

# A coding-independent function of gene and pseudogene mRNAs regulates tumour biology

Laura Poliseno<sup>1\*†</sup>, Leonardo Salmena<sup>1\*</sup>, Jiangwen Zhang<sup>2</sup>, Brett Carver<sup>3</sup>, William J. Haveman<sup>1</sup> & Pier Paolo Pandolfi<sup>1</sup>

The canonical role of messenger RNA (mRNA) is to deliver protein-coding information to sites of protein synthesis. However, given that microRNAs bind to RNAs, we hypothesized that RNAs could possess a regulatory role that relies on their ability to compete for microRNA binding, independently of their protein-coding function. As a model for the protein-coding-independent role of RNAs, we describe the functional relationship between the mRNAs produced by the *PTEN* tumour suppressor gene and its pseudogene *PTENP1* and the critical consequences of this interaction. We find that *PTENP1* is biologically active as it can regulate cellular levels of *PTEN* and exert a growth-suppressive role. We also show that the *PTENP1* locus is selectively lost in human cancer. We extended our analysis to other cancer-related genes that possess pseudogenes, such as oncogenic *KRAS*. We also demonstrate that the transcripts of protein-coding genes such as *PTEN* are biologically active. These findings attribute a novel biological role to expressed pseudogenes, as they can regulate coding gene expression, and reveal a non-coding function for mRNAs.

In human cancers, monoallelic mutation of *PTEN* without loss or mutation of the second allele is prevalent at presentation, whereas complete loss is observed at low frequencies with the exception of advanced cancers<sup>1</sup>. In mouse models, heterozygosity for *Pten* leads to multiple cancers<sup>2</sup>, and serial reduction of *Pten* dosage has critical consequences for the incidence and severity of epithelial cancers<sup>3,4</sup>, together indicating that *PTEN* is a functionally haploinsufficient tumour suppressor gene. The identification and validation of numerous *PTEN*-targeting microRNAs (miRNAs) demonstrates that post-transcriptional regulation has a pivotal role in determining *PTEN* abundance in cancer cells<sup>5–11</sup>. Cells are ultrasensitive to even subtle decreases in *PTEN* abundance, thus highlighting the importance of miRNA-mediated *PTEN* regulation in cancer<sup>4</sup>. Therefore, we reasoned that the relationship between *PTEN* and its pseudogene *PTENP1* (also called *PTH2* or  $\psi$ *PTEN*)<sup>12</sup> could represent a compelling test for our hypothesis (Fig. 1a).

Pseudogenes are defined as genomic loci that resemble real genes, yet are considered to be biologically inconsequential because they harbour premature stop codons, deletions/insertions and frameshift mutations that abrogate their translation into functional proteins. Nevertheless, nucleotide sequences contained within pseudogenes are well preserved, suggesting that selective pressure to maintain these genetic elements exists, and that they may indeed have an important cellular role.

Pseudogenes exist as either processed or non-processed genetic elements. Although non-processed pseudogenes arose from genetic duplications, processed pseudogenes were generated through retrotransposition; thus, they contain no introns yet they commonly share 5' and 3' untranslated region (UTR) sequences with their ancestral genes<sup>13</sup>. Pseudogenes are almost as numerous as coding genes and represent a significant proportion of the 'transcriptome'<sup>14</sup>. Despite

lacking canonical promoters, processed pseudogenes use proximal regulatory elements to mediate their transcription<sup>15</sup>. Pseudogene transcription exhibits tissue-specificity<sup>16</sup> and is aberrantly activated in cancer<sup>17</sup>, indicating that pseudogenes may contribute to carcinogenesis, although the mechanisms still remain elusive. Very few pseudogenes have been functionally characterized thus far<sup>13</sup>.

miRNAs, a large class of small non-coding RNAs, have emerged as a critical element in cellular biology and pathophysiology, and have been demonstrated to have an impact on almost all cellular processes and cell types from plants to humans<sup>18</sup>. miRNAs function by annealing to complementary sites on coding sequences or 3' UTRs of target gene transcripts, where they promote the recruitment of protein complexes that impair translation and/or decrease the stability of mRNA, leading to a decrease in target protein abundance<sup>18–21</sup>. Physiologically, aberrant expression of miRNAs has been causally linked to human diseases and cancer<sup>22</sup>.

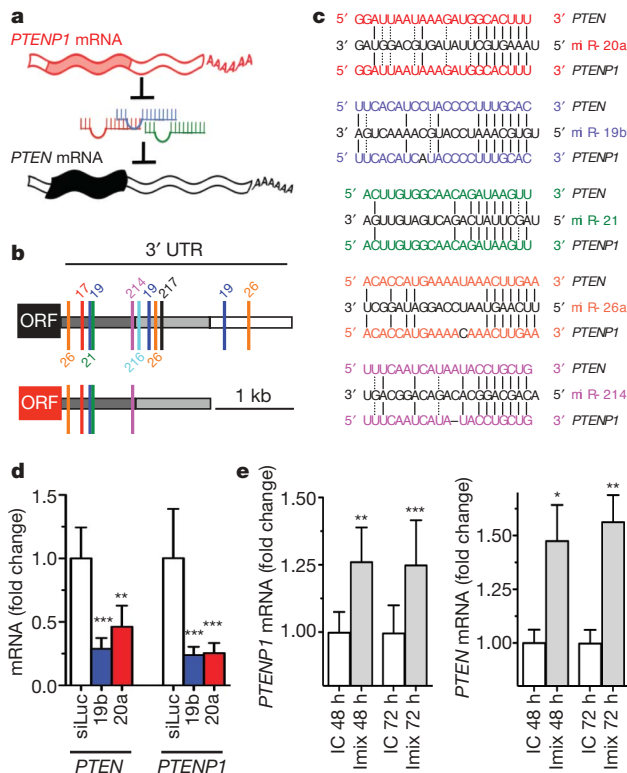
We have tested whether pseudogene-derived RNA transcripts and mRNA transcripts possess an active biological role in cancer that is independent of their protein-coding function but would rely upon their ability to compete for miRNA binding, thereby modulating the derepression of miRNA targets (Fig. 1a).

## *PTENP1* is targeted by *PTEN*-targeting miRNAs

*PTENP1* is a processed pseudogene located at 9p13.3; it is highly homologous to *PTEN*, with only 18 mismatches throughout the coding sequence. A missense mutation of the initiator methionine codon prevents translation<sup>12</sup>. *PTENP1* possesses a 3' UTR that is ~1 kilobase shorter than that of *PTEN* (Fig. 1b). It can be divided into two regions relative to its homology with the *PTEN* 3' UTR: a high homology (~95%) 5' region and a low homology (<50%) 3' region (Fig. 1b and Supplementary Fig. 1). Within the high homology

<sup>1</sup>Cancer Genetics Program, Beth Israel Deaconess Cancer Center, Departments of Medicine and Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA. <sup>2</sup>FAS Research Computing & FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>3</sup>Human Oncology and Pathogenesis Program, Department of Surgery, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA. †Present address: Department of Dermatology, New York University Medical Center, New York, New York 10016, USA.

\*These authors contributed equally to this work.



**Figure 1** *PTENP1* is targeted by *PTEN*-targeting miRNAs. **a**, Working hypothesis: *PTEN* is protected from miRNA binding by *PTENP1*. miRNAs are indicated by red, blue and green structures. 5' and 3' UTRs, open rectangles; open reading frames, filled rectangles. **b**, *PTEN* (top) and *PTENP1* (bottom) 3' UTRs contain a highly conserved (dark grey) followed by a poorly conserved (light grey) region. *PTEN*-targeting miRNA seed matches within the high homology region are conserved between *PTEN* and *PTENP1*. **c**, Binding of *PTEN*-targeting miRNAs to *PTENP1*. Seeds and seed matches, bold; canonical pairings, solid lines; non-canonical pairings (G•U), dotted lines. **d**, *PTEN*-targeting miR-19b and miR-20a decrease *PTEN* and *PTENP1* mRNA abundance. **e**, miR-17 and miR-19 family inhibitors (Imix) derepress *PTENP1* abundance (left). *PTEN* is used as positive control (right). IC, miRNA inhibitor negative control. **d**, **e**, mean  $\pm$  s.d.,  $n \geq 3$ . \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ .

region, we found perfectly conserved seed matches for the *PTEN*-targeting miR-17, miR-21, miR-214, miR-19 and miR-26 families (Fig. 1c and Supplementary Fig. 1). To measure the function of these miRNAs on both *PTEN* and *PTENP1* expression, we designed specific PCR primer sets in the non-homologous 3' UTR regions (Supplementary Fig. 2a, b). In DU145 prostate cancer cells, *PTEN*-targeting miRNAs miR-19b and miR-20a suppressed both *PTEN* and *PTENP1* mRNA abundance (Fig. 1d and Supplementary Fig. 3a). In these cells, a pool of inhibitors of endogenously expressed *PTEN*-targeting miRNAs (Supplementary Fig. 3b) derepressed both *PTEN* and *PTENP1* transcript levels (Fig. 1e). The use of chimaeric luciferase plasmids indicated that the miRNA–*PTENP1* interaction was direct (Supplementary Fig. 4a–c). These data indicate that *PTENP1* and *PTEN* are subjected to the same miRNA-mediated, post-transcriptional regulation.

### The 3' UTR of *PTENP1* has tumour suppressive activity

We examined the ability of *PTENP1* 3' UTR to function as a decoy of *PTEN*-targeting miRNAs using a retroviral vector expressing this 3' UTR (Supplementary Fig. 5a). The 3' UTR can be transcribed, but it cannot code for protein; however, it still may exert a biological role. Indeed, *PTENP1* 3' UTR overexpression resulted in a derepression of both *PTEN* transcript and protein (Fig. 2a and Supplementary Figs 5b and 10c). Consistent with elevated *PTEN*, AKT phosphorylation was reduced upon stimulation of cells with EGF (Fig. 2b). These molecular

observations were accompanied by growth inhibition (Fig. 2c and Supplementary Figs 5c and 10d) and a significant reduction in the number of colonies generated in semisolid medium (Fig. 2d).

The derepression of *PTEN* abundance by *PTENP1* 3' UTR overexpression was blunted in HCT116 *DICER*<sup>-/-</sup> colon carcinoma cells (Fig. 2e). In these cells, the disruption of *DICER*—the enzyme that catalyses the last step of miRNA maturation—leads to reduced levels of mature miRNAs compared to parental HCT116 cells<sup>23</sup>. This in turn supports the notion that the 3' UTR of *PTENP1* requires mature miRNAs for its function towards *PTEN*.

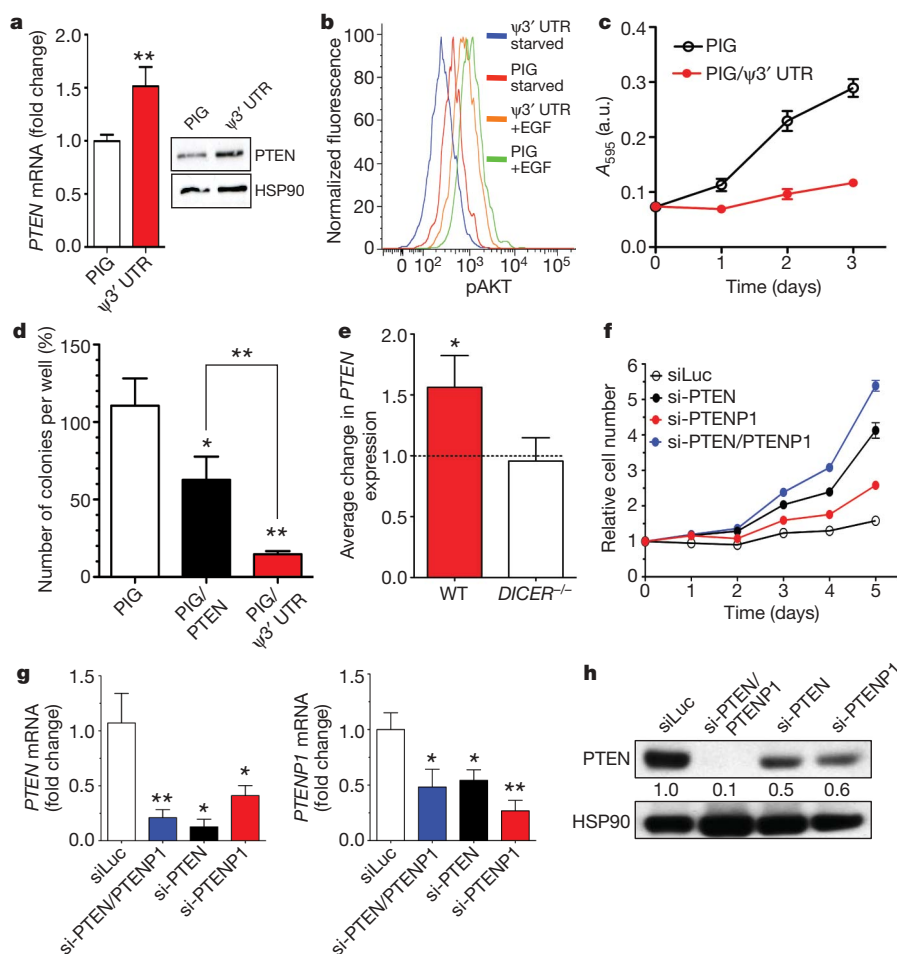
To examine the phenotypic consequences of *PTENP1* downregulation, we designed custom small interfering RNA (siRNA) pools (Dharmacon) to target specifically either *PTENP1* (si-*PTENP1*) or *PTEN* (si-*PTEN*) expression (Supplementary Fig. 6) because commercially available siRNA pools for *PTEN* (si-*PTEN*/*PTENP1*) bind to common sequences in *PTEN* and *PTENP1* (Supplementary Fig. 7a). si-*PTENP1* transfection accelerated cell proliferation, indicating that *PTENP1*, although expressed at lower relative levels, can exert a biological activity in DU145 cells (Fig. 2f). si-*PTEN*/*PTENP1*, which silences both *PTEN* and *PTENP1*, showed the strongest effect, indicating that *PTEN* and its pseudogene may have additive roles for growth suppression. *PTENP1* knockdown resulted in decreased *PTEN* mRNA and protein abundance (Fig. 2g, h), mirroring the results obtained with overexpression of *PTENP1* 3' UTR (Fig. 2a).

In DU145 cells *PTENP1* 3' UTR is a more potent growth suppressor compared to *PTEN* (Fig. 2d and Supplementary Fig. 5c). This result may be explained by the fact that miRNAs for which *PTENP1* functions as a decoy also bind other targets with tumour suppressive activities. For instance, the miR-17 family targets E2F1 and p21<sup>24</sup>, and miR-21 targets PDCD4<sup>25</sup>. Accordingly, miR-17 and miR-21 mimics increase proliferation of *PTEN*-null PC3 cells (Supplementary Fig. 8a), indicating *PTEN* independency. Indeed, si-*PTENP1* resulted in a dose-dependent downregulation not only of *PTEN* but also of p21 (Supplementary Fig. 8b). Additionally, both si-*PTENP1* and si-*PTEN*/*PTENP1* were able to suppress *PTENP1* and increase proliferation in a dose-dependent manner in *PTEN*-null PC3 cells (Supplementary Fig. 8c, d). Conversely, stable infection of *PTENP1* 3' UTR in PC3 cells suppressed foci formation (Supplementary Fig. 8e), supporting the notion that *PTENP1* and its 3' UTR exert a tumour suppressive role that goes beyond the regulation of *PTEN* abundance alone.

### Expression and losses of *PTENP1* in human cancer

*PTEN* and *PTENP1* expression was explored in normal human tissues and prostate tumour samples, using custom Taqman probes (see Methods and Supplementary Fig. 2c). In both the normal tissue and prostate tumour samples, the direct correlation ( $r = 0.8087$ ,  $P < 0.0001$  and  $r = 0.7538$ ,  $P < 0.0001$ , respectively) between *PTEN* and *PTENP1* expression suggests that they may be co-regulated (Fig. 3a, b). This finding supports our molecular observations that *PTENP1* can regulate *PTEN* expression. *PTENP1* was found to be variably abundant, and in some cases expressed at higher levels than *PTEN*.

Next, we examined alterations of the *PTENP1* genomic locus. Several array-based databases were mined including The Cancer Workbench (<https://cgwb.nci.nih.gov/cgi-bin/heatmap>) and NCBI GEO (<http://www.ncbi.nlm.nih.gov/geo/>) (Supplementary Fig. 9a, b and Supplementary Table 1). Remarkably, in a data set of sporadic colon cancer (<http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE16125>) (Fig. 3c–e), hierarchical clustering identified a clear population of samples with detectable copy number losses occurring specifically at the *PTENP1* locus (Fig. 3c). Notably, these copy number losses were focal, not associated with large losses of 9p, and independent of losses at the *CDKN2A* locus (Supplementary Fig. 9c). This data set formally demonstrates the existence of independent genomic copy number losses at the *PTENP1* locus, supporting the notion that *PTENP1* exerts tumour suppressive functions and is under selective pressure to undergo copy number losses in cancer.



**Figure 2** | *PTEN1* 3' UTR exerts a tumour suppressive function by acting as a decoy for *PTEN*-targeting miRNAs. **a–c**, PIG/ $\psi$ 3' UTR-infected DU145 cells show increased *PTEN* mRNA and protein levels (**a**) reduced phospho-AKT levels upon EGF stimulation (**b**) and decreased proliferation rate (**c**). **d**, Growth in semisolid medium of DU145 cells infected with PIG, PIG/ $\psi$ 3' UTR or PIG/*PTEN*. **e**, *PTEN* mRNA levels 24 h after the transfection of pCMV/ $\psi$ 3' UTR in parental HCT116 or HCT116 *DICER*<sup>-/-</sup> cells. Data are

normalized using pCMV empty-transfected cells. **f**, Growth curve of DU145 cells transfected with control siLuc, si-*PTEN*/*PTENP1*, si-*PTEN* or si-*PTENP1*. **g**, mRNA levels of *PTEN* (left) and *PTENP1* (right) 24 h after the transfection of siLuc (white), si-*PTEN*/*PTENP1* (blue), si-*PTEN* (black) and si-*PTENP1* (red). **h**, Western blot of *PTEN* 48 h after the transfection of the indicated siRNAs. Quantification of *PTEN* protein is reported. Data in **a**, **c–g** indicate mean  $\pm$  s.d.,  $n \geq 3$ . \* $P < 0.05$ ; \*\* $P < 0.01$ .

In the same patient sample set, cluster analysis of *PTEN* expression showed that it was downregulated compared to normal colon samples ( $P = 0.0008156$ ; Fig. 3d). Regression analysis of *PTENP1* copy number variation with the expression levels of *PTEN* identified two discrete populations of patients in which *PTENP1* copy number variation and *PTEN* expression were directly and significantly correlated (population 1:  $r = 0.6105$ ,  $P = 0.0092$ ; population 2:  $r = 0.6056$ ,  $P = 0.0129$ ) (Fig. 3e). The existence of a direct relationship between *PTENP1* copy number and *PTEN* expression supports our hypothesis that *PTENP1* transcript levels can regulate *PTEN* expression. Together, these findings constitute a proof of principle for the onc-suppressive activity of *PTENP1*.

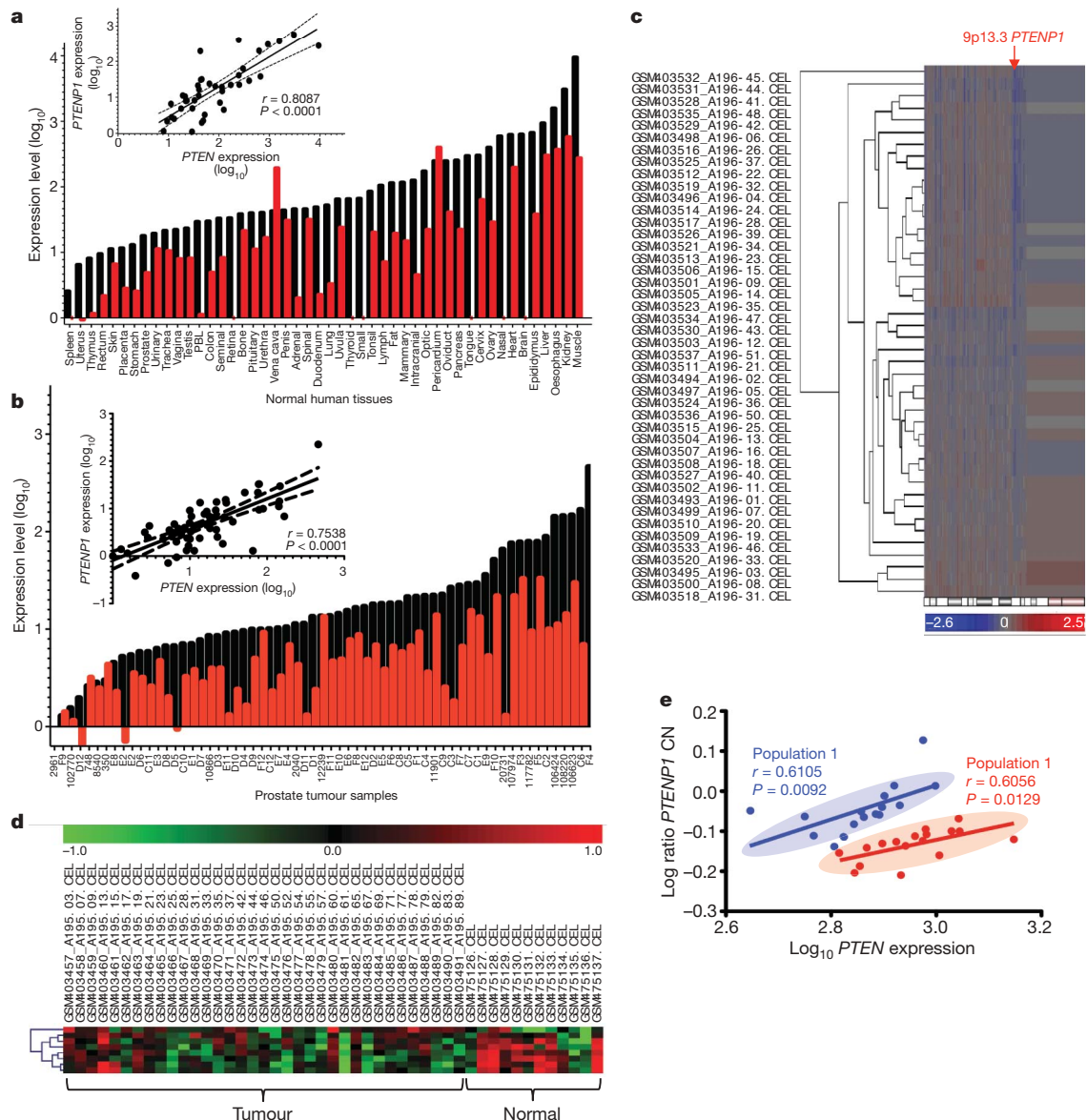
### A general model for endogenous mRNA-mediated biology

On the basis of our results, we expected that the *PTEN* 3' UTR would also have biological activity. Indeed, we found that *PTEN* 3' UTR can derepress *PTENP1* abundance, as *PTENP1* does on *PTEN* (Fig. 4a, left and Supplementary Fig. 10a–c). Importantly, *PTEN* 3' UTR overexpression was accompanied by growth inhibition, suggesting that *PTEN* exerts its tumour suppressive activity, at least in part, through its 3' UTR (Fig. 4a, right and Supplementary Fig. 10d).

To extend our studies beyond *PTEN* and its pseudogene, we examined other cancer-related pseudogenes and genes (Supplementary Tables 2 and 3 and Supplementary Figs 11–17). Alignments of gene and pseudogene sequences show that miRNA-binding sites are

well conserved: for example, the miR-145 binding site on *OCT4* (also called *POU5F1*) and its pseudogenes *OCT4-pg1*, *OCT4-pg3*, *OCT4-pg4* and *OCT4-pg5* (also called *POU5F1P1*, *POU5F1P3*, *POU5F1P4* and *POU5F1P5*, respectively; Supplementary Fig. 11a); miR-1 family binding sites on connexin 43 (*CX43*, also called *GJA1*) and its pseudogene (Supplementary Fig. 11b). Notably, *OCT4-pg1* and *OCT4-pg5* are exclusively expressed in cancer tissues and not in normal tissues<sup>17</sup>. Furthermore *OCT4-pg5* is truncated at the 5' end and expresses only a partial open reading frame region followed by the 3' UTR<sup>26</sup>. Further examples of such conservation include: miR-34 family binding site on *CDK4PS* (Supplementary Fig. 12); miR-182 binding site on *FOXO3B* (Supplementary Fig. 13); miR-17 family binding site on *E2F3P1* (Supplementary Fig. 14); and miR-143 and let-7 family binding sites on *KRASIP* (Supplementary Fig. 15).

Because the 3' UTR of *PTENP1* was growth suppressive like its parental gene *PTEN*, we hypothesized a similar relationship between *KRAS* and its pseudogene *KRASIP*. Indeed, *KRASIP* 3' UTR overexpression in DU145 cells resulted in increased *KRAS* mRNA abundance (Fig. 4b, left and Supplementary Fig. 18a, b) and accelerated cell growth (Fig. 4b, right). We also found that the *KRAS* and *KRASIP* transcript levels are positively correlated in prostate cancer (Supplementary Fig. 18c). Notably, the *KRASIP* locus at 6p11-12 is amplified in different human tumours, including neuroblastoma, retinoblastoma and hepatocellular carcinoma<sup>27–29</sup>. Together these findings point to a putative proto-oncogenic role for *KRASIP*, and



**Figure 3** | Loss of *PTENP1* in cancer. **a**, **b**, Expression level of *PTEN* (black) and *PTENP1* (red) in a panel of normal human tissues (**a**) and prostate tumour samples (**b**). Prostate sample expression values derived from both commercial and clinical sources were independently normalized and combined for correlation analysis. Linear regression of *PTEN* versus *PTENP1* expression is shown in the upper left corner. **c**, Cluster analysis of 48 sporadic colon cancer samples interrogated by Affymetrix Human SNP

support the notion that pseudogene functions mirror the functions of their cognate genes as explained by a miRNA decoy mechanism.

## Discussion

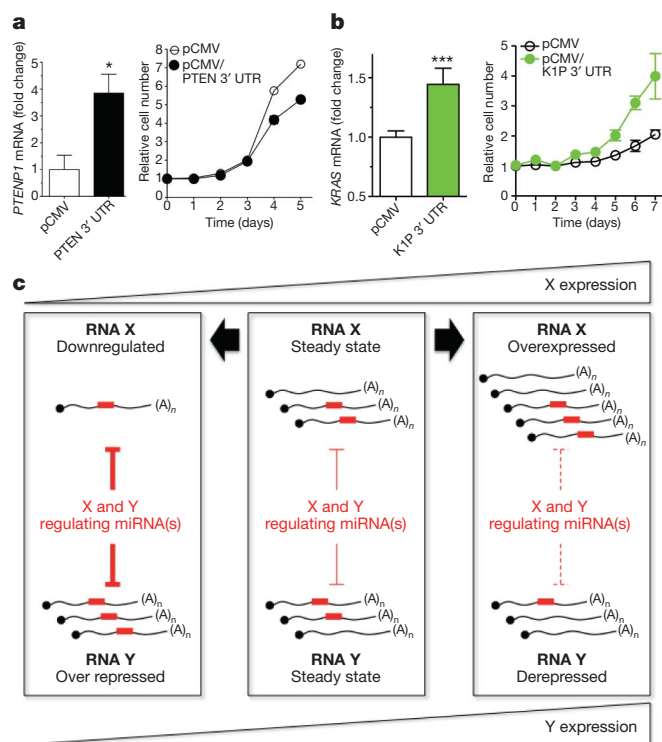
The findings presented in this study have allowed us to reach a number of conclusions. First, the discovery of an miRNA decoy function for pseudogenes identifies these transcripts as biologically active units. We show that *PTENP1* and *KRASIP* affect the levels of their cognate genes and are possibly involved in disease pathogenesis. Thus, the analysis of pseudogene expression level and genomic status in tumorigenesis needs to be undertaken systematically to further our understanding of disease progression.

Processed pseudogenes in mouse oocytes have been previously reported to generate endogenous siRNAs (endo-siRNAs) that down-regulate the expression of cognate genes through conventional RNA interference<sup>30</sup>. However, endo-siRNA production has yet to be identified in somatic human cells<sup>31</sup>. Notably, although only few

Array. **d**, Heat map and cluster analysis of Affymetrix Human Exon 1.0 ST Array for normalized *PTEN* intensity values. **e**, Plot of log ratio of *PTENP1* copy number (CN) against log<sub>10</sub> *PTEN* expression intensity. Lines of best fit represent regression analyses of two populations. The correlation coefficient ( $r$ ) measures the reliability and the  $P$ -value measures the statistical significance of the correlation between the  $x$  and  $y$  variables.

pseudogenes undergo antisense transcription, all transcribed pseudogenes can in principle compete with cognate genes for miRNA binding. Similarly, the miRNA decoy capacity of pseudogenes is likely to be more widespread than their cleavage into Piwi-interacting RNAs (piRNAs), which have been recently discovered only in germ cells of many organisms, including mouse<sup>32</sup>.

We also demonstrate that pseudogenes such as *PTENP1* can depress their cognate genes, even when expressed at lower levels (Supplementary Fig. 3a and Fig. 2f–h). We propose that pseudogenes are ‘perfect decoys’ for their ancestral genes, because they retain many of the miRNA binding sites and can compete for the binding of many miRNAs at once. It has been hypothesized that suboptimal ‘pseudo-targets’ may compete with authentic targets for miRNA binding<sup>33</sup>. By contrast, we propose that pseudogenes have an intrinsic biological activity in miRNA networks because they are legitimate miRNA targets and compete with other legitimate targets for miRNA binding (Fig. 2e). This notion is corroborated by the ‘target mimicry’ process



**Figure 4 | *PTEN* 3' UTR and *KRAS1P* 3' UTR function as decoys and a general model for endogenous miRNA decoy mechanism. **a**, *PTEN1* mRNA level 24 h after the transfection of the empty pCMV or pCMV/*PTEN*3' UTR plasmid in DU145 cells (left) and growth curve (right). **b**, *KRAS* mRNA level 24 h after the transfection of the empty pCMV or pCMV/*KIP3'* UTR plasmid in DU145 cells (left) and growth curve (right). **c**, Model: X and Y are different transcripts targeted by the same miRNA(s). In the steady state (middle), equilibrium exists between the miRNA molecules and their targets X and Y. Downregulation of X (left) leads to increased availability of miRNA molecules to bind to Y, thus decreasing its abundance. By contrast, overexpression of X (right) leads to fewer miRNA molecules free to bind to Y, and thus Y abundance increases. Red rectangles, miRNA molecules. X and Y can be a pseudogene and its cognate protein-coding gene. Data in **a** and **b** indicate mean  $\pm$  s.d.,  $n \geq 3$ . \* $P < 0.05$ ; \*\*\* $P < 0.001$ .**

that in plants is achieved by the expression of non-protein coding genes that sequester miRNAs<sup>34</sup>.

Exogenously administered miRNA sponges have recently emerged as effective and specific inhibitors of miRNAs<sup>35,36</sup>. Pseudogenes act like 'endogenous competitors', able to affect the distribution of miRNA molecules on all their targets. They may be particularly effective precisely because they are non-coding, thus active translation does not interfere with miRNA binding<sup>37</sup>.

The ability of pseudogenes to regulate the biology of a cell goes beyond their ability to modulate the levels of their cognate gene (Supplementary Fig. 8). This phenomenon is consistent with the fact that each miRNA has multiple targets and can lead to widespread homeostatic effects. Also, given that a single gene often has numerous differentially regulated pseudogenes (for example, *OCT4*, *NPM1* (Supplementary Figs 11a and 17) and ribosomal protein pseudogenes<sup>38</sup>), such networks can become intricately dynamic.

Cellular miRNA abundance is dictated by total genomic copy number and by their biogenesis process<sup>39</sup>. Less is known about the regulation of mature miRNA activity. miRNAs can increase their spectrum of targetable mRNAs by undergoing deamination<sup>40</sup>, whereas shortening of 3' UTRs<sup>41</sup> and polymorphisms can prevent miRNA binding to mRNAs<sup>42</sup>. Pseudogene-mediated miRNA decoys offer a new dimension regulating the cross-talk between miRNAs and their targets. Indeed, the greater the number of pseudogenes that a protein-coding gene has, the more it is protected from miRNAs.

Our discovery of a functional role for *PTEN1* is relevant to *PTEN* biology as minute changes in *PTEN* can have tumorigenic consequences<sup>3</sup>. In our analysis we found that *PTEN1* positively regulates *PTEN* levels. Furthermore, we found that the *PTEN1* locus undergoes copy number losses in human cancer and this correlates with a decrease in *PTEN*; thus we propose that *PTEN1* is a bona fide tumour suppressor gene. In light of this, better tools must be developed to detect pseudogene abundance in cancer. For instance, pseudogenes including *PTEN1* have been overlooked to such an extent that pseudogene-specific probes are absent in some microarray platforms (Supplementary Fig. 7b).

An important implication of our findings is that the decoy mechanism may not be limited to pseudogenes, but may include other long non-coding RNA transcripts including ribosomal RNAs, large intergenic noncoding RNAs (lincRNAs) and coding gene mRNAs<sup>38,43</sup> (Fig. 4a and Supplementary Fig. 10). Beyond their function as *cis* regulatory elements that impact the stability of their own transcripts, UTRs are also *trans* modulators of gene expression through miRNA binding. Furthermore, as binding sites for miRNAs are also located in open reading frame sequences<sup>20</sup>, the entire transcript of coding genes, and not only the 3' UTR, may possess decoy function (see working model, Fig. 4c).

As our model suggests, mRNA introduced into a cell can potentially perturb the interaction between miRNAs and their multiple targets and thus have a biological activity independent of the translation of the protein they encode<sup>44</sup>. Notably, the same applies to the transcriptional induction or repression of endogenous mRNA levels, which can lead to changes in the number of mRNA molecules present within a cell in the scale of several orders of magnitude<sup>45</sup>.

Our findings indicate that when studying specific nonsense or frameshift mutations and genomic alterations leading to 'read-through' or fusion transcripts one must consider this new RNA-regulated biological dimension. For example, chromosomal fusion events such as the t(15;17) translocation of APL which generates *PML-RAR $\alpha$*  and *RAR $\alpha$ -PML* fusion transcripts or recurrent read-through transcripts in melanoma such as *CDK2-RAB5B* could exert oncogenic activities through aberrant competition for microRNAs<sup>46,47</sup>. This phenomenon could also occur as a consequence of somatic genomic rearrangements, which are emerging as grossly unappreciated events in many cancers<sup>48</sup>. Moreover, the shortening of 3' UTRs as observed in human cancer cells<sup>41</sup> would not solely affect miRNA-dependent regulation but also alter the 'competing' capacity of a given RNA transcript. Finally, in the case of *PTEN*-loss-associated cancers, little is known of the molecular consequences of *PTEN* mutations where the *PTEN* transcript is retained, compared to complete genetic loss of *PTEN* where no transcript remains<sup>1</sup>. Although these events were previously thought to alter protein abundance, protein signalling and protein networks, they will also have a significant impact on cellular RNA and miRNA homeostasis. We have therefore identified a novel dimension by which cellular and tumour biology can be regulated.

## METHODS SUMMARY

Cell lines were cultured under standard conditions. miRNA overexpression was obtained by transient transfection (si-miRNAs). *PTEN1/PTEN/KRAS1P* 3' UTR overexpression was achieved by transient transfection (pCMV expression vectors) or stable infection with MSCV-PIG retroviral constructs. miRNA-target interaction was measured by luciferase reporter assay. *PTEN1*, *PTEN*, *KRAS*, *KRAS1P* and miRNA expression level was detected by real-time PCR. Proliferation, foci and soft agar assays were performed according to standard protocols.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 21 September 2009; accepted 22 April 2010.

- Salmena, L., Carracedo, A. & Pandolfi, P. P. Tenets of PTEN tumor suppression. *Cell* **133**, 403–414 (2008).
- Di Cristofano, A. *et al.* Impaired Fas response and autoimmunity in *Pten*<sup>+/-</sup> mice. *Science* **285**, 2122–2125 (1999).

3. Trotman, L. C. *et al.* Pten dose dictates cancer progression in the prostate. *PLoS Biol.* **1**, e59 (2003).
4. Alimonti, A. *et al.* Subtle variations in Pten dose determine cancer susceptibility. *Nature Genet.* (2010).
5. Xiao, C. *et al.* Lymphoproliferative disease and autoimmunity in mice with increased miR-17-92 expression in lymphocytes. *Nature Immunol.* **9**, 405–414 (2008).
6. Takakura, S. *et al.* Oncogenic role of miR-17-92 cluster in anaplastic thyroid cancer cells. *Cancer Sci.* **99**, 1147–1154 (2008).
7. Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
8. Meng, F. *et al.* Involvement of human micro-RNA in growth and response to chemotherapy in human cholangiocarcinoma cell lines. *Gastroenterology* **130**, 2113–2129 (2006).
9. Huse, J. T. *et al.* The PTEN-regulating microRNA miR-26a is amplified in high-grade glioma and facilitates gliomagenesis *in vivo*. *Genes Dev.* **23**, 1327–1337 (2009).
10. Kato, M. *et al.* TGF- $\beta$  activates Akt kinase through a microRNA-dependent amplifying circuit targeting PTEN. *Nature Cell Biol.* **11**, 881–889 (2009).
11. Yang, H. *et al.* MicroRNA expression profiling in human ovarian cancer: miR-214 induces cell survival and cisplatin resistance by targeting PTEN. *Cancer Res.* **68**, 425–433 (2008).
12. Fujii, G. H., Morimoto, A. M., Berson, A. E. & Bolen, J. B. Transcriptional analysis of the PTEN/MMAC1 pseudogene,  $\Psi$ PTEN. *Oncogene* **18**, 1765–1769 (1999).
13. D'Errico, I., Gadaleta, G. & Saccone, C. Pseudogenes in metazoa: origin and features. *Brief. Funct. Genomics Proteomics* **3**, 157–167 (2004).
14. Harrison, P. M., Zheng, D., Zhang, Z., Carriero, N. & Gerstein, M. Transcribed processed pseudogenes in the human genome: an intermediate form of expressed retrosequence lacking protein-coding ability. *Nucleic Acids Res.* **33**, 2374–2383 (2005).
15. Birney, E. *et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007).
16. Bristow, J., Gitelman, S. E., Tee, M. K., Staels, B. & Miller, W. L. Abundant adrenal-specific transcription of the human P450c21A “pseudogene”. *J. Biol. Chem.* **268**, 12919–12924 (1993).
17. Suo, G. *et al.* Oct4 pseudogenes are transcribed in cancers. *Biochem. Biophys. Res. Commun.* **337**, 1047–1051 (2005).
18. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
19. Baek, D. *et al.* The impact of microRNAs on protein output. *Nature* **455**, 64–71 (2008).
20. Tay, Y., Zhang, J., Thomson, A. M., Lim, B. & Rigoutsos, I. MicroRNAs to *Nanog*, *Oct4* and *Sox2* coding regions modulate embryonic stem cell differentiation. *Nature* **455**, 1124–1128 (2008).
21. Lal, A. *et al.* miR-24 inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to “seedless” 3'UTR microRNA recognition elements. *Mol. Cell* **35**, 610–625 (2009).
22. Ventura, A. & Jacks, T. MicroRNAs and cancer: short RNAs go a long way. *Cell* **136**, 586–591 (2009).
23. Cummins, J. M. *et al.* The colorectal microRNAome. *Proc. Natl Acad. Sci. USA* **103**, 3687–3692 (2006).
24. Petrocca, F. *et al.* E2F1-regulated microRNAs impair TGF $\beta$ -dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer Cell* **13**, 272–286 (2008).
25. Lu, Z. *et al.* MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene. *Oncogene* **27**, 4373–4379 (2008).
26. Pain, D., Chirn, G. W., Strassel, C. & Kemp, D. M. Multiple retrosequences from pluripotent cell-specific gene expression indicates a potential signature for novel gene identification. *J. Biol. Chem.* **280**, 6265–6268 (2005).
27. van der Wal, J. E. *et al.* Comparative genomic hybridisation divides retinoblastomas into a high and a low level chromosomal instability group. *J. Clin. Pathol.* **56**, 26–30 (2003).
28. Zimonjic, D. B., Keck, C. L., Thorgerisson, S. S. & Popescu, N. C. Novel recurrent genetic imbalances in human hepatocellular carcinoma cell lines identified by comparative genomic hybridization. *Hepatology* **29**, 1208–1214 (1999).
29. Plantaz, D. *et al.* Gain of chromosome 17 is the most frequent abnormality detected in neuroblastoma by comparative genomic hybridization. *Am. J. Pathol.* **150**, 81–89 (1997).
30. Tam, O. H. *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**, 534–538 (2008).
31. Okamura, K., Chung, W. J. & Lai, E. C. The long and short of inverted repeat genes in animals: microRNAs, mirtrons and hairpin RNAs. *Cell Cycle* **7**, 2840–2845 (2008).
32. Robine, N. *et al.* A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **19**, 2066–2076 (2009).
33. Seitz, H. Redefining microRNA targets. *Curr. Biol.* **19**, 870–873 (2009).
34. Franco-Zorrilla, J. M. *et al.* Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature Genet.* **39**, 1033–1037 (2007).
35. Ebert, M. S., Neilson, J. R. & Sharp, P. A. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nature Methods* **4**, 721–726 (2007).
36. Lee, D. Y. *et al.* A 3'-untranslated region (3'UTR) induces organ adhesion by regulating miR-199a\* functions. *PLoS ONE* **4**, e4527 (2009).
37. Gu, S., Jin, L., Zhang, F., Sarnow, P. & Kay, M. A. Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nature Struct. Mol. Biol.* **16**, 144–150 (2009).
38. Balasubramanian, S. *et al.* Comparative analysis of processed ribosomal protein pseudogenes in four mammalian genomes. *Genome Biol.* **10**, R2 (2009).
39. Winter, J., Jung, S., Keller, S., Gregory, R. I. & Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nature Cell Biol.* **11**, 228–234 (2009).
40. Kawahara, Y. *et al.* Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **315**, 1137–1140 (2007).
41. Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673–684 (2009).
42. Kim, J. & Bartel, D. P. Allelic imbalance sequencing reveals that single-nucleotide polymorphisms frequently alter microRNA-directed repression. *Nature Biotechnol.* **27**, 472–477 (2009).
43. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
44. Frieth, M. C. *et al.* Pseudo-messenger RNA: phantoms of the transcriptome. *PLoS Genet.* **2**, e23 (2006).
45. Jiang, S. L., Lozanski, G., Samols, D. & Kushner, I. Induction of human serum amyloid A in Hep 3B cells by IL-6 and IL-1 beta involves both transcriptional and post-transcriptional mechanisms. *J. Immunol.* **154**, 825–831 (1995).
46. Scaglioni, P. P. & Pandolfi, P. P. The theory of APL revisited. *Curr. Top. Microbiol. Immunol.* **313**, 85–100 (2007).
47. Berger, M. F. *et al.* Integrative analysis of the melanoma transcriptome. *Genome Res.* **20**, 413–427 (2010).
48. Stephens, P. J. *et al.* Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462**, 1005–1010 (2009).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank Pandolfi laboratory members for critical discussions, in particular A. Carracedo for critical input; S. Feng for technical assistance. We thank B. Vogelstein for *DICER*<sup>-/-</sup> cells; T. Yuan for assistance with FACS analysis; A. Tuccoli for assistance with miRNA RT-PCR; I. Osman for support and suggestions. L.P. was supported by fellowships from the Istituto Toscano Tumori and the American Italian Cancer Foundation. L.S. was supported by fellowships from the Human Frontier Science Program and the Canadian Institutes of Health Research. This work was supported by NIH grant R01 CA-82328-09 to P.P.P.

**Author Contributions** P.P.P. spearheaded and supervised the project; L.P., L.S. and P.P.P. designed experiments; L.P., L.S. and W.J.H. performed experiments; B.C. provided prostate cancer patient sample cDNAs. J.Z. performed all bioinformatic analyses. L.P., L.S. and P.P.P. analysed the data and wrote the paper. All authors critically discussed the results and the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to P.P.P. ([ppandolf@bidmc.harvard.edu](mailto:ppandolf@bidmc.harvard.edu)).

## METHODS

**Reagents.** Anti-HSP90 antibody 61041 (Becton Dickinson); anti-PTEN antibody 9559, anti-p21 antibody 2947, anti-Tubulin antibody 2125 (Cell Signaling); siGENOME non-targeting siRNA 2 (siLuc), si-PTEN, si-PTENP1, si-PTEN/PTENP1, siGLO RISC-free control siRNA, si-miR-17, si-miR-19b, si-miR-20a, si-miR-21, si-miR-26b, si-miR-214, miRNA inhibitor negative control 1 (IC), miR-19b inhibitor, miR-93 inhibitor, miR-106b inhibitor, Dharmafect 1 (Dharmacon); lipofectamine 2000, Trizol reagent, DNaseI amplification grade, SuperScript II reverse transcriptase, Dulbecco's Modified Eagle Medium (DMEM), RPMI-1640, fetal bovine serum (FBS) (Invitrogen); tissue scan normal tissue qPCR arrays: human major tissue (HMRT103); tissue scan disease tissue qPCR arrays: prostate cancer II (HPRT102) (Origen); pGL3-control, pRL-TK, dual-luciferase reporter assay (Promega); polybrene, puromycin (Sigma); QuantiTect Sybr green PCR kit, Effectene (Qiagen); EGF (R&D); QuikChange II XL site-directed mutagenesis kit, Herculase Taq polymerase (Stratagene).

**Plasmids.** The 3' UTR of *PTENP1* (NR\_023917.1) was amplified by PCR from the genomic DNA of PC3 cells and cloned into the BamHI and XhoI sites of pCMV-MCS expression plasmid. In this way, pCMV/ $\psi$ 3' UTR plasmid was obtained. The primers used for PCR amplification were: forward 5'-GAGGAG CCGTCAAATCCAGAG-3' and reverse 5'-TCGTCAATGTGTGAGGTTCC-3'. The 3' UTR of *PTENP1* was then subcloned into the BglII and XhoI sites of MSCV-PIG retroviral vector<sup>49</sup> to obtain PIG/ $\psi$ 3' UTR plasmid.

The 3' UTR of *PTEN* (NM\_000314.4) was amplified by PCR from the genomic DNA of HeLa cells and cloned into the BamHI and XhoI sites of pCMV-MCS expression plasmid. In this way, pCMV/PTEN3' UTR plasmid was obtained. The primers used for PCR amplification were: forward 5'-TAGAGGA GCCGCAAATCCA-3' and reverse 5'-TGGACATCTGATGGGATGA-3'.

The 3' UTR of *KRASIP* (NC\_000006.11) was amplified by PCR from the genomic DNA of HeLa cells and cloned into the BamHI and XhoI sites of pCMV-MCS expression plasmid. In this way, pCMV/KIP3' UTR plasmid was obtained. The primers used for PCR amplification were: forward 5'-AACC AGCAAAGACAGGGTGT-3' and reverse 5'-GTTCAATTGCTCAACGCA GA-3'. The homology between wild-type *KRAS* and its pseudogene is very high (>90%) across the whole mRNA sequence. The primers used for the amplification of *KRASIP* 3' UTR contain many mismatches which made them specific for the pseudogene.

To construct pGLU/ $\psi$ 3' UTR chimaeric luciferase plasmid, the multicloning site of pGL3-control plasmid was removed from its original position and inserted into the XbaI site located downstream of Luciferase stop codon (pGLU). *PTENP1* 3' UTR was then subcloned from pCMV/ $\psi$ 3' UTR plasmids using the SmaI and XhoI sites of pGLU. The QuikChange II XL site-directed mutagenesis kit was used to generate the mutated version of this plasmid (pGLU/ $\psi$ 3' UTRmut).

**Cells and culture conditions.** Phoenix A, 293T, DU145<sup>50</sup> and PC3 cells were grown in DMEM plus 10% FBS. RWPE-1, PWR-1E and VCaP were grown in keratinocyte medium plus EGF and BPE. 22Rv1, and LnCaP were grown in RPMI 1640 plus 10% FBS. All cell lines were obtained from ATCC and grown in penicillin/streptomycin and glutamine containing medium, at 37 °C in a humidified atmosphere with 6% CO<sub>2</sub>.

**Transient transfection.** For the transfection of siRNAs/si-miRNAs/miRNA inhibitors, DU145 ( $1.5 \times 10^5$ ) or PC3 ( $1 \times 10^5$ ) were seeded in 12-well dishes. The following day they were transfected with 100 nM siRNAs/si-miRNAs or 400 nM miRNA inhibitors using Dharmafect 1 according to the manufacturer's recommendations. With this protocol more than 90% of cells were positive to the fluorescent siGLO RISC-free control siRNA (data not shown).

For plasmid transfection, 293T and DU145 were seeded in 6-cm dishes ( $2.5$  and  $3.5 \times 10^5$ , respectively) and the day after they were transfected with Effectene.

Six hours later, cells were trypsinized and seeded for the various assays.

**Dual luciferase reporter assay.** DU145 cells were seeded at a density of  $6 \times 10^4$  cells per 24-well dish. Twenty-four hours later, 720 ng of pGLU/ $\psi$ 3' UTR or pGLU/ $\psi$ 3' UTRmut were co-transfected with 80 ng of pRL-TK. Lipofectamine 2000 was used as transfectant. Twenty-four hours after transfection, luciferase activity was measured and normalized as in ref. 7.

**Retroviral infection.** Phoenix A cells were plated in 10-cm poly-D-lysine coated dishes ( $3 \times 10^6$  per dish) and 24 h later were transfected with PIG retroviral plasmids using Lipofectamine 2000. Sixteen hours later, the medium was changed and after additional forty-eight hours, the virus-containing medium (10 ml) was filtered, mixed with 5 ml of freshly prepared medium, supplemented with  $4 \mu\text{g ml}^{-1}$  polybrene and added to  $5 \times 10^5$  DU145 or PC3 cells plated in a 10-cm dish the day before. Puromycin ( $2 \mu\text{g ml}^{-1}$ ) was administered 48 h after infection. The cells were selected for 2 days and then used for the various assays. Selection medium was changed every day.

**PCR analysis.** Total RNA was extracted using Trizol reagent according to the manufacturer's instructions. It was then subjected to DNase treatment and retro-

transcription (1  $\mu\text{g}$  RNA per vial). Regular PCR was performed using Herculase Taq Polymerase.

Real-time PCR of *PTEN*, *PTENP1*, *KRAS* and *KRASIP* was carried out using Sybr green fluorescence. Two microlitres of RT were used in a 20- $\mu\text{l}$  reaction. Actin was used as an internal standard. Relative quantification of gene expression was performed with the comparative  $C_T$  method<sup>51</sup>. *PTEN* primers: forward 5'-GTTTACCGGCAGCATCAAAT-3', reverse 5'-CCCCACTTTAGTGACAC AGT-3'. *PTENP1* primers: forward 5'-TCAGAACATGGCATAACACAA-3', reverse 5'-TGATGACGTCCGATTTTCA-3'. *KRAS* primers: forward 5'-ATTG TGAATGTTGGTGT-3', reverse 5'-GAAGGTCTCAACTGAAAT-3'. *KRASIP* primers: forward 5'-AAGTTTCTCCAGTTCT-3', reverse 5'-ATTTGGGA ATTTTGTGAG-3'. Actin primers: forward 5'-CATGTACGTTGCTATCCA GGC-3', reverse 5'-CTCCTTAATGTCACGCACGAT-3'.

The real-time PCR of mature miRNAs was performed according to ref. 52 with some modifications. Briefly, an independent retrotranscription reaction was set up for each miRNA using 0.05  $\mu\text{M}$  of the miRNA-specific RT primer and 0.05  $\mu\text{M}$  of *SNORD44* RT primer (5'-GTCGTATCCAGTGCAGGG TCCGAGGTATTCCGACTGGATACGACagtcag-3'). Real-time PCR of both the miRNA and *SNORD44*, which was used as an internal standard, were then carried out using Sybr green fluorescence (2  $\mu\text{l}$  of RT in a 20  $\mu\text{l}$  reaction). For the miRNA, a specific forward primer and the universal reverse primer (5'-GTG CAGGGTCCGAGGT-3') were used. For *SNORD44*, 5'-CGGCGTggcgatga ggaggtacc-3' forward primer and the universal reverse primer were used. The miRNA-specific RT and forward PCR primers are listed in Supplementary Fig. 19. The portion of the RT and forward PCR primers that recognizes the miRNAs or *SNORD44* is in lower case. The real-time PCR reaction comprised 40 cycles of 95 °C for 15 s followed by 60 °C for 1 min. Relative quantification of gene expression was performed with the comparative  $C_T$  method as described above.

TaqMan RT-PCR was performed at the HMS Biopolymers Facility using an Applied Biosystems 7900 HT Fast instrument.

**Western blot.** Cells were collected and lysed (50 mM Tris pH 8.0, 1 mM EDTA, 1 mM MgCl<sub>2</sub>, 150 mM NaCl, 1% NP-40, 1 mM  $\beta$ -glycerophosphate, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 1 mM NaF, protease inhibitors). Proteins (30  $\mu\text{g}$  per lane) were separated on 10% SDS-polyacrylamide gel and transferred to nitrocellulose membrane. Immunoblotting of the membranes was performed using the following primary antibodies: anti-PTEN (1:1,000), anti-p21 (1:1,000), anti-HSP90 (1:1,000), anti-Tubulin (1:2,000). Signals were revealed after incubation with recommended secondary antibody coupled to peroxidase by using enhanced chemiluminescence. Scanned images were quantified using ImageJ software.

**FACS analysis.** After 10-min treatment with 50 ng ml<sup>-1</sup> EGF, cells were scraped from 10-cm dishes, immediately fixed in 4% PFA and permeabilized with ice cold Methanol. After rehydrating with 0.1% BSA in PBS, cells were stained with Phospho-AKT (Thr 308) rabbit monoclonal antibody Alexa Fluor 647 conjugate (Cell Signaling). Cells were analysed on an LSRII flow cytometer (BD).

**Cell proliferation.** At the end of the selection period (infection) or 6 h after transfection,  $2 \times 10^5$  DU145 cells were trypsinized, resuspended in 50 ml and seeded in 8 sets of 3 wells of a 12-well plate. Starting from the following day (d0), one set of wells per day was washed once with PBS, fixed in 10% formalin solution for 10 min at room temperature and then kept in PBS at 4 °C. At day 7, all the wells were stained with crystal violet. After lysis with acetic acid 10%, optical density was read at 590 nm.

**Foci assay.** At the end of the selection period (infection) or 6 h after transfection, DU145 or PC3 cells were trypsinized. A total of  $5 \times 10^5$  cells were plated on 10-cm dishes. Fourteen to twenty-one days later, the plates were stained with crystal violet and the foci were counted.

**Growth in semisolid medium.** The bottom layer was obtained by covering 6-well dishes with 3 ml of 0.6% agar in DMEM. The day after,  $5 \times 10^4$  infected cells were seeded on top in triplicate in 2 ml of 0.3% agar in DMEM + 10% FBS. Colonies were counted after 3–4 weeks at 40 $\times$  magnification.

**Analysis of *PTEN* and *PTENP1* genomic status and expression.** For breast cancer, Affymetrix GeneChip Human Mapping 500K Array data sets GSE7545 and GSE16619 were downloaded from NCBI GEO and analysed with the Partek Genomic Suite (Partek Inc.) for detection of genomic regions with alterations and data visualization. Copy number aberrations were scored with the Partek segmentation algorithm with default parameters: *P*-value cutoff at 0.001 for neighbouring regions with significantly different means, 10 minimum number of probe sets required for any candidate region, 0.3 signal to noise difference as minimum magnitude of change, and *P*-value threshold 0.01 for one-sided *t*-test of probes in each region to be considered as significantly deviated from the expected normal. All aberrations were calculated with respect to a set of 270 HapMap-normal persons. A total of 118 breast cancer samples and 44 normal samples were included in the study.

For colon cancer, GSE16125 sporadic colon cancer raw data sets were downloaded from NCBI GEO (<http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE16125>). There are two chip platforms used in this data set: 48 sporadic colon cancer samples interrogated by Affymetrix GeneChip(r) Human Mapping 250K Nsp SNP Array and 36 of them analysed by Affymetrix Human Exon 1.0 ST Array. The SNP array raw data sets were analysed with the Partek Genomic Suite (Partek Inc.) for detection of genomic regions with alterations and data visualization (Partek smoothing algorithm was based on 46 probes). Forty-eight normal samples from the HapMap project supplied by Affymetrix were used as an un-paired reference set ([http://www.affymetrix.com/support/technical/sample\\_data/500k\\_data.affx](http://www.affymetrix.com/support/technical/sample_data/500k_data.affx)). Raw exon array intensity CEL files for colon cancer and normal colon were analysed by Affymetrix Power Tools (APT, v. 1.12.0). Normalized intensity value for *PTEN* was calculated as an average of eight probes corresponding to two *PTEN*-specific exon probe sets '3256703' and '3256704'. These values were extracted out by APT software. Affymetrix Human Exon 1.0 ST Array data set for normal colon epithelial cells was downloaded from NCBI GEO GSE1916 data set (<http://www.ncbi.nlm.nih.gov/geo/query/acc>.

[cgi?acc=GSE19163](http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE19163)). The correlation plot with  $r$  and  $P$  values between  $\log_{10}$  *PTEN* expression intensity and log ratio of *PTENP1* copy number was generated in GraphPad Prism (GraphPad Software, Inc.). The log ratio of *PTENP1* was based on an average of 14 SNP probes flanking the *PTENP1* gene.

**Statistical analysis.** *In vitro* data were analysed using unpaired  $t$ -test (GraphPad Prism, GraphPad Software, Inc.). Values of  $P < 0.05$  were considered statistically significant. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ . The mean  $\pm$  s.d. of three or more independent experiments is reported. Regression analyses and correlation coefficients were generated using GraphPad Prism, GraphPad Software, Inc.

49. Maeda, T. *et al.* Role of the proto-oncogene *Pokemon* in cellular transformation and ARF repression. *Nature* **433**, 278–285 (2005).
50. Myers, M. P. *et al.* P-TEN, the tumour suppressor from human chromosome 10q23, is a dual-specificity phosphatase. *Proc. Natl Acad. Sci. USA* **94**, 9052–9057 (1997).
51. Drabkin, H. A. *et al.* Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. *Leukemia* **16**, 186–195 (2002).
52. Chen, C. *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* **33**, e179 (2005).