HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2019 August 10.

Published in final edited form as:

Nat Genet. 2018 June ; 50(6): 783–789. doi:10.1038/s41588-018-0118-8.

3′ UTR shortening represses tumor-suppressor genes in trans by disrupting ceRNA crosstalk

Hyun Jung Park[1,2,3,9], **Ping Ji**[4,9], **Soyeon Kim**[5], **Zheng Xia**[1,2], **Benjamin Rodriguez**[1,2], **Lei Li**[1,2], **Jianzhong Su**[1,2], **Kaifu Chen**[1,2], **Chioniso P. Masamha**[6], **David Baillat**[4], **Camila R. Fontes-Garfias**[4], **Ann-Bin Shyu**[7], **Joel R. Neilson**[8], **Eric J. Wagner**[4,10,*], **Wei Li**[1,2,10,*]

[1]Division of Biostatistics, Dan L Duncan Cancer Center, Baylor College of Medicine, Houston, TX, USA.

[2]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, USA.

[3]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA.

[4]Department of Biochemistry & Molecular Biology, University of Texas Medical Branch, Galveston, TX, USA.

[5]Center for Precision Health, School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA.

[6]Department of Pharmaceutical Sciences, College of Pharmacy and Health Sciences, Butler University, Indianapolis, IN, USA.

[7]Department of Biochemistry and Molecular Biology, University of Texas, McGovern Medical School, Houston, TX, USA.

[8]Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston,TX, USA.

[9]These authors contributed equally: Hyun Jung Park, Ping Ji.

[*]**Correspondence and requests for materials** should be addressed to E.J.W. or W.L., ejwagner@utmb.edu; WL1@bcm.edu.

Author contributions

H.J.P. and W.L. conceived the project, designed the experiments and performed the data analysis. S.K. performed the regression analysis. Z.X. performed the APA analysis. L.L., J.S. and K.C. helped with data analysis. C.P.M., E.J.W. and A.-B.S. obtained the miRNA-Seq data. P.J., C.R.F.-G. and D.B. performed the *NUDT21*-knockdown experiments. H.J.P, PJ., E.J.W. and W.L. wrote the manuscript with input from B.R., A.-B.S., C.PM. and J.R.N.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588–018-0118–8.

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

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

Code availability. The open source MAT3UTR program (version 0.9.2) is freely available at https://github.com/thejustpark/MAT3UTR with necessary example data for this analysis.

Data availability. Raw and processed miRNA-Seq data for the *NUDT21*-depletion experiment have been deposited to GEO under the accession number GSE78198.

[10]These authors jointly supervised this work: Eric J. Wagner, Wei Li.

Abstract

Widespread mRNA 3′ UTR shortening through alternative polyadenylation[1] promotes tumor growth in vivo[2]. A prevailing hypothesis is that it induces proto-oncogene expression in cis through escaping microRNA-mediated repression. Here we report a surprising enrichment of 3′UTR shortening among transcripts that are predicted to act as competing-endogenous RNAs (ceRNAs) for tumor-suppressor genes. Our model-based analysis of the trans effect of 3′ UTR shortening (MAT3UTR) reveals a significant role in altering ceRNA expression. MAT3UTR predicts many trans-targets of 3′ UTR shortening, including *PTEN*, a crucial tumor-suppressor gene[3] involved in ceRNA crosstalk[4] with nine 3′UTR-shortening genes, including *EPS15* and *NFIA*. Knockdown of *NUDT21*, a master 3′ UTR-shortening regulator[2], represses tumor-suppressor genes such as *PHF6* and *LARP1* in trans in a miRNA-dependent manner. Together, the results of our analysis suggest a major role of 3′ UTR shortening in repressing tumor-suppressor genes in trans by disrupting ceRNA cross-talk, rather than inducing proto-oncogenes in cis.

Widespread 3′ UTR shortening (3′US) through alternative polyadenylation (APA) occurs during enhanced cellular proliferation and transformation[1,5–8]. Recently, we reported that *NUDT21*-mediated 3′US promotes glioblastoma growth, further underscoring its significance to tumorigenesis[2]. A prevailing hypothesis is that a shortened 3′ UTR results in activation of proto-oncogenes in cis through escaping microRNA (miRNA)-mediated repression. Indeed, several well-characterized oncogenes, such as *CCND1*, have been shown to use 3′US to increase their protein levels, but mostly in cell lines[5]. However, in recent PolyA sequencing[7] and our TCGA RNA sequencing (RNA-Seq) APA analysis[1] (5 and 358 tumor/normal pairs, respectively), most oncogenes with 3′US previously identified in vitro[5] displayed almost no changes in their 3′ UTR lengths in tumors (Fig. 1a). For example, we identified 1,346 recurrent (occurrence rate >20%) 3′US genes in 358 tumor/normal pairs[1]. However, *CCND1* is not on that list as its 3′ US occurred in only a very small portion (8 out of 358; 2.2%) of tumors (Fig. 1b). Furthermore, similar to random genes, 3′US genes from all 5 previous APA studies have little overlap with the top 500 ($P < 0.01$) high-confidence oncogenes as defined on the basis of distinct somatic mutational patterns of >8,200 tumor/normal pairs[9] (Fig. 1c). These results challenge the previous hypothesis and suggest a different function of 3′ US for tumorigenesis.

Aside from regulating its cognate transcript in cis, the 3′ UTR has also been implicated in competing-endogenous RNA (ceRNA) regulation in trans[10]. Although the scope is not fully understood, ceRNA is generally thought to form global regulatory networks (ceRNETs) controlling important biological processes[11]. For example, the tumor suppressor *PTEN's* ceRNAs, *CNOT6L* and *VAPA*, have been shown to regulate *PTEN* and phenocopy its tumor-suppressive properties[12]. As the ceRNA's regulatory axis is mostly based on miRNA-binding sites on 3′ UTRs, we hypothesize that when genes with shortened 3′ UTRs no longer sequester miRNAs, the released miRNAs would then be directed to repress their ceRNA partners, such as tumor-suppressor genes, in trans, thereby contributing to tumorigenesis.

To test this hypothesis, we first used well-established strategies to reconstruct two ceRNETs from 97 TCGA breast tumors and their matched normal tissues, respectively, based on miRNA-binding-site overlap and co-expression[13,14] between genes of active ceRNA regulation (Methods). In general, transcripts are less correlated between each other in tumors than in normal tissues, partially due to tumor heterogeneity[15] and global reduction of miRNA expression in tumors[16] (Fig. 2a). As expected, the loss of co-expression results in a much smaller (tenfold reduced) ceRNET for tumors than for normal tissues (Fig. 2b).

To investigate the role of 3′ US in ceRNET disruption, we focused on estrogen-receptor-positive (ER^+) breast tumors, which comprise the majority (68/97) of TCGA breast tumor samples. We built normal and tumor ceRNETs using the same procedure as above. Using the DaPars algorithm[1], we identified 427 3′ US genes recurring in >20% of tumors. Close inspection indicates that 3′ US is associated with ceRNET disruption. For example, we identified *PTEN* and *EPS15* as a ceRNA pair in normal ceRNET (4 miRNA-binding-site overlap and $\rho = 0.63$ co-expression). However, since *EPS15* underwent 3′ US in 23 (33.8%) out of 68 tumors, thereby losing its capability to compete with *PTEN* for miRNAs, it lost ($\rho = 0.32$) the co-expression (and ceRNA) relationship with *PTEN* in tumors (Fig. 2c). Globally, the top 100 ceRNAs with the most significant 3′US genes all lost their interactions in tumors, while 12 out of 100 ceRNAs lacking 3′ US retained ($P = 0.0002$) their interactions. Furthermore, in separate ceRNETs from 30 tumor/normal pairs with the least and most amount of 3′ US (upper panel in Fig. 2d), more 3′ US is clearly associated with more ceRNET loss (38.6 versus 16.4 in fold decrease, $P < 1 \times 10^{-16}$, lower panel in Fig. 2d). From these findings, we conclude that 3′ US is strongly associated with ceRNA network disruption in tumors.

To understand the function of 3′ US-mediated ceRNET disruption, we selected 381 3′ US genes and 2,131 of their ceRNA partner genes (3′ US ceRNAs), including 591 3′ US ceRNA hub and 1,540 3′ US ceRNA non-hub genes, in the normal ceRNET (Supplementary Table 1, Methods). We hypothesized that 3′ US genes released their miRNAs to repress their ceRNA partners in trans. Consistent with our hypothesis, expression changes of 2,131 3′ US ceRNA genes in tumors are anti-correlated ($r = -0.21$; $P = 5 \times 10^{-24}$) with the degree of 3′ US of the associated 3′ US genes (Supplementary Fig. 1a). As a result, among 976 genes in normal ceRNET downregulated in tumors, 816 (83.6%) are ceRNAs of 3′ US genes. Surprisingly, 3′ US ceRNA hub genes are enriched in tumor-suppressor genes ($P \sim 1 \times 10^{-20}$) but not in oncogenes (Fig. 3a), suggesting that the 3′ US represses tumor suppressors in trans. For example, 3′ US of *EPS15* would contribute to downregulating its ceRNA partner *PTEN* in tumors (Fig. 2c). Globally, 160 expressed tumor-suppressor genes from 3′ US ceRNAs are more likely downregulated than 226 control tumor-suppressor genes not in ceRNET ($P = 8 \times 10^{-3}$, Fig. 3b), indicating a significant association between 3′ US and tumor-suppressor gene repression.

Additional analyses on sequence features partially explain why 3′ US genes, but not tumor suppressors in their ceRNA partners, are likely to have alternative proximal polyadenylation sites, leading to 3′ US (Supplementary Note). We have also analyzed TCGA 450K methylation array data and found that the 3′ US-mediated ceRNA repression is independent of promoter hypermethylation (Supplementary Note).

To better quantify the trans effects of 3′ US, we developed a mathematical model (MAT3UTR) based on its 3′ US gene(s) expression, 3′ US level, miRNA-binding site(s) and miRNA expression(s) (Methods). In 1,548 differentially expressed 3′ US ceRNAs, MAT3UTR can explain 47.6% of variation in gene expression (Supplementary Fig. 3c). In contrast, the MAT3UTR-control model, which considers miRNA expression but not 3′ US, explains only 27.2% of variation (Supplementary Fig. 3d), consistent with previous reports17 that miRNA alone has a weak role in regulating gene expression. The results suggest that the trans effects of 3′ US plays a major role in regulating ceRNA gene expression.

MAT3UTR predicts many trans-target genes of 3′ US, including *PTEN*, in ceRNA crosstalk[11–13] (top 1% MAT3UTR score, Supplementary Table 2). In normal ceRNET, *PTEN* is predicted to be a ceRNA of nine 3′ US genes (Fig. 3c). When we ranked 97 breast tumor/normal pairs by the amount of 3′ US across these nine genes (upper panel in Fig. 3d), tumors with more 3′ US showed more down-regulation of *PTEN* ($P = 0.03$, lower panel in Fig. 3d). Furthermore, MAT3UTR can explain 86.9% of the variation in *PTEN's* expression across tumors (Supplementary Fig. 3g), suggesting that the trans effects of 3′ US play a major role in downregulating *PTEN*

To empirically test the hypothesis that 3′ US can downregulate *PTEN* in trans, we focused on *EPS15* among the nine 3′ US genes(Methods). We observed that depletion of *EPS15* by siRNA in MCF7 cells reduces *PTEN* expression (Fig. 3e). To ascertain whether this effect depends on miRNA-based targeting of the *PTEN* 3′ UTR, we used a luciferase reporter vector with the *PTEN* 3′ UTR (pLight-Switch-PTEN 3′ UTR) to test the effect of *EPS15* knockdown on its expression. We observed that reduction of *EPS15* reduces *PTEN* 3′ UTR luciferase activity (Fig. 3f). To further understand whether the crosstalk is miRNA-dependent, we depleted *DICER1* to abolish miRNA biogenesis and found that loss of *DICER1* can relieve the influence of *EPS15* knockdown on *PTEN* 3′ UTR expression (Fig. 3g). Finally, overexpression of the *EPS15* 3′ UTR increased the number of PTEN-positive cells (Fig. 3h,i). Thus, *EPS15* 3′ US may impact *PTEN* expression.

To gain insights into the global cause-and-effect relationship between 3′ US and the repression of tumor-suppressor genes, we revisited our previous data from *NUDT21*-knockdown HeLa cells, since NUDT21 is one of the master regulators of 3′ US[2]. We identified 1,168 3′ US ceRNAs in *NUDT21*-knockdown cells solely on the basis of significant miRNA-binding-site overlap with 1,450 3′ US genes, since co-expression cannot be effectively estimated from two replicates of our experiments. With 9,914 expressed RefSeq genes with no significant miRNA-binding-site overlap with 3′ US genes as random controls, the tumor-suppressor genes remain strongly enriched in 3′ US ceRNAs ($P \sim 1 \times 10^{-38}$, Fig. 4a). Among 57 tumor-suppressor genes in 3′ US ceRNAs, 33 (57.9%) showed repression in *NUDT21*-knockdown samples; whereas a smaller portion (44.5%) of 339 control tumor-suppressor genes showed repression ($P \sim 0.03$, Fig. 4b), suggesting that *NUDT21*-mediated 3′ US represses tumor-suppressor genes in trans. In spite of potentially higher false positives due to lack of co-expression in ceRNA identification, these results are highly consistent with our observations in TCGA breast cancer. On the basis of these results, we posit that repression of tumor-suppressor ceRNAs would correlate with increased

occupancy of AGO2 in the RISC complex. To formally test this hypothesis, we isolated cytoplasmic fractions from control or *NUDT21*-knockdown cells and conducted RNA immunoprecipitation (RIP) using anti-AGO2 antibodies. On average, we observed ~200-fold enrichment of ceRNAs in Ago2 RIP complexes relative to control IgG (Supplementary Fig. 4b). Reduced expression of *NUDT21* does not impact AGO2/DICER1 expression and *GAPDH* messenger RNA binding to AGO2 (Fig. 4c,d and Supplementary Fig. 4b). Furthermore, we sequenced miRNAs from control and NUDT21-knockdown cells, and found that miRNAs are equally likely to be upregulated or downregulated (Supplementary Fig. 4d), ruling out a general effect on miRNA biogenesis. Importantly, we could detect increased association of multiple tumor-suppressor ceRNAs with AGO2 following *NUDT21* depletion that ranged from 1.5-fold to nearly 7-fold (Fig. 4d). These results demonstrate that 3′ US can lead to reduction of tumor-suppressor genes through their increased association with repressive AGO2 complexes.

To further validate the miRNA-dependent, repressive trans effects of 3′ US, we monitored expression of the tumor-suppressor genes *PHF6* and *LARP1* and their ceRNA partners, *YOD1* and *LAMC1* (Supplementary Table 3). We consistently observed that *PHF6* and *LARP1* expression levels were decreased in *NUDT21*-knockdown cells while both *YOD1* and *LAMC1* expression levels were increased (Fig. 5a). To determine whether the 3′ UTR mediated this effect, we transfected luciferase reporters containing the 3′ UTR of either *PHF6* or *LARP1* into control or *NUDT21*-knockdown cells and measured luciferase activity. We found that both reporters were downregulated after *NUDT21* knockdown (Fig. 5b). Both *PHF6* and *LARP1* have been shown as tumor-suppressor genes[9,18,19] and downregulation of *PHF6* or *LARP1* in HeLa cells increases cell growth, confirming their tumor suppressive activity (Supplementary Fig. 5).

To further investigate the mechanism of tumor-suppressor ceRNA downregulation, we chose *PHF6* on the basis of MAT3UTR analysis and experimental results (Methods). We selected two miRNAs targeting *PhF6* (Fig. 5c), which were released by 3′ US of *YOD1* (miR-3187–3p as the highest and miR-549 as the sixth highest in terms of $\beta_j \mathrm{miR}_j$ in equation (3); Methods and Supplementary Table 4). Neither of these miRNAs was found to change its expression following *NUDT21* knockdown (Supplementary Fig. 4d). However, *PHF6* expression was partially rescued by an antagomir blocking the activity of miR-549 and completely rescued by an antagomir targeting miR-3187–3p (Fig. 5d). Moreover, *PHF6* 3′ UTR-mediated luciferase activity was partially rescued by the miR-3187–3p antagomir or *YOD1* siRNA (Fig. 5e). To understand whether reduced expression of *PHF6* depends on *YOD1* levels, we transfected *YOD1* cDNA into cells depleted of *YOD1* and found that re-expression of *YOD1* could not restore either the expression of endogenous *PHF6* (Fig. 5f) or the expression of the *PHF6* 3′ UTR-mediated luciferase (Fig. 5g), suggesting that the trans effect on *PHF6* is due to the 3′ UTR of *YOD1* Finally, to determine whether the crosstalk between *PHF6* and *YOD1* is miRNA-dependent, we also showed that depletion of *DICER1* abolishes *PHF6* and *YOD1* crosstalk (Fig. 5h). Collectively, the data strongly suggest that *NUDT21*-mediated 3′ US causes tumor-suppressor gene repression in trans in a miRNA-dependent manner.

Although analyzing ceRNA crosstalk in light of 3′ US has been briefly suggested[20–22], our MAT3UTR analysis of 97 breast cancer RNA-Seq data followed by functional validation suggests a wide-spread causal role of 3′ US in repressing tumor-suppressor genes in trans. While the trans effect further emphasizes the importance of APA in tumor progression, it also provides an additional layer of gene regulation and underscores the need for further investigation into other potential mechanisms[23,24] that could per-turb ceRNA crosstalk, such as RNA editing and competition with RNA-binding proteins.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41588–018-0118–8.

Tumor-suppressor genes and oncogenes.

The tumor-suppressor genes and oncogenes used in this study were defined by the TUSON algorithm from genome sequencing of >8,200 tumor/normal pairs[9], namely residue-specific activating mutations for oncogenes and discrete inactivating mutations for tumor-suppressor genes. TUSON is a computational method that analyzes patterns of mutation in tumors and predicts the likelihood that any individual gene functions as a tumor-suppressor gene or oncogene. We ranked genes by their TUSON prediction P values from the most to the least significant and used the top 500 genes ($P < 0.01$) as the reference tumor-suppressor genes or oncogenes. After removing 30 genes in common, 470 tumor-suppressor genes and oncogenes were used for the enrichment analysis. Note that there were very few breast tumor-specific tumor-suppressor genes and oncogenes (36 and 3 with breast q-value ≤ 0.5, respectively) and 90% of them were found in the top 500 pan-cancer predictions.

Previously identified 3′ US genes in cancers.

Xia et al. identified 1,187 3′ US genes across 7 TCGA cancer types[1]. Mayr and Bartel selected 23 3′ US genes from 27 cancer cell lines[5]. Fu et al. identified 428 3′ US genes in human breast cancer cell lines[6]. Lin et al. reported 120 3′ US genes in major cancers and tumor cell lines[7]. Morris et al. found 286 3′ US genes in human colorectal tumor samples[8]. The 3′ US genes of Xia et al. were randomly sampled to 100 genes for a fair comparison.

Selection of miRNA-binding sites.

Predicted miRNA-binding sites were obtained from TargetScanHuman version 6.2[25]. Only those with a preferentially conserved targeting score (Pct) more than 0 were used[1]. Experimentally validated miRNA-binding sites were obtained from TarBase version 5.0[26], miRecords version 4[27] and miRTarBase version 4.5[28]. The binding sites found in indirect studies such as microarray experiments and high-throughput proteomics measurements were filtered out[29]. Another source is the microRNA target atlas composed of public AGO-CLIP data[30] with significant binding sites (q-value <0.05). The predicted and validated binding site information was then combined to use in this study.

TCGA breast tumor RNA-Seq and miRNA-Seq data.

Quantified gene expression files (RNASeqV1) for primary breast tumors (TCGA sample code 01) and their matching solid normal samples (TCGA sample code 11) were downloaded from the TCGA Data Portal[31]. We used 97 breast tumor samples that have matched normal tissues. A total of 10,868 expressed RefSeq genes (fragments per kilobase of transcript per million mapped reads (FPKM) ≥ 1 in >80% of all samples) were selected for downstream analyses. To better quantify gene expression in the presence of 3′ US, we used only coding regions (CDS) to quantify mRNA expression. Exon and CDS annotation for TCGA data and miRNA expressions (syn1445790) were downloaded from Sage Bionetworks' Synapse database.

CeRNA identification in TCGA breast tumors.

CeRNAs were identified by miRNA-binding-site overlap and expression correlation[13,14]. Only microRNAs with intermediate expression (between 0.01 and 100 in averaged fragments per million mapped fragments (FPM)) were used to capture dynamic interactions[14]. After removing genes with fewer than six such miRNA-binding sites, gene pairs with significant miRNA-binding-site overlap (<0.05 in Benjamini–Hochberg-corrected *P* value) were selected. Among them, pairs correlated (>0.6 in Pearson's correlation coefficient) ($P < 1 \times 10^{-10}$) in gene expression were defined as ceRNAs. To account for mRNAs with variable 3′ UTRs, we used only CDS to quantify mRNA expression. Genes that are connected with >500 ceRNAs were defined as hub genes.

Model-based analysis of trans effect of 3′ US (MAT3UTR).

Suppose transcript x has a constitutive proximal 3′ UTR (pUTR) and a distal 3′ UTR that might be shortened in tumors (dUTR) (Supplementary Fig. 3a). We define MiRs(x, miR$_j$) as the amount of binding sites for miRNA miR$_j$ in x.

$$\mathrm{MiRs}\left(x, \mathrm{miR}_j\right) = \left(\mathrm{pUTR}\left(x, \mathrm{miR}_j\right) + \mathrm{dUTR}\left(x, \mathrm{miR}_j\right) \times \mathrm{PDUI}(x)\right) \times \mathrm{FPKM}(x) \quad (1)$$

where pUTR(x, miR$_j$) and dUTR(x, miR$_j$) are the numbers of miR_j binding sites in pUTR and dUTR of x, and $FPKM(x)$ is expression of x. PDUI indicates the percentage of dUTR usage index[1]. Note that equation (1) can also estimate for genes with no distal 3′ UTR by setting PDUI = 1.

To estimate the trans effect of 3′ US on gene y', we define X to be a set of 3′ US genes that are ceRNA partners of y' (Supplementary Fig. 3b) and Y to be a set of ceRNA partners to $x \in X$, including y'. Only moderately expressed miRNAs are considered, since they are likely to bind all possible binding sites. Thus, we can roughly use the amount of miRNA-binding sites to represent the miRNA function. The miR$_j$-binding effect on each copy of y' can be defined as follows:

$$\mathrm{TransE}\left(y', \mathrm{miR}_j\right) = \frac{\mathrm{FPM}\left(\mathrm{miR}_j\right)}{\sum_{x \in X} \mathrm{MiRs}\left(x, \mathrm{miR}_j\right) + \sum_{y \in Y} \mathrm{MiRs}\left(y, \mathrm{miR}_j\right)} \quad (2)$$

where FPM(miR$_j$) is the miR$_j$ expression level. Since miRNA can bind to any binding sites in the genes connected by the ceRNA relationship ($X \cup Y$), both X and Y need to be considered.

The high-dimensional MAT3UTR input data are often highly correlated with each other (for example, 588 miRNAs in equation (2)). Therefore, MAT3UTR employs the ridge regression that is known to address the dimensionality and collinearity[32,33] in biological data. Indeed, the ridge regression yields a remarkably higher prediction power than classical linear regression. For example, MAT3UTR has a much smaller mean square error (0.38) than classical linear regression (mean square error = 10.84) (Supplementary Fig. 3f).

$$\mathrm{MAT3UTR}(y') = \sum_{\mathrm{miR}_j \in 3'\mathrm{UTR}(y')} \beta_{\mathrm{miR}_j} \times \log \frac{\mathrm{transE}\left(y', \mathrm{miR}_j\right)_{\mathrm{tumor}}}{\mathrm{transE}\left(y', \mathrm{miR}_j\right)_{\mathrm{normal}}} + \epsilon_{y'} \quad (3)$$

subject to $\sum_{\mathrm{miR}_j \in 3'\mathrm{UTR}(y')} \beta_{\mathrm{miR}_j} \leq t$, the ridge regression penalty. $MAT3UTR(y')$ is the trans effect of 3′ US; β_{miRj} is the regression coefficient of miR$_j$; $\epsilon_{y'}$ is the gene-specific error term. We use R^2 to show how much variation in gene expression can be explained by the *MAT3UTR* model. We also used 10-fold cross-validation (CV) to choose the optimal regularization parameter t with 75% of data for training and the remaining 25% for testing. CV error is measured by mean-squared error. Then, to estimate β, we fit the ridge regression with the entire data set using the selected regularization parameter as chosen by CV.

As a result, y' would be more repressed following 3′US, if: y' contains more miRNA-binding sites in its 3′ UTR; X and Y contain fewer miRNA-binding sites; and more transcripts in X undergo 3′ US. The MAT3UTR-control model, which considers miRNA expression but not 3′ US, is defined as:

$$\mathrm{MAT3UTR-control}(y') = \sum_{\mathrm{miR}_j \in 3'\mathrm{UTR}(y')} \beta_{\mathrm{miR}_j} \times \log \frac{\mathrm{FPM}\left(\mathrm{miR}_j\right)_{\mathrm{tumor}}}{\mathrm{FPM}\left(\mathrm{miR}_j\right)_{\mathrm{normal}}} + \epsilon_{y'} \quad (4)$$

where FPM(miR$_j$) is the miR$_j$ expression level. For model comparison between MAT3UTR and MAT3UTR-control, we randomly selected 75% of data for training and the remaining 25% for testing. We perform random division 100 times to evaluate the performance of the MAT3UTR and MAT3UTR-control models, where 10-fold CV also confirms that MAT3UTR has a 2-fold higher prediction power on gene expression variation than the MAT3UTR-control model (Supplementary Fig. 3e).

Selecting genes for experimental validations.

To test the trans repressive effect of 3'US on PTEN, we chose EPS15 on three grounds. First, its expression is easily detected in MCF-7 cells; second, analysis of RNA-Seq from MCF-7 cells[34] indicates distal polyA site usage of the *EPS15* transcript; third, the *EPS15* 3′ UTR contains four microRNA target sites that compete with the *PTEN* 3′ UTR.

To investigate the tumor-suppressor ceRNA downregulation mechanism, we chose *PHF6*, because among 57 tumor-suppressor genes in 3′ US ceRNAs, *PHF6* was predicted as a strong (sixth highest in MAT3UTR score, Supplementary Table 3) trans-target of 3′ US, was significantly downregulated (second highest in gene expression) and was the most enriched in AGO2 RIP complexes of the ceRNA tested (Fig. 4d).

Statistical analyses.

Differential expression analyses were carried out by edgeR (version 3.8.6)[35] (tumor samples versus normal samples) with false discovery rate (FDR) control at 0.05. The significance of observed values for a particular class compared to its control is calculated from one-tailed Pearson's χ^2 test. Each variable follows either a binomial or multinomial distribution and each case consists of at least five counts, which meets the assumption of Pearson's χ^2 test. To test whether there is a significant enrichment of tumor-suppressor genes or oncogenes among a gene list of our interest, we conducted hypergeometric tests with normalized overlap counts, since assessing overlap between sets meets all criteria to use hypergeometric tests, including trials without replacement. To compare means of two groups that have different variances, we used Welch's *t*-test, which does not assume equal population variance. To check the normality assumption for the t-test, we conducted a Shapiro-Wilk normality test for small samples ($n < 50$). All statistical computations were performed in the Python scipy stats package (version 0.15.1) or R (version 3.1.1).

RNA-Seq for *NUDT21* depletion experiment.

We previously sequenced two control and two *NUDT21* depletion samples of HeLa cells by HiSeq 2000 (LC Sciences)[2]. After trimming adaptors using Trim Galore (version 0.4.1), paired-end RNA-Seq reads of 101 base pairs in each end were used to reconstruct the transcriptome in the Tuxedo protocol[36] (TopHat 2.0.6 and Cufflinks 2.1.1). The resulting FPKM values were normalized for comparison using Cuffdiff 2.2.0. Further analyses are based on 10,681 expressed (FPKM ≥ 1 in >3 samples) RefSeq genes. We sequenced miRNAs from control and *NUDT21*-knockdown cells to utilize only miRNAs with intermediate expression in ceRNA identification.

CeRNA identification in the *NUDT21*-knockdown experiment in the HeLa cell line.

Due to the small sample size (two for each condition wild-type and *NUDT21* knockdown), ceRNAs were identified solely on the basis of miRNA-binding-site overlap. We considered only binding sites for miRNAs with intermediate expression (between 0.01 and 100 in averaged FPM). A total of 1,450 3′ US genes identified by DaPars had significant miRNA-binding-site overlap with 1,168 ceRNA genes (3′ US ceRNA partners).

MiRNA-Seq for the *NUDT21* depletion experiment.

HeLa cells were transfected with control or *NUDT21* siRNA. *NUDT21* depletion was validated as previously described[2]. Small RNA libraries were generated from one control and one *NUDT21* depletion sample using the Illumina Truseq Small RNA Preparation kit, and sequenced on Illumina GAIIx. Raw sequencing reads (40 nucleotides) were obtained using Illumina's Sequencing Control Studio software following image analysis and base-calling by Illumina's Real-Time Analysis (v 1.8.70). Then a script ACGT101-miR v 4.2 (LC Sciences) was used for data analysis, where reads are mapped to the reference database (miRbase). The script also normalizes the counts by a library size parameter for comparison.

CeRNA tumor-suppressor repression in HeLa cells with *NUDT21* knockdown.

Parental HeLa cells were purchased from ATCC (cat. no. CCL-2) and maintained in Eagle's minimum essential medium (Lonza, cat. no. 12–604F) with 10% fetal bovine serum. The cells were made mycoplasma free by incubating with Plasmocin (InvivoGen, cat. no. ant-MPT) for two weeks before transfection with three different siRNAs for *NUDT21* (Sigma Aldrich, ID: SASI_Hs01_00146875~77) and negative control siRNA (Sigma Aldrich, ID:SIC002) using previously established approaches[2]. Western blotting was also performed as described in our previous work[2] using antibodies raised against: PHF6 (Santa Cruz, cat. no. sc-271767), YOD1 (abcam, ab178979), NUDT21 (Proteintechlab, cat. no. 10322–1-AP) and GAPDH (Sigma, G9545). To block miRNA function, we selected two miRNAs with a strong trans effect targeting *PHF6* (miR-3187–3p and miR-549) and HeLa cells were co-transfected with siRNA for *NUDT21* and the two antagomirs, to block the two predicted miRNAs, miR-549 and miR-3187–3p in the *PHF6* 3′ UTR. The two antagomirs were designed[37] and synthesized from Sigma-Genosys: Antagomir-3187–3p: 5′-[mU]s[mU]s[mG]mG] [mC][mC][mA][mU][mG][mG][mG][mG][mC][mU][mG] [mC][mG]s[mC]s[mG]s[mG]s-chol-3′; and Antagomir-549: 5′-[mU]s[mG]s[mA][mC][mA][mA][mC][mU][mA][mU][mG][mG][mA][mU][mG][mA][mG][mC]s[mU]s[mC]s[mU]s-chol-3′. PHF6 and YOD1 expression were detected by western blotting and quantified by Image Lab software (version 5.2.1) from Bio-Rad.

Detection of ceRNA tumor-suppressor gene enrichment by RIP with quantitative PCR.

HeLa cells were seeded in a 6-well plate at 4×10^5 cells per well and transfected with a *Cas9* and single-guide RNA (sgRNA) plasmid targeting *NUDT21* or with *Cas9* and *GFP* as a control. sgRNAs for *NUDT21* (top, ccggccgcccaatcgctcgcagac; bottom, aaacgtctgcgagcgattg ggcgg) were synthesized (Sigma), and the annealing double-stranded DNA was cloned into pGL3-U6-sgRNA-PGK-puromycin. The transfected cells from three wells were combined and then selected with 10 $\mu g ml^{-1}$ blasticidin for three days. *NUDT21*-knockdown efficiency was determined by western blot with NUDT21 antibody. RIP was performed with anti-AGO2 antibody, and AGO2-associated RNAs were purified and measured by quantitative real-time PCR[38]. Briefly, the cells were harvested and lysed with 100 μl polysome lysis buffer (100 mM KCl, 5 mM $MgCl_2$, 10 mM Hepes pH 7.0, 0.5% NP50, 1 mM DTT and 1×PI cocktail). The cell lysate was centrifuged at 10,000*g* for 15 min and added to magnetic beads (A+G) with 5 μg anti-Ago2 antibody or normal mouse IgG suspended in 900 μl of NET2 buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$,

0.05% NP-40, 17.5 mM EDTA pH 8.0, 1 mM DTT and 100 units ml^{-1} RNaseOUT). The beads were washed six times with NT2 buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% NP-40). Beads were resuspended in 150 μl proteinase K buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% NP-40 and 1% SDS) with 9 μl proteinase K. Samples were incubated at 55 °C for 30 min and isolate total RNAs with 150 μl phenol–chloroform. The total RNA was reverse transcribed and the candidate ceRNAs were determined by quantitative real-time PCR using primers described in Supplementary Table 5 (Bio-Rad real-time PCR system).

LightSwitch luciferase reporter assay with *PTEN*, *PHF6* and *LARP1* 3′ UTR.

LightSwitch luciferase reporter constructs with *PTEN, PHF6* and *LARP1* 3′ UTR were purchased from SWITCHGEAR genomics. Briefly, HeLa cells were seeded in a 96-well white TC plate in 100 μl total volume to yield ≥80% confluence at the time of transfection. For each transfection, the following reagents were combined: 50 nM siRNA and/or miRNAs and/or antagomir RNA, individual GoClone reporter (30 ng μl^{-1}) 3.33 μl and 1 ng Rluc reporter. Lipofectamine 2000 was diluted in OPTI-MEM medium at 1:10 and incubated at room temperature for 5 min and then added to each tube. Following a 20-min incubation at room temperature, 80 μl of pre-warmed (37 °C) OPTI-MEM medium per replicate was added for a total of 100 ul per replicate transfection. All 100 μl of the transfection mixture was added to each well and incubated overnight. The luciferase reporter assays were performed according to the manufacturer's protocol (Invitrogen).

Immunofluorescence staining for PTEN in MCF7 cells with *EPS* 3′ UTR.

pLightSwitch-*EPS15* 3′ UTR construct was purchased from SWITCHGEAR genomics and transfected into MCF7 cells. PTEN expression was detected by immunofluorescence staining with anti-PTEN antibody from Cell Signaling. Briefly, 1×10^5 MCF7 cells were seeded in 4-well chamber slides overnight, and transfected with pLightSwitch-*EPS15* 3′ UTR/GFP constructs at 10:1 or pLightSwitch-3′ UTR/GFP constructs as a control. One day after transfection, the cells were fixed with 90% cold methanol at −20 °C overnight. The next day, 0.5% Triton X-100 in PBS was added and incubated at room temperature for 30 min. Samples were blocked in 3% BSA in PBS at room temperature for 1 h. *PTEN* antibody was used at 1:200 dilution in 3% BSA/PBS and 200 μl per well was added to the chamber slides and incubated for 1 h at room temperature. After washing three times, the cells were incubated with Alexa-594-conjugated secondary antibody in 3% BSA/PBS for 1 h at room temperature, in the dark. The cells were rinsed three times with PBS, with the third wash containing DAPI. The coverslips were mounted in anti-fade mounting medium and detected by immunofluorescence microscopy. Both *PTEN*- and *GFP*-positive cells were counted in *EPS15* 3′ UTR/GFP cells and pLightSwitch-3′ UTR/GFP control cells.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by US National Institutes of Health (NIH) grants R01HG007538, R01CA193466 and U54CA217297, Cancer Prevention Research Institute of Texas (CPRIT) grant RP150292 to W.L., CPRIT RP100107 to E.J.W. and A.-B.S., CPRIT RP140800 and Welch Foundation AU-1889 to E.J.W., and NIH R01GM046454 and the Houston Endowment, Inc. to A.-B.S.

References

1. Xia Z et al. Dynamic analyses of alternative polyadenylation from RNA-Seq reveal landscape of 3′ UTR usage across 7 tumor types. Nat. Commun. 5, 5274, (2014). [PubMed: 25409906]
2. Masamha CP et al. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. Nature 510, 412–416 (2014). [PubMed: 24814343]
3. Zhang S & Yu D PI(3)king apart PTEN's role in cancer. Clin. Cancer Res.16, 4325–4330 (2010). [PubMed: 20622047]
4. Poliseno L et al. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature 465, 1033–1038 (2010). [PubMed: 20577206]
5. Mayr C & Bartel DP Widespread shortening of 3′ UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 138, 673–684 (2009). [PubMed: 19703394]
6. Fu Y et al. Differential genome-wide profiling of tandem 3′ UTRs among human breast cancer and normal cells by high-throughput sequencing. Genome Res. 21, 741–747 (2011). [PubMed: 21474764]
7. Lin Y et al. An in-depth map of polyadenylation sites in cancer. Nucleic Acids Res. 40, 8460–8471 (2012). [PubMed: 22753024]
8. Morris AR et al. Alternative cleavage and polyadenylation during colorectal cancer development. Clin. Cancer 18, 5256–5266 (2012).
9. Davoli T et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. Cell 155, 948–962 (2013). [PubMed: 24183448]
10. Salmena L, Poliseno L, Tay Y, Kats L & Pandolfi PP A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 146, 353–358 (2011). [PubMed: 21802130]
11. Thomson DW & Dinger ME Endogenous microRNA sponges: evidence and controversy. Nat. Rev. Genet. 17, 272–283 (2016). [PubMed: 27040487]
12. Tay Y et al. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell 147, 344–357 (2011). [PubMed: 22000013]
13. Sumazin P et al. An extensive microRNA-mediated network of RNA–RNA interactions regulates established oncogenic pathways in glioblastoma. Cell 147, 370–381 (2011). [PubMed: 22000015]
14. Ala U et al. Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments.Proc. Natl Acad. Sci. USA 110, 7154–7159 (2013). [PubMed: 23536298]
15. Swanton C Intratumor heterogeneity: evolution through space and time. Cancer Res. 72, 4875–4882 (2012). [PubMed: 23002210]
16. Lu J et al. MicroRNA expression profiles classify human cancers. Nature 435, 834–838 (2005). [PubMed: 15944708]
17. Hausser J & Zavolan M Identification and consequences of miRNA–target interactions — beyond repression of gene expression. Nat. Rev. Genet. 15, 599–612 (2014). [PubMed: 25022902]
18. Mets E et al. MicroRNA-128–3p is a novel oncomiR targeting PHF6 in T-cell acute lymphoblastic leukemia. Haematologica 99, 1326–1333 (2014). [PubMed: 24895337]
19. Selcuklu SD et al. MicroRNA-9 inhibition of cell proliferation and identification of novel miR-9 targets by transcriptome profiling in breast cancer cells. J. Biol. Chem. 287, 29516–29528 (2012). [PubMed: 22761433]
20. Tian B & Manley JL Alternative cleavage and polyadenylation: the long and short of it. Trends Biochem. Sci. 38, 312–320 (2013). [PubMed: 23632313]

21. Mueller AA, Cheung TH & Rando TA All's well that ends well: alternative polyadenylation and its implications for stem cell biology. Curr. Opin. Cell Biol. 25, 222–232 (2013). [PubMed: 23357469]
22. Li L et al. 3′ UTR shortening identifies high-risk cancers with targeted dysregulation of the ceRNA network. Sci. Rep. 4, 5406 (2014). [PubMed: 24953077]
23. Tay Y, Rinn J & Pandolfi PP The multilayered complexity of ceRNA crosstalk and competition. Nature 505, 344–352 (2014). [PubMed: 24429633]
24. Wang Y et al. The emerging function and mechanism of ceRNAs in cancer. Trends Genet. 32, 211–224 (2016). [PubMed: 26922301]
25. Lewis BP, Burge CB & Bartel DP Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15–20 (2005). [PubMed: 15652477]
26. Papadopoulos GL, Reczko M, Simossis VA, Sethupathy P & Hatzigeorgiou AG The database of experimentally supported targets: a functional update of TarBase. Nucleic Acids Res. 37, D155–D158 (2009). [PubMed: 18957447]
27. Xiao F et al. miRecords: an integrated resource for microRNA–target interactions. Nucleic Acids Res. 37, D105–D110 (2009). [PubMed: 18996891]
28. Hsu S-D et al. miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res. 42, D78–D85 (2014). [PubMed: 24304892]
29. Dvinge H et al. The shaping and functional consequences of the microRNA landscape in breast cancer. Nature 497, 378–382 (2013). [PubMed: 23644459]
30. Hamilton MP et al. Identification of a pan-cancer oncogenic microRNA superfamily anchored by a central core seed motif. Nat. Commun. 4, 2730 (2013). [PubMed: 24220575]
31. Goldman M et al. The UCSC Cancer Genomics Browser: update 2013. Nucleic Acids Res. 41, D949–D954 (2013). [PubMed: 23109555]
32. Friedman JM, Hastie T & Tibshirani R Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33, 1–22 (2010). [PubMed: 20808728]
33. Kim S, Baladandayuthapani V & Lee JJ Prediction-oriented marker selection (PROMISE): with application to high-dimensional regression. Stat. Biosci 9, 217–245 (2016). [PubMed: 28785367]
34. Bayerlová M et al. Newly constructed network models of different WNT signaling cascades applied to breast cancer expression data. PLoS ONE 10, 1–19 (2015).
35. Robinson MD, McCarthy DJ & Smyth GK edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010). [PubMed: 19910308]
36. Trapnell C et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–578 (2012). [PubMed: 22383036]
37. Krutzfeldt J et al. Silencing of microRNAs in vivo with 'antagomirs'. Nature 438, 685–689 (2005). [PubMed: 16258535]
38. Tenenbaum SA, Lager PJ, Carson CC & Keene JD Ribonomics: identifying mRNA subsets in mRNP complexes using antibodies to RNA-binding proteins and genomic arrays. Methods 26, 191–198 (2002). [PubMed: 12054896]

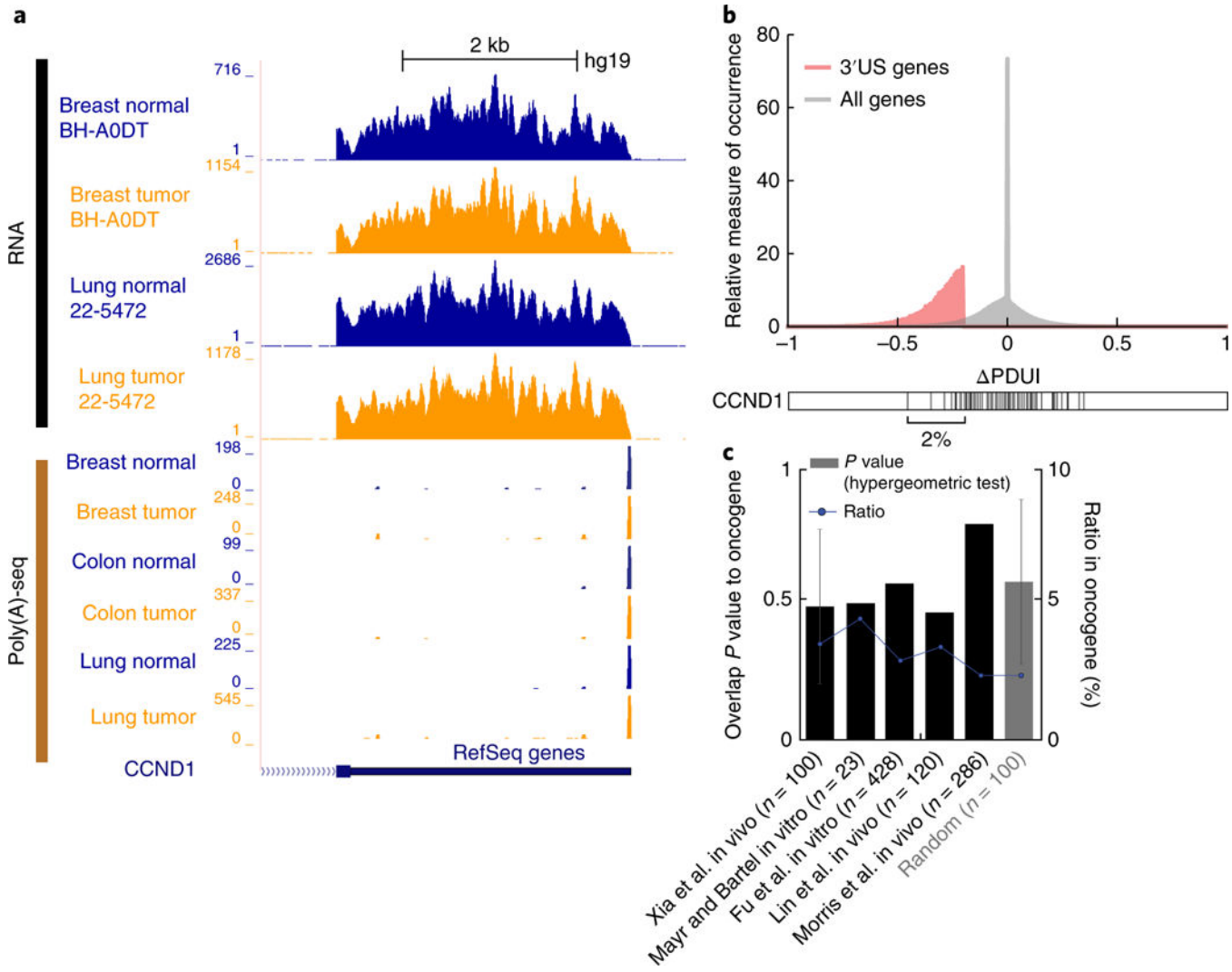

Fig. 1 |. 3′ US genes are not strongly associated with oncogenes.

a, TCGA RNA-Seq data for *CCND1* demonstrates no change in 3′ UTR usage between tumors (yellow) and matched normal samples (blue). A similar pattern was also observed in PolyA-seq[7] of *CCND1*. **b**, ΔPDUI values for 3′ US genes (red) and all genes (gray) in 358 TCGA tumor/normal pairs[1] (upper panel). A negative ΔPDUI represents 3′ UTR shortening. The lower panel shows ΔPDUI values for CCND1 across 358 tumor/normal pairs[1]. Significant *CCND1* 3′ UTR shortening occurred only in a very small portion (8 out of 358; 2.2%) of tumors. **c**, Overlap *P* values and the ratios between previously identified 3′ US genes and oncogenes. 'Random ($n = 100$)' represents the averaged *P* value from 100 random sampling of 100 RefSeq genes. The error bar represents standard variation values of *P* values from 100 random trials.

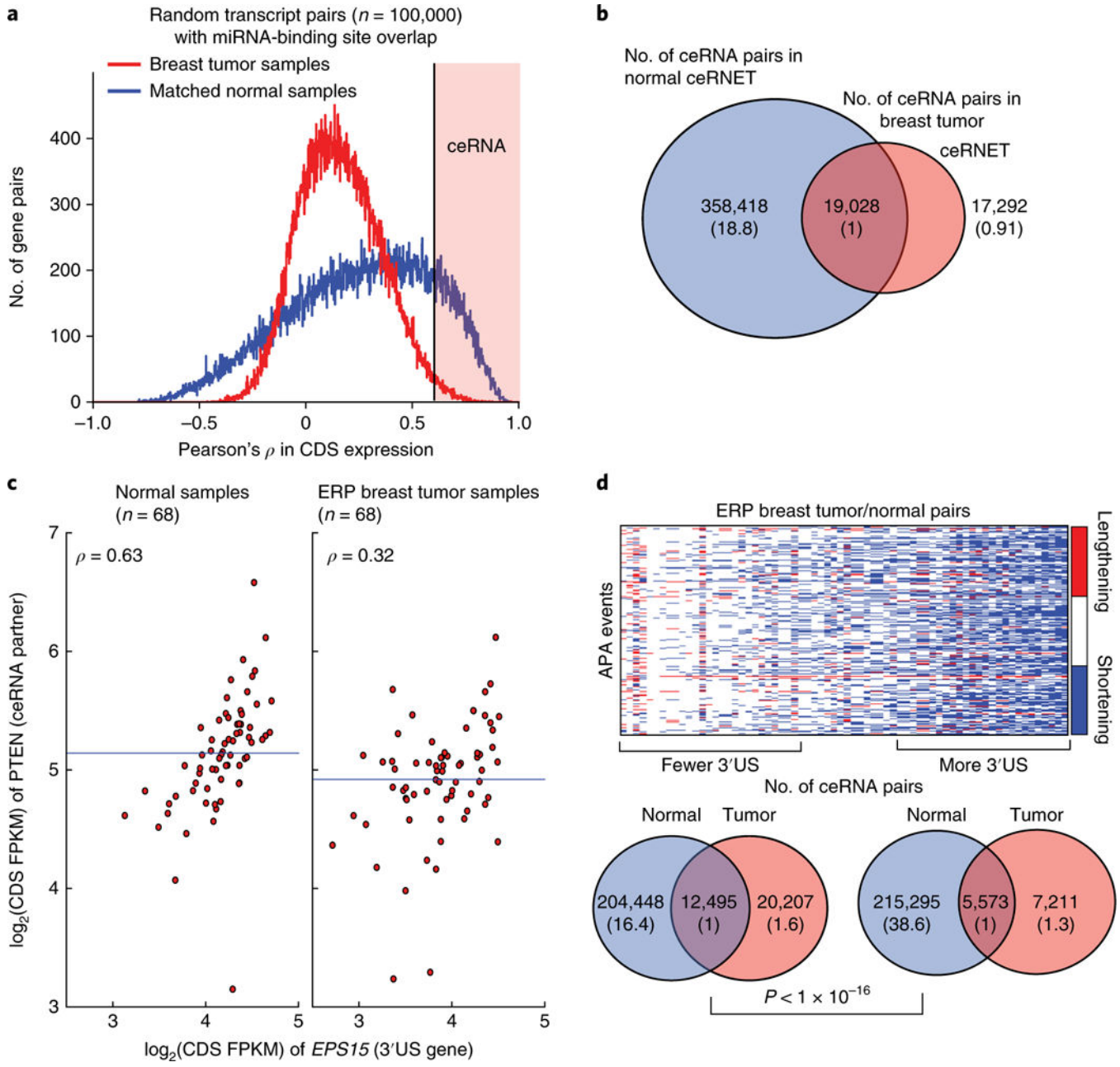

Fig. 2 |. 3′UTR shortening contributes to ceRNET disruption.

a, Pearson's correlation coefficients of 100,000 randomly selected transcript pairs with significant miRNA-binding-site overlap in breast tumors and matched normal tissues. **b**, The number of ceRNA pairs in breast tumor and the matched normal ceRNETs. The numbers in parentheses are normalized to the number of edges shared between tumor and normal. **c**, Gene expression of *EPS15* (3′ US gene) and *PTEN* (ceRNA partner) on 68 estrogen-receptor-positive (ERP) breast tumors and matched normal samples. The horizontal lines represent the mean expression values of *PTEN*, which is decreased in tumors (FDR = 2.1×10^{-10}). **d**, The upper heatmap exhibits significant APA events (rows) across 68 ERP tumor/normal pairs (columns), ranked by the number of 3′ US genes. The Venn diagrams show the

number of ceRNA pairs in the normal and tumor ceRNET. The numbers in parentheses are normalized to the number of edges shared between tumor and normal tissues. The *P* value was calculated from a one-tailed Pearson's chi-squared test.

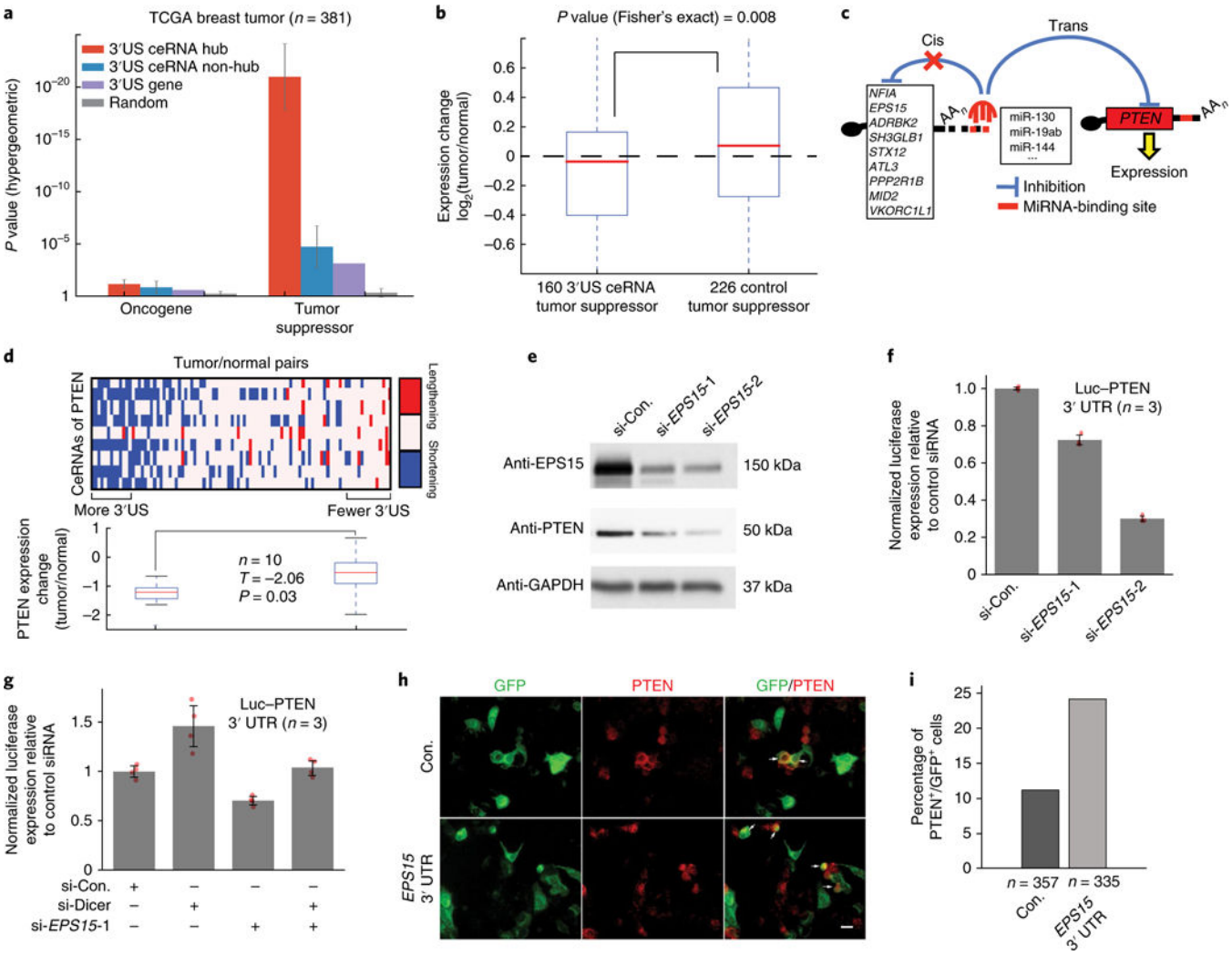

Fig. 3 |. 3′ UTR shortening represses tumor-suppressor genes in TCGA breast cancer.

a, Functional enrichment of 3′ US ceRNA hub genes (red), 3′ US ceRNA non-hub genes (blue), 3′ US genes (purple) and random RefSeq genes (gray). We randomly sampled each gene category to the same number (381) 100 times; averaged *P* values with standard deviation are plotted. **b**, Relative expression (tumor/normal) of tumor-suppressor genes that are 3′US ceRNAs ($n = 160$, left box) is lower than for those that are not in ceRNET ($n = 226$, right box) ($P = 8 \times 10^{-3}$). **c**, 3′ US genes (left) might repress their ceRNA partner *PTEN* in trans through miRNAs (middle) commonly released through 3′ UTR shortening. **d**, A heatmap in the top panel showing APA events for the nine 3′ US genes (rows). The boxplots in the bottom panel show *PTEN* expression levels in 10 tumors with the most (left) or least (right) 3′ UTR shortening. **e**, Western blot analysis of lysates from MCF7 cells treated with control (Con.) or *EPS15*-targeting siRNAs. The image is representative of three independent experiments. **f**, Quantification of luciferase activity from cell lysates derived from MCF7 cells transfected with a luciferase reporter containing the *PTEN* 3′ UTR and *EPS15*-targeting siRNAS. Data are the average luciferase activity ± standard deviation from three independent experiments ($P = 0.011$ and $P < 0.001$, two-sided *t*-test). **g**, *PTEN* 3′ UTR

luciferase reporter activity in MCF7 cells transfected with *EPS15*- and *DICER*-targeting siRNAS. Data are the average luciferase activity ± standard deviation from three independent experiments ($P = 0.045$, $P = 0.003$ and $P = 0.645$, two-sided *t*-test). **h**, Indirect immunofluorescence of MCF7 cells transfected with either a heterologous reporter containing a vector-derived 3′ UTR (Con.) or the *EPS15* 3′ UTR together with a *GFP* construct. PTEN was detected by anti-PTEN antibody conjugated with Alexa Fluor-594. The arrows highlight $PTEN^+$ transfected cells. A representative image is shown from three independent experiments. Scale bar, 20 μM. **i**, The number of PTEN-positive cells in the transfected cells with either the *EPS15* 3′ UTR ($n = 335$) or the control 3′ UTR ($n = 357$) from three images.

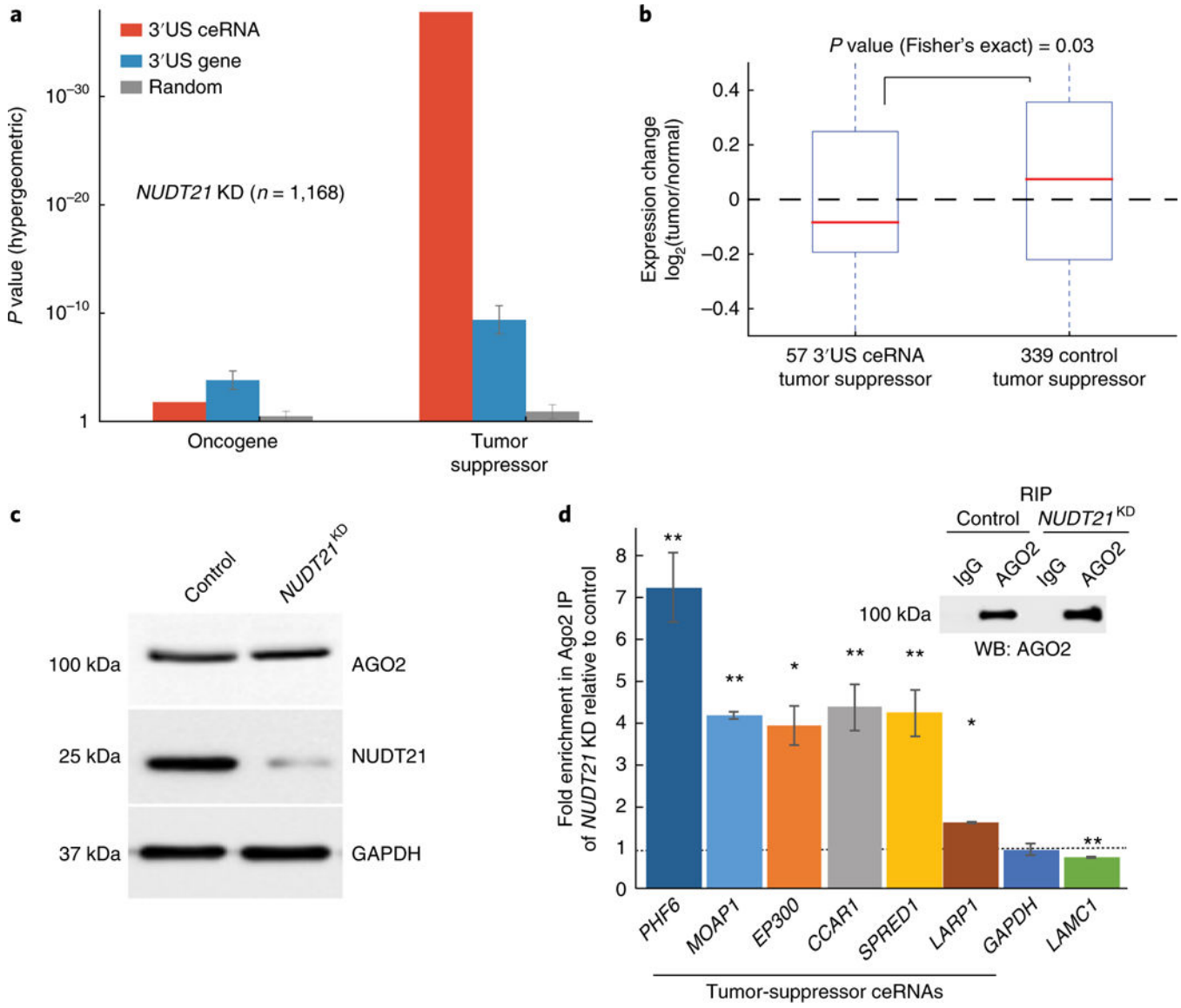

Fig. 4 |. *NUDT21*-mediated 3′ UTR shortening causes tumor-suppressor repression in trans.

a, Oncogene or tumor-suppressor gene enrichment of 3′ US ceRNAs (red), 3′ US genes (blue) and RefSeq genes (gray), in the *NUDT21*-knockdown (KD) experiment. We randomly sampled each gene category to the same number (n = 1,168) 100 times, and reported the averaged P values with standard deviation (error bar). **b**, Expression change of tumor-suppressor genes that are 3′US ceRNAs (n = 57, left box) or that are not connected to 3′US genes (n = 339, right box). 3′US ceRNA tumor-suppressor genes showed lower expression in *NUDT21*-knockdown samples (P = 0.03). **c**, Knockdown of *NUDT21* in HeLa cells using CRISPR/Cas9 and reduced NUDT21 was detected by western blot analysis in three independent experiments. **d**, RIP was performed with anti-AGO2 antibody; normal mouse IgG served as a control. The RIP complexes were detected by western blot with a distinct AGO2 antibody from rat (inset). The indicated ceRNAs associated with AGO2 enrichment in *NUDT21*-knockdown cytoplasmic lysates versus the control are shown with

average fold change ± standard deviation from three independent assays ($P = 0.0002$, $P = 5.2 \times 10^{-6}$, $P = 0.0004$, $P = 0.0005$, $P = 0.0006$, $P = 0.01$ and $P = 5.47 \times 10^{-7}$, two-sided *t*-test, $^{**}P < 0.001$, $^{*}P < 0.01$).

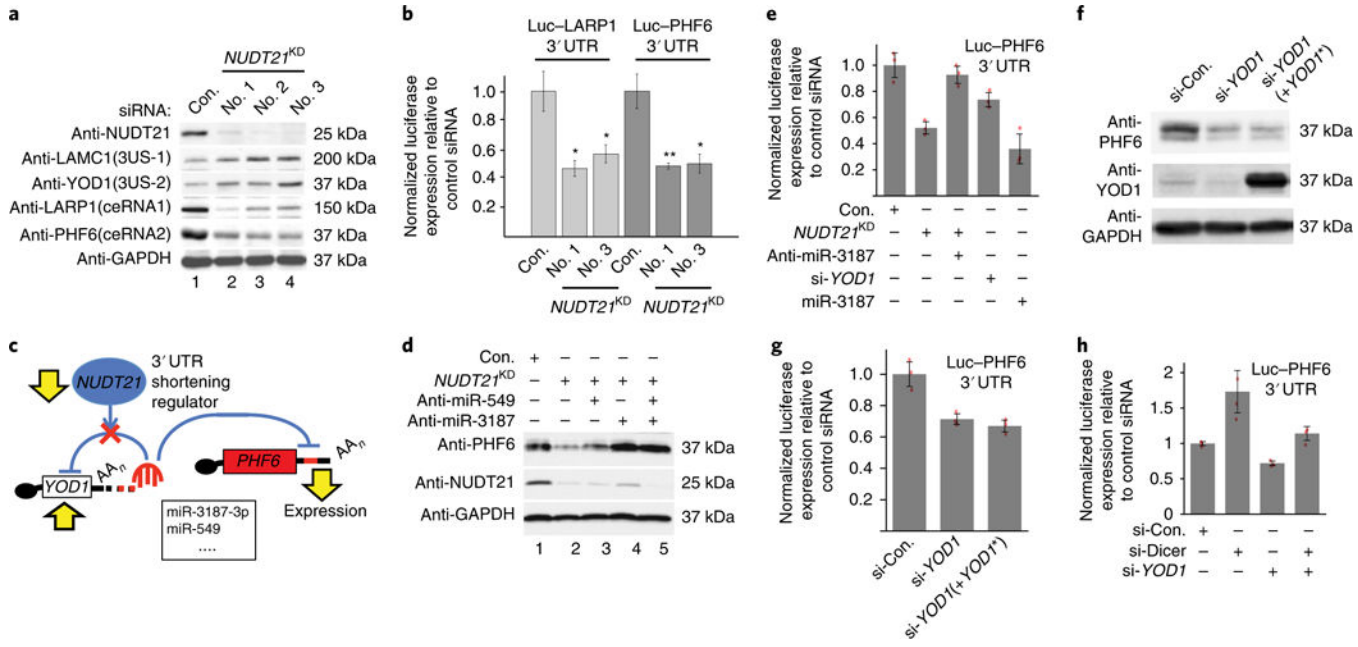

Fig. 5 |. *NUDT21*-mediated 3′ UTR shortening represses the tumor-suppressor genes *PHF6* and *LARP1*.

a, Western blot of 3′ US ceRNA tumor-suppressor genes (*PHF6*/*LARP1*) and 3′ US genes (*LAMC1*/*YOD1*) in *NUDT21*-knockdown cells. A representative image is shown from three independent experiments. **b**, Activity of the *PHF6* 3′ UTR and *LARP1* 3′ UTR luciferase reporter constructs in *NUDT21*-knockdown cells relative to control siRNA-transfected cells. The data are the average of luciferase activity ± standard deviation from three independent experiments ($P = 0.037$ and 0.05; $P = 0.016$ and 0.025, two-sided *t*-test). **c**, *NUDT21* knockdown induces 3′ UTR shortening and upregulation of *YOD1*, allowing miR-3187–3p and miR-549 to repress *PHF6*. **d**, Western blot analysis using the indicated antibodies on lysates from HeLa cells transfected with siRNA for *NUDT21* (si-*NUDT21*-4) and two antagomirs that block miR-549 and miR-3187–3p. The image is representative of two independent experiments. **e**, Activity of the *PHF6* 3′ UTR luciferase reporter construct in HeLa cells with the indicated siRNAs, miRNAs or antagomirs. The data are the average of luciferase activity ± standard deviation from three independent experiments ($P = 0.004$, $P = 0.90$, $P = 0.018$ and $P = 0.015$, two-sided *t*-test). **f**, Western blot analysis of cell lysates from cells transfected with either control siRNA or *YOD1* siRNA. In the third lane, the cells were transfected with *YOD1* siRNA and then transfected with *YOD1* cDNA. The data are representative of three independent experiments. **g**, Activity of the *PHF6* 3′ UTR luciferase reporter in cells treated with the same experimental design as in **f**. The data are the average of luciferase activity ± standard deviation from three independent experiments ($P = 0.016$ and $P = 0.01$, two-sided *t*-test). **h**, Activity of the *PHF6* luciferase reporter in cells transfected with the indicated siRNAs. The data are the average of luciferase activity ± standard deviation from three independent experiments ($P = 0.025$, $P = 0.009$ and $P = 0.99$, two-sided *t*-test).