁻¹ Mitomycin C. Images were taken using an Olympus CX41 microscope and analyzed 4 to 5 using the ImageJ 1.49r package (http://imagej.nih.gov/ij/index.html). Average distance of wound obtained from six microscopic fields was used for the calculation of percent wound healed. Experiments were performed three times in triplicate.

Immunofluorescence. Either mock NMuMG cells or stable transfectants over-expressing either EPR-FLAG, EPR-PSTOE-FLAG or EPRSTOE-PAGE were plated on glass coverslips in 24-well plates (60,000 cells/well). Immunofluorescence was carried out 2 days after plating essentially as reported in ref. 47. Rabbit polyclonal anti-cingulin antisemur (C532) was used at a 1:5000 dilution while anti-Anti-FLAG antibody (F1804, Sigma) was used at a 1:500 dilution. Secondary antibodies were diluted in IF buffer and incubated for 30 min at 37 °C, Alexa488 anti-rabbit (711-545-152, Jackson Laboratory) dilution 1:400, Cy3 Goat anti-mouse (715-165-151, Jackson Laboratory) dilution 1/400. Pictures were taken using a Zeiss Axioskop widefield fluorescent microscope (X-Cite 120Q mercury lamp light source, Excitlas Technologies; retix EII, cooled mono 12-bit, Qiming camera; x63 oil objective; Openlab software). Images were imported into ImageJ to split and merge channels, cropped and adjusted for resolution and for intensity level range using Photoshop (scale bar = 10 μm).

Orthotopic 4T1 injection in BALB/c mice. BALB/c 8–10-week-old female mice (Envigo) were anesthetized using 100 mg kg⁻¹ ketamine and 10 mg kg⁻¹ xylazine intraperitoneal. Eye lubricant was applied, hair around the abdominal and inguinal fat pads were trimmed and the skin was sterilized. With the aid of magnifying glass, a 1 mm diameter incision was made at the site of mammary gland fat pad below was located and 100 μl of a suspension of either
mM NaCl, 0.5% Triton X-100, 1× Complete (Roche). Total RNA was prepared were washed three times with a buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 100 mM KCl, 2 mM Dithiothreitol (DTT), 5% glycerol, 0.5% NP-40, yeast RNA (1 μg), and heparin (1 μg). The labeled RNA was transcribed using Sp6 polymerase from a template generated by inserting into pC12 vector 46–47. A C12 product corresponding to nucleotides from 276 to 407 of murine BCO306700.

RNA isolation from cytoplasm, nucleoplasm, and chromatin. We followed the protocol recently published by Corey and coworkers 48 starting from 10^7 cells. Both cytoplasmic and nucleoplasmic RNAs were precipitated and washed with ice-cold 70% (vol/vol) ethanol prior to be dissolved in QIAzol Lysis Reagent (Qiagen) while the total RNA were immediately precipitated in QIAzol. Controls of 0.5 M EDTA was added to all the samples in QAIOAl that were heated to 65 °C with vortexing until dissolved (~10 min). The preparation of RNA was continued as described below. In parallel to RNA, protein extracts were prepared as described by Corey and coworkers 49.

qRT-PCR, analysis of nascent transcripts, and mRNA decay. Total RNA was isolated using either the miRNeasy mini kit or QIAzol (Qiagen) and retro-transcribed (50–100 ng) using Transcriptase Reverse Transcriptase (Roche) and random hexamers in most experiments according to the manufacturers’ instructions. In order to verify if EPR is polyadenylated, qPCR amplification was performed using as template the product of reverse transcription reactions performed with oligo-dT (that pairs with the poly-A tail). Quantitative PCR was performed using the Precision 2× QPCR master mix (Primer Design), and the Realplex II Mastercycler (Eppendorf) according to the manufacturers’ instructions. The sequence-specific primers utilized for PCR reactions are listed in Supplementary Data 3. In order to analyze gene expression changes among the pool of nascent miRNA, we adopted the Click-it Nascent RNA Capture kit (ThermoFischer) and performed experiments according to the manufacturer’s instructions. NMuMG cells were pulsed with 0.5 mM 5-ethynyl Uridine (EU) for 1 h. In order to analyze mRNA decay we either blocked transcription by treating cells with 100 μM 5,6-Dichlorobenzimidazole-β-D-riboribonuclease (DRB, Sigma-Aldrich) and isolating total RNA, we adopted the Click-IT Nascent RNA Capture kit (ThermoFischer) and performed experiments according to the manufacturer’s instructions. NMuMG cells were weighted and resuspended in 1.0 ml of complete Lysis Buffer (50 mM Tris-Cl, pH 7.0, 1 mM EDTA, 15% formamide, 1× Complete, 1000 U RNase Inhibitor) and hybridized with 1 μl (100 pmol) of either OVEN or ODD pools of 20-mer 3′ Bio-TAGE DNA oligonucleotides designed with single-molecule FISH online designer (Stellaris) (see Supplementary Data 3), respectively. Hybridization was carried out at 37 °C for 4 h under continuous shaking. Seventy microliters of prewashed C-1 magnetic beads (ThermoFisher) were added to each hybridization mixture for 30 min at 37 °C under continuous shaking. beads (power was immobilized and washed four times for 5 min at 37 °C with shaking (wash buffer: 2× NaCl and Sodium citrate (SSC), 0.5% SDS, 1× Complete). While one aliquot (10% of the material) was utilized for RNA extraction, the remaining 90% was subject to DNA purification by incubating two times each bead pellet with 150 μl Complete DNA Elution Buffer (50 mM NaHCO3, 1% SDS, 25 μg ml⁻¹ RNaseA, 100 U ml⁻¹ RNaseH) for 30 min at 37 °C with shaking. Eluted DNA was incubated with Proteinase K (1 mg ml⁻¹ final dilution) for 45 min at 50 °C with shaking, extracted with Phenol/Chloroform/Isomylalcohol, and ethanol-purified, and aliquots were analyzed by qPCR.

ChIP-qPCR. ChIP experiments were performed according to the protocol published by Ghisletti et al. 39 Briefly, ChIP lysates were generated from 4 × 10^6 cells. Each lysate was immunoprecipitated with 10 μg of anti-Pol II, anti-H3K27me3, and anti-SMAD3 antibodies (and the corresponding control IgG).

Antibodies were prebound overnight to 100 μl of A/G protein-coupled paramagnetic beads (ThermoFisher) in PBS/BSA 0.5%. Beads were then added to lysates (the preclearing step was omitted), and incubation was allowed to proceed overnight. Beads were washed six times in a modified RIPA buffer (50 mM HEPES (pH 7.6), 500 mM LiCl, 1 mM EDTA, 1% NP-40, and 0.7% Na-deoxycholate) and once in TE containing 50 mM NaCl. DNA was eluted in TE containing 2% SDS and crosslinks reversed by incubation overnight at 65 °C. DNA was then purified by Qiaquick columns (Qiagen) and quantified with PicoGreen (ThermoFisher).

Sucrose-gradient fractionation and polysome profiling. Experiments were performed as described 48. NMuMG cells (~70% confluence) were treated with cycloheximide (0.1 mg ml⁻¹) for 5 min at 37 °C, washed twice with PBS supplemented with 0.01 mg ml⁻¹ cycloheximide, scraped in PBS 1× with 0.01 mg ml⁻¹ cycloheximide, pelletized by centrifugation, lysed in 500 μl of ice-cold Lysis Buffer (50 mM Tris-Cl, 1% Triton X-100, 1% Na-deoxycholate, 0.2 μl ml⁻¹ RNase Inhibitor, 1 mM DTT, 0.01 mg ml⁻¹ cycloheximide), centrifuged for 5 min at 16,000 × g at 4 °C, and supernatants were loaded onto sucrose gradients. One milliliter fractions were collected monitoring the absorbance at 260 nm using a Density Gradient Fractionation System by Teledyne ISCO with sensitivity set to 0.2. Using the profile of the 260 nm absorbance, fractions corresponding to free ribosomal subunits (40S and 60S) and monosomes (80S, considered as not translating), separately from fractions corresponding to light polysomes (2–5 ribosomes) and
heavy polysomes (>6 ribosomes) were pooled together and processed for RNA extraction and RNA was quantified by Nanodrop (ThermoFisher).

**Cell cycle analyses by flow cytometry.** NMuMG cells (either mock or EPR- or EPRSTOPE-transfected) were seeded in six-well plates. For the analysis by the Cyclets™ Plus DNA Kit (BD Medical Technology), cells are detached by trypsinization and centrifuged in Eppendorf tubes at 300 × g for 5 min at room temperature. Supernatant is removed and the pellet is resuspended in 1 ml 1× PBS followed by centrifugation. Cells are then resuspended in PBS and counted using Countess® Automated Cell Counter and the cell concentration is adjusted to 7 × 10^6 cells ml^{-1} using the same buffer. The DNA staining procedure is performed using 0.5 ml of cell suspension (7 × 10^6 cells). Cells are pelleted by centrifugation (400 × g for 5 min at RT). After carefully removing the supernatant, cells are mixed in Solution A (provided by the kit, containing trypsin in a spermine tetra-hydrochloride buffer for digestion of cell membranes and cytoskeleton), without using a vortex. Two hundred microliters of solution B (provided by the kit, containing Propidium Iodide and spermine tetrahydrochloride in citrate-stabilizing buffer). The sample is incubated in the dark on ice for 10 min and then filtered by cell strainer caps and analyzed by flow cytometer (BD FACS Canto™). Data on at least 10,000 events for sample were processed using ModFit LT 4.2 software. The experiment was repeated twice.

To estimate more precisely the fraction of cells in S phase, the Click-iTTM Plus EdU Flow Cytometry Assay (Invitrogen) was used. EdU (10 mM stock in DMSO) was added directly to the culture medium at the 20 μM final concentration and incubated for 40 min. Cells were then harvested by trypsinization and washed using 3 ml PBS containing 1% BSA. Pellets were resuspended in PBS+1% BSA, counted using Countess® Automated Cell Counter and 1.5 × 10^6 cells are transferred to flow tubes, washed again with 3 ml of PBS containing 1% BSA, pelleted by centrifugation followed by removal of the supernatant. Cells are resuspended in 100 μl of Click-iTTM fixative mixing well with a pipette and incubated for 15 min at RT in the dark. Cells are then washed as performed in the previous step, resuspended and incubated for 10 min in 100 μl of 1× Click-iTTM saponin-based reagent. Samples are then processed for the Click-iTTM reaction, preparing the Click-iTTM Plus reaction cocktail according to the manufacturer’s guidelines, adding 0.5 ml of it to each sample, to reach a final volume of 600 μl containing 1.5 × 10^6 cells and incubating for 30 min at room temperature, in the dark. Cells are then washed once more with 1× Click-iTTM saponin-based reagent, washed and resuspended in 600 μl of the same solution to which the propidium iodide staining solution is added to stain DNA. Propidium iodide solution contains 50 μg ml^{-1} PI and 100 μg ml^{-1} RNase. Samples are then analyzed by flow cytometer (counting 20,000 events, BD FACS Canto™). As controls, cell aliquots incubated with EdU and processed by the same protocol, but skipping the Click-iTTM reaction or the PI staining, or both.

**QM phase cell sorting.** Cells (mock, EPR, EPRSTOPE) were harvested by trypsinization when they reached ~90% confluence, washed once in 1× PBS and resuspended in DMEM without serum at a concentration of 10^5 cells ml^{-1}. Hoechst 33342 (ThermoFisher Scientific) was added to the media at the concentration of 10 μg ml^{-1} and cells were incubated for 1 h at 37 °C. Cells were then centrifuged to remove Hoechst-containing media and resuspended in 1× PBS. Sorting was performed by BD Air A3™ cytometer (BD Bioscience) using a 100 μm nozzle and setting a gate on the population of cells in G1. At least 90 K events for every sample were sorted in 1× PBS at room temperature. After sorting, purity was assessed by re-running the samples. Sorted cells were pelleted and immediately stored at −80 °C. RNA was extracted and analyzed by qRT-PCR as described above.

**Cell proliferation analysis by high-content image analysis.** The proliferation of NMuMG cells (either mock or EPR- or EPRSTOPE-transfected) was quantified using Click-iTTM High Content Imaging System, acquiring images at different time points by digital phase contrast with a ×40 objective. Images were analyzed using Harmony® High Content Imaging and Analysis Software. Five hundred cells were seeded in 96-well plates in triplicates. Pictures were taken at different time points, by automatically acquiring eight fields for each well. Data were analyzed in Excel and plotted as average and standard deviations of replicates.

**Quantification of cell proliferation by crystal violet.** For some experiments cell proliferation was assessed by crystal violet staining. At the indicated time after plating, cells were fixed (10% formalin) and stained (0.1% crystal violet) with crystal violet solution. After two washes with water, crystal violet staining was measured by spectrophotometer at a wavelength of 590 nm.

**Clonogenic and anchor-dependent cell growth essays.** For the clonogenic assays, cells were plated in triplicate on six-well plates at 500 cells per well and left to grow for 4–6 days. Cells were fixed and stained with crystal violet solution. Anchor-dependent cell growth assays were assessed according to the protocol published by Borowicz et al. with minor modifications. Briefly, 2500 cells were seeded in 0.3% agar in complete medium and placed on a layer of 0.5% of bottom agar in 12-well plates. Each cell line was seeded in sextuplicates and fed every 3 days. After 21 days cells were colored with crystal violet and photographs were taken.

**Quantification and statistical analysis of RNA-Seq.** Raw FASTQ reads were trimmed at the ends to remove low-quality reads with FASTX (http://hannonlab. dbrc.uh.edu/fastx_toolkit). Paired-end reads were aligned to indexed mm10 genome with STAR (v. 2.3.0e-r291).

To quantify expression levels mapped reads were counted from BAM files with HTSeq counts version 1.2.1, in intersection-strict mode, feature type exon and id attribute gene_name against reference annotation Ensembl GRCh38.74.

**Quantification of transcript differential expression analysis.** In addition to gene-level analysis with STAR-HTSeq, the transcript abundance was further estimated using an alignment-free approach based on Kallisto 0.43.1 software, using Gencode Mouse vM15 transcripts as reference.

Abundance files were imported in R 3.3.1 with TxlImport.1.2.0 with option txOut = TRUE to quantify alternatively spliced transcripts. edgeR_3.16.5 and limma_3.30.13 were used to log2-transform transcripts count in Count Per Million (cpm). Only transcript with | log2 cpm | in at least three samples were retained. Cpm were transformed by library size and normalized by mean variance with limma-voom. Statistics and log-ratio were calculated with limma-ebayes, by fitting data to a single-factor linear-model with three different levels (mock, EPR, EPRSTOPE).

**Venn diagram and box plots.** We kept differentially expressed transcripts when the observed Bayesian statistic was significant (Benjamini and Hochberg corrected p value < 0.01; logFC > 0.5 |). The functions limma-vennDiagram and pheatmap were used to cluster and visualize the significant genes. For box plots, summary statistics and plots were calculated and rendered with R software (version 3.5.0, https://www.R-project.org) through package ggpubr (https://CRAN.R-project.org/package=ggpubr) and ggplot2 (https://ggplot2.tidyverse.org). Data distributions and normality have been evaluated using the Shapiro–WilK and Mann–Whitney tests for unpaired nonparametric data.

**Gene ontology and pathway enrichment.** Significant transcripts were summarized at gene level (tximport-summaryToGene), annotated by Gene Ontology and enriched by statistically over-represented term with the Enrichr web-application using a nominal p value (Student’s t test) threshold of p < 0.01. The Enrichr p values refer to the Fisher Exact Test statistics, which is a proportion test that assumes a binomial distribution and independence for probability of any gene belonging to any set.

**Protein alignment.** Multiple alignment of mammalian EPR sequences was conducted by using the ClustalW2 package (https://www.ebi.ac.uk).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** Raw data from RNA deep-sequencing analyses have been published on the GEO archive under the accession GSE131178. Human EPR expression in different subpopulations of FACS-sorted normal breast cells and in different human organs was inferred through either the Expression Atlas (https://www.ebi.ac.uk/gxa/home) or the GEOPIA web server (http://geozioma.cancer-pku.cn). Proteins interacting with EPRp were unambiguously identified by searching a comprehensive nonredundant protein database of the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) and the Mass Spectrometry protein sequence Database (MSDB, http://msdb.msdemon.com/en-us/library/ms187112.aspx).

Received: 19 May 2018 Accepted: 27 March 2019
Published online: 29 April 2019

**References.**

1. Jandura, A. & Krause, H. M. The new RNA world: growing evidence for long noncoding RNA functionality. Trends Genet. 33, 665–676 (2017).
2. Kopp, F. & Mendell, J. T. Functional classification and experimental dissection of long noncoding RNAs. Cell 172, 393–407 (2018).
3. Schmitt, A. M. & Chang, H. Y. Long noncoding RNAs in cancer pathways. Cancer Cell 29, 452–463 (2016).
4. Arun, G., Diermeier, S. D. & Spector, D. L. Therapeutic targeting of long noncoding RNAs in cancer. Trends Mol. Med. 24, 257–277 (2018).
5. Long, Y., Wang, X., Youmans, D. T. & Cech, T. R. How do IncRNAs regulate transcription? Sci. Adv. 3, eaao2110 (2017).
6. Gupta, R. A. et al. LncRNA HOAT1 reprograms chromatin state to promote cancer metastasis. Nature 464, 1071–1076 (2010).
7. Yoon, J. H. et al. LincRNA-p21 suppresses target mRNA translation. Mol. Cell 47, 648–655 (2012).
8. Cao, L., Zhang, P., Li, J. & Wu, M. LAST, a c-MyC-inducible long noncoding RNA complexes with CNTP to promote CCND1 mRNA stability in human cells. Elife 6, e30433 (2017).
9. Geisler, S. & Coller, J. mRNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat. Rev. Mol. Cell. Biol. 14, 699–712 (2013).
10. Yuan, J. H. et al. A long noncoding RNA activated by TGF-β promotes the invasion-metastasis cascade in hepatocellular carcinoma. Cancer Cell 25, 666–681 (2014).
11. Dhamija, S. & Diederichs, S. From junk to master regulators of invasion: lncRNAs in migration, EMT and metastasis. Int. J. Cancer 139, 269–280 (2016).
12. Shi, Y. & Massague, J. Mechanisms of TGF-β signaling from cell membrane to the nucleus. Cell 113, 685–700 (2003).
13. Budi, E. H., Duan, D. & Derynck, R. Transforming growth factor-β receptors and Smad: regulatory complexity and functional versatility. Trends Cell Biol. 7, 658–672 (2017).
14. Briata, P. et al. Diverse roles of the nucleic acid-binding protein KHSRP in cell differentiation and disease. Wiley Interdiscip. Rev. RNA 7, 227–240 (2016).
15. Puppo, M. et al. mRNA-mediated KHSRP silencing rewire distinct post-transcriptional programs during TGF-β-induced epithelial-to-mesenchymal transition. Cell Rep. 16, 967–978 (2016).
16. Amman, J. et al. SB-431542 is a potent and specific inhibitor of transforming growth factor-β superfamily type I activator-agonist-like kinase (ALK) receptors ALK4, ALK5, and ALK7. Mol. Pharmacol. 62, 65–74 (2002).
17. Hill, C. S. Transcriptional control by the SMADs. Cold Spring Harb. Perspect. Biol. 8, 402079 (2016).
18. Shirakihara, T., Saitoh, M. & Miyazono, K. Differential regulation of epithelial to mesenchymal transition induced by TGF-β. Mol. Biol. Cell 18, 3533–3544 (2007).
19. Pellacani, D. et al. Analysis of normal human mammary epigenomes reveals cell-specific active enhancer states and associated transcription factor networks. Cell Rep. 17, 2060–2074 (2016).
20. Matsumoto, A. et al. (2017). mTORC1 and muscle regeneration are regulated by CBP/p300 and HDAC1/2. Mol. Cell 63, 927–938 (2016).
21. Bogani, M., Derynck, R. & Miyazono, K. TGF-β- and the TGF-β family: context-dependent roles in cell and tissue physiology. Cold Spring Harb. Perspect. Biol. 8, 402183 (2016).
22. Anderson, D. M. et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell 160, 595–606 (2015).
23. Nelson, B. R. et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science 351, 271–275 (2016).
24. Neveu, E. et al. Long non-coding RNA Linc-RAM enhances myogenic differentiation by interacting with MysD. Nat. Commun. 8, 14016 (2017).
25. Cardoso, D. et al. Tissue-specific expression of a subset of microRNAs. Nat. Commun. 13, 674–690 (2012).
26. Sharma, S. et al. Acetylation-dependent control of global Poly(A) RNA degradation by CBP/p300 and HDAC1/2. Mol. Cell 63, 927–938 (2016).
27. Bjerre, B. & Moses, H. L. Tumour microenvironment: TGFbeta the molecular Jekyll and y Frankenstein. Cancer. Nat. Rev. Cancer 6, 506–520 (2006).
28. Morikawa, M., Derynck, R. & Miyazono, K. TGF-β and the TGF-β family: context-dependent roles in cell and tissue physiology. Cold Spring Harb. Perspect. Biol. 8, 402183 (2016).
29. Briata, P. et al. p38-dependent phosphorylation of the mRNA decay-promoting factor KSRP controls the stability of select myogenic transcripts. Mol. Cell 20, 891–903 (2005).
30. Gupta, R. A. et al. Nuclear and YB-1 are required for JNK-mediated interleukin-2 mRNA stabilization during T-cell activation. Genes Dev. 14, 2136–2148 (2006).
31. Viskov, A. F., Liu, J., Janowsky, B. A. & Corey, D. R. Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonauta loading. Nat. Protoc. 9, 2045–2060 (2014).
32. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal. Chem. 68, 850–858 (1996).
33. Wolf, R. B. et al. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature 459, 1010–1014 (2009).
34. Briata, P. et al. p38-dependent phosphorylation of the mRNA decay-promoting factor KSRP controls the stability of select myogenic transcripts. Mol. Cell 20, 891–903 (2005).
35. Chen, C. Y. et al. Nucleolin and YB-1 are required for JNK-mediated interleukin-2 mRNA stabilization during T-cell activation. Genes Dev. 14, 2136–2148 (2006).
36. Nishizuka, A., Kaler, A., Kolker, E. & Aebi, R. A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75, 4646–4658 (2003).
37. Cui, R., Zhao, C. J., Chang, H. Y. Chromatin isolation by RNA purification (CHIRP). J. Vis. Exp. 25, 3912 (2012).
38. Ghielletti, P. A. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity 32, 317–328 (2010).
39. Zaccara, S. et al. p53-directed translational control can shape and expand the universe of p53 target genes. Cell Death Differ. 21, 1522–1534 (2014).
40. Borowicz, S. et al. The soft agar colony formation assay. J. Vis. Exp. 92, e51998 (2014).

Acknowledgements

We thank Pier Lorenzo Puri (SBP Medical Discovery Institute, La Jolla, CA) for discussions, Angela Cattaneo and Angela Bach for discussions related to the MS analysis of EPRp interactions performed at the Proteinic Unit of IFOM (Milano), Sandro Poggi, Francesco Alessandrini, Federico Villa and Aldo Pagan (IRCCS San Martino) for discussions, Angela Cattaneo and Angela Bachi for discussions related to the MS analysis of EPRp interactions performed at the Proteinic Unit of IFOM (Milano), Sandro Poggi, Francesco Alessandrini, Federico Villa and Aldo Pagan (IRCCS San Martino) for sharing reagents and expertise, Claudia Bagni (University of Leuven), Paola Menichini and Gilberto Fonza (IRCCS San Martino) for sharing cells and reagents. This project has been supported, in part, by grants from the Associazione Italiana per la Ricerca sul Cancro (AIRC I.G. grant 10090 and 21541) to R.G., the Italian Ministry of Health with 5 × 1000 funds 2013 to P.B., and the Italian Ministry of Health with 5 × 1000 funds 2014 and 2015 to R.G. The work of G.B. was supported by the Italian Ministry of Health with 5 × 1000 funds 2014 and 2015 to R.G. We wish to thank Michael Pancher of the High Throughput Screening facility, CIBIO, for assistance with high-content protein assay and Isabella Pesce of the Cell Analysis & Separation facility for assistance with the cell cycle analyses. We are grateful to people at TIB Mol.Biol. (Genova branch) for their professional assistance.

Author contributions

P.B. and R.G. conceived the study and performed some experiments. M.R. performed most experiments, G.B. performed the bioinformatics analysis of the RNA-Seq data. D.B. performed polysome profiling analyses of cell cycle distribution, and high-content image analysis. M.J.M. analyzed hEPR expression in human samples. D.B. performed...
bioinformatics analyses related to EPR peptide. A.F. and D.S. performed immunofluorescence experiments. M.P. produced the recombinant peptide EPRp and some recombinant plasmids. I.E. and M.C. performed the orthotopic transplants in mice. F.N., S. C. and A.I. analyzed and discussed the data obtained in their laboratories with P.B. and R.G. P.B. and R.G. wrote the manuscript. A.I. discussed the revision strategy with P.B. and R.G.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-09754-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: Nature Communications thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019