Molecular and Cellular Biology®

An Evolutionarily Conserved AU-Rich Element in the 3' Untranslated Region of a Transcript Misannotated as a Long Noncoding RNA Regulates RNA Stability

Emily A. Dangelmaier,[a] Xiao Ling Li,[a] Corrine Corrina R. Hartford,[a] Julianna C. King,[b] Meira S. Zibitt,[a] Raj Chari,[b] Ioannis Grammatikakis,[a] Ashish Lal[a]

[a]Regulatory RNAs and Cancer Section, Genetics Branch, Center for Cancer Research (CCR), National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, Maryland, USA
[b]Genome Modification Core, Frederick National Lab for Cancer Research, National Cancer Institute, Frederick, Maryland, USA

Emily A. Dangelmaier and Xiao Ling Li contributed equally to this article. The order was determined by the corresponding author based on their contribution.

ABSTRACT   One of the primary mechanisms of post-transcriptional gene regulation is the modulation of RNA stability. We recently discovered that *LINC00675*, a transcript annotated as a long noncoding RNA (lncRNA), is transcriptionally regulated by FOXA1 and encodes a highly conserved small protein that localizes to the endoplasmic reticulum, hence renamed as *FORCP* (FOXA1-regulated conserved small protein). Here, we show that the endogenous *FORCP* transcript is rapidly degraded and rendered unstable as a result of 3′UTR-mediated degradation. Surprisingly, although the *FORCP* transcript is a canonical nonsense-mediated decay (NMD) and microRNA (miRNA) target, we found that it is not degraded by NMD or miRNAs. Targeted deletion of an evolutionarily conserved region in the *FORCP* 3′UTR using CRISPR/Cas9 significantly increased the stability of the *FORCP* transcript. Interestingly, this region requires the presence of an immediate downstream 55-nt-long sequence for transcript stability regulation. Functionally, colorectal cancer cells lacking this conserved region expressed from the endogenous *FORCP* locus displayed decreased proliferation and clonogenicity. These data demonstrate that the *FORCP* transcript is destabilized via conserved elements within its 3′UTR and emphasize the need to interrogate the function of a given 3′UTR in its native context.

KEYWORDS ARE, AU-rich, mRNA stability, mRNA decay, NMD, miRNA, lncRNA, *LINC00675*, *TMEM238L*, *FORCP*, CRISPR/Cas9, micropeptide, 3′UTR, FOXA1, RNA stability, conserved, lincRNA, misannotated

Despite sharing an identical genetic code, cells differentiate and attain specialized functions through shifting patterns of gene expression. At the post-transcriptional level, differential gene expression is regulated and fine-tuned by RNA binding proteins (RBPs) and regulatory RNAs that modulate RNA processing, stability, localization, and translation (1). Post-transcriptional gene regulation plays well-established roles in diverse cellular processes, such as cell proliferation, differentiation, metabolism, apoptosis, and senescence, and aberrant post-transcriptional gene regulation has been implicated in numerous human diseases, including cancer (1–4). One of the primary mechanisms controlling gene expression post-transcriptionally is the regulation of mRNA stability. Intriguingly, multiple studies have demonstrated that mRNA steady-state levels do not always directly correlate with transcription rates (5–10), suggesting a critical role for mRNA stability in the regulation of gene expression.

Several RNA decay pathways and features are essential for mRNA quality control and prevent the translation of aberrant transcripts. One example is nonsense-mediated decay (NMD), a translation-dependent mechanism that primarily serves to prevent the translation of

transcripts containing premature translation-termination codons (PTCs) (11, 12). Additional features that sensitize mRNAs to NMD include an upstream open reading frame (uORF), a splice site >50 nucleotides (nt) downstream of a stop codon, a long 3′-untranslated region (3′UTR), and a short ORF (sORF) (13–16).

The 3′UTR is a key feature in controlling mRNA stability, which is exemplified by the established correlation of 3′UTR length with mRNA instability (17). One mechanism through which 3′UTRs influence mRNA stability is through interaction with microRNAs (miRNAs), a class of conserved, small noncoding RNAs that promote RNA degradation or translation inhibition through direct interaction with complementary sequences within the 3′UTRs of targeted transcripts (18, 19). miRNAs regulate diverse biological processes, including those involved in tumorigenesis, and aberrant expression of miRNAs contributes to cancer progression (20–22). In addition to miRNA target sites, AU-rich elements (AREs) represent another principal 3′UTR regulatory element (23, 24). These elements are regulated by a class of RBPs called ARE-binding proteins (ARE-BPs), which have heterogeneous functions in promoting the degradation or stability of mRNAs (23, 24). Importantly, AREs and ARE-BPs have significant implications for the post-transcriptional regulation of genes governing various hallmarks of cancer (25). For example, the expression of HuR, an ARE-BP that promotes RNA stability, is increased in many cancer types, and correlates with reduced patient survival (25–29). In contrast, the RNA decay-promoting ARE-BP TTP (tristetraprolin) is frequently deficient in human cancers (30–32). Because of this intrinsic association between AREs and cancer, understanding the role of AREs in cancer-related genes may help elucidate the regulatory pathways that contribute to tumorigenesis and cancer progression.

RNA processing is also integral for regulation at the post-transcriptional level. Noncoding RNA processing is often distinct from mRNA processing. For example, miRNA maturation depends on cleavage and processing by specific factors, such as Dicer and Drosha (33). In contrast, mRNAs and long noncoding RNAs (lncRNAs), a class of noncoding RNAs >200 nucleotides long, share similar patterns of biogenesis. Like mRNAs, many lncRNAs undergo splicing, polyadenylation, and 7-methylguanylate (m7G) capping, processes which regulate transcript stability (34, 35). Despite their similar structure, mRNAs are exclusively cytoplasmic, whereas lncRNAs can be nuclear and/or cytoplasmic (36), and subcellular localization is typically directly linked to lncRNA function (37). The major differentiating factor between mRNAs and lncRNAs is the ability to code for protein, however, this barrier may not be so easily defined as previously thought (38–41).

Emerging evidence has revealed that many cytoplasmic lncRNAs containing sORFs are translated to produce micropeptides, or small proteins less than 100 amino acids in length, many of which have important biological functions (41, 42). The presence of a sORF can lead to targeted regulation of these RNA transcripts by specific cellular pathways. For example, transcripts containing sORFs can be subjected to RNA quality control pathways like NMD as mentioned above (43). Therefore, understanding how micropeptide-encoding transcripts are regulated is imperative to improving our understanding of biological processes and human diseases.

Recently, we discovered that *FORCP* (FOXA1-regulated conserved small protein), previously annotated as a long noncoding RNA, *LINC00675* or *TMEM238L*, encodes a highly conserved, 79-amino acid small protein that localizes to the endoplasmic reticulum (ER) and inhibits cell proliferation, clonogenicity, and tumorigenesis in well-differentiated colorectal cancer (CRC) cell lines (40). In addition, *FORCP* depletion results in decreased apoptosis in response to endoplasmic reticulum (ER) stress (40). Due to the role of *FORCP* in regulating tumorigenicity and apoptosis in CRC cells, we aimed to uncover mechanisms that govern *FORCP* expression. At the transcriptional level, we recently reported that *FORCP* expression is driven by the pioneer transcription factor FOXA1 (40, 44). Building upon this prior work, here we investigate mechanisms through which *FORCP* is post-transcriptionally regulated. We report that the *FORCP* transcript is rapidly degraded in CRC cells through repression of its 3′UTR, interestingly not via NMD or miRNAs. Additionally, our findings suggest that an evolutionarily conserved region in the 3′UTR promotes decay of the *FORCP* transcript, and that this effect is dependent on AU-rich sequences within this region. Interestingly, this

region functionally interacts with a 55-nt-long sequence directly downstream, and deletion of both these segments is required for *FORCP* transcript stability. Functionally, deletion of this region using the CRISPR/Cas9 technology led to decreased proliferation and clonogenicity of CRC cells. Our results highlight the role of specific conserved sequences within the 3'UTR in regulating the stability of the *FORCP* transcript and associated cancer phenotypes.

RESULTS

***FORCP* mRNA is unstable but is not degraded by nonsense-mediated decay.** To investigate if *FORCP* mRNA could be post-transcriptionally regulated, we first determined the half-life of the *FORCP* transcript by treating CRC cells with actinomycin D (ActD) to inhibit transcription and used RT-qPCR to quantify RNA levels. Since *FORCP* is highly expressed in well-differentiated CRC cell lines (40), we assessed *FORCP* mRNA stability in the well-differentiated SW1222, LS180, and LS174T CRC cell lines. We found that *FORCP* mRNA has a relatively short half-life of ~0.8-1.5 h (Fig. 1A to C). *MYC* mRNA and the lncRNA *MALAT1* served as positive controls for unstable and stable RNAs, respectively (Fig. 1A to C). To test whether this instability is specific to well-differentiated CRC cell lines, we stably overexpressed the *FORCP* full-length mature transcript in the poorly differentiated CRC cell line, HCT116, using a lentiviral-based expression vector (pLVX-puro). Although the half-life of the overexpressed *FORCP* mRNA was higher in HCT116 cells (Fig. 1D) compared to that of endogenous *FORCP* mRNA in well-differentiated CRC cells (Fig. 1A to C), it was relatively unstable in HCT116 cells (Fig. 1D), suggesting that the factor(s) regulating *FORCP* mRNA instability is not specifically expressed in well-differentiated CRC cells. Additionally, endogenous *FORCP* mRNA showed a similar pattern of instability when the well-differentiated LS180 cells were treated with another transcription inhibitor, 5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) (Fig. 1E). To measure the stability of nascent *FORCP* RNA in LS180 cells, we used the Click-IT Nascent RNA capture kit. LS180 cells were pulse-labeled and then incubated in non-labeling media for 4 hours to assay for nascent RNA decay. The results show that the *FORCP* transcript is unstable, comparable to the unstable RNA *MYC* (Fig. 1F), which is consistent with the results from ActD and DRB treatments.

We next sought to determine which factor(s) are responsible for the observed instability of the *FORCP* mRNA. We first investigated whether *FORCP* mRNA is a target of NMD because in addition to harboring a sORF that is translated, *FORCP* mRNA contains other features of a classical NMD target; an upstream start codon (uAUG) with lower coding potential than the downstream start codon, a long 3'UTR relative to the sORF, and a splice site >50 nucleotides downstream of the stop codon (Fig. 2A) (40, 41). To inhibit NMD, we used two approaches. The first involves treating CRC cells with cycloheximide (CHX) to block translation, as NMD is a translation-dependent RNA decay pathway (12). Following inhibition of NMD upon treating LS180 or SW1222 cells with CHX, we observed significantly increased expression of the canonical NMD target mRNAs *ATF3* and *GADD45B* (45) (Fig. 2B and C). However, we did not observe any significant change in *FORCP* mRNA levels in LS180 (Fig. 2B). Interestingly, there was a significant decrease in *FORCP* mRNA levels in SW1222, suggesting regulation through a mechanism not directly involving NMD (Fig. 2C). In the second approach, we knocked down *UPF1* a highly conserved RNA helicase and ATPase that plays a central role in NMD (46). Knocking down *UPF1* with siRNAs in LS180 and SW1222 cells significantly upregulated the NMD target mRNA, *ATF3*, but the levels of *FORCP* mRNA did not change significantly (Fig. 2D and E). Together, these data suggest that while *FORCP* mRNA is unstable, this instability is likely not the result of degradation by NMD.

***FORCP* mRNA instability is mediated through its 3'UTR.** Because *FORCP* has a relatively long 3'UTR (Fig. 2A), we next sought to determine whether *FORCP* mRNA instability is the result of the regulation of its 3'UTR. To test this, we first inserted the *FORCP* 3'UTR in the multiple cloning site in the 3'UTR of the *Renilla* luciferase gene in the dual luciferase reporter psiCHECK-2, which also independently transcribes a firefly luciferase reporter. The presence of the *FORCP* 3'UTR led to significantly decreased (~40%) *Renilla* luciferase activity compared to the empty vector following transfection in the poorly differentiated HCT116, and in well-differentiated LS180 and LS174T cells (Fig. 3A), suggesting that instability of *FORCP* mRNA may be mediated by negative regulation of its 3'UTR. These data showing that the
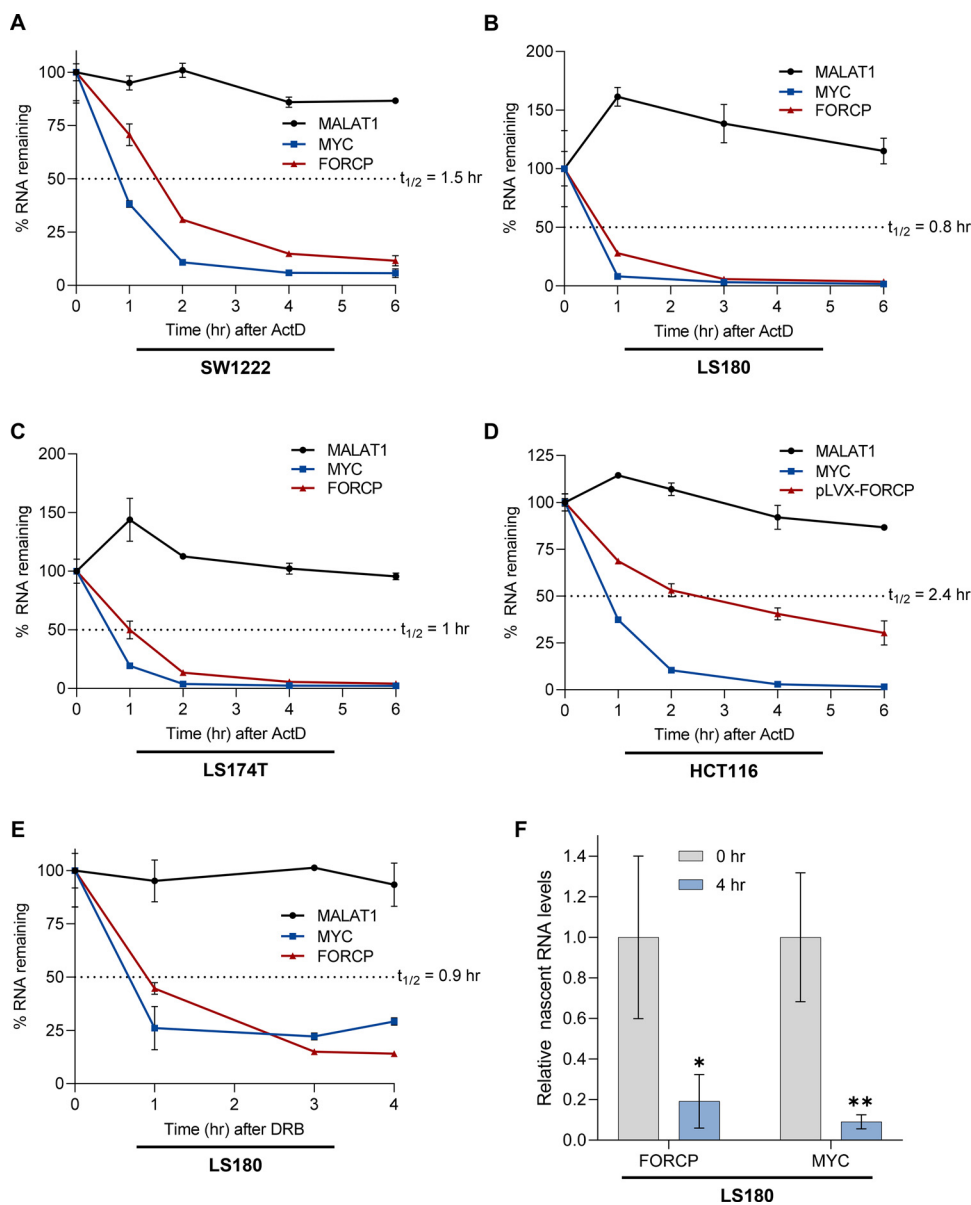
FIG 1 *FORCP* encodes a rapidly degraded RNA transcript. (A–D) RNA stability assays were performed for endogenous *FORCP* mRNA by measuring its levels by RT-qPCR after ActD treatment for the indicated time points in SW1222 (A), LS180 (B) and LS174T (C) cells or for exogenous *FORCP* mRNA overexpressed in HCT116 (D) cells. (E) LS180 cells were treated with DRB for 0, 1, 3, and 4 h and *FORCP* mRNA levels were measured by RT-qPCR. The half-life of *FORCP* is indicated as $t_{1/2}$ (A–E). In these experiments (A–E), *MALAT1* served as a stable RNA control and *MYC* as an unstable RNA control. *GAPDH* was used as a loading control. The graphs show the average of two biological replicates (N = 2). (F) Nascent *FORCP* RNA was measured after pulse-labeling with 5-ethynyl uridine (EU). LS180 cells were incubated with EU-containing media and harvested at time point 0 h or incubated with non-EU-containing media for 4 h before harvest and RNA extraction. EU-labeled RNA was isolated and *FORCP* mRNA levels were measured by RT-qPCR. *MYC* was used as unstable RNA control. Error bars in panels 1A-1E are from two biological replicates; panel 1F was from three independent experiments. *$P < 0.05$, **$P < 0.01$.

FORCP 3′UTR has a repressing effect on luciferase in HCT116 cells that do not express endogenous *FORCP* suggest that the factor(s) targeting the *FORCP* 3′UTR is/are not specifically expressed in well-differentiated CRC cells.

To ensure that the observed repressive effect of the *FORCP* 3′UTR was not unique to the luciferase reporters that we had transiently transfected, we next generated a retroviral reporter (https://www.addgene.org/91975/) in which the *FORCP* 3′UTR was inserted in the 3′UTR of *EGFP*; we named this reporter EGFP-FORCP-3′UTR. We stably transduced this reporter or the EGFP reporter lacking the *FORCP* 3′UTR (EGFP) in HCT116 and
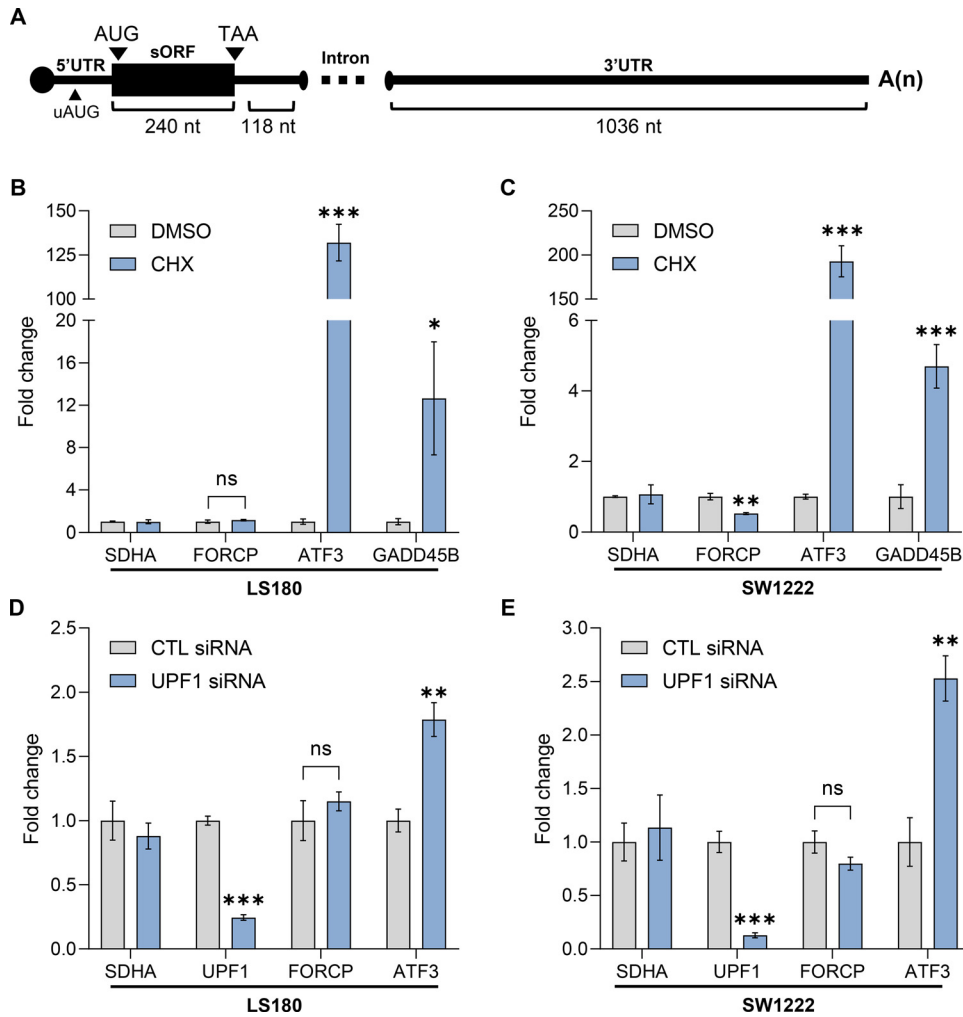
FIG 2 The NMD pathway does not regulate *FORCP* mRNA levels. (A) Diagram showing the features of *FORCP* mRNA indicating the nucleotide lengths of the ORF and 3′UTR. The 5′UTR, the intronic sequences, the upstream start codon (uAUG), and the translation start and stop codons are indicated. (B–C) RT-qPCR assays were performed for the indicated mRNAs upon 4 h of CHX treatment of LS180 (B) and SW1222 (C) cells. DMSO was used as vehicle control. (D–E) RT-qPCR assays were performed for the indicated mRNAs upon *UPF1* knockdown in LS180 (D) and SW1222 (E) cells. *SDHA* served as a negative control. *ATF3* and *GADD45B* are known NMD targets. *GAPDH* served as a loading control. Error bars in panels 2B-2E are from three independent experiments. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

LS174T cells and examined the effect of the *FORCP* 3′UTR on EGFP expression. Consistent with our findings from the luciferase assays, we observed decreased fluorescence of EGFP-FORCP-3′UTR compared to EGFP alone by both flow cytometry and fluorescence microscopy in HCT116 and LS174T cells (Fig. 3B and C). Similarly, we also observed decreased relative expression of EGFP-FORCP-3′UTR by RT-qPCR and immunoblotting (Fig. 3D to F). Finally, to determine whether the difference in expression between EGFP-FORCP-3′UTR and EGFP alone was due to decreased stability of the *EGFP* mRNA, we performed mRNA stability assays using ActD. We found that *EGFP-FORCP-3′UTR* mRNA was less stable than *EGFP* mRNA lacking the *FORCP* 3′UTR (Fig. 3G). Interestingly, although the *EGFP* mRNA containing the *FORCP* 3′UTR was unstable (Fig. 3G), it was more stable than the full-length *FORCP* mRNA (Fig. 1D) indicating that the coding region and/or 5′UTR of the *FORCP* mRNA may also be important in destabilizing *FORCP* mRNA. Alternatively, the *EGFP* coding region might be responsible for a relatively more stable *EGFP-FORCP-3′UTR* transcript compared to full-length *FORCP* mRNA. Nevertheless, these reporter-based assays suggest that *FORCP* mRNA instability is mediated in part through negative regulation of its 3′UTR.
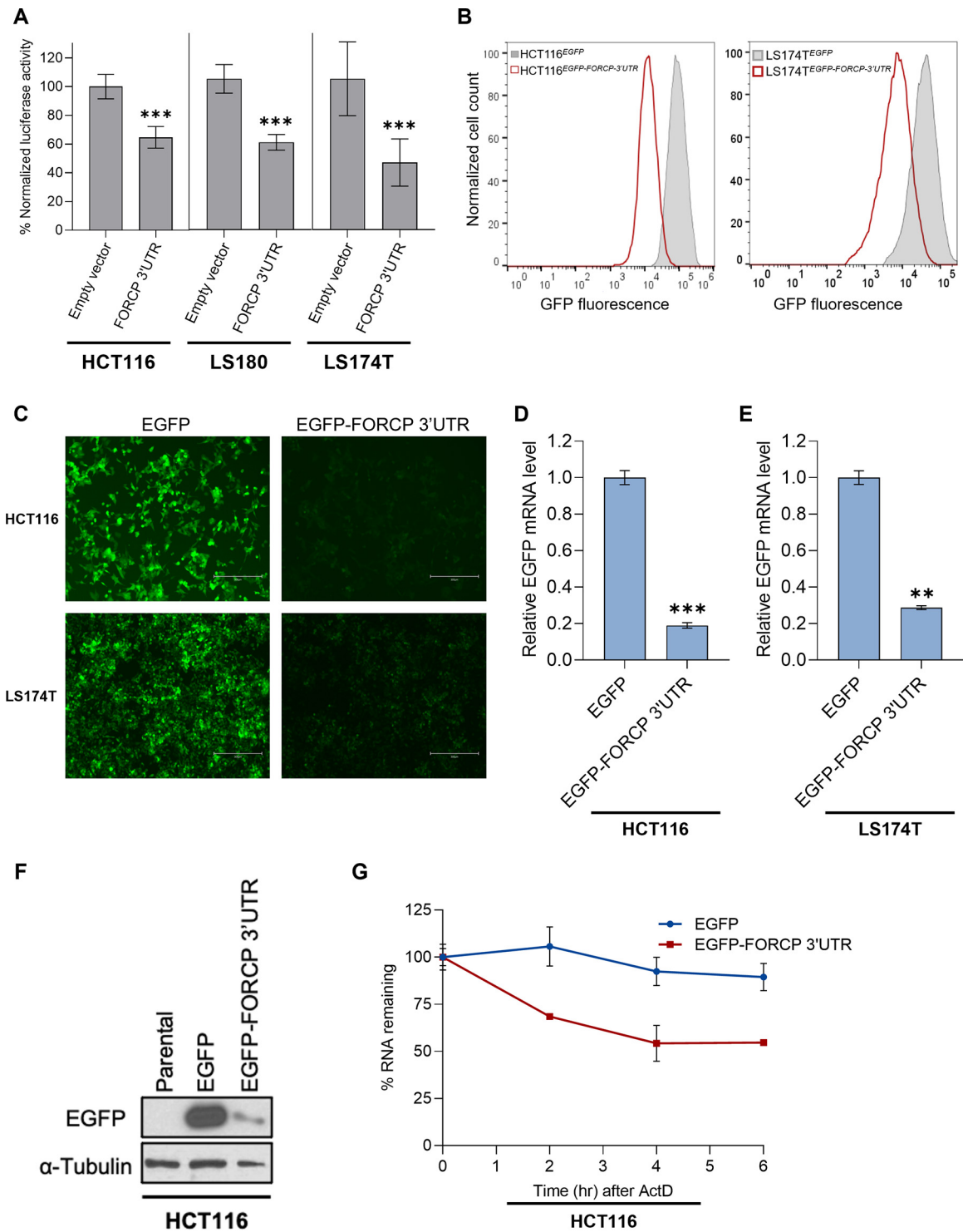
FIG 3 The 3'UTR of *FORCP* mRNA represses reporter expression. (A) Luciferase assays showing percent normalized luciferase activity of a construct containing the *FORCP* 3'UTR compared to empty vector from HCT116, LS180 and LS174T cells. (B) Flow cytometry analysis showing GFP fluorescence of HCT116 (left) or LS174T (right) cells stably expressing an *EGFP* transgene in the presence or absence of the *FORCP* 3'UTR. (C) Microscopy images of HCT116 (upper panel) and LS174T (lower panel) for the data shown in panel B. (D-E) RT-qPCR assays were performed from HCT116 (D) and LS174T (E) cells expressing *EGFP* or *EGFP-FORCP*-3'UTR mRNAs. *GAPDH* was used as a loading control. (F) Western blotting was performed from parental HCT116 cells or HCT116 expressing *EGFP* or *EGFP-FORCP*-3'UTR mRNAs. α-Tubulin was used as loading control. (G) RNA stability assays were performed after 0, 2, 4, and 6 h of ActD treatment of HCT116 cells expressing *EGFP* or *EGFP-FORCP*-3'UTR mRNAs. *GAPDH* was used as a loading control (N = 2). Error bars in panels 3A and 3D are from three independent experiments; panel 3G is from two biological replicates. **$P < 0.01$, ***$P < 0.001$.

FORCP expression is not regulated by miRNAs. miRNAs represent a major class of post-transcriptional regulatory factors that function to promote RNA degradation largely through direct interaction with 3′UTRs of targeted transcripts (18, 19). Therefore, we next asked whether degradation of *FORCP* mRNA is due to targeting by miRNAs. To determine this, we first performed a knockdown of *DICER1*, an essential component of the miRNA processing pathway (47, 48). Following transfection of SW1222 with *DICER1* siRNAs, we observed decreased expression of *DICER1* and miR-200a via RT-qPCR (Fig. 4A and B). However, *FORCP* mRNA levels did not significantly change upon *DICER1* knockdown (Fig. 4A). As with SW1222 cells, knockdown of *DICER1* in LS180 cells did not significantly change *FORCP* mRNA levels (Fig. 4C). In these experiments, depletion of Dicer protein was confirmed by immunoblotting (Fig. 4D). Next, to specifically determine whether the *FORCP* 3′UTR is targeted by miRNAs we performed dual-luciferase assays for the *FORCP* 3′UTR in both *DICER1* knockout (KO) 293T cells generated previously (49) and following DICER1 knockdown in parental 293T cells. As a positive control, we generated a luciferase reporter containing multiple bulged binding sites for miR-19 inserted within the 3′UTR of *Renilla* luciferase, previously used in a recent study in a EGFP reporter (50, 51). As expected, luciferase activity was strongly repressed by the presence of miR-19 binding sites and was significantly rescued in *DICER1* KO cells (Fig. 4E). However, luciferase activity of the *FORCP* 3′UTR reporter did not change significantly upon deletion or knockdown of DICER1 (Fig. 4E and F), suggesting that miRNAs are unlikely to play a role in repressing *FORCP* expression.

An evolutionarily conserved AU-rich element in the *FORCP* 3′UTR regulates *FORCP* mRNA stability. To identify sequence elements within the *FORCP* 3′UTR that may have important regulatory functions, we performed a sequence alignment of the human *FORCP* 3′UTR and the 3′UTR of the mouse homolog (*9130409J20Rik*) using BLASTN (52). The alignment revealed a 56-nt region in the *FORCP* 3′UTR that was evolutionarily conserved between humans and mice (Fig. 5A). Further analysis revealed that this region contains AU-rich elements (AREs), which have been largely implicated in regulating RNA stability (23–25, 53). When we inserted the mouse *FORCP* 3′UTR downstream of *Renilla* luciferase in psiCHECK-2 and performed luciferase assays, we observed a significant reduction in luciferase activity compared to the empty vector (Fig. 5B). These data suggest that both the human and mouse *FORCP* 3′UTRs repress its expression.

To determine if this conserved region in the *FORCP* 3′UTR plays a role in destabilizing the endogenous *FORCP* transcript, we used the CRISPR/Cas9 technology to delete this region in LS174T cells. To do this, we transfected LS174T cells with an all-in-one plasmid expressing GFP, Cas9, and a pair of guide RNAs (sgRNAs) that target sites upstream and downstream of this conserved region (Fig. 5A). Using fluorescence-activated cell sorting (FACS), we selected GFP-positive cells and then performed clonal selection. We were able to isolate one clone that had deletions in both *FORCP* alleles and confirmed the sequence deletion after genomic DNA extraction. This LS174T clone, termed *FORCP* DEL, had a 119 bp deletion in the *FORCP* 3′UTR encompassing the conserved region on one allele. The other allele had a 6 bp deletion at the end of the *FORCP* 3′UTR, but the conserved region was intact (Fig. 5C). We therefore termed the allele with the 6 bp deletion as wild type (WT). Using primers that specifically amplify the DEL or WT allele (Fig. 5A), we next performed RT-qPCR after treating the cells with ActD. We found that the *FORCP* DEL transcript was significantly more stable than the *FORCP* WT transcript ($t_{1/2}$ = 1.45 h versus $t_{1/2}$ = 0.85 h) in the *FORCP* heterozygous LS174T cells (Fig. 5D), suggesting that the instability of the *FORCP* transcript is mediated in part by this 119 nt region.

To determine the effect of this deletion on the steady-state levels of the *FORCP* transcript, we performed RT-qPCR assays that utilized primers that recognize a region that is either unique or common to the *FORCP* WT or *FORCP* DEL transcripts as shown in Fig. 5A. Compared to control LS174T cells, we found a modest increase in total *FORCP* transcript levels and an ∼2-fold decrease in the *FORCP* WT transcript (Fig. 5E). As expected, *FORCP* DEL was dramatically upregulated in the LS174T *FORCP* heterozygous cells (Fig. 5E). We next cloned the *FORCP* DEL 3′UTR into the 3′UTR of *Renilla* luciferase in psiCHECK-2 and performed luciferase assays after transfecting LS180 cells. We found that unlike the *FORCP* WT 3′UTR that significantly repressed luciferase expression, the *FORCP* DEL 3′UTR luciferase
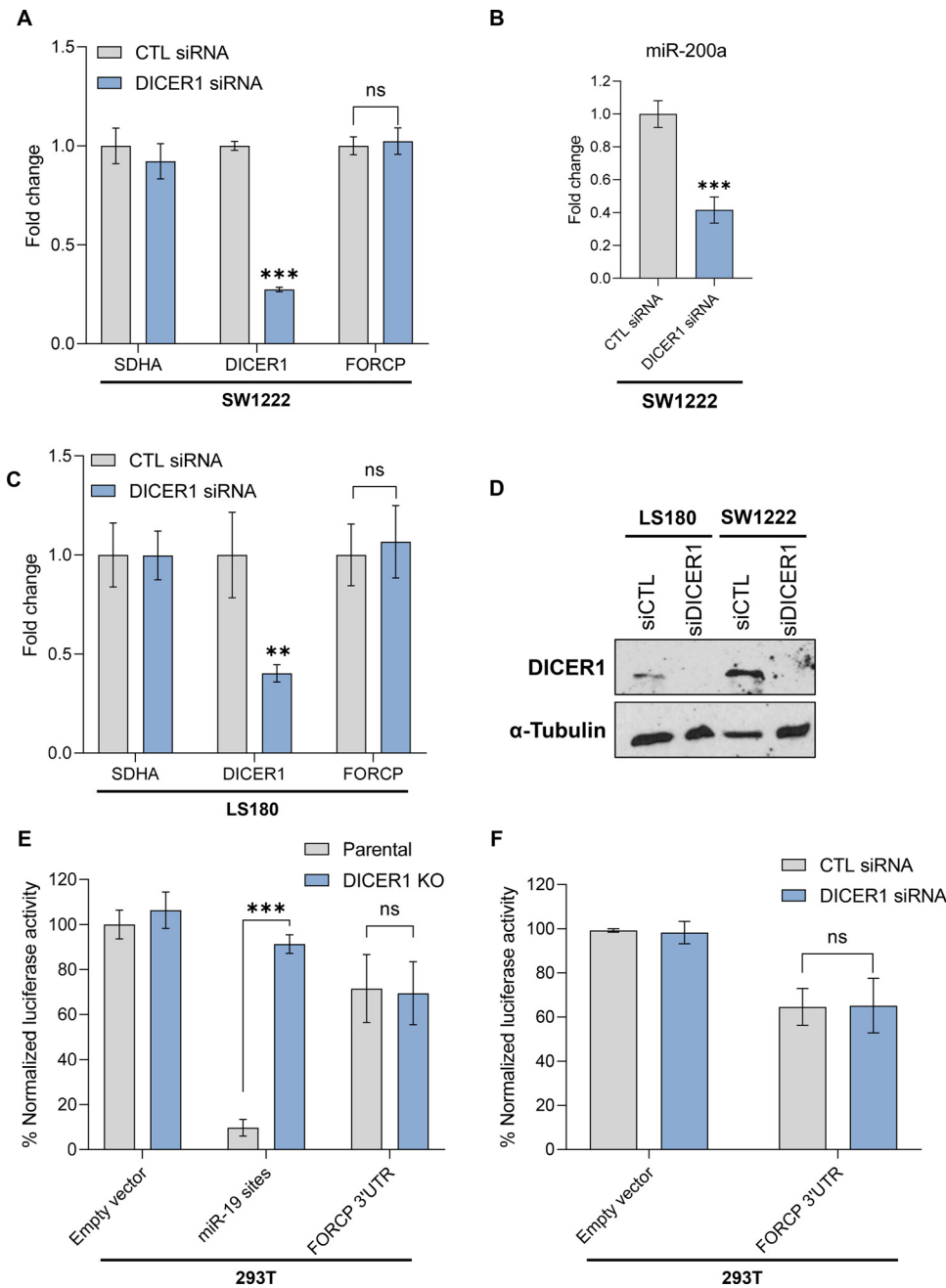
FIG 4 The microRNA pathway does not regulate *FORCP* expression. RT-qPCR was performed for the indicated mRNAs upon *DICER1* knockdown in SW1222 (A) and LS180 (C) cells. *SDHA* serves as a negative control. (B) The relative levels of miR-200a were also measured upon *DICER1* siRNA treatment to confirm *DICER1* knockdown. *U6* was used for loading control. (D) Western blotting was performed from whole-cell lysates showing DICER1 knockdown at the protein level in LS180 and SW1222 cells. α-Tubulin was used for the loading control. (E-F) Luciferase assays were performed from parental and in isogenic *DICER1* knockout 293T cells (E) or upon DICER1 knockdown in 293T cells using siRNAs (F). *FORCP* 3'UTR was cloned downstream of the luciferase gene but showed no effect by *DICER1* knockout (E) or knockdown (F) (N = 2). An EGFP construct containing bulged miR-19 sites was used as a positive control in *DICER1* KO cells. *GAPDH* served as a loading control for all the RT-qPCR assays. Error bars in panels 4A-4C and 4E-4F are from three independent experiments. **$P < 0.01$, ***$P < 0.001$.

expression was not significantly different from the empty vector (Fig. 6A). Since the deletion contains not only the conserved region but also a 55-nt-long region downstream of the conserved sequences, we aimed to dissect the role of these two segments by performing luciferase assays. Surprisingly, we found that deletion of either of these short segments (DEL-conserved or DEL-55 nt) showed an effect similar to that of *FORCP* WT 3'UTR on luciferase
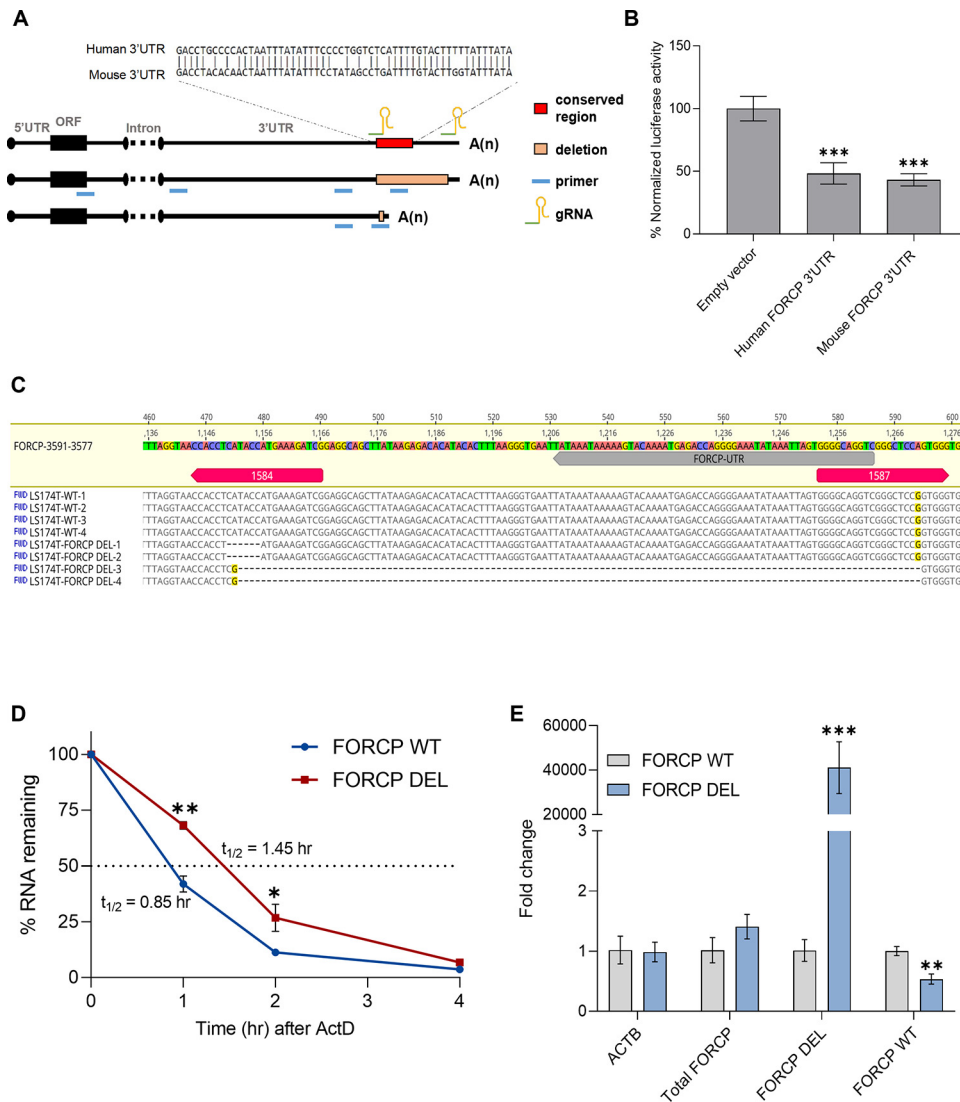
FIG 5 The *FORCP* 3'UTR harbors an evolutionarily conserved AU-rich element that regulates *FORCP* mRNA stability. (A) Upper panel: Alignment of the 56 nt long sequence of the conserved region on *FORCP* 3'UTR between human (NR_036581.1) and mouse (NR_130645.1) genes using BLASTN. Bottom panel: Diagram showing the location of the conserved region (red box) and the target sites of the sgRNAs used for CRISPR are highlighted (top). The middle diagram depicts the 119-nt-long deleted region (orange box) upon CRISPR. The location of primers (blue bars) used in RT-qPCR experiments for detecting Total *FORCP* (middle-left), *FORCP* WT allele (middle-right), and *FORCP* upon deletion (bottom) of the conserved region are indicated. (B) Luciferase assays were performed from HCT116 cells transfected with psiCHECK-2 empty vector or psiCHECK-2 containing the human or mouse *FORCP* 3'UTR. (C) CRISPR/Cas9 technology was used to delete the conserved region on the 3'UTR of *FORCP* in LS174T cells. The panel shows the sequence of the WT *FORCP* gene and the sequence upon deletion (DEL-3 and DEL-4). A 6 bp deletion is observed on the other allele (DEL-1 and DEL-2). (D) RNA stability assays were performed 0, 1, 2, and 4 h after ActD treatment of *FORCP* heterozygous LS174T cells expressing FORCP WT from one allele and FORCT DEL from the other allele. The half-life of the transcripts is indicated as $t_{1/2}$. *GAPDH* was used as a loading control. (E) RT-qPCR assays were performed to determine the relative levels of the indicated transcripts in LS174T after using sgRNAs to delete the conserved region in the 3'UTR of *FORCP* compared to WT cells. *GAPDH* was used as a loading control. Error bars in panels 5B, 5D and 5E are from three independent experiments. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

expression (Fig. 6A). Based on these data, we conclude that the presence of either of these short segments could be sufficient in regulating *FORCP* transcript stability. It may be the case that these two regions physically or functionally interact to mediate transcript degradation. Together, these data suggest that the conserved region in the *FORCP* 3'UTR plays an important role in destabilizing the *FORCP* transcript.

Next, to determine whether the reduced stability of *FORCP* mRNA is a result of reduced translation potential, we performed polysome fractionation experiments from LS174T FORCP
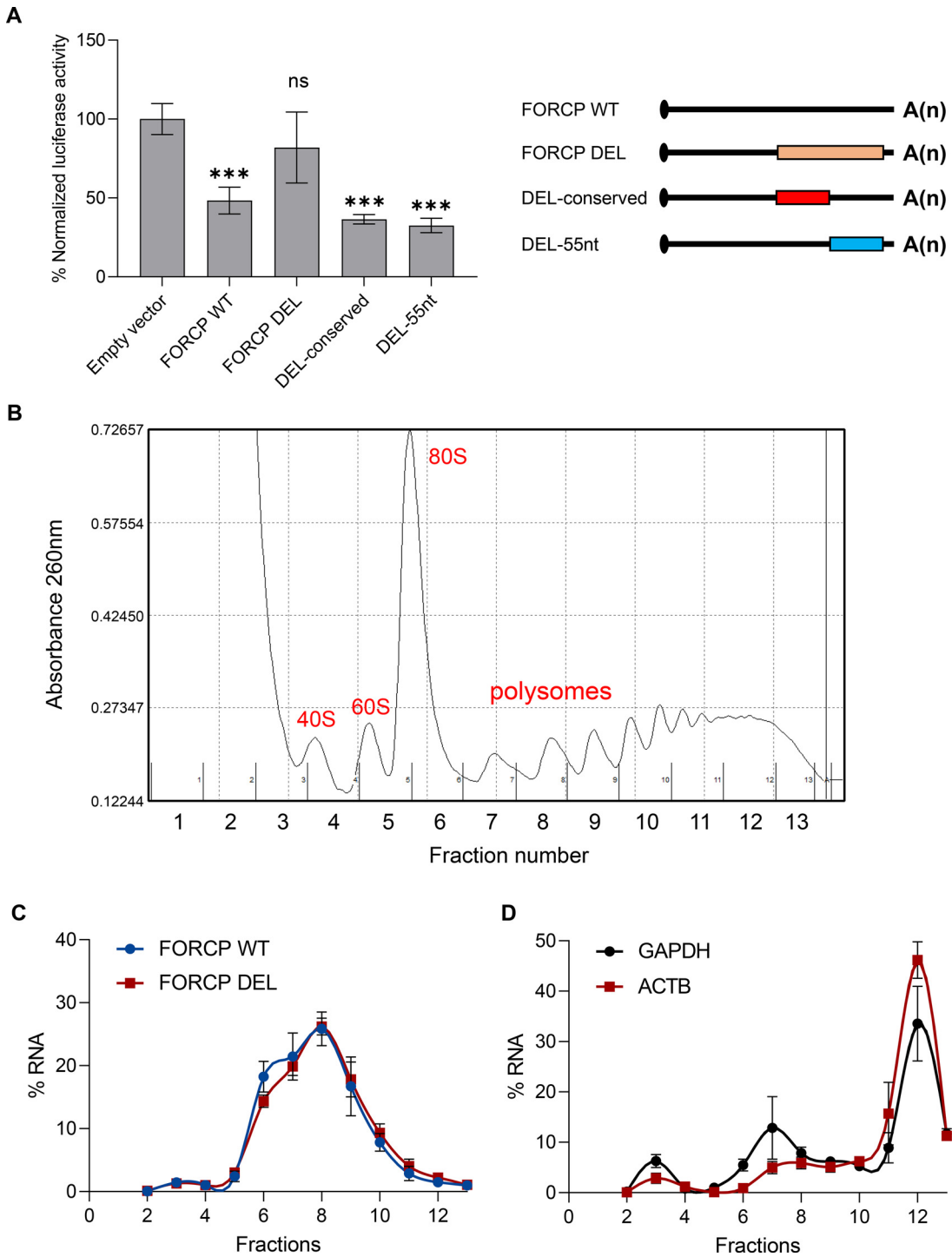
A

B

C

D

FIG 6 Deletion of the evolutionary conserved region in the *FORCP* 3′UTR does not alter translation of *FORCP* mRNA. (A) Luciferase assays showing normalized luciferase activity of a construct containing *FORCP* WT, *FORCP* DEL, a construct with the conserved region deleted (DEL-conserved) and a construct with the downstream 55 nt-long region deleted (DEL-55 nt) compared to the empty vector in LS180 cells. The right panel shows diagrams of the deletions made within each construct in colored boxes. (B) Representative polysome profile from cytoplasmic lysates from LS174T FORCP DEL cells fractionated through a sucrose gradient. Peaks corresponding to the 40S and 60S, small and large ribosomal subunits, respectively, are indicated, as well as peaks corresponding to 80S monosomes and polysomes. (C-D) Percentage fraction distribution of *FORCP* WT and *FORCP* DEL transcripts (C) and *GAPDH* and *ACTB* (D) from lysates fractionated as in panel B. Error bars in panels 6A, 6C and 6D are from three independent experiments. ***$P < 0.001$.

DEL cells (Fig. 6B). RT-qPCR from individual fractions showed that the distribution of *FORCP* WT and the *FORCP* DEL transcripts do not show any significant differences in polysome profile distribution (Fig. 6C), suggesting that their difference in stability is not a result of impaired translational potential. As internal controls, we also determined the profiles of two mRNAs encoding the housekeeping genes *ACTB* and *GAPDH* (Fig. 6D).

Deletion of the evolutionary conserved *FORCP* 3'UTR region results in growth defects. Our recent study showed that in well-differentiated CRC cells, *FORCP* knockdown results in increased proliferation (40). To determine if the increased *FORCP* expression in the *FORCP* DEL LS174T cells results in altered growth, we performed cell count and Incucyte live-cell proliferation assays. We found that the *FORCP* DEL cells displayed growth disadvantage compared to *FORCP* WT cells (Fig. 7A and B). Moreover, in colony formation assays we found a significant decrease in clonogenicity in *FORCP* DEL cells compared to *FORCP* WT cells (Fig. 7C and D). Together, the data presented in this study suggest that the *FORCP* transcript is unstable, and this instability is partly mediated through a conserved region containing AU-rich sequences in the *FORCP* 3'UTR. Deletion of this conserved region using CRISPR/Cas9 results in upregulation of the *FORCP* transcript leading to growth disadvantage in well-differentiated CRC cells.

DISCUSSION

Regulation of mRNA expression and localization via the 3'UTR has been shown to play important roles during development (54–56), senescence (57, 58), proliferation (59, 60), and cancer (61, 62). Our current work focusing on the *FORCP* transcript demonstrates that in addition to conventional mRNAs, a transcript containing a sORF expressed in a cell-type and tissue-specific manner can also be post-transcriptionally regulated. Our data reveal that the *FORCP* transcript is unstable as the result of 3'UTR-mediated degradation. In addition, we found that this regulation is likely not due to degradation by NMD or miRNAs. Finally, our findings using CRISPR/Cas9-mediated targeting of the *FORCP* 3'UTR suggest that an evolutionarily conserved AU-rich region in the *FORCP* 3'UTR serves to destabilize the *FORCP* transcript. This region requires the presence of a downstream 55-nt-long element to exert its function on mRNA stability implying a structural element or an RNA-protein complex that could potentially be responsible for *FORCP* regulation. This work aimed to expand upon prior investigations of the regulation of *FORCP*, which we have recently shown to regulate processes implicated in tumorigenesis and cancer progression in CRC (40).

The activity of a 3'UTR has typically been determined using reporter systems. Although reporter-based assays are efficient, a limitation of these reporter assays is that they do not consider interactions between the 3'UTR with the coding region or 5'UTR of the mRNA. To circumvent this problem, some recent studies have employed genetic approaches to delete the 3'UTR of a gene of interest in its native context and determine the effects on the regulation of the gene at the post-transcriptional level. In a recent study (63), the authors unexpectedly found that in both human cell lines and in mice, deleting the 3'UTR of *p53* mRNA did not affect p53 expression. These results contrast previous reports where the *p53* 3'UTR was shown to be regulated by specific RBPs or miRNAs (64). Furthermore, although the authors found that the 3'UTR of *p53* mRNA repressed reporter gene expression, which is consistent with previous studies, when the coding region of *p53* mRNA was added to the reporters it eliminated any 3'UTR-dependent expression differences. Discrepancies between reporter-based assays and gene expression from native contexts were also found when endogenous 3'UTRs of nine cytokine genes were deleted using CRISPR/Cas9 (65). The authors found that, although six of the nine cytokine 3'UTRs had a repressive effect on the respective cytokine gene expression in both 3'UTR reporter assays and CRISPR/Cas9-mediated deletion of the 3'UTR, the remaining three cytokine 3'UTRs showed decreased reporter gene expression but increased expression of the endogenous gene when their 3'UTRs were deleted. In another recent study, deleting the 3'UTR of *mTOR* in mice revealed a role of the *mTOR* 3'UTR in mRNA localization (66). These studies and our work on post-
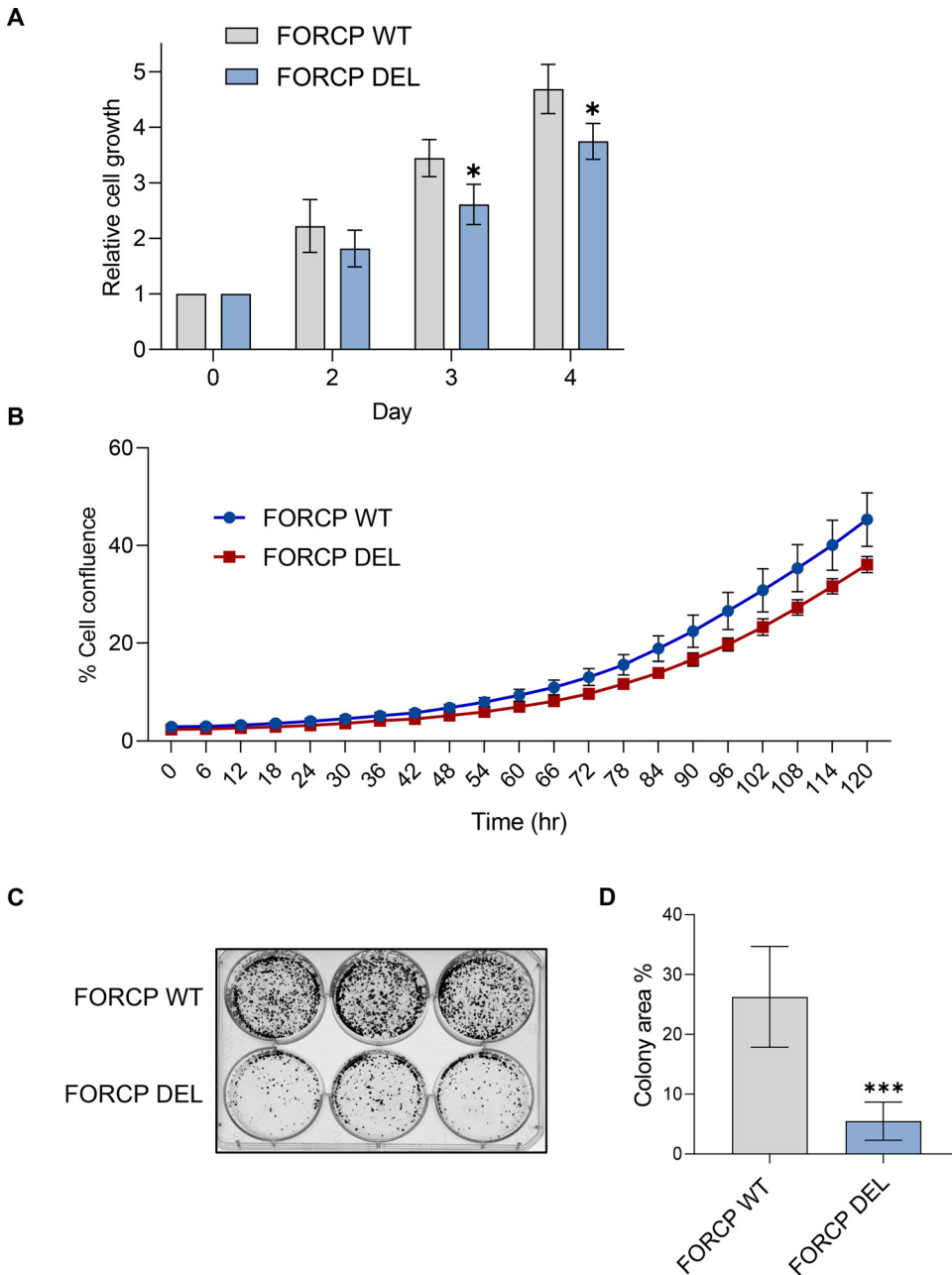
FIG 7 Deletion of the evolutionary conserved region in the *FORCP* 3'UTR region results in growth defects. (A) Cell growth assays showing the relative growth of LS174T expressing either *FORCP* WT or *FORCP* DEL. The cells were plated and counted for the days indicated. (B) Incucyte data analysis showing the cell confluence of LS174T expressing *FORCP* WT or *FORCP* DEL over time as indicated. (C-D) *In vitro* colony formation assays showing reduced clonogenicity of LS174T expressing *FORCP* DEL compared to *FORCP* WT (C). The graph (panel D) shows the percentage (%) area of the wells covered by the colonies. Error bars in panels 7A, 7B and 7D are from three independent experiments. *$P < 0.05$, ***$P < 0.001$.

transcriptional regulation of the *FORCP* transcript emphasize the need to study the roles of 3'UTRs in both reporter assays and their native setting.

What could be the physiological significance of regulating *FORCP* expression post-transcriptionally? Previously, we showed that *FORCP* functions to suppress CRC cell proliferation and tumor growth and is transcriptionally activated by the pioneer transcription factor FOXA1 (40, 44). Our present study shows that the *FORCP* transcript is unstable and a conserved region in the 3'UTR of the *FORCP* transcript regulates its stability and expression resulting in growth defects. These findings suggest that the post-

transcriptional regulation of *FORCP* could allow cells to proliferate. Moreover, because we have previously shown that *FORCP* is upregulated in response to ER stress, we also investigated whether *FORCP* mRNA stability was altered following induction of ER stress (data not shown). However, we observed no effect of ER stress on *FORCP* mRNA stability. It is possible, however, that *FORCP* mRNA stability is altered in a yet-to-be-characterized biological context. Identifying the specific regulatory factors that control *FORCP* mRNA stability may, therefore, uncover novel pathways in which FORCP functions.

In addition to identifying new mechanisms controlling *FORCP* expression, this study expands upon the current knowledge of mechanisms regulating transcripts encoding small proteins. Typically, investigations into the regulation of unstable transcripts containing sORFs have focused on translation-coupled RNA decay pathways, such as NMD (41, 43). In contrast, we saw no effect of NMD inhibition by *UPF1* knockdown nor inhibition of translation by cycloheximide on *FORCP* mRNA steady-state levels. We did, however, observe a robust increase in canonical NMD targets following cycloheximide or *UPF1* knockdown, suggesting that these processes were effectively inhibited. This implied that despite its sORF, *FORCP* mRNA instability is not due to translation-dependent RNA decay pathways. Interestingly, we saw a decrease in *FORCP* levels upon CHX treatment in SW1222 cells, suggesting a mechanism of indirect regulation, such as the inhibition of a cell line-specific RBP targeted by NMD. Other studies also found that RNAs containing sORFs tend to be more unstable (67–69). For example, the *PLN* gene encodes phospholamban, a 52-amino acid transmembrane micropeptide in the sarcoplasmic reticulum of mouse cardiac cells (70). The 3'UTR of *PLN* has a length of approximately 2 kb and cloning it into a reporter leads to decreased luciferase activity implying a role in transcript instability (71).

A major component of the post-transcriptional gene regulation process is the class of short noncoding RNAs, miRNAs (72). miRNAs exert their function by binding to specific sequences on the 3'UTRs, leading to mRNA loading on the RNA-induced silencing complex (RISC) and eventually to transcript degradation (18, 19). Thus, we also investigated the role of miRNAs in the post-transcriptional regulation of *FORCP*. Since Dicer is the major component of miRNA processing and maturation (47, 48), we sought to identify the effect of DICER1 depletion on *FORCP* levels. Our results showed that miRNA maturation was not responsible for regulating *FORCP* RNA degradation, suggesting that miRNAs are not involved in this process.

Similar to miRNAs, RBPs regulate transcript stability and translation by binding to specific sequences in 3'UTRs (73–75). During our study, we used CRISPR/Cas9 to delete the conserved region in the *FORCP* 3'UTR that also created a 55-nt deletion downstream of the conserved site. To specifically dissect the role of each segment, we generated deletion mutants for each and performed luciferase reporter assays. Interestingly, our results indicated that both regions are required for transcript stability regulation, implying a physical or functional interaction of the two sequences. We speculate that their simultaneous presence may be required to regulate accessibility of the targeted RNA region and induce or inhibit the formation of a complex, which may include certain RBPs regulating RNA degradation. The presence of the entire 119-nt-long sequence is necessary for *FORCP* transcript regulation, and further studies will shed light as to whether this region is responsible for specific RBP recruitment or inhibition.

In summary, the discovery of post-transcriptional mechanisms that regulate *FORCP* expression, together with other studies, stresses the importance of post-transcriptional regulation in controlling the expression of genes with vital biological functions, particularly in cancer. Because of the functional role of *FORCP* in regulating CRC pathogenesis, elucidating processes regulating *FORCP* expression remains critical as these pathways may generally apply to other integral cancer genes or genes related to ER stress. Specific investigations into mechanisms regulating the expression of micropeptides are necessary to expand our understanding of this emerging field of study. Further characterization of RBPs regulating *FORCP* expression may reveal regulatory networks that target genes with cancer-related functions.

MATERIALS AND METHODS

Cell culture. HCT116 (ATCC # CCL-247), LS180 (ATCC #: CL-187), 293T (ATCC #: CRL-11268) and LS174T cell lines were purchased from ATCC. SW1222 (ECACC 12022910) cells were purchased from

Millipore Sigma and 293T parental cells and *DICER1* KO derivative cell line used in Fig. 4E were a gift from Dr. Bryan R. Cullen (49). All cell lines were cultured in Dulbecco's modified Eagle medium (DMEM) (Thermo Fisher Scientific) containing 10% (vol/vol) fetal bovine serum (Thermo Fisher Scientific) and 1% (vol/vol) penicillin-streptomycin (Thermo Fisher Scientific) at 37°C in a humidified atmosphere containing 5% CO_2. All cell lines were routinely tested for mycoplasma using the Venor GeM Mycoplasma detection kit (Millipore Sigma-Aldrich).

RNA extraction and RT-qPCR. Total RNA was isolated using TRIzol reagent (Invitrogen) or RNeasy Mini Plus kit (Qiagen). For RT-PCR, 500 ng RNA was reverse-transcribed using iScript Reverse Transcription Supermix (Bio-Rad). Real-time-qPCR was performed using FastStart SYBR green Master Mix (Millipore Sigma, 04913914001) according to the manufacturer's instructions. miRNA quantitative PCR was performed using TaqMan MicroRNA Assay (Applied Biosystems) per the manufacturer's instructions. Primer sequences for each gene are indicated below: ACTB: TGACCCAGATCATGTTTGAGA and AGGGCATACCCCTCGTAGAT; ATF3: GTTTG CCATCCAGAACAAGC and GTCGCCTCTTTTTCCTTTCATC; DICER1: TTCCTCACCAATGGGTCCTTT and GCTTCAAGC AGTTCAACCTGAT; EGFP: AGAACGGCATCAAGGTGAAC and TGCTCAGGTAGTGGTTGTCG; Total FORCP: GAGG AGAAGAGACGCAGGTG and GTATTGCAGCCTCTCGTTCC; FORCP DEL: ACCTTAAGGGTGGGACTGTT and TAGGT AACCACCTCAGTGGG; FORCP WT: ACCTTAAGGGTGGGACTGTT' and CAAAATGAGACCAGGGGAAA; GADD45B: ACCCATGAACTCCCAGTTTG and ATCTCGCTCTCAGTGGTTCG; GAPDH: TGCACCACCAACTGCTTAGC and GGCA TGGACTGTGGTCATGAG; MALAT1: GACGGAGGTTGAGATGAAGC and ATTCGGGGGCTCTGTAGTCCT; MYC: CCACA GCAAACCTCCTCACAG and GCAGGATAGTCCTTCCGAGTG; SDHA: TGGGAACAAGAGGGCATCTG and CCACCACT GCATCAAATTCAT; UPF1: GCTGTCCCAGTATTAAAAGGTG and AGCAGTGGAAAACAGGTATCC.

RNA stability assay using transcription inhibition by Actinomycin D or DRB. LS180, SW1222, and LS174T cells were seeded at 4×10^5 cells/well and HCT116 cells were seeded at 2.5×10^5 cells/well in 35 mm dishes. The following day, cells were treated with 2.5 μg/mL Actinomycin D (ActD) or 20 μg/mL 5,6-dichloroben-zimidazole 1-β-D-ribofuranoside (DRB) (Millipore Sigma-Aldrich) and collected at the indicated time points following addition of TRIzol after PBS wash. RNA levels at various time points were assessed using RT-qPCR as described above.

Nascent RNA labeling and measurement. For nascent RNA measurement, we used the Click-iT Nascent RNA Capture kit according to the manufacturer's protocol (Thermo Fisher Scientific). Briefly, LS180 cells were grown in a 10-cm dish to confluence and were pulse-labeled with 0.2 mM 5-ethynyl uridine (EU) for 4 h. Then, they were either harvested at time point 0 h or incubated with non-EU-containing media for an additional 4 h. Total RNA was then extracted as described above and EU-RNA was biotinylated, isolated using strep-tavidin beads and measured with RT-qPCR.

siRNA transfections. Allstars Negative (CTL) siRNAs were purchased from Qiagen and UPF1 and *DICER1* siRNAs (SMARTpool siRNAs) were purchased from Horizon Discovery. Cells were reverse trans-fected with 20 nM siRNAs using Lipofectamine RNAiMAX (Invitrogen) in Opti-MEM I reduced serum me-dium (Thermo Fisher Scientific). For reverse transfection, cells were seeded at 4×10^5 cells/well in 6-well plates and collected 48 h post-transfection for analysis by RT-qPCR.

Cycloheximide treatment. LS180 and SW1222 cells were seeded at 4×10^5 cells/well in 6-well plates and treated the following day with DMSO or 100 μg/mL cycloheximide (CHX) (Millipore Sigma-Aldrich). Four hours post-treatment, cells were harvested with TRIzol reagent (Invitrogen) for RNA extrac-tion and RT-qPCR as described above.

Dual luciferase reporter assays. For the luciferase assays, the *FORCP* 3'UTR was cloned into the 3' UTR of *Renilla* luciferase in the psiCHECK-2 dual luciferase vector (Promega). The construct containing the full-length human *FORCP* 3'UTR and *FORCP* 3'UTR with deletion of the conserved region (FORCP DEL-conserved) were generated by GenScript. For the constructs containing the 119-nt *FORCP* 3'UTR deletion (FORCP DEL) or the downstream 55-nt deletion (FORCP DEL-55 nt) as well as the mouse 3'UTR, the inserts were generated with gBlocks purchased from IDT. The construct containing bulged binding sites for miR-19 was generated to serve as a positive control in *DICER1* KO cells. The insert was amplified from the pMSCV-EGFP-Bulged-miR-19 vector (Addgene #91976) (primers AATTCTCGAGCTGGTTAACGACGGGTCC and AATTGCGGCCGCGTCGTTTAA ACGTCGGGA) and was cloned into psiCHECK2 using the restriction enzymes XhoI and NotI (New England Biolabs). Ligation was performed with T4 DNA Ligase (New England Biolabs) and ligation products were subse-quently transformed into DH5α competent *E. coli* (Thermo Fisher Scientific). Transformed DH5α were incu-bated overnight at 37°C on LB agar plates containing 100 μg/mL ampicillin. Colonies were inoculated over-night at 37°C in liquid cultures containing 100 μg/mL ampicillin. Plasmid DNA was isolated from liquid culture using the QIAprep Spin Miniprep kit (Qiagen).

For luciferase assays, cells were seeded at 1×10^5 cells/well in 24-well plates 1 day prior to transfection. Cells were transfected the next day with 100 ng plasmid DNA using Lipofectamine 2000 (Invitrogen) in Opti-MEM I reduced serum medium (Thermo Fisher Scientific). For cotransfection of luciferase reporter constructs with *DICER1* siRNAs, transfection was performed using the Amaxa Cell Line Nucleofector kit V (Lonza Bioscience). Luciferase activity was measured 48 h post-transfection using the dual luciferase reporter system (Promega).

Generation of stable cell lines. 293T cells were seeded at 5.0×10^5 cells/well in a 6-well plate. To generate cells that stably overexpress *FORCP*, 293T cells were transfected after fully reattaching to the plate with 1200 ng of pLVX-puro or pLVX-FORCP (40) along with lentiviral packaging vectors using Lipofectamine 2000 (Life Technologies Invitrogen), as directed by the manufacturer. Alternatively, to generate EGFP reporter lines, *FORCP* 3'UTR was cloned into the pMSCV-EGFP vector (Addgene #91975). The insert was amplified with primers GGCCGAATTCTTTGGCCATGGGAAGAGG and GGCCGAATTCCATAAAAACAAAGGCCAAATCT. 293T cells were transfected with 1 μg pMSCV-EGFP or pMSCV-EGFP-FORCP-3'UTR along with retroviral packaging vec-tors. Medium containing packaged viral particles was harvested and replenished 48, 56, and 72 h following transfection and stored at −80°C until further use. Virus titer was determined by serial dilution method, and a multiplicity of infection (MOI) of ~1 was used to generate all stable cell lines. For HCT116 pLVX-puro and

pLVX-FORCP cells, selection was performed using 1 μg/mL puromycin. For pMSCV-EGFP and pMSCV-EGFP-FORCP-3'UTR cells, EGFP positive cells were selected by FACS using the BD FACSAria II Cell Sorter. Cell sorting was performed by the Center for Cancer Research (CCR) Flow Cytometry Core Facility at the National Cancer Institute (NCI) in Bethesda, MD.

Flow cytometry. Cells were prepared for flow cytometry by trypsinization in 0.25% trypsin-EDTA (Thermo Fisher Scientific), resuspended in medium, then pelleted and washed twice with PBS. Following the second wash with PBS, cells were resuspended in PBS and analyzed for GFP fluorescence using the BDFACS Canto II flow cytometer.

Immunoblotting. Total cell lysates were prepared using RIPA buffer containing 1X protease inhibitor cocktail (Roche). Protein concentration was determined using the Bicinchoninic Acid (BCA) protein quantitation kit (Thermo Fisher Scientific). 10 μg whole-cell lysate per lane was loaded onto a 10% SDS-PAGE gel and transferred to nitrocellulose membrane (Thermo Fisher Scientific). The membrane was blocked with 5% skim milk (Oxoid) in TBST (0.05% TWEEN). The following antibodies were used: anti-GFP (1:1000, rabbit) (Cell Signaling), anti-α-tubulin (1:1000, rabbit) (Cell Signaling), anti-DICER1 (1:1000, mouse) (Abcam), and secondary anti-mouse and anti-rabbit (1:5000) from Cell Signaling.

CRISPR-Cas9 mediated partial deletion of the *FORCP* 3′ UTR. Genomic DNA sequence encompassing the FORCP 3′ UTR region was downloaded from the UCSC Genome Browser and subsequently used as input for sgRNA Scorer 2.0 to identify candidate guide RNA target sites (76). Six sites, three sequences on each side of the region, were assessed for activity using a previously described approach (77) and CGATCTTTCATGGTATGAGG**TGG** and GGGGCAGGTCGGGCTCCAGT**GGG** (PAM sequences in bold) were found to be most effective and subsequently used for targeted deletion experiments. Oligonucleotides corresponding to each guide RNA were annealed and cloned into pDG458 (Addgene #100900) using golden gate ligation (78). pDG458 was a gift from Paul Thomas (Addgene #100900; http://n2t.net/addgene:100900; RRID: Addgene_100900).

Genomic DNA extraction and sequencing. To confirm the partial deletion on *FORCP* 3'UTR after CRISPR, genomic DNA was extracted using the Qiagen DNeasy blood and tissue kit. 50 ng of gDNA was used as the template in a Q5 Hot Start PCR (New England Biolabs) with primers GACGGAGTTTCATCATGTTGG and TCCTAATGGGATCTCCTGCG. Gel extracted PCR products of a region encompassing the guide RNA target sites were generated from both wild type and edited cells, cloned into the pCR-Blunt II-TOPO vector (Thermo Fisher Scientific), and transformed into NEB 5-alpha cells (New England BioLabs). For each PCR product, four colonies were picked and miniprepped and sent for Sanger sequence validation (Genewiz). Sanger data were then aligned and analyzed using Geneious software.

Polysome fractionation and analysis. LS174T FORCP DEL cells were grown on a 10-cm dish to confluence and incubated with 100 μg/mL cycloheximide (Sigma) for 10 min. Cytoplasmic lysates were extracted with Polysome Extraction Buffer (100 mM KCl, 5 mM MgCl$_2$, 0.5% NP-40, 100 μg/mL cycloheximide, 2 mM DTT, 40 U/mL RNase OUT (Invitrogen), 1x Complete protease inhibitor cocktail (Roche)) and fractionated by centrifugation in a SW40Ti swinging bucket rotor (Beckman Coulter) through a 10%–50% linear sucrose gradient. A fraction collector (Biocomp Instruments) was used to obtain 13 fractions while monitored by optical density measurement at 260 nm. 250 μL from each fraction was used to extract RNA using TRIzol LS (Thermo Fisher) and the RNA was used for RT-qPCR analysis, as described above. The first fraction was omitted from the analysis due to absence of RNA.

Cell growth and clonogenicity assays. To measure cell growth as in Fig. 7A, 3×10^5 cells were plated in 6-well plates, harvested each day and counted using a cell counter (Bio-Rad). Cell confluence was assayed by plating 3000 cells in 96-well plates. Live cell proliferation assays were then performed using the Incucyte (Essen BioScience) system for the indicated times. Confluence on captured images was measured using the Incucyte Base Analysis Software. For the clonogenicity assay, 5000 cells/well were plated on 6-well plates. 2 weeks later, colonies were fixed with ice-cold 100% methanol for 5 min, stained with crystal violet and colony area was counted and analyzed using ImageJ.

Statistical analysis. In all figures, error bars represent standard deviation, and all statistical analysis was performed in Prism using the Student's t test.

REFERENCES

1. Corbett AH. 2018. Post-transcriptional regulation of gene expression and human disease. Curr Opin Cell Biol 52:96–104. https://doi.org/10.1016/j.ceb.2018.02.011.
2. Audic Y, Hartley RS. 2004. Post-transcriptional regulation in cancer. Biol Cell 96:479–498. https://doi.org/10.1016/j.biolcel.2004.05.002.
3. Garcia-Cardenas JM, Guerrero S, Lopez-Cortes A, Armendariz-Castillo I, Guevara-Ramirez P, Perez-Villa A, Yumiceba V, Zambrano AK, Leone PE, Paz YMC. 2019. Post-transcriptional regulation of colorectal cancer: a focus on RNA-binding proteins. Front Mol Biosci 6:65. https://doi.org/10.3389/fmolb.2019.00065.
4. Jewer M, Findlay SD, Postovit LM. 2012. Post-transcriptional regulation in cancer progression: microenvironmental control of alternative splicing and translation. J Cell Commun Signal 6:233–248. https://doi.org/10.1007/s12079-012-0179-x.

5. Keene JD, Tenenbaum SA. 2002. Eukaryotic mRNPs may represent post-transcriptional operons. Mol Cell 9:1161–1167. https://doi.org/10.1016/s1097-2765(02)00559-2.

6. Hieronymus H, Silver PA. 2004. A systems view of mRNP biology. Genes Dev 18:2845–2860. https://doi.org/10.1101/gad.1256904.

7. Moore MJ. 2005. From birth to death: the complex lives of eukaryotic mRNAs. Science 309:1514–1518. https://doi.org/10.1126/science.1111443.

8. Keene JD. 2007. RNA regulons: coordination of post-transcriptional events. Nat Rev Genet 8:533–543. https://doi.org/10.1038/nrg2111.

9. Vogel C, Marcotte EM. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet 13:227–232. https://doi.org/10.1038/nrg3185.

10. Garcia-Martinez J, Aranda A, Perez-Ortin JE. 2004. Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. Mol Cell 15:303–313. https://doi.org/10.1016/j.molcel.2004.06.004.

11. Brogna S, Wen J. 2009. Nonsense-mediated mRNA decay (NMD) mechanisms. Nat Struct Mol Biol 16:107–113. https://doi.org/10.1038/nsmb.1550.

12. Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol 16:665–677. https://doi.org/10.1038/nrm4063.

13. Colombo M, Karousis ED, Bourquin J, Bruggmann R, Muhlemann O. 2017. Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA 23:189–201. https://doi.org/10.1261/rna.059055.116.

14. Muhlrad D, Parker R. 1994. Premature translational termination triggers mRNA decapping. Nature 370:578–581. https://doi.org/10.1038/370578a0.

15. Muhlrad D, Parker R. 1999. Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. RNA 5:1299–1307. https://doi.org/10.1017/s1355838299990203.

16. Nagy E, Maquat LE. 1998. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem Sci 23:198–199. https://doi.org/10.1016/s0968-0004(98)01208-0.

17. Matoulkova E, Michalova E, Vojtesek B, Hrstka R. 2012. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. RNA Biol 9:563–576. https://doi.org/10.4161/rna.20231.

18. Chekulaeva M, Filipowicz W. 2009. Mechanisms of miRNA-mediated posttranscriptional regulation in animal cells. Curr Opin Cell Biol 21:452–460. https://doi.org/10.1016/j.ceb.2009.04.009.

19. Lai EC. 2002. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet 30:363–364. https://doi.org/10.1038/ng865.

20. Visone R, Croce CM. 2009. MiRNAs and cancer. Am J Pathol 174:1131–1138. https://doi.org/10.2353/ajpath.2009.080794.

21. Baer C, Claus R, Plass C. 2013. Genome-wide epigenetic regulation of miRNAs in cancer. Cancer Res 73:473–477. https://doi.org/10.1158/0008-5472.CAN-12-3731.

22. Peng Y, Croce CM. 2016. The role of MicroRNAs in human cancer. Signal Transduct Target Ther 1:15004. https://doi.org/10.1038/sigtrans.2015.4.

23. Otsuka H, Fukao A, Funakami Y, Duncan KE, Fujiwara T. 2019. Emerging Evidence of Translational Control by AU-Rich Element-Binding Proteins. Front Genet 10:332. https://doi.org/10.3389/fgene.2019.00332.

24. Tebo J, Der S, Frevel M, Khabar KS, Williams BR, Hamilton TA. 2003. Heterogeneity in control of mRNA stability by AU-rich elements. J Biol Chem 278:12085–12093. https://doi.org/10.1074/jbc.M212992200.

25. Khabar KS. 2017. Hallmarks of cancer and AU-rich elements. Wiley Interdiscip Rev RNA 8. https://doi.org/10.1002/wrna.1368.

26. Nabors LB, Gillespie GY, Harkins L, King PH. 2001. HuR, a RNA stability factor, is expressed in malignant brain tumors and binds to adenine- and uridine-rich elements within the 3' untranslated regions of cytokine and angiogenic factor mRNAs. Cancer Res 61:2154–2161.

27. Young LE, Sanduja S, Bemis-Standoli K, Pena EA, Price RL, Dixon DA. 2009. The mRNA binding proteins HuR and tristetraprolin regulate cyclooxygenase 2 expression during colon carcinogenesis. Gastroenterology 136:1669–1679. https://doi.org/10.1053/j.gastro.2009.01.010.

28. Zhu H, Berkova Z, Mathur R, Sehgal L, Khashab T, Tao RH, Ao X, Feng L, Sabichi AL, Blechacz B, Rashid A, Samaniego F. 2015. HuR suppresses fas expression and correlates with patient outcome in liver cancer. Mol Cancer Res 13:809–818. https://doi.org/10.1158/1541-7786.MCR-14-0241.

29. Giaginis C, Alexandrou P, Delladetsima I, Giannopoulou I, Patsouris E, Theocharis S. 2014. Clinical significance of histone deacetylase (HDAC)-1, HDAC-2, HDAC-4, and HDAC-6 expression in human malignant and benign thyroid lesions. Tumour Biol 35:61–71. https://doi.org/10.1007/s13277-013-1007-5.

30. Sohn BH, Park IY, Lee JJ, Yang SJ, Jang YJ, Park KC, Kim DJ, Lee DC, Sohn HA, Kim TW, Yoo HS, Choi JY, Bae YS, Yeom YI. 2010. Functional switching

31. Sharma A, Bhat AA, Krishnan M, Singh AB, Dhawan P. 2013. Trichostatin-A modulates claudin-1 mRNA stability through the modulation of Hu antigen R and tristetraprolin in colon cancer cells. Carcinogenesis 34:2610–2621. https://doi.org/10.1093/carcin/bgt207.

32. Sobolewski C, Sanduja S, Blanco FF, Hu L, Dixon DA. 2015. Histone deacetylase inhibitors activate tristetraprolin expression through induction of early growth response protein 1 (egr1) in colorectal cancer cells. Biomolecules 5:2035–2055. https://doi.org/10.3390/biom5032035.

33. Li S, Patel DJ. 2016. Drosha and dicer: slicers cut from the same cloth. Cell Res 26:511–512. https://doi.org/10.1038/cr.2016.19.

34. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25:1915–1927. https://doi.org/10.1101/gad.17446611.

35. Ni T, Yang Y, Hafez D, Yang W, Kiesewetter K, Wakabayashi Y, Ohler U, Peng W, Zhu J. 2013. Distinct polyadenylation landscapes of diverse human tissues revealed by a modified PA-seq strategy. BMC Genomics 14:615. https://doi.org/10.1186/1471-2164-14-615.

36. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigo R. 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 22:1775–1789. https://doi.org/10.1101/gr.132159.111.

37. Grammatikakis I, Lal A. 2021. Significance of lncRNA abundance to function. Mamm Genome https://doi.org/10.1007/s00335-021-09901-4.

38. Ingolia NT, Brar GA, Stern-Ginossar N, Harris MS, Talhouarne GJ, Jackson SE, Wills MR, Weissman JS. 2014. Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. Cell Rep 8:1365–1379. https://doi.org/10.1016/j.celrep.2014.07.045.

39. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324:218–223. https://doi.org/10.1126/science.1168978.

40. Li XL, Pongor L, Tang W, Das S, Muys BR, Jones MF, Lazar SB, Dangelmaier EA, Hartford CC, Grammatikakis I, Hao Q, Sun Q, Schetter A, Martindale JL, Tang B, Jenkins LM, Robles AI, Walker RL, Ambs S, Chari R, Shabalina SA, Gorospe M, Hussain SP, Harris CC, Meltzer PS, Prasanth KV, Aladjem MI, Andresson T, Lal A. 2020. A small protein encoded by a putative lncRNA regulates apoptosis and tumorigenicity in human colorectal cancer cells. Elife 9. https://doi.org/10.7554/eLife.53734.

41. Hartford CCR, Lal A. 2020. When long noncoding becomes protein coding. Mol Cell Biol 40. https://doi.org/10.1128/MCB.00528-19.

42. Ji Z, Song R, Regev A, Struhl K. 2015. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. Elife 4:e08890. https://doi.org/10.7554/eLife.08890.

43. Wery M, Descrimes M, Vogt N, Dallongeville AS, Gautheret D, Morillon A. 2016. Nonsense-mediated decay restricts LncRNA levels in yeast unless blocked by double-stranded RNA structure. Mol Cell 61:379–392. https://doi.org/10.1016/j.molcel.2015.12.020.

44. Lazar SB, Pongor L, Li XL, Grammatikakis I, Muys BR, Dangelmaier EA, Redon CE, Jang SM, Walker RL, Tang W, Ambs S, Harris CC, Meltzer PS, Aladjem MI, Lal A. 2020. Genome-wide analysis of the FOXA1 transcriptional network identifies novel protein-coding and long noncoding RNA targets in colorectal cancer cells. Mol Cell Biol 40. https://doi.org/10.1128/MCB.00224-20.

45. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. 2004. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet 36:1073–1078. https://doi.org/10.1038/ng1429.

46. Kim YK, Maquat LE. 2019. UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. RNA 25:407–422. https://doi.org/10.1261/rna.070136.118.

47. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409:363–366. https://doi.org/10.1038/35053110.

48. Jaskiewicz L, Filipowicz W. 2008. Role of Dicer in posttranscriptional RNA silencing. Curr Top Microbiol Immunol 320:77–97. https://doi.org/10.1007/978-3-540-75157-1_4.

49. Bogerd HP, Whisnant AW, Kennedy EM, Flores O, Cullen BR. 2014. Derivation and characterization of Dicer- and microRNA-deficient human cells. RNA 20:923–937. https://doi.org/10.1261/rna.044545.114.

50. Kluiver J, Gibcus JH, Hettinga C, Adema A, Richter MK, Halsema N, Slezak-Prochazka I, Ding Y, Kroesen BJ, van den Berg A. 2012. Rapid generation of microRNA sponges for microRNA inhibition. PLoS One 7:e29275. https://doi.org/10.1371/journal.pone.0029275.

51. Golden RJ, Chen B, Li T, Braun J, Manjunath H, Chen X, Wu J, Schmid V, Chang TC, Kopp F, Ramirez-Martinez A, Tagliabracci VS, Chen ZJ, Xie Y, Mendell JT. 2017. An Argonaute phosphorylation cycle promotes microRNA-mediated silencing. Nature 542:197–202. https://doi.org/10.1038/nature21025.

52. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

53. Chen CY, Shyu AB. 1995. AU-rich elements: characterization and importance in mRNA degradation. Trends Biochem Sci 20:465–470. https://doi.org/10.1016/s0968-0004(00)89102-1.

54. Ulitsky I, Shkumatava A, Jan CH, Subtelny AO, Koppstein D, Bell GW, Sive H, Bartel DP. 2012. Extensive alternative polyadenylation during zebrafish development. Genome Res 22:2054–2066. https://doi.org/10.1101/gr.139733.112.

55. Agarwal V, Lopez-Darwin S, Kelley DR, Shendure J. 2021. The landscape of alternative polyadenylation in single cells of the developing mouse embryo. Nat Commun 12:5101. https://doi.org/10.1038/s41467-021-25388-8.

56. Ji Z, Lee JY, Pan Z, Jiang B, Tian B. 2009. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci U S A 106:7028–7033. https://doi.org/10.1073/pnas.0900028106.

57. Chen M, Lyu G, Han M, Nie H, Shen T, Chen W, Niu Y, Song Y, Li X, Li H, Chen X, Wang Z, Xia Z, Li W, Tian XL, Ding C, Gu J, Zheng Y, Liu X, Hu J, Wei G, Tao W, Ni T. 2018. UTR lengthening as a novel mechanism in regulating cellular senescence. Genome Res 28:285–294. https://doi.org/10.1101/gr.224451.117.

58. Zheng D, Wang R, Ding Q, Wang T, Xie B, Wei L, Zhong Z, Tian B. 2018. Cellular stress alters 3'UTR landscape through alternative polyadenylation and isoform-specific degradation. Nat Commun 9:2268. https://doi.org/10.1038/s41467-018-04730-7.

59. Elkon R, Drost J, van Haaften G, Jenal M, Schrier M, Oude Vrielink JA, Agami R. 2012. E2F mediates enhanced alternative polyadenylation in proliferation. Genome Biol 13:R59. https://doi.org/10.1186/gb-2012-13-7-r59.

60. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. 2008. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science 320:1643–1647. https://doi.org/10.1126/science.1155390.

61. Park HJ, Ji P, Kim S, Xia Z, Rodriguez B, Li L, Su J, Chen K, Masamha CP, Baillat D, Fontes-Garfias CR, Shyu AB, Neilson JR, Wagner EJ, Li W. 2018. 3' UTR shortening represses tumor-suppressor genes in trans by disrupting ceRNA cross-talk. Nat Genet 50:783–789. https://doi.org/10.1038/s41588-018-0118-8.

62. Mayr C, Bartel DP. 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 138:673–684. https://doi.org/10.1016/j.cell.2009.06.016.

63. Mitschka S, Mayr C. 2021. Endogenous p53 expression in human and mouse is not regulated by its 3'UTR. Elife 10. https://doi.org/10.7554/eLife.65700.

64. Haronikova L, Olivares-Illana V, Wang L, Karakostis K, Chen S, Fahraeus R. 2019. The p53 mRNA: an integral part of the cellular stress response. Nucleic Acids Res 47:3257–3271. https://doi.org/10.1093/nar/gkz124.

65. Zhao W, Siegel D, Biton A, Tonqueze OL, Zaitlen N, Ahituv N, Erle DJ. 2017. CRISPR-Cas9-mediated functional dissection of 3'-UTRs. Nucleic Acids Res 45:10800–10810. https://doi.org/10.1093/nar/gkx675.

66. Terenzio M, Koley S, Samra N, Rishal I, Zhao Q, Sahoo PK, Urisman A, Marvaldi L, Oses-Prieto JA, Forester C, Gomes C, Kalinski AL, Di Pizio A, Doron-Mandel E, Perry RB, Koppel I, Twiss JL, Burlingame AL, Fainzilber M. 2018. Locally translated mTOR controls axonal local translation in nerve injury. Science 359:1416–1421. https://doi.org/10.1126/science.aan1053.

67. Campalans A, Kondorosi A, Crespi M. 2004. Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localization of a nuclear RNA binding protein in Medicago truncatula. Plant Cell 16:1047–1059. https://doi.org/10.1105/tpc.019406.

68. Cabrera-Quio LE, Herberg S, Pauli A. 2016. Decoding sORF translation - from small proteins to gene regulation. RNA Biol 13:1051–1059. https://doi.org/10.1080/15476286.2016.1218589.

69. Bazin J, Baerenfaller K, Gosai SJ, Gregory BD, Crespi M, Bailey-Serres J. 2017. Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. Proc Natl Acad Sci U S A 114: E10018–E10027. https://doi.org/10.1073/pnas.1708433114.

70. Anderson DM, Makarewich CA, Anderson KM, Shelton JM, Bezprozvannaya S, Bassel-Duby R, Olson EN. 2016. Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides. Sci Signal 9:ra119. https://doi.org/10.1126/scisignal.aaj1460.

71. Hu H, Jiang M, Cao Y, Zhang Z, Jiang B, Tian F, Feng J, Dou Y, Gorospe M, Zheng M, Zheng L, Yang Z, Wang W. 2020. HuR regulates phospholamban expression in isoproterenol-induced cardiac remodelling. Cardiovasc Res 116:944–955. https://doi.org/10.1093/cvr/cvz205.

72. Mayr C. 2017. Regulation by 3'-Untranslated Regions. Annu Rev Genet 51: 171–194. https://doi.org/10.1146/annurev-genet-120116-024704.

73. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW. 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149:1393–1406. https://doi.org/10.1016/j.cell.2012.04.031.

74. Baltz AG, Munschauer M, Schwanhausser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C, Landthaler M. 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell 46:674–690. https://doi.org/10.1016/j.molcel.2012.05.021.

75. Gerstberger S, Hafner M, Tuschl T. 2014. A census of human RNA-binding proteins. Nat Rev Genet 15:829–845. https://doi.org/10.1038/nrg3813.

76. Chari R, Yeo NC, Chavez A, Church GM. 2017. sgRNA Scorer 2.0: a Species-Independent Model To Predict CRISPR/Cas9 Activity. ACS Synth Biol 6: 902–904. https://doi.org/10.1021/acssynbio.6b00343.

77. Gooden AA, Evans CN, Sheets TP, Clapp ME, Chari R. 2021. dbGuide: a database of functionally validated guide RNAs for genome editing in human and mouse cells. Nucleic Acids Res 49:D871–D876. https://doi.org/10.1093/nar/gkaa848.

78. Adikusuma F, Pfitzner C, Thomas PQ. 2017. Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. PLoS One 12:e0187236. https://doi.org/10.1371/journal.pone.0187236.