Published online -

Nucleic Acids Research, -, Vol. -, No. - ***1–13***
doi:10.1093/nar/gkn000

Charting DENR-dependent translation reinitiation uncovers predictive uORF features and links to circadian timekeeping via Clock

Violeta Castelo-Szekely [1], Mara De Matos [1], Marina Tusup [2,3], Steve Pascolo [2,3], Jernej Ule [4,5], and David Gatfield [1,*]

[1]Center for Integrative Genomics, University of Lausanne, Genopode, 1015 Lausanne, Switzerland. [2]Department of Dermatology, University Hospital of Zurich, Gloriastrasse 31, 8091 Zurich, Switzerland. [3]Faculty of Medicine, University of Zurich, 8091 Zurich, Switzerland. [4]Department of Neuromuscular Diseases, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK. [5]The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK.

ABSTRACT

The non-canonical initiation factor DENR promotes translation reinitiation on mRNAs harbouring upstream open reading frames (uORFs). Moreover, DENR depletion shortens circadian period in mouse fibroblasts, suggesting involvement of uORF usage and reinitiation in clock regulation. To identify DENR-regulated translation events transcriptome-wide and, in particular, specific core clock transcripts affected by this mechanism, we have used ribosome profiling in DENR-deficient NIH3T3 cells. We uncovered 240 transcripts with altered translation rate, and used linear regression analysis to extract 5' UTR features predictive of DENR dependence. Among core clock genes, we identified Clock as a DENR target. Using Clock 5' UTR mutants, we mapped the specific uORF through which DENR acts to regulate CLOCK protein biosynthesis. Notably, these experiments revealed an alternative downstream start codon, likely representing the bona fide CLOCK N-terminus. Our findings provide insights into uORF-mediated translational regulation that can regulate the mammalian circadian clock and gene expression at large.

INTRODUCTION

Initiation is considered the rate-limiting step for eukaryotic translation (1) and can be regulated by multiple mechanisms including through upstream open reading frames (uORFs), which are short translated sequences defined by a start codon within the 5' untranslated region (5' UTR) of the transcript and an in-frame stop codon. Canonical translation initiation involves association of the small ribosomal subunit with an mRNA's 5' cap structure and scanning of the 5' UTR until a start codon is recognised. In cases when this translation event is on a uORF, the ribosome can resume scanning after termination, allowing reinitiation at the downstream protein coding sequence (CDS). Alternatively, a scanning ribosome might skip certain uORFs, in a process known as leaky scanning. Translation reinitiation and leaky scanning are poorly understood mechanisms that may involve canonical and non-canonical initiation factors and that are likely regulated by specific features of the uORF and within the transcript (reviewed in (2, 3)).

Density regulated protein (DENR) and Malignant T-cell amplified sequence 1 (MCTS1) form a heterodimeric non-canonical initiation factor that has been described to bind to the small ribosomal subunit and to promote the eIF2-independent recruitment of the initiator tRNA to viral mRNAs, as well as to participate in the recycling of post-termination 40S subunits (4, 5). In *Drosophila*, DENR-MCTS1 was identified as a reinitiation factor required for efficient CDS translation on transcripts containing short and strong uORFs (6). Furthermore, based on reporter assays with synthetic constructs containing uORFs of varying lengths and Kozak strengths, endogenous DENR targets have recently been predicted in human cells (7). However, an unbiased identification of DENR targets *in vivo* and a transcriptome-wide quantification of the effects exerted by DENR on translational regulation are still lacking.

uORF usage is pervasive in the mammalian transcriptome, with up to 40% of transcripts predicted to contain at least one uORF (8). It generally reduces CDS translation efficiency (TE) because scanning ribosomes are diverted from the main ORF and unlikely all reinitiate (3). We have previously reported that several transcripts encoding components of the circadian clock machinery contained translated uORFs (9). Moreover, DENR loss-of-function – as a means of globally interfering with uORF regulation – shortened the period of free-running circadian oscillations in mouse fibroblasts by up to 1.5-hours (9). However, it has remained elusive on which circadian clock transcripts and through which uORFs DENR can act.

Here, we have used ribosome profiling in mouse NIH3T3 fibroblasts to identify and quantify DENR-mediated translational regulation transcriptome-wide, and in particular within core clock genes. We have used comprehensive

*To whom correspondence should be addressed. Tel: +41 216923994; Fax: +41 216923955; Email: David.Gatfield@unil.ch

© - The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

regression analysis to characterise 5' UTR features associated with DENR-dependent translational regulation and validate targets identified from our translatome-wide approaches. Finally, we show that DENR regulates efficient CLOCK biosynthesis involving a strong uORF and a so-far unannotated CDS initiation codon.

MATERIALS AND METHODS

Cell culture.

NIH3T3 and HEK293FT cells were cultured under standard conditions (DMEM; 10% FCS, 1% penicillin/streptomycin, all from Invitrogen; 37°C; 5% CO_2).

Lentiviral particle production in HEK293FT cells using envelope pMD2.G and packaging psPAX2 plasmids, and viral transduction of NIH3T3 cells, were performed following published protocols (10), with puromycin selection at 5 μg/ml for 4 days.

Cloning and plasmids.

For the generation of lentiviral shRNA expression vectors, two different sequences targeting *Denr* were cloned into pLKO.1puro backbone vector (Addgene no. 10878 (11)): shRNA1: GTACCACAGAAGGTCACGATA, corresponding to clone TRCN0000308443 of the TRC shRNA Library from the Broad Institute; and shRNA2: GTGCCAAGTTAGATGCGGATT, corresponding to clone TRCN0000098826. pLKO.1puro constructs containing *Gfp*, and Scramble (Addgene no. 1864) shRNAs served as controls.

For the generation of dual luciferase (Firefly/Renilla) reporter plasmids, fragments containing the 5' UTR and first 5-14 codons of the selected DENR targets/controls were amplified by PCR from genomic DNA or cDNA, cloned into the BamHI site of the prLV1 dual luciferase reporter plasmid (12), and validated by sequencing. The following primers were used for PCR: *Etaa1*, forward (for): aaaggatccGACTTGCAAAGATGGCGCTGCAC, reverse (rev): tttggatccactagtGTCCTTCAGCTGCATTCACATTT; *Map2k5*, for: aaaggatccGGTTCCGGAGTAACAGCGGTCTAAC, rev: tttggatccactagtGGCCAGCCACAGCATTACAGGTTAA; *Lrrc28*, for: aaaggatccGCCGCTGAGTGCCGGTCAGCGGGC, rev: tttggatccactagtGATTTCGGATGCCATGACTGAACA; *Klhdc8a*, for: aaaggatccGTCTCCGACCCTGTAGACACTGCAG, rev: tttggatccactagtATTGGGCACTTCCATGGCAGCCCGG; *Clock*, for: aaaggatccGGGGAGGAGCGCGGCGGTAGCGGTG, rev: tttggatccactagtAACAATTGAGCTCATTTTACTACAGCTTACGG; *Fth1*, for: aaaaggatcCAGACGTTCTCGCCCAGAGTCGC, rev: tttggatcccatGGTGGCGGCGGGGCGAGGCGC; *Rpl30*, for: aaggatcCAGCGGCGGGTGGACAGAGATGCC, rev: tttggatccacgCGTCTTCTTTGCGGCCACCATCTTC; *Deaf1*, for: aaggatccGGGCGTCAGACTGCCTCTTCGGTC, rev: tttggatccacgcgtGGCAGCGGCCGCCACAGCTGCCGC; *Cry1*, for: aaaaagatctactagtCACCCGGGCAGCCTCGGGACCG, rev: tttagatctTCGGAACCAGTGCACGGCGTTCAC.

5' UTR mutants were generated by site-directed mutagenesis of the corresponding plasmids, using the following primers (mutated nucleotides in small case with mutated start/stop codon underlined in the forward sequence): *Klhdc8a*-Mut, forward (for): GGCAGGGAGCCCGTGCAGCGCGgtaccGAGGCTGAGAGAGGGGACGCGCC, reverse (rev): GGCGCGTCCCCTCTCTCAGCCTCggtacCGCGCTGCACGGGCTCCCTGCC; *Map2k5*-Mut, for: CTCTTACTCACAAGGACTACAGgctagcTTGTGGGTTCTGTTTTGTCCAC, rev: GTGGACAAAACAGAACCCACAAgctagcCTGTAGTCCTTGTGAGTAAGAG; *Etaa1*-Mut, for: cgaaggatcCGACTTGCAAAGgtaccGCTGCACGGTGGCGCGCGGGCTC, rev: GAGCCCGCGCGCCACCGTGCAGCggtacCTTTGCAAGTCGgatccttcg; *Clock*-uORF-Stop-Mut, for: GTTATGGTGTTTACCGTcAGCTGTAGTAAAATGAG, rev: CTCATTTTACTACAGCTgACGGTAAACACCATAAC; *Klhdc8a*-uORF-Stop-Mut, for: GCCGACCGGCGCCGTGcacCCCGCGGGCCCCGGGGCT, rev: AGCCCCGGGGCCCGCGGGgtgCACGGCGCCGGTCGGC; *Etaa1*-uORF-Stop-Mut (small case correspond to an insertion that changes uORF frame), for:TGCTCCGCTGTAGCAgcAAATGTGAATGCAGC, rev: GCTGCATTCACATTTgcTGCTACAGCGGAGCA; *Clock*-uORF1-Mut, for: CCCAAAATCACCAGCAAGAGTTgctagcGTCAGTCACACAGAAGACGGCC, rev: GGCCGTCTTTCTGTGTGACTGACgctagcAACTCTTGCTGGTGATTTTGGG; *Clock*-uORF2-Mut, for: AAAGTGAAAGAGGAGAAGTACAggtacCTACCACAAGACGAAAACATAAT, rev: ATTATGTTTTCGTCTTGTGGTAGgtaccTGTACTTCTCCTCTTTCACTTT; *Clock*-AnnotCDS-Mut, for: CAAGACGAAAACATAATGTGTTtcGGTGTTTACCGTAAGCTGTAG, rev: CTACAGCTTACGGTAAACACCgaAACACATTATGTTTTCGTCTTG; *Clock*-AlterCDS-Mut, for: GTTTACCGTAAGCTGTAGTAAAcaGAGCTCAATTGTTactagtg, rev: cactagtAACAATTGAGCTCtgTTTACTACAGCTTACGGTAAAC.

iCLIP experiments and analyses.

For iCLIP experiments, we either used NIH3T3 cells without overexpression to immunoprecipitate the endogenous DENR, or NIH3T3 cells transiently transfected with *Denr* cloned into a FLAG-Venus-pLenti6 vector (Addgene 36948). Cells were grown to confluency in a 15-cm dish. One dish was used for overexpression, and six for non-transfected experiments.

iCLIP was performed as described (13, 14). RNA-protein complexes were crosslinked with UV-C (254 nm) at 150 mJ/cm^2 in a Stratalinker 2400. After cell lysis (50 mM Tris-HCl, pH 7.4; 100 mM NaCl, 1% Igepal (Sigma), 0.1% SDS, 0.5% Na deoxycholate, EDTA-free protease inhibitors), RNA was partially digested with 0.2 U RNase I (Thermo Fisher). DENR-RNA complexes were pulled down with anti-DENR antibody (Abcam 108221), pre-immobilised on protein G-coated magnetic beads (Dynal). An IgG immunoprecipitation was used as control. Immuno-complexes were washed twice with High Salt Buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1 mM EDTA, 1% Igepal, 0.1% SDS, 0.5% sodium deoxycholate) and twice with PNK Wash Buffer (20 mM Tris-HCl pH 7.4, 10 mM $MgCl^2$, 0.2% Tween-20). An infrared adaptor was ligated to RNA 3' ends, and protein-RNA complexes were visualised in a LiCor CLx Imager (Odyssey), after denaturing PAGE and nitrocellulose membrane transfer. 60-130 kDa DENR-RNA complexes were excised from the membrane and liberated by proteinase K (Roche) digestion. This size range likely reflects DENR-MCTS1 complexes (26 kDa and 22 kDa, respectively) bound to RNA, rather than DENR-RNA alone, indicating that the DENR-MCTS1

interaction is strong. During library preparation, cDNAs were ligated on their 5' end to a 5-nt sample barcode to allow demultiplexing. The barcode is preceded by a 4-nt and followed by a 3-nt random sequence that serves as unique molecular identifier (UMI) to mark individual cDNA molecules. cDNA libraries were then PCR amplified (24 cycles) and subjected to high-throughput sequencing (Illumina HiSeq2500 platform). Experiments were done in triplicate for DENR-overexpression, and in duplicate for non-transfected and IgG IP conditions.

iCLIP data was processed and analysed using the iCount platform (https://github.com/tomazc/iCount), as described (15). Briefly, the sample barcodes were removed, and the UMI was excised and added to the read ID prior to mapping with Bowtie 0.12.7 to the mouse genome (mm10), allowing two mismatches. Sequences with the same UMI and starting at the same genomic position – likely representing PCR duplicates – were excluded from subsequent analyses. High-confidence crosslink sites were calculated with the iCount function *peaks*, with options *RegionsAsOne* and 15-nt flank.

Ribosome profiling.

One 15-cm dish of confluent NIH3T3 cells was used per replicate. Ribosome profiling and parallel RNA-seq were performed in triplicate for *Denr* shRNA1, *Denr* shRNA2 and scr shRNA, and in duplicate for *Gfp* shRNA, following the ART-Seq protocol by Epicentre-Illumina (Cat.no. RPHMR12126) with minor modifications as described below. Cells were treated with cycloheximide (100 μg/ml) for 2 min at 37°C to arrest translation elongation, and immediately harvested, followed by lysate preparation. For the generation of ribosome-protected fragments, cell lysates were treated with 5 units RNase I (ART-Seq) per OD260 and monosomes were purified in MicroSpin S-400 columns (GE Healthcare). For both RPF- and RNA-seq, 5 μg of RNA were used for rRNA depletion following the Ribo-Zero Magnetic Kit instructions from Epicentre-Illumina (Cat.no. MRZH11124). Finally, libraries were amplified using 12 PCR cycles, and sequenced (HiSeq2500 platform).

Sequencing pre-processing, alignment and quantification.

Initial processing, mapping and quantification of mRNA and footprint abundance were performed as previously (16). Briefly, adaptor sequence was trimmed with Cutadapt (17) and trimmed sequences were filtered using a custom python script to keep 26-35 nt and 21-60 nt reads for RPF- and RNA-seq, respectively. Reads were then sequentially aligned to mouse rRNA, tRNA and cDNA (Ensembl mouse release 84) using Bowtie2 (18) and mouse genome using Tophat v2.0.11 (19). Trimmed and filtered sequences from RNA-seq were also directly mapped against the mouse genome with Tophat to estimate expressed transcript isoforms (using Cufflinks v2.2.1 (20)). Transcriptome-mapping reads from the sequential alignment were counted towards their location on 5' UTR, CDS or 3' UTR per gene, using only the set of expressed isoforms as estimated above. The location of the putative A-site of the footprint was used for RPF-seq counting (position +15 from the 5' for reads $\leq$30 nt and position +16 for reads >30 nt), and the 5' end was used for RNA-seq reads. For genes with multiple expressed isoforms, reads that did not map unambiguously to a single feature were assigned to one in preference order: CDS, 5' UTR, 3' UTR.

RPF- and RNA-seq read counts were normalized separately with the upper quartile normalization from edgeR (21) and then transformed to RPKM values as read count per gene kilobase per normalized library size. Genes with less than 10 reads in 2/3 of samples and mitochondrially-encoded genes were removed before normalization.

Translation efficiencies were then calculated per sample as the ratio of RPF RPKM to RNA RPKM of the region of interest (5' UTR, CDS or uORF), log2 transformed and averaged over all replicates.

Differential translation efficiency analysis.

Significant changes in TE between knockdown and control cells were assessed with several available algorithms: Xtail (22) with options *bins=10000*, RiboDiff (23) with options *-d 1 -p 1*, Riborex (24) with options *engine="DESeq2"*, and Babel (25) with options *nreps=100000*.

Briefly, all four methods, with the options implemented in our study for maximal consistency among them, model raw read counts by following a negative binomial distribution, and use global normalization methods for sequencing depth that would not account for global changes in TE if existing. RiboDiff and Riborex use DESeq2 to estimate a single mean and dispersion for both RPF- and RNA-seq, and detect differentially translated genes by fitting RPF- and RNA-seq read counts to a single GLM, whereas Xtail uses DESeq2 for mean and dispersion estimation separately for RPF- and RNA-seq and fits two independent GLMs. Xtail then derives probability distributions for fold changes of RNA or RPF, or for RPF-to-RNA ratios, and evaluates statistical significance between conditions from one of the two (the most conservative) pipelines. Babel relies on edgeR for parameter estimation and uses an errors-in-variables regression to test for significant changes in translation within and between conditions.

Differential TE was set at FDR < 0.1 for all four methods and performance for false positive discovery rate was further assessed by randomly label-shuffling knockdown and control datasets, which should not yield differential TE genes.

uORF annotation and translation analyses.

For uORF detection, a set of transcripts was used for which reads can be unambiguously assigned to the 5' UTR, i.e. transcripts that are the only protein-coding isoform expressed (N = 4637). In order to include as many transcripts as possible for the annotation, genes with multiple protein-coding isoforms were also considered if 1) one isoform represented $>$ 65% of all expressed (as derived from our transcript expression estimation, see "Sequencing pre-processing, alignment and quantification", N = 3211) or 2) all expressed isoforms had the same CDS start, in which case the transcript with the highest expression was selected (N = 1826). For the selected transcripts, uORFs were annotated and considered as translated with the following criteria: 1) started with AUG, CUG, GUG or UUG, 2) had an in-frame stop codon within the 5' UTR or overlapping/within the CDS, 3) were at least 9 nt long (including stop codon), 4) had a coverage $>$ 25%, 5) showed a frame preference and 6)

the preferred frame was the one corresponding to the uORF 5' end/start codon. Importantly, we considered only uORFs coding for at least 2 amino acids (i.e. 9 nt, including the stop codon), as for shorter uORFs frame bias assessment would be highly prone to false positives, and Kozak context scoring would be compromised (for 6 nt uORFs, the position +4 would always be the T of the stop codon, and thus cannot be a G of the strongest Kozak context).

Multiple regression analysis.

For the modelling, all transcripts containing translated uORFs in *Denr* knockdown or control cells were included (N = 4755) and the following 5' UTR features were used as predictors of the fold change in CDS TE across conditions: 1) the number of uORFs, 2) the uORFs length, using the median length in transcripts with several uORFs, 3) the proportion of AUG uORFs, 4) the uORF Kozak context, scored relative to the consensus GccA/Gcc###G, where upper-case letters denote highly conserved bases (scored +3), lower-case letters indicate the most common nucleotides (scored +1) and ### represents the start codon (not scored). The median Kozak score was used when several uORFs were present on a transcript. 5) the CDS Kozak context, scored in the same manner, 6) the distance of the first uORF start codon from the 5' cap, 7) the distance of the last uORF stop codon to the CDS start, 8) the uORF GC content, using the median for transcripts with several uORFs, 9) the presence of overlapping uORFs in the transcript (binary predictor). Predictors with skewed distributions (number of uORFs, uORF length, distance from the cap and distance to the CDS start) were log2 transformed. Partial regression plots were generated with the R package *'visreg'*. Absence of multicollinearity between predictors was assessed with the variance inflation factor using R function *vif* (*car* package).

Functional enrichment analysis.

Functional enrichment analysis was performed using g:Profiler (26), including Gene Ontology Biological Process and Molecular Function terms and biological pathway information from Reactome. Functional categories of more than 2 and less than 1500 members were used and p-values were corrected by Benjamini-Hochberg FDR. Categories with FDR < 0.05 were considered as significant and were used for visualization with Cytoscape v.3.5.1 EnrichmentMap tool (27).

Dual luciferase assay.

NIH3T3 cells were first lentivirally transduced with the dual luciferase reporters containing the 5' UTR construct of interest and then transduced with *Denr*-shRNA/scramble-shRNA. After puromycin selection, cells were lysed in passive lysis buffer and luciferase activities were quantified using DualGlo luciferase assay system and a GloMax 96 Microplate luminometer (all from Promega). Firefly/Renilla luciferase (FL/RL) activities were first normalized to that of the empty vector in scramble shRNA-transduced cells. For DENR target experiments, FL/RL signals were also normalized to the scramble shRNA-transduced, non-mutated construct. For *Clock* mutational experiments, FL/RL signals were secondly normalized to the scramble shRNA-transduced *Clock*-WT construct.

***In vivo* bioluminescence imaging.**

mRNAs were produced using in vitro transcription at the "ivt mRNA production and formulation platform" in Zurich (http://www.cancer.uzh.ch/en/Research/mRNA-Platform.html). The 5' end consisted of a CleanCapTM (Trilink) followed by the WT or mutated *Klhdc8a* 5' UTR and by a codon optimised firefly luciferase open reading frame. The 3' end consisted of a tandem repeat of the mouse beta-globin 3' UTR and a poly-A tail (28). Uridine residues were fully substituted by 1-methyl pseudouridine residues (this modification can prevent immunostimulation, avoiding production of type I interferon and consecutive inhibition of translation *in vivo* (28)).

Animal procedures were approved by the Veterinary Office and its research ethics review committee of the University of Zurich (Kanton Zurich, Health Direction, Veterinary Office, Zollstrasse 20; 8090 Zurich; license number ZH215/17). Animals were purchased from Envigo (Netherlands). 4 to 8-week-old mice were injected intravenously with 1 μg of mRNA formulated with TransIT transfection reagent following the manufacturer's suggestions (Mirus). The procedure for one mouse was as follows: the nucleic acid was diluted in 38 μl of Opti-MEM (Gibco) before the addition of 0.72 μl of the "mRNA Boost Reagent". After mixing, 1.12 μl of the TransIT reagent were added. The solution was thoroughly homogenated and injected immediately: 40 μl per mouse. At different time points after mRNA injection, *in vivo* bioluminescence imaging was performed on an IVIS Lumina instrument (PerkinElmer). Immediately before each measurement, D-luciferin (Synchem) dissolved in PBS (15 mg/ml stock) and sterile filtered was injected (150 μg/g intraperitoneally). Emitted photons from live animals were quantified 10-20 min post luciferin injections, with an exposure time of 3 min. Regions of interest (ROI) were quantified for average radiance (photons s-1 cm-2 sr-1) (IVIS Living Image 3.2).

Western blotting.

Total protein extracts from NIH3T3 cells were prepared according to the NUN procedure (29). Antibodies used were rabbit anti-CLOCK 1:3000 (ab3517, abcam), mouse anti-U2AF65 1:5000 (Sigma), rabbit anti-DENR 1:4000 (ab108221, abcam), and anti-rabbit-HRP or anti-mouse-HRP 1:10000 (Promega). Blots were visualised with the Fusion Fx7 system and quantified with ImageJ 1.48k.

RESULTS

Redistribution of ribosomes from CDS to 5' UTRs upon *Denr* knockdown

We hypothesized that if DENR is required for efficient protein production from the main CDS of mRNAs that undergo reinitiation after uORF translation, we would be able to identify DENR-regulated transcripts by comparing ribosome occupancies transcriptome-wide in DENR loss-of-function vs. control cells by ribosome profiling. Conceivably, DENR target mRNAs should exhibit decreased translation efficiencies (TEs) on their CDS when DENR is absent (Fig. 1A). We therefore performed shRNA-mediated knockdown

of *Denr* in mouse NIH3T3 fibroblasts, a cell line in which we had previously observed that *Denr* depletion led to a shortening of the period of circadian oscillations (9). Cells were lentivirally transduced with two different *Denr*-targeting shRNAs (shRNA1, shRNA2; in triplicate) or control RNAs (scramble shRNA, in triplicate; *Gfp* shRNA in duplicate), grown to confluency and harvested in the presence of cycloheximide (CHX) to stabilize elongating ribosomes. After preparation and sequencing of ribosome-protected mRNA fragment (RPF-seq) and total RNA libraries (RNA-seq) as previously described (9), we obtained ribosome footprints that were predominantly 29-30 nt in length (Supplementary Fig. 1A), that preferentially mapped to protein coding sequences (Supplementary Fig. 1B) and that showed the characteristic triplet periodicity expected from high-quality footprint data (Supplementary Fig. 1C). Across replicates, CDS-mapping normalised reads were highly correlated ($R^2 > 0.97$, Supplementary Fig. 1D), indicating good reproducibility. The RNA-seq and RPF-seq data confirmed the efficiency of *Denr* knockdown (77-90% reduction in *Denr* transcript levels), whereas the abundance of *Mcts1* mRNA, encoding the DENR heterodimerization partner, or of *Eif2d*, encoding a structurally and functionally related homologue of the DENR-MCTS1 dimer (4), remained unchanged (Fig. 1B, C and Supplementary Fig. 1E).

A reduction in *Denr* levels that alters reinitiation efficiencies should engender differences in ribosome occupancies on the CDS and, possibly, on uORFs. We first compared CDS footprint levels for all expressed genes (N = 10516) and noted no major global effects of *Denr* knockdown (Fig. 1D, grey). We then selected the subset of transcripts that also displayed footprint coverage within the 5' UTR (N = 5148). Of note, the quantification of footprints from the full 5' UTR can serve as a useful proxy for uORF translation (30, 31), given that uORFs are considered the major source of 5' UTR-mapping reads. This group of genes showed an overall similar distribution as all genes (Fig. 1D), albeit with a slight shift to lower levels of CDS-mapping footprints upon *Denr* knockdown ($p < 2.2e-16$, K-S test). We concluded that also for this subset of genes, the global effect on translational output was relatively small. We next calculated the ratio of 5' UTR-mapping to CDS-mapping footprints per gene, and determined how this measure of "relative 5' UTR translation" was affected in *Denr*-depleted vs. control cells. Our analysis revealed a striking increase in relative translation on 5' UTRs in *Denr*-deficient cells, indicative of a redistribution of translating ribosomes within transcripts (Fig. 1E and Supplementary Fig. 1F). This effect was absent from the RNA-seq data (Supplementary Fig. 1G). Conceivably, these changes on the translational landscape may reflect complex alterations in initiation efficiency and uORF usage, and are in line with a role of DENR in translation reinitiation (6).

Translational remodelling that affects a sizable proportion of the transcriptome is most consistent with a scenario in which DENR interacts with ribosomes at mRNAs with little, or no, sequence specificity. To address this question, we examined the direct interactions of DENR with RNAs *in vivo* by performing individual nucleotide resolution crosslinking and immunoprecipitation (iCLIP) (15) of DENR (Supplementary Fig. 2A). Over 85% of iCLIP reads mapped to ribosomal RNA, in particular to 18S rRNA (Supplementary Fig. 2B, C), with the five most prominent binding sites on helices 24-26a (Supplementary Fig. 2D, E). This outcome is in line with structural data that places the DENR-MCTS1 complex in proximity to the mRNA channel within the small ribosomal subunit and that identifies helix 24 as a potential interface of MCTS1 with 18S rRNA (32, 33). Moreover, helix 26 has previously been reported to make contacts with sequences upstream of restarting AUG codons on viral RNAs (34). In contrast, only 10.5% of iCLIP reads mapped to mRNAs, which was not sufficient to identify specific mRNA binding sites. In summary, while we cannot exclude transient direct interactions of DENR-MCTS1 with mRNAs, our iCLIP data are consistent with structural studies and support a model in which the complex binds close to the mRNA channel of initiating ribosomes at a location that could facilitate reinitiation when uORFs are present, but travels indiscriminately along mRNAs.

Decreased CDS translation efficiency, uORF enrichment, and pathway analysis of DENR-regulated transcripts

From the shift in the distribution of 5' UTR vs. CDS footprints that was affecting a significant proportion of the transcriptome in *Denr*-deficient cells (Fig. 1E), we wished to narrow down to the most-DENR-sensitive transcripts that were potentially direct and functionally relevant targets. We reasoned that changes in CDS translation efficiency (TE; defined as the ratio of CDS-mapping RPF-seq to RNA-seq reads) would represent an appropriate measure to identify such targets, as TE differences would alter the amount of protein produced from the transcript. We therefore calculated transcriptome-wide TEs and, given that there is currently no established consensus on the most suitable analytical tool for differential translation analysis, we then applied several available algorithms to detect significant TE changes. Briefly, the four methods we employed, Xtail (22), RiboDiff (23), Riborex (24), and Babel (25), are all conceived for the quantification of translational changes that deviate from possible variations already occurring at the level of mRNA abundances. They differ in parameter estimation and modeling approaches, as well as in sensitivity, false discovery rates and computational speed. When applied using settings that maximize comparability of results (see Materials and Methods), we identified 240, 78, 41 and 30 genes that had a significant change in TE between *Denr*-deficient and control cells with Xtail, RiboDiff, Riborex, and Babel, respectively (FDR < 0.1 for all four algorithms). Importantly, differential TE genes detected by Riborex, RiboDiff and Xtail completely overlapped, while only about half of the genes detected by Babel were also found by any of the other methods (Supplementary Fig. 3A). Moreover, most genes detected by Xtail, RiboDiff and Riborex showed decreased TE in *Denr*-depleted cells, as expected for the knockdown of a reinitiation factor, and they spanned the full range of expression levels (Fig. 2A, and Supplementary Fig. 3B, C). By contrast, Babel was biased towards low abundance transcripts and identified similar numbers of up- and down-regulated TE genes (Supplementary Fig. 3B, C). We concluded that Xtail likely represented the most sensitive method to identify DENR targets, and further estimated the false discovery rate of the algorithm by randomly shuffling the dataset labels, which

eliminated nearly all differential TE genes (Supplementary Fig. 3D). Please note that, for simplicity, we shall further use the term "DENR targets" for the 240 differential TE genes identified by Xtail, although in the absence of information on direct DENR-mRNA interactions, the term *"most DENR-sensitive transcripts"* would be more precise.

Among the 240 differential TE genes identified by Xtail (see Supplementary file 1) were several for which the presence of regulatory uORFs has been described previously, such as *Atf5* (35), *Rnaseh1* (36, 37), *Jak2* and *Map2k6* (38). In order to investigate whether the TE decrease among DENR-sensitive transcripts was associated with uORF translation, we annotated uORFs transcriptome-wide. Briefly, we included AUG and non-canonical, near-AUG (i.e., CUG, GUG, UUG) initiation codons located in 5' UTRs, and assessed the active translation of the corresponding ORFs by their footprint coverage, frame preference and triplet codon periodicity (Supplementary Fig. 4A). Using this approach, we detected translated uORFs in nearly half of all expressed genes (N = 4755 genes, Fig. 2B), consistent with previous predictions (8). For the majority of genes, the uORFs were fully contained in the 5' UTRs (51.8%, N = 2464 genes), but genes with uORFs overlapping the CDS (19.5%, N = 926 genes), or that had both overlapping and non-overlapping uORFs (28.7%, N = 1365 genes) were frequent, too (Supplementary Fig. 4B). Notably, DENR targets were highly enriched for genes with translated uORFs (72.7%, p = 5.73e-4, Fisher test, Fig. 2B) that were non-overlapping uORFs (Supplementary Fig. 4B). The depletion of overlapping uORFs (Supplementary Fig. 4B) – indicating that they are insensitive to DENR regulation – is consistent with the idea that after terminating on an overlapping uORF, a ribosome would presumably not be able to backtrack to reinitiate on an upstream CDS start codon. Moreover, relative 5' UTR translation in *Denr*-depleted vs. control cells calculated over the full 5' UTR – as a proxy of uORF translation – was significantly higher for DENR targets than for the global set of uORF-containing genes (Supplementary Fig. 4C, D). In summary, these observations were fully consistent with an involvement of DENR in reinitiation and in efficient CDS translation of a subset of uORF-containing mRNAs.

We next analysed whether specific functional pathways were over-represented among the 240 differential TE genes. GO term analysis revealed enrichment of categories associated with translation and mRNA stability, as well as cellular functions related to protein phosphorylation and catabolism, kinase binding and cell cycle (Fig. 2C and Supplementary Fig. 4E), in line with previous observations in *Drosophila* S2 cells (6) and indicating evolutionary conservation of DENR functions.

Among the DENR-regulated transcripts were several relevant to cancer, such as *Klhdc8a*, *Etaa1*, *Map2k5*, *Slc20a1* or *Vegfd* (39, 40, 41, 42, 43). Individual inspection confirmed their reduced CDS translation efficiency (Fig. 3A), as well as a shift of ribosome occupancy from the CDS to the 5' UTR in *Denr*-depleted cells (Fig. 3B). This signature was characteristic of many differential TE genes (see Supplementary Fig. 5A, B), and correlated with the presence of translated uORFs (Supplementary Fig. 5C). Among the strongly DENR-regulated cases, we also noted *Atf5* (Supplementary File 1), whose translation (similar to that of *Atf4*) is known to respond to cellular stress signals via two uORFs, the first of which is reinitiation-permissive whereas the second overlaps the main CDS (35, 44). In *Denr*-depleted cells, translation of the first uORF (on *Atf5* and, to lesser extent, on *Atf4*) remained unchanged compared to control cells, while translation of the second uORF and the CDS were decreased (Supplementary Fig. 5D, E). These findings validate that our approach indeed delivered reinitiation-prone transcripts.

In order to confirm the DENR- and uORF-dependence of CDS translation for selected presumed targets, we cloned the 5' UTR and immediate start codon context of several transcripts of interest in frame and upstream of Firefly luciferase (FL) in a lentiviral vector that also expresses *Renilla* luciferase (RL, for internal normalisation) from the same bidirectional promoter (Fig. 3C, upper panel). *Denr* depletion led to a 18-44% reduction in FL/RL signals for five different control reporters that contained either no cloned 5' UTR, or 5' UTRs of genes that were not among the differential TE genes (of these, *Cry1* and *Fth1* harbour translated uORFs as well, whereas *Deaf1*, and *Rpl30* do not) (Fig. 3C, lower panel). By contrast, we observed a stronger reduction in FL/RL signals – of 49% to 90% – for four 5' UTR reporters of presumed DENR target genes (*Etaa1, Klhdc8a, Map2k5* and *Lrrc28*), indicating that DENR was indeed involved in their regulation. Mutating uORF start codons increased reporter signal and abolished most DENR-dependence (Fig. 3C, bottom), indicating that uORF translation was necessary for DENR dependance. Furthermore, in line with the view of DENR as a reinitiation factor, rather than a uORF skipping factor, the presence of a functional stop codon was required for DENR regulation (Supplementary Fig. 5F). uORF-dependent regulation of *Klhdc8a* translation was particularly striking, with a >4-fold translation upregulation in WT cells with a uORF-less construct (Fig. 3C, bottom). We next wished to investigate whether uORF-mediated translational regulation could also be recapitulated *in vivo* and selected *Klhdc8a* for this experiment, as this mRNA contained the strongest, reinitiation-dependent uORF in the cellular assay. We injected mice intravenously with *in vitro*-synthesized mRNAs comprising the *Klhdc8a* 5' UTR (WT or uORF start codon-mutated) fused to Firefly luciferase CDS and carrying a synthetic polyA tail. We then monitored expression by bioluminescence live imaging in mice. Bioluminescence signal was primarily detected in the livers of the animals, and was 2 - 4.3 times higher in the uORF-less compared to the WT constructs (Supplementary Fig. 6A, B), demonstrating the strong inhibitory activity of the *Klhdc8a* uORF also *in vivo*.

Taken together, we concluded that DENR was involved in the regulation of protein biosynthetic output, and that a set of >200 transcripts containing uORFs was particularly sensitive to loss of this reinitiation factor.

Linear regression analysis correlates DENR-dependence with uORF strength and with intercistronic distance

uORFs can operate through a bewildering diversity of mechanisms that act on CDS translation (by affecting start site selection/leaky scanning/re-initiation/ribosome shunting), transcript abundance (by eliciting nonsense-mediated mRNA decay), or in some cases through uORF-encoded bioactive

peptides (reviewed in (3)). However, rules to predict specific mechanisms from specific uORF features have not yet been established. Our datasets could potentially enable the identification of signatures that render certain uORF-containing transcripts DENR-dependent. Previous studies have suggested that, at least in the context of synthetic reporters in *Drosophila* S2 (6) and human HeLa (7) cells, very short uORFs (coding for 1-2 amino acids) with favourable Kozak contexts were most strongly dependent on DENR for efficient reinitiation at a downstream reporter CDS.

We used multiple regression analysis to determine whether DENR-dependent translation events were associated with specific transcript features. We modeled the CDS TE change observed upon *Denr* knockdown using nine relevant transcript features, namely the number of uORFs, uORF start codon identity (AUG vs. near-cognate) and context, the length, GC content and location of the uORF within the 5' UTR, the presence of overlapping uORFs, and CDS start codon context (Fig. 4A). The model was applied to all translated uORF-containing genes (N = 4741) and yielded several features with significant predictive value for DENR-sensitivity of TE (Fig. 4B-J; adjusted R^2 = 0.057, F-statistic = 33.21 on 4731 DF, $p <$ 2.2e-16). Three of the identified features reflected the expected positive relationship between efficient uORF translation and a greater need for reinitiation to subsequently translate the CDS as well. These features were: the number of uORFs present on a transcript (Fig. 4B, b = -0.019, $p <$ 2e-16), the presence of AUG, rather than non-canonical uORF start codons (Fig. 4C, b = -0.05, $p <$ 1.15e-9), and the strength of Kozak contexts around uORF start codons (Fig. 4D, b = -0.0026, p = 0.013). Relatedly, DENR targets were enriched for AUG uORFs as compared to the ensemble of translated uORF-containing transcripts (p = 0.0044, Fisher test, Supplementary Fig. 7A), and genes harboring exclusively AUG uORFs showed the strongest effect on CDS TE in *Denr*-deficient cells (Supplementary Fig. 7B). By contrast, the Kozak context around the CDS start codon bore no predictive value (Fig. 4E, b = -0.0006, p = 0.38), suggesting no DENR involvement in its recognition. Moreover, the presence of CDS-overlapping uORFs was not associated with DENR-dependent TE change (Fig. 4F, b = 0.005, p = 0.44), as expected for a reinitiation factor, given that overlapping uORFs would be reinitiation-incompetent unless the transcripts contained a downstream alternative CDS start codon.

Two noteworthy outcomes from the model were that uORF length lacked a relationship (Fig. 4G, b = 0.0006, p = 0.81), whereas the uORF-CDS intercistronic distance had predictive value for DENR-sensitivity of CDS TE (Fig. 4H, b = -0.0101, p = 7.89e-7). Briefly, already early studies on a model transcript had suggested that the location of a uORF with respect to the main ORF could dictate the efficiency of reinitiation, as longer spacing would generally allow the reacquisition of initiation factors to the scanning ribosome (45). Our model provides an unbiased and transcriptome-wide validation of this notion for DENR-regulated transcripts, where larger distances from the stop codon of the 3'-most uORF to the start codon of the main ORF were associated with downregulation of TE (Fig. 4H); uORFs located further upstream from the CDS were thus more likely to reinitiate in a DENR-dependent fashion. Importantly, this effect of uORF-CDS distance was not caused by longer 5' UTRs *per se*, given that the distance from the 5' cap to the first uORF was not associated with DENR regulation (Fig. 4I, b = -0.0018, p = 0.18). Of note, the uORF-CDS distance was only weakly correlated with 5' UTR length (Spearman rho = 0.16), whereas the correlation of the 5' cap-first uORF distance with 5' UTR length was more prominent (Spearman rho = 0.27, Supplementary Fig. 7C, D).

No linear relationship was found between uORF length and CDS TE (Fig. 4G, b = 0.0006, p = 0.81), in agreement with previous results in human cells using synthetic uORFs encoding >2 amino acids (9 nt) (7). Of note, a length dependence for extremely short uORFs of 1-2 amino acids, as reported by Schleich et al. 2017 (7), would go unnoticed in our analyses, as we excluded them due to the uncertainties associated with their annotation (see Materials and Methods).

Finally, the uORF GC content showed a significant relationship with TE fold change (Fig. 4J, b = 0.058, p = 0.005), suggesting that potential weaker secondary structures (low GC) favoured reinitiation involving DENR.

For the above model, we used the CDS TE fold change as response, which has the benefit of accounting for DENR regulation in a continuous manner. Of note, however, also a logistic regression model based on a binary response (i.e. DENR target as defined in Fig. 2 vs. non-target) recapitulated the main findings (Supplementary Fig. 7E), further supporting our predictions for the strongest cases of DENR translational control. In summary, although the model explained only a relatively modest proportion of the total variance (see Discussion), our results indicated that DENR was important to enable efficient protein synthesis when the transcripts contained uORFs initiating with AUG in strong Kozak contexts (i.e. uORFs expected to be highly translated), and that the distance between the most downstream uORF and the CDS start was an important parameter for efficient reinitiation.

DENR-mediated reinitiation regulates CLOCK protein biosynthesis

We recently found that uORF translation was pervasive within circadian core clock transcripts, and that the depletion of *Denr* in NIH3T3 cells engendered robust circadian period shortening by 1-1.5 hours (9). We therefore sought to uncover which core clock transcripts were DENR-regulated, and to map the relevant uORFs. Conceivably, these would represent likely contributors to the observed circadian phenotype.

The expression of several core clock genes was altered in *Denr*-depleted vs. control cells. In most cases (*Per1/2/3*, *Cry1/2*, *Nr1d1*, *Ror*α) the observed changes occurred already at the level of mRNA abundance and subsequently led to comparable changes at the footprint level as well (Fig. 5A, and Supplementary Fig. 8A). None of these transcripts (nor *Bmal1*, whose lower RNA abundance was partially compensated at the footprint level) was associated with marked TE reduction (Fig. 5B, and Supplementary Fig. 8B), and we thus considered them unlikely candidates for DENR targets. For *Clock*, however, the reduction in expression was mostly translational, resulting in a $\sim$20% decrease in protein biosynthesis, as quantified from *Clock* RPFs (Figs. 5A,B, and Supplementary Figs. 8A,B). In our initial analysis of differential TE (Fig. 2A), *Clock* had, however, not passed the significance threshold (Supplementary File 1). The reason

was that the applied algorithm, Xtail, evaluates statistical significance of TE differences through two separate tests – one regarding differences between the RNA and the RPF fold-change distributions, and a second one evaluating differences in TE distributions directly – and then derives the final p-value from the most conservative of the two tests (22). While the additional test might prevent false positives, it can discard differential TE genes close to the statistical threshold. Importantly, for one of the tests, the TE fold change for *Clock* was significant (p = 0.025 in test 1; p = 0.19 for test 2; Supplementary File 1); *Clock* thus may represent a DENR target whose detection is at the borderline of statistical significance.

We annotated two translated AUG-initiated uORFs in *Clock* (Fig. 5C, and Supplementary Fig. 8C-E). The second uORF overlapped with the main CDS in a different frame – a configuration that would render the two translation events (uORF-2 and annotated CDS) mutually exclusive. However, downstream of the uORF-2 stop codon and 10 amino acids into the annotated *Clock* CDS, we noted an additional AUG in the *Clock* coding frame. We thus hypothesised that, when uORF-2 was translated, initiation at this alternative AUG could still allow for the biosynthesis of a CLOCK protein variant that was N-terminally truncated by 9 amino acids (Fig. 5C). In *Denr*-deficient cells, translation was increased on both uORFs (Fig. 5C-D) and, in particular, uORF-2 showed ∼23% higher translation rate upon *Denr* depletion (Fig. 5E), suggesting that DENR-mediated reinitiation after uORF-2 was crucial for the translation of the *Clock* CDS.

Using GWIPS-viz, an online genome browser for the visualisation of ribosome profiling data from published studies (46), we looked for further evidence of uORF-2 and alternative CDS translation across mammalian datasets (of note, both annotated and alternative methionine residues are conserved across mammals, see Supplementary Fig. 9). Interestingly, initiating ribosomes can be detected on both uORF-2 and alternative CDS start codons (but not the annotated *Clock* start codon) across human and mouse datasets, and elongating ribosome profiling studies showed good footprint coverage on uORF-2 and the characteristic accumulation of ribosomes at the start and stop codons (Supplementary Fig. 9). Together, the published and our new data thus suggested abundant translation of uORF-2 and of *Clock* CDS from the alternative AUG, indicating that reinitiation has an important role in ensuring CLOCK protein biosynthesis.

In order to investigate uORF usage, effect, and DENR-dependence for CLOCK biosynthesis from the annotated and alternative CDS, we cloned the *Clock* 5' UTR and first 14 codons (containing both the annotated and alternative AUGs) in frame and upstream of *Firefly* luciferase in our dual luciferase vector (Fig. 5F, upper panel). We then systematically mutated the uORF-1, uORF-2, and CDS annotated/alternative start codons, or combinations thereof (Supplementary Fig. 8F), and compared the effects on luciferase output in *Denr*-depleted vs. control cells. Of note – and in contrast to Fig. 3C – we also adjusted for the effect of *Denr* depletion on the control reporter (whose levels were set to 1) to display the isolated effects of the *Clock* variants.

Denr depletion resulted in an almost 30% reduction in expression for the *Clock WT* construct (Fig. 5F, lower panel), in line with the decrease in relative translation quantified for the endogenous transcript (Fig. 5D) and with reduced levels of endogenous CLOCK protein (Supplementary Fig. 10). Surprisingly, mutation of the annotated CDS initiation codon did not substantially affect reporter levels and DENR dependence when compared to *Clock WT*. However, mutation of the alternative AUG decreased luciferase signal by ∼95% and abolished DENR dependence (Fig. 5F, lower panel), suggesting that most CLOCK biosynthesis in NIH3T3 cells involves the alternative AUG. Mutation of uORF-1 led to a 19% increase in luciferase signal and fully retained DENR dependence. By contrast, uORF-2 mutation led to a ∼50% increase in protein synthesis and eliminated DENR regulation. These findings indicated that uORF-2 translation acts to significantly reduce CLOCK biosynthesis and to render it DENR-dependent. This model was further supported by the combinatorial mutations. Mutation of both uORF-1+2 start codons thus resulted in an additive, >90% increase in protein synthesis, and in a loss of DENR-dependence. In line with minor levels of synthesis of the long CLOCK isoform, mutation of the annotated CDS start codon in combination with uORF mutations led to similar results as mutation of the uORFs alone. Furthermore, mutation of the alternative AUG in combination with uORF mutations led to strongly decreased luciferase signals, consistent with the predominant synthesis of the short CLOCK isoform. In particular, the case where both the alternative CDS and uORF-2 (or uORF-1+2) start codons were mutated – a scenario that would allow unhindered CLOCK synthesis from the annotated AUG – resulted in a >80% reduction in CLOCK protein production (Fig. 5F, lower panel). Taken together, we concluded that *Clock* required DENR for efficient protein synthesis from a CDS that is 9 codons shorter than the annotated one, and that uORF-2 plays a key role in coordinating the complex interactions that occur at the translational level on the *Clock* transcript (Fig. 5G). Moreover, we propose that the CLOCK N-terminus is likely wrongly annotated not only in NIH3T3 cells, but based on our analysis of published ribosome profiling data, also in other cell types and in tissues (Supplementary Fig. 9).

DISCUSSION

The annotation of translated uORFs has remained a particular challenge, because they are short, frequently initiate from non-AUG codons, and do not obey the same rules of evolutionary conservation as most protein coding sequences (3). This has changed with the availability of ribosome profiling data, which has made it possible to systematically identify translated uORFs from footprints. The resulting burst in interest in uORFs and the rising appreciation of how prevalent their usage is in mammals, is, however, not yet matched by an understanding at the mechanistic level. In this study, we have investigated one particularly intriguing mechanism that is implicated in the regulation through uORFs, namely translation reinitiation. As an entry point, we have used the loss-of-function of DENR, a non-canonical initiation factor recently associated with reinitiation both in *Drosophila* (6) and in HeLa cells (7), with the aim of identifying rules that explain why certain uORFs trigger DENR-dependent reinitiation while others do not. The conclusions we were able to draw from experiments in mouse NIH3T3 fibroblasts showed similarities and striking differences to the previous

studies in fly and in human cells (6, 7). Species-specificity in uORF biology may explain some of the disparities, but differences in experimental and computational approaches likely have an important influence as well.

We view our quantitative modeling approach of DENR-dependent changes in CDS TE as particularly unbiased as it relies on translational changes detected on endogenous mRNAs transcriptome-wide. Several features carry predictive value for the magnitude of DENR-mediated CDS TE regulation. Of these, three (number of uORFs, start codon identity, start codon sequence context) likely reflect a straightforward logic between a stronger requirement for reinitiation when the usage of uORFs is efficient. This outcome is expected and fully consistent with human (7) and fly data (6) based on reporter assays. Two other outcomes from the model are noteworthy. First, uORFs located further away from the main ORF are more prone to trigger DENR-dependent reinitiation. Such intercistronic distance effects were reported early on in the uORF field, typically using distinct uORF-containing model transcripts (45, 47, 48). They had generally been assumed to reflect the kinetics of association and dissociation of initiation and termination factors. Our data validate this idea in a transcriptome-wide fashion and are in line with the hypothesis that longer distances allow for extended re-scanning times, which would be favourable for the loss of termination factors and/or the re-acquisition of the initiation machinery, on time for CDS translation. Second, it is noteworthy that uORF length lacks predictive value in our model. The hypothesis that increased uORF length impairs reinitiation efficiency hence does not appear valid, at least for DENR-regulated reinitiation events. As a caveat, the nature of our approach (transcriptome-wide from endogenous transcripts) did not allow to confidently annotate uORFs down to only 1-2 amino acids in length. We may thus be missing DENR effects that have recently been observed in reporters carrying such ultra-short synthetic uORFs (7). We can evidently also not exclude that the time required for uORF translation in determining scanning resumption is relevant nevertheless, as linear length *per se* may simply not be a good proxy for translation time (2, 8, 49).

Overall, our model explains a relatively modest proportion of the total variance (adjusted R^2 of 0.057). This is reminiscent of analogous models that have been developed to predict, for example, miRNA target sites, whose adjusted R^2s range from 5-13%, but that have nevertheless proven biologically useful (50). In our case, the presence of multiple uORFs on a transcript, often overlapping each other, and each with its own, distinct features (start codon, length, distance to the CDS, etc), and the resulting combinatorial effects, clearly complicated their scoring and likely masked some of the origins of translational regulation. We deliberately wished to keep our model simple, aiming at identifying the most significant predictors, but at the expense of overall explanatory power. In the future, the use of other regression types and models that include interaction terms could potentially explain higher biological variability, although at the cost of becoming considerably more difficult to interpret. Overall, our study constitutes a first attempt to analyse signatures associated with reinitiation through DENR. It was already able to recapitulate previous observations (6, 7), and as we continue to identify more molecular features that can regulate non-canonical translation initiation and that can be integrated into the model in the future, a larger proportion of biological variance will be explained.

We have identified 240 transcripts as particularly DENR-sensitive. Although we lack formal proof that they are direct DENR targets, the observations that (1) most of them show TE downregulation and few upregulation in the absence of DENR, that (2) they are depleted for CDS-overlapping and enriched for non-overlapping uORFs, and that (3) a DENR-dependent inhibitory effect of their 5' UTRs can be rescued by uORF start codon mutations in reporter assays, all make a direct activity of DENR on these transcripts very likely. Moreover, these observations are all compatible with a role of DENR in reinitiation, but not in uORF skipping (which could have been an alternative interpretation of the redistribution of ribosomes from CDS to 5' UTRs seen in Fig. 1E). In particular, we would not expect CDS-overlapping uORFs to be depleted (Supplementary Fig. 4B) if DENR acted to regulate targets via uORF skipping. Moreover, mutating the uORF stop codon should not have had an influence on the DENR-effect in our reporter assays, if the activity of DENR consisted in regulating the efficiency of uORF start codon selection (Supplementary Fig. 5F).

Many identified DENR targets are involved in proliferation and associated with cancer. Of note, DENR-MCTS1 have themselves been identified as oncogenes (51) that are upregulated, for example, in prostate cancer (52). Moreover, translational reprogramming, especially with regard to upstream initiation sites, appears to represent an important cancer driver (30), and understanding how the translation of cancer-relevant DENR targets is regulated may thus lead to important insights of diagnostic or therapeutic value. We found, for example, very strong DENR requirement for the translation of Kelch Domain Containing 8A (KLHDC8A). It has been reported that KLHDC8A overexpression in human gliomas that become resistant to epidermal growth factor receptor (EGFR) silencing (a commonly used anticancer therapeutic approach), allows tumours to maintain aggressiveness through an unknown EGFR-independent pathway (39). It is conceivable that upon loss of EGFR sensitivity, KLHDC8A levels increase through DENR-mediated translational control. Similarly, other candidates from our DENR target list (Supplementary file 1) are upregulated in (and possibly causally contributing to) different cancers, such as Vascular endothelial growth factor D (VEGF-D), or Mitogen-Activated Protein Kinase 5 (MAP2K5) (41, 43). By understanding how DENR regulates its targets, it may thus be possible to reduce reinitiation activity and thus the oncogenic levels of the modulated proteins in proliferating, but not in quiescent cells.

A comparison of DENR targets found in our and in previous studies returned little overlap. Of the eight targets validated in HeLa cells (7), only one (*Tmem60*) was expressed in NIH3T3 cells. *Tmem60* was predicted to undergo a 25% translational downregulation upon *Denr* loss-of-function and indeed showed a decrease in that range in reporter assays in HeLa cells (7), whereas in our system a $\sim$8% decrease in CDS TE of endogenous *Tmem60* was quantified. Moreover, for the ensemble of 104 predicted targets in human cells (7), $>$90% were not expressed in NIH3T3 cells or lacked a mouse orthologue; others (e.g. *Lca5*, *Mllt11* or *Dsel*) were

indeed downregulated upon *Denr* knockdown, though to a lesser extent. Conceivably, DENR targets are species- and cell type-specific; moreover, it is likely that the differences in the magnitude of translational regulation observed across studies are the result of different experimental and analytical approaches.

How do DENR and circadian rhythms mechanistically connect? The DENR target list did not reveal core clock genes that would explain the circadian clock phenotype, i.e. the previously reported ∼1.5-hour period shortening of free-running oscillations (9). Individual inspection of core clock gene transcripts, however, suggested that *Clock* was a DENR target just below the significance threshold. Follow-up experiments confirmed *Clock* regulation through DENR and allowed mapping the responsible uORF. These findings establish uORF translation and re-initiation as a novel, potentially regulated mechanism acting on core clock gene expression, which opens several intriguing questions. First, is it possible that reduced biosynthesis of *Clock* would be responsible for the observed phenotype? Loss-of-function of a general gene expression machinery component typically goes along with sizeable distortions in global RNA and protein profiles, and *Denr* knockdown is no exception in this respect. Any phenotype thus likely represents a composite effect, rather than the result of a single misregulated gene expression event. We therefore do not view reduced CLOCK biosynthesis as a satisfactory hypothesis that alone would be able to explain the substantial period shortening. Nevertheless, it is reassuring that $Clock^{-/-}$ fibroblasts have recently been shown to have shorter periods as well (53), which suggests that downregulation of *Clock* translation upon *Denr* depletion may be a factor contributing to the phenotype.

A second intriguing question relates to the alternative initiation site that would generate an N-terminally truncated variant of CLOCK. We propose that the 9 amino acid shorter CLOCK is the main variant produced in NIH3T3 cells, and likely in other cell types and species as well. What are the functional implications? The bHLH domain responsible for DNA binding of CLOCK starts at amino acid 26, and the very N-terminus has not yet been associated with any function of the protein and is also not included in the crystal structure of the CLOCK-BMAL1 complex (54). However, it is noteworthy that for several myogenic bHLH transcription factors, N-terminal motifs are critical for cooperative DNA binding with cofactors (55). Much of our current knowledge on CLOCK stems from *in vivo* and *in vitro* experiments that have used *Clock* cDNA constructs devoid of the endogenous uORF regulation, and thus defining CLOCK's N-terminus to the annotated one. It will be exciting to re-evaluate selected DNA-binding and activity assays using both CLOCK isoforms side by side.

Finally, a speculative question to ask is whether DENR-mediated control is a regulated, rather than constitutive step in CLOCK biosynthesis, serving for example to uncouple CLOCK protein abundance from mRNA levels in specific cell types or under certain physiological conditions. Of note, there is precedence from *Neurospora crassa* for uORF translation that mediates temperature-dependent regulation of biosynthesis of the core clock component FRQ (56). In the case of DENR, it is interesting that this protein was originally identified because its expression increased in cultured cells at high density, but not during growth arrest *per se*, which led to its name Density Regulated Protein (57). First links between cell density and circadian rhythmicity have indeed been described in recent years (58, 59), and it would be interesting to assess whether CLOCK uORFs and DENR play roles in determining density-dependent differences of clock parameters.

DATA AVAILABILITY

High-throughput sequencing data was deposited to GEO, under accession number: GSE124793 (SuperSeries composed of ribosome profiling data: GSE116221 and iCLIP data: GSE124790). Code for ribosome profiling pre-processsing, aligning and quantification, and for uORF annotation, is available online at http://doi.org/10.5281/zenodo.521199.

SUPPLEMENTARY DATA

Supplementary data is available.

FUNDING

DG acknowledges support by: Swiss National Science Foundation [grants 157528, 179190], National Centre of Competence in Research (NCCR) RNA & Disease; Olga Mayenfisch Stiftung; Foundation Herbette; University of Lausanne. SP acknowledges support by the University of Zurich (URPP "Translational Cancer Research").

ACKNOWLEDGEMENTS

We thank the Lausanne Genomics Technologies Facility for high-throughput sequencing infrastructure. We thank Angelica Liechti and Georgia Katsioudi who were involved in initial steps of reporter plasmid clonings. We are grateful to Igor Ruiz de los Mozos for assistance with the iClip analyses through iCount.

Conflict of interest statement. None declared.

REFERENCES

1. Jackson R. J. , Hellen C. U. T. , and Pestova T. V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol*, 11(2):113–127.
2. Barbosa C. , Peixeiro I. , and Romao L. (2013) Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. *PLOS Genetics*, 9(8):e1003529.
3. Wethmar K. (2014) The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip Rev RNA*, 5(6): 765–778.
4. Skabkin M. A. , Skabkina O. V. , Dhote V. , Komar A. A. , Hellen C. U. T. , and Pestova T. V. (2010) Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev.*, 24 (16):1787–1801.
5. Young D. J., Makeeva D. S., Zhang F., Anisimova A. S., Stolboushkina E. A., Ghobakhlou F., Shatsky I. N., Dmitriev S. E., Hinnebusch A. G., and Guydosh N. R. (2018) Tma64/eIF2d, Tma20/MCT-1, and Tma22/DENR Recycle Post-termination 40s Subunits InVivo. *Molecular Cell*, 71(5):761–774.e5.
6. Schleich S. , Strassburger K. , Janiesch P. C. , Koledachkina T. , Miller K. K. , Haneke K. , Cheng Y.-S. , Kuchler K. , Stoecklin G. , Duncan K. E. , and Teleman A. A. (2014) DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. *Nature*, 512 (7513):208–212.
7. Schleich S. , Acevedo J. M. , Hohenberg K. C. v. , and Teleman A. A. (2017) Identification of transcripts with short stuORFs as targets for DENR-MCTS1-dependent translation in human cells. *Scientific Reports*, 7(1):3722.
8. Calvo S. E. , Pagliarini D. J. , and Mootha V. K. (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. U.S.A.*, 106(18): 7507–7512.
9. Janich P. , Arpat A. B. , Castelo-Szekely V. , Lopes M. , and Gatfield D. (2015) Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames. *Genome Res.*, 25(12):1848–1859.
10. Salmon P. and Trono D. (2007) Production and titration of lentiviral vectors. *Curr Protoc Hum Genet*, Chapter 12:Unit 12.10.
11. Moffat J. , Grueneberg D. A. , Yang X. , Kim S. Y. , Kloepfer A. M. , Hinkle G. , Piqani B. , Eisenhaure T. M. , Luo B. , Grenier J. K., Carpenter, A. E., Foo, S. Y., Stewart, S. A., Stockwell, B. R., Hacohen, N., Hahn, W. C., Lander, E. S., Sabatini, D. M., and Root, D. E. (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell*, 124(6):1283–1298.
12. Du N.-H. , Arpat A. B. , De Matos M. , and Gatfield D. (2014) MicroRNAs shape circadian hepatic gene expression on a transcriptome-wide scale. *Elife*, 3:e02510.
13. Huppertz I., Attig J., D'Ambrogio A., Easton L. E., Sibley C. R., Sugimoto Y., Tajnik M., Konig J, and Ule J. (2014) iCLIP: Protein-RNA interactions at nucleotide resolution. *Methods*, 65(3):274–287.
14. Zarnegar B. J., Flynn R. A.,Shen Y, Do B. T., Chang H. Y., and Khavari P. A. (2016) iCLIP platform for efficient characterization of protein-RNA interactions. *Nat Meth*, 13(6):489–492.
15. Konig J., Zarnack K., Rot G., Curk T., Kayikci M., Zupan B., Turner D. J., Luscombe N. M., and Ule J. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol*, 17(7):909–915.
16. Castelo-Szekely V. , Arpat A. B. , Janich P. , and Gatfield D. (2017) Translational contributions to tissue specificity in rhythmic and constitutive gene expression. *Genome Biology*, 18:116.
17. Martin M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1):pp. 10–12.
18. Langmead B. and Salzberg S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Meth*, 9(4):357–359.
19. Trapnell C. , Pachter L. , and Salzberg S. L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, 25(9):1105–1111.
20. Trapnell C. , Williams B. A. , Pertea G. , Mortazavi A. , Kwan G. , Baren M. J. van , Salzberg S. L. , Wold B. J. , and Pachter L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.*, 28(5): 511–515.
21. Robinson M. D. and Oshlack A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11(3):R25.
22. Xiao Z. , Zou Q. , Liu Y. , and Yang X. (2016) Genome-wide assessment of differential translations with ribosome profiling data. *Nat Commun*, 7: 11194.
23. Zhong Y. , Karaletsos T. , Drewe P. , Sreedharan V. T. , Kuo D. , Singh K. , Wendel H.-G. , and Ratsch G. (2017) RiboDiff: detecting changes of mRNA translation efficiency from ribosome footprints. *Bioinformatics*, 33(1):139–141.
24. Li W. , Wang W. , Uren P. J. , Penalva L. O. F. , and Smith A. D. (2017) Riborex: fast and flexible identification of differential translation from Ribo-seq data. *Bioinformatics*, 33(11):1735–1737.
25. Olshen A. B. , Hsieh A. C. , Stumpf C. R. , Olshen R. A. , Ruggero D. , and Taylor B. S. (2013) Assessing gene-level translational control from ribosome profiling. *Bioinformatics*, 29(23):2995–3002.
26. Reimand J. , Arak T. , Adler P. , Kolberg L. , Reisberg S. , Peterson H. , and Vilo J. (2016) g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.*, 44(W1):W83–89.
27. Merico D. , Isserlin R. , Stueker O. , Emili A. , and Bader G. D. (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE*, 5(11):e13984.
28. Tusup M. , French L. E. , Guenova E. , Kundig T. , and Pascolo S. . (2018). Optimizing the Functionality of *in vitro*-Transcribed mRNA. *BJSTR*, 7(2): 001–006.
29. Lavery D. J. and Schibler U. (1993) Circadian transcription of the cholesterol 7 alpha hydroxylase gene may involve the liver-enriched bZIP protein DBP. *Genes Dev.*, 7(10):1871–1884.
30. Sendoel A. , Dunn J. G. , Rodriguez E. H. , Naik S. , Gomez N. C. , Hurwitz B. , Levorse J. , Dill B. D. , Schramek D. , Molina H. , Weissman J. S. , and Fuchs E. (2017) Translation from unconventional 5' start sites drives tumour initiation. *Nature*, 541(7638):494.
31. Blair J. D. , Hockemeyer D. , Doudna J. A. , Bateup H. S. , and Floor S. N. (2017) Widespread Translational Remodeling during Human Neuronal Differentiation. *Cell Reports*, 21(7):2005–2016.
32. Lomakin I. B. , Stolboushkina E. A. , Vaidya A. T. , Zhao C. , Garber M. B. , Dmitriev S. E. , and Steitz T. A. . (2017). Crystal Structure of the Human Ribosome in Complex with DENR-MCT-1. *Cell Reports*, 20(3): 521–528.
33. Weisser M. , Schafer T. , Leibundgut M. , Bohringer D. , Aylett C. H. S. , and Ban N. . (2017). Structural and Functional Insights into Human Re-initiation Complexes. *Molecular Cell*, 67(3):447–456.e7.
34. Zinoviev A. , Hellen C. U. T. , and Pestova T. V. . (2015). Multiple Mechanisms of Reinitiation on Bicistronic Calicivirus mRNAs. *Molecular Cell*, 57(6):1059–1073.
35. Watatani Y. , Ichikawa K. , Nakanishi N. , Fujimoto M. , Takeda H. , Kimura N. , Hirose H. , Takahashi S. , and Takahashi Y. (2008) Stress-induced Translation of ATF5 mRNA Is Regulated by the 5'-Untranslated Region. *J. Biol. Chem.*, 283(5):2543–2553.
36. Suzuki Y. , Holmes J. B. , Cerritelli S. M. , Sakhuja K. , Minczuk M. , Holt I. J. , and Crouch R. J. (2010) An Upstream Open Reading Frame and the Context of the Two AUG Codons Affect the Abundance of Mitochondrial and Nuclear RNase H1. *Mol. Cell. Biol.*, 30(21):5123–5134.
37. Liang X.-h. , Shen W. , Sun H. , Migawa M. T. , Vickers T. A. , and Crooke S. T. (2016) Translation efficiency of mRNAs is increased by antisense oligonucleotides targeting upstream open reading frames. *Nature Biotechnology*, 34(8):875–880.
38. Schulz J. , Mah N. , Neuenschwander M. , Kischka T. , Ratei R. , Schlag P. M. , Castanos-Velez E. , Fichtner I. , Tunn P.-U. , Denkert C. , Klaas O. , Berdel W. E. , Kries J. P. , Makalowski W. , Andrade-Navarro M. A. , Leutz A. , and Wethmar K. (2018) Loss-of-function uORF mutations in human malignancies. *Scientific Reports*, 8(1):2395.
39. Mukasa A. , Wykosky J. , Ligon K. L. , Chin L. , Cavenee W. K. , and Furnari F. (2010) Mutant EGFR is required for maintenance of glioma growth *in vivo*, and its ablation leads to escape from receptor dependence. *PNAS*, 107(6):2616–2621.
40. Childs E. J. , Mocci E. , Campa D. , Bracci P. M. , Gallinger S. , Goggins M. , Li D. , Neale R. E. , Olson S. H. , Scelo G., Amundadottir, L. T., Bamlet, W. R., Bijlsma, M. F., Blackford, A., Borges, M., Brennan, P., Brenner, H., Bueno-de Mesquita, H. B., Canzian, F., Capurso, G., Cavestro, G. M., Chaffee, K. G., Chanock, S. J., Cleary, S. P., Cotterchio, M., Foretova, L., Fuchs, C., Funel, N., Gazouli, M., Hassan, M., Herman, J. M., Holcatova, I., Holly, E. A., Hoover, R. N., Hung, R. J., Janout, V., Key, T. J., Kupcinskas, J., Kurtz, R. C., Landi, S., Lu, L., Malecka-Panas, E., Mambrini, A., Mohelnikova-Duchonova, B., Neoptolemos, J. P., Oberg, A. L., Orlow, I., Pasquali, C., Pezzilli, R., Rizzato, C., Saldia,

A., Scarpa, A., Stolzenberg-Solomon, R. Z., Strobel, O., Tavano, F., Vashist, Y. K., Vodicka, P., Wolpin, B. M., Yu, H., Petersen, G. M., Risch, H. A., and Klein, A. P. (2016) Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nature Genetics*, 47(8):911–916.
41. Diao D. , Wang L. , Wan J. , Chen Z. , Peng J. , Liu H. , Chen X. , Wang W. , and Zou L. (2016) MEK5 overexpression is associated with the occurrence and development of colorectal cancer. *BMC Cancer*, 16:302.
42. Sato K. and Akimoto K. (2017) Expression Levels of KMT2c and SLC20a1 Identified by Information-theoretical Analysis Are Powerful Prognostic Biomarkers in Estrogen Receptor-positive Breast Cancer. *Clinical Breast Cancer*, 17(3):e135–e142.
43. Thelen A. , Scholz A. , Benckert C. , Marschall Z. von , Schroder M. , Wiedenmann B. , Neuhaus P. , Rosewicz S. , and Jonas S. (2008) VEGF-D promotes tumor growth and lymphatic spread in a mouse model of hepatocellular carcinoma. *Int. J. Cancer*, 122(11):2471–2481.
44. Vattem K. M. and Wek R. C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *PNAS*, 101 (31):11269–11274.
45. Kozak M. (1987) Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes. *Mol. Cell. Biol.*, 7(10):3438–3445.
46. Michel A. M. , Fox G. , M. Kiran A. , De Bo C. , O'Connor P. B. F. , Heaphy S. M. , Mullan J. P. A. , Donohue C. A. , Higgins D. G. , and Baranov P. V. (2014) GWIPS-viz: development of a ribo-seq genome browser. *Nucleic Acids Res*, 42(D1):D859–D864.
47. Kozak M. (2001) Constraints on reinitiation of translation in mammals. *Nucleic Acids Res*, 29(24):5226–5232.
48. Luukkonen B. G. , Tan W. , and Schwartz S. (1995) Efficiency of reinitiation of translation on human immunodeficiency virus type 1 mRNAs is determined by the length of the upstream open reading frame and by intercistronic distance. *J Virol*, 69(7):4086–4094.
49. Poyry T. A. A. , Kaminski A. , and Jackson R. J. (2004) What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes Dev.*, 18(1):62–75.
50. Agarwal V. , Bell G. W. , Nam J.-W. , and Bartel D. P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *eLife Sciences*, 4:e05005.
51. Reinert L. S. , Shi B. , Nandi S. , Mazan-Mamczarz K. , Vitolo M. , Bachman K. E. , He H. , and Gartenhaus R. B. (2006) MCT-1 Protein Interacts with the Cap Complex and Modulates Messenger RNA Translational Profiles. *Cancer Res*, 66(18):8994–9001.
52. Cerami E. , Gao J. , Dogrusoz U. , Gross B. E. , Sumer S. O. , Aksoy B. A. , Jacobsen A. , Byrne C. J. , Heuer M. L. , Larsson E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., and Schultz, N. (2012) The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov*, 2(5):401–404.
53. Landgraf D. , Wang L. L. , Diemer T. , and Welsh D. K. (2016) NPAS2 Compensates for Loss of CLOCK in Peripheral Circadian Oscillators. *PLOS Genetics*, 12(2):e1005882.
54. Huang N. , Chelliah Y. , Shan Y. , Taylor C. A. , Yoo S.-H. , Partch C. , Green C. B. , Zhang H. , and Takahashi J. S. (2012) Crystal Structure of the Heterodimeric CLOCK:BMAL1 Transcriptional Activator Complex. *Science*, 337(6091):189–194.
55. Knoepfler P. S. , Bergstrom D. A. , Uetsuki T. , Dac-Korytko I. , Sun Y. H. , Wright W. E. , Tapscott S. J. , and Kamps M. P. (1999) A conserved motif N-terminal to the DNA-binding domains of myogenic bHLH transcription factors mediates cooperative DNA binding with pbx-Meis1/Prep1. *Nucleic Acids Res.*, 27(18):3752–3761.
56. Diernfellner A. C. , Schafmeier T. , Merrow M. W. , and Brunner M. (2005) Molecular mechanism of temperature sensing by the circadian clock of Neurospora crassa. *Genes Dev*, 19(17):1968–1973.
57. Deyo J. E. , Chiao P. J. , and Tainsky M. A. (1998) drp, a novel protein expressed at high cell density but not during growth arrest. *DNA Cell Biol.*, 17(5):437–447.
58. Noguchi T. , Wang L. L. , and Welsh D. K. (2013) Fibroblast PER2 circadian rhythmicity depends on cell density. *J. Biol. Rhythms*, 28(3): 183–192.
59. Guenthner C. J. , Luitje M. E. , Pyle L. A. , Molyneux P. C. , Yu J. K. , Li A. S. , Leise T. L. , and Harrington M. E. (2014) Circadian rhythms of Per2::Luc in individual primary mouse hepatocytes and cultures. *PLoS ONE*, 9(2):e87573.

FIGURE LEGENDS

Figure 1. *Denr* knockdown engenders ribosome redistribution from CDS to 5' UTR. A) Experimental approach of the study. Ribosome profiling (RPF-seq) and parallel RNA-seq were performed on NIH3T3 mouse fibroblasts depleted of *Denr* via lentivirally expressed shRNAs, with scramble (Scr) and *Gfp* shRNAs as control conditions. Profiling was carried out in biological triplicates for the two *Denr* shRNAs (shRNA1, shRNA2) and for Scr shRNA, and in duplicate for *Gfp* shRNA. We hypothesized that if DENR is required for efficient reinitiation after uORF translation, knocking down *Denr* would lead to reduced translation efficiency on the main coding sequence of DENR targets. **B)** Distribution of normalized RNA-seq read counts along the *Denr* transcript show high *Denr* knockdown efficiency. Different orange color shadings correspond to the different replicates. **C)** RPKM values for CDS-mapping RNA-seq read counts for *Denr* (left), *Mcts1* (middle), and *Eif2d* (right) indicate specific and high knockdown efficiency. Each dot corresponds to a sample and replicate and error bars indicate mean$\pm$sd. **D)** Distribution of CDS translation (from RPF-seq) differences between *Denr* knockdown (Kd) and control cells. In grey, all detected genes (N = 10516); in blue, genes with sufficient coverage in 5' UTRs (i.e. a minimum of 20 RPF-seq reads in the 5' UTR in 3/4 of samples; N = 5148). Note that both distributions are centered around zero, indicating that *Denr* knockdown does not strongly affect global protein synthesis *per se*. **E)** Distribution of 5' UTR-to-CDS translation differences between *Denr* Kd and control cells, for genes with sufficient 5' UTR coverage (N = 5148). Relative 5' UTR translation was significantly higher in *Denr* Kd cells (median = 0.32, $p < 2.2e-16$, Wilcoxon signed rank test), suggesting a ribosome footprint redistribution to 5' UTRs upon *Denr* knockdown.

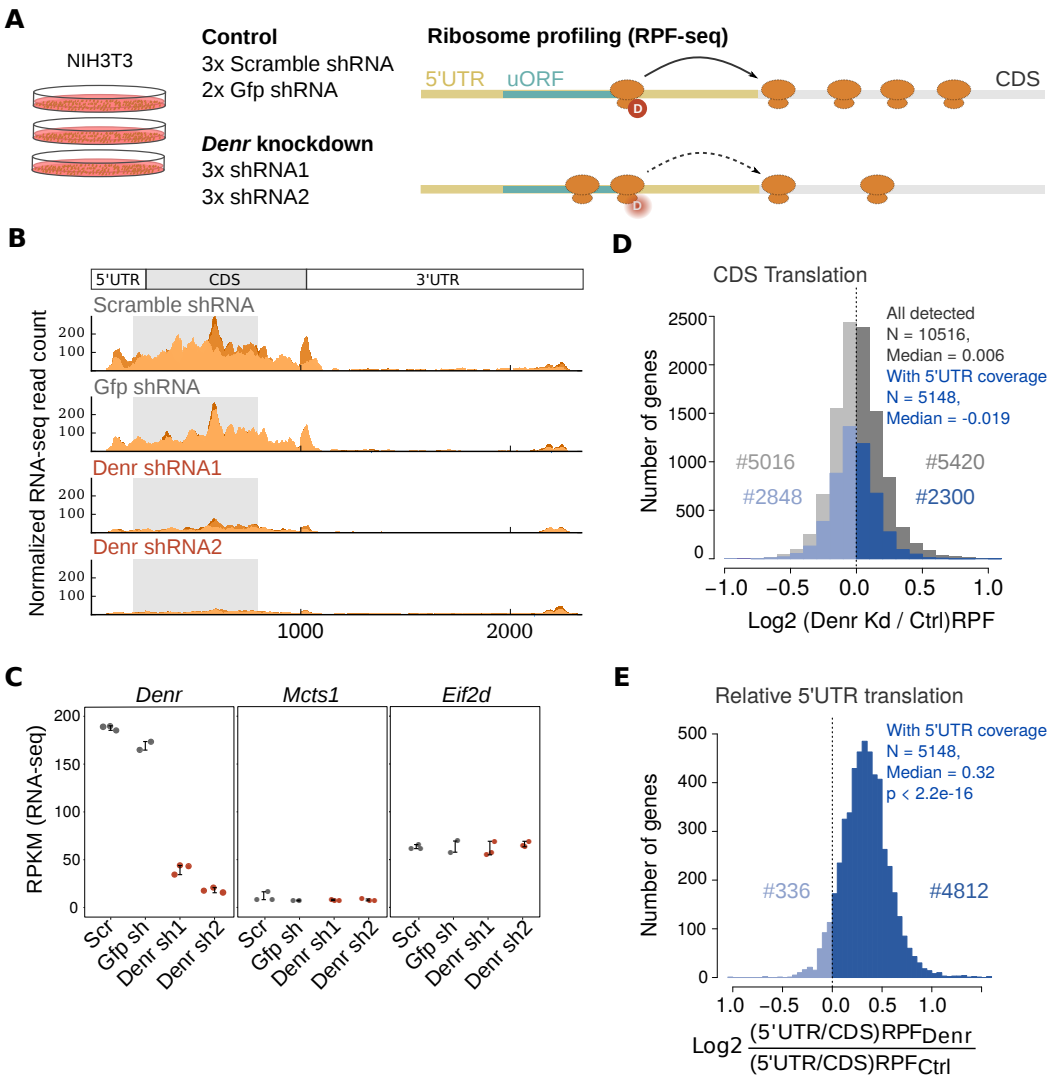
Figure 2. Characteristics of DENR targets: decreased CDS TE; enrichment for uORFs; association with functional categories. A) Scatter plot of *Denr* Kd-to-control ratio of mRNA abundance vs. CDS translation efficiency (TE) for all expressed genes (N = 10516). Highlighted in orange are genes with differential TE (detected by Xtail (22), FDR < 0.1, N =240), the majority of which showed decreased TE in *Denr* Kd cells, in line with a role of DENR in regulating translation reinitiation. **B)** uORF enrichment in DENR targets. Proportion of genes with translated uORFs transcriptome-wide compared to that among the differential TE genes detected in (A). DENR targets showed enrichment for translated uORF-containing genes (p = 5.73e-4, Fisher test). **C)** Enrichment map for GO-Biological process, GO-Metabolic function and Reactome functional categories of genes with differential TE. Only categories with an FDR < 0.05 in the GO analysis were considered. Node size is proportional to the number of genes associated with the GO category, shading corresponds to the statistical significance, and edge width is proportional to the number of genes shared between functional categories.

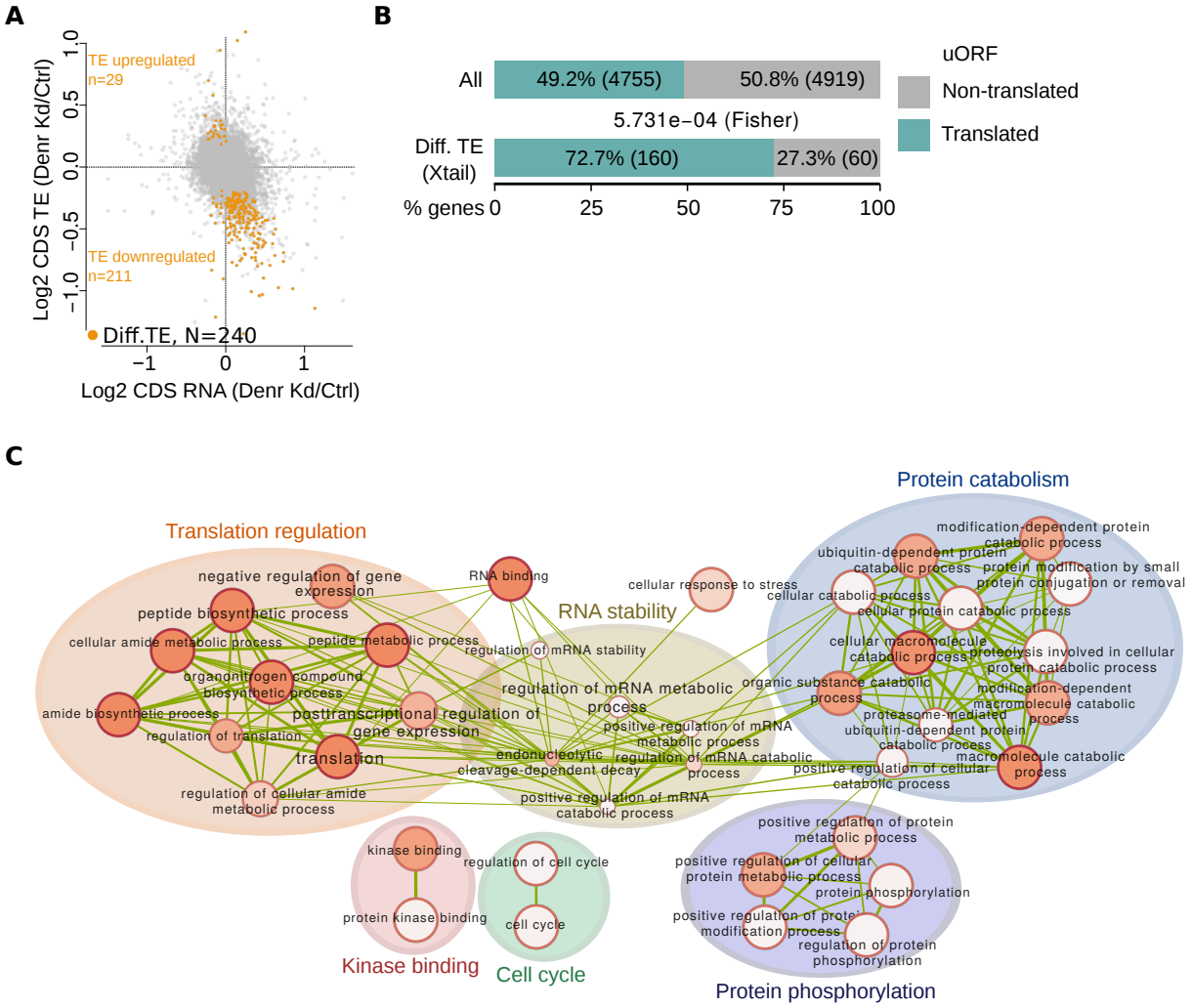
Figure 3. Mutation of uORF start codons relieves DENR dependence. A) Representative examples of genes with differential TE. Panels show average CDS translation efficiency in control (grey) and *Denr* Kd (red) cells. **B)** *Denr* Kd-to-control translation on the 5' UTR (brown) and on the CDS (grey) for the same examples as in (A). DENR

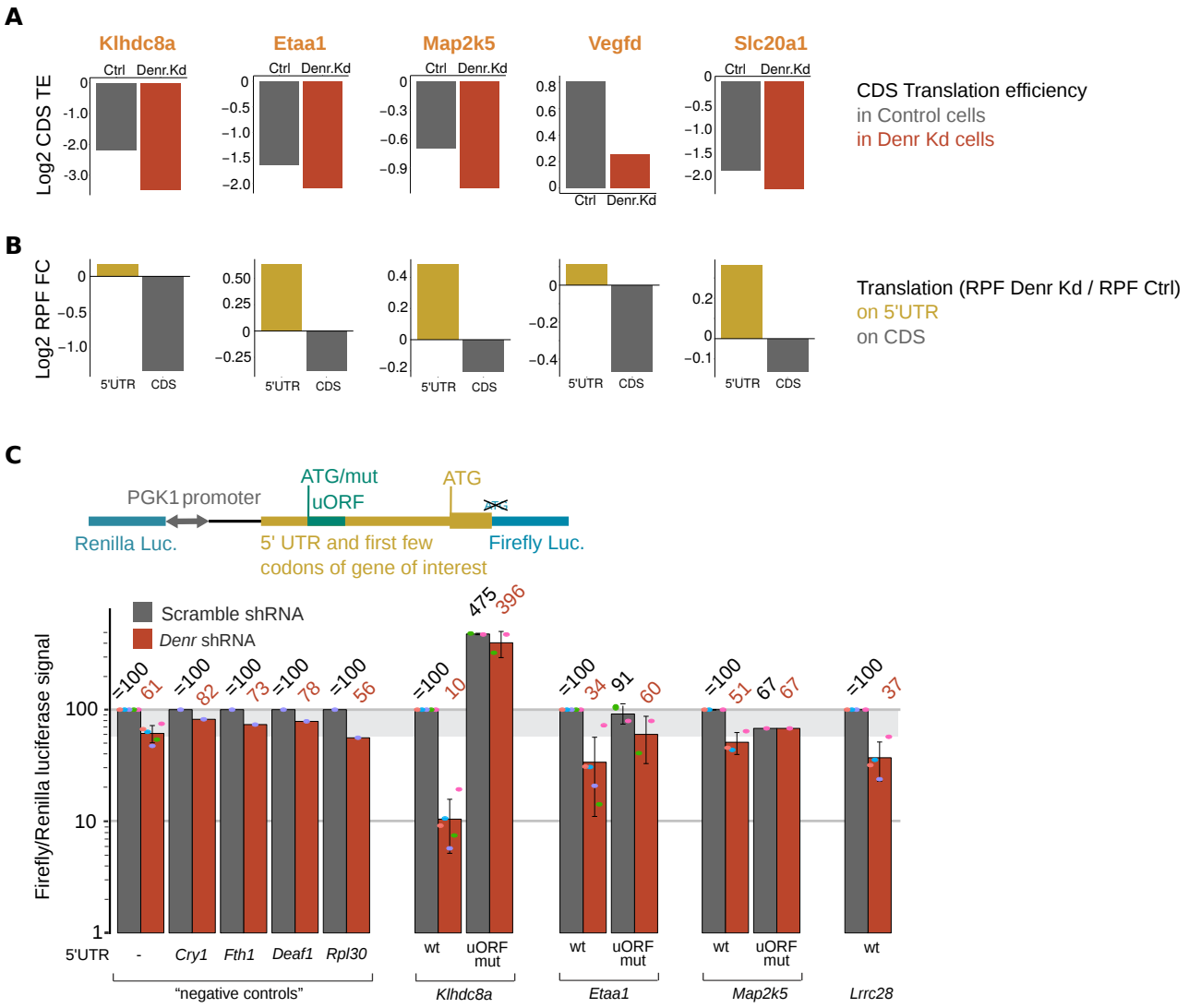
targets showed decreased CDS TE (A) and increased 5' UTR translation concomitant with reduced CDS translation in *Denr*-deficient cells. **C)** *(Upper)* Schematic representation of the dual luciferase reporter construct, where the 5' UTR of the gene of interest – with the WT or mutated uORF start – and first few codons of the CDS were cloned upstream of firefly luciferase (FL). *(Lower)* Results of dual luciferase assays, where the influence of the 5' UTR (WT and uORF mutant) on the expression of FL was measured in control and *Denr* knockdown cells, and internally normalised to RL (expressed from the same bidirectional promoter). Note that for all targets tested, depletion of *Denr* reduced the FL/RL luciferase signal, indicating DENR-dependence of reporter protein expression. Mutation of uORF start codons relieved the inhibition and partially abolished DENR-dependence. For each 5' UTR assayed, FL/RL signals were normalised to the Scr-transduced non-mutated construct. The area of grey shading marks the non-specific effect of *Denr* loss-of-function in this assay, as judged from the reduction in signal seen for negative control reporters that contained no cloned 5' UTR (-), or irrelevant 5' UTRs selected from non-targets.

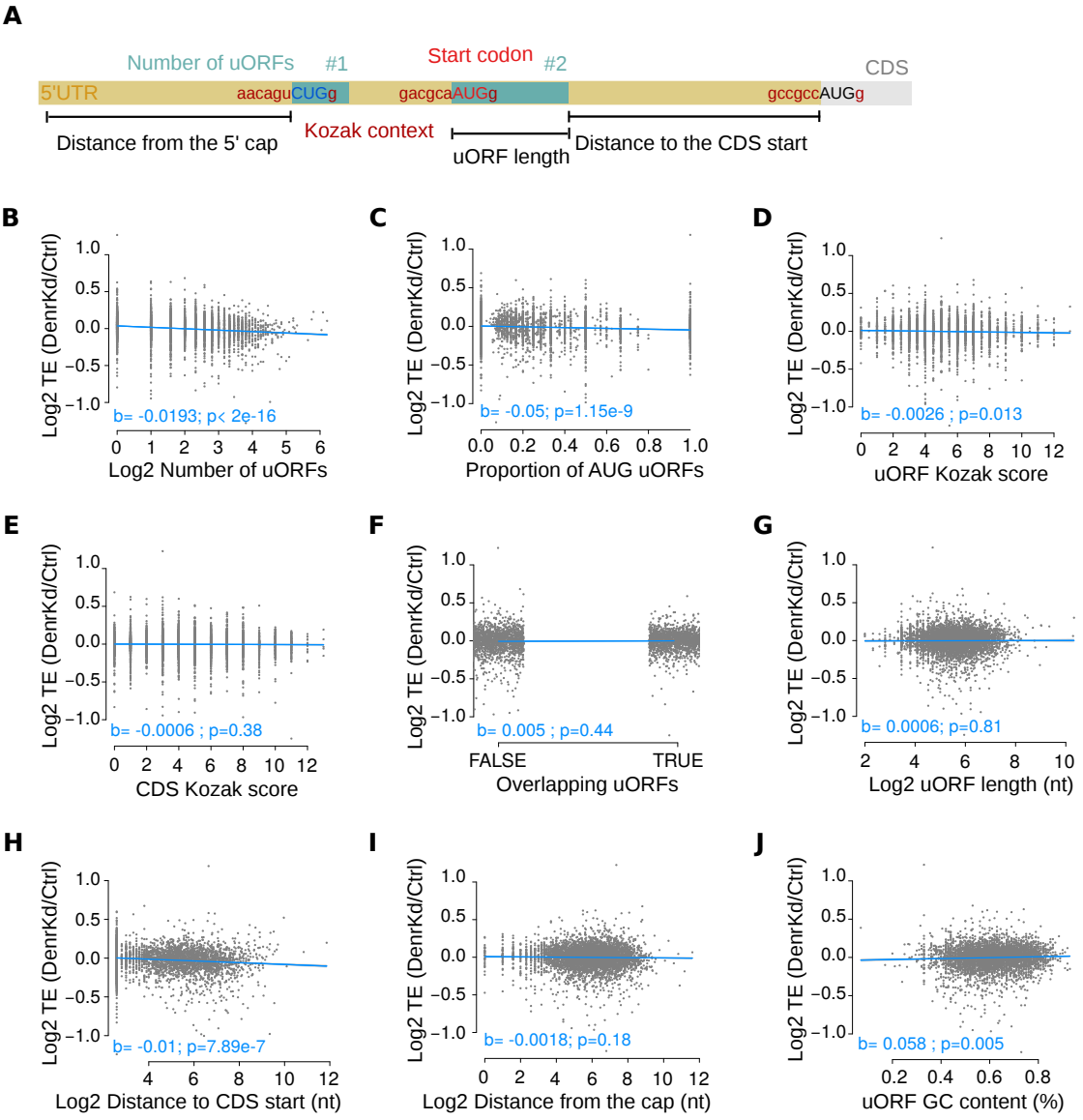
Figure 4. 5' UTR features predictive of translation efficiency regulation by DENR. A) Schematic of the 5' UTR features used as predictors for the modeling of CDS TE change. The number of uORFs, the start codon identity and context, the presence of overlapping uORFs, the uORF length and GC content, and the relative location within the 5' UTR were considered. Distance to the 5' cap was computed for the first uORF, and distance to the CDS start was quantified for the last uORF. **B-J)** Partial regression plots for the modeling of translation efficiency fold change (*Denr* Kd vs. Ctrl) on each uORF feature. Indicated in blue are the slope coefficients for the partial regression line and their significance levels (t-test). The number of uORFs (B), the proportion of AUG uORFs (C) and the distance of the last uORF to the CDS start (H) were the strongest predictors.

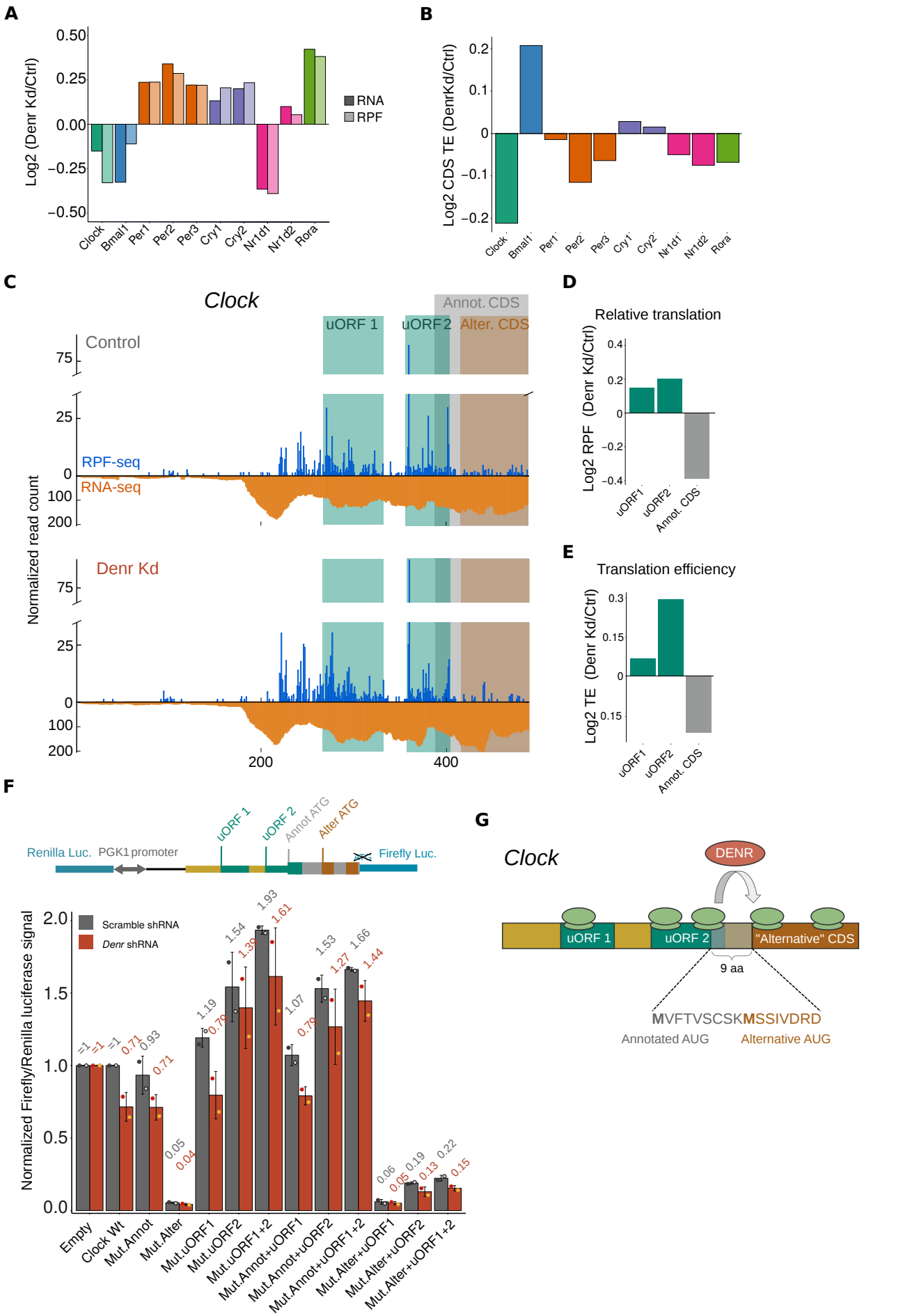
Figure 5. DENR-mediated reinitiation regulates CLOCK protein abundance and N-terminus. A) Bar graph of the average RPKM ratio (*Denr* Kd-to-control) for the main circadian core clock genes, at the level of mRNA abundance (dark shadings) and footprint abundance (light shadings). **B)** CDS translation efficiency ratio between *Denr* Kd and control cells showed that *Clock* had a reduced TE upon *Denr* depletion. **C)** Distribution of normalised RPF-seq (blue) and RNA-seq (orange) reads along the 5' UTR and first 100 nt of *Clock* CDS, in control (top) and *Denr* Kd (bottom) cells. Green shaded areas indicate the two AUG-initiated uORFs; grey area corresponds to the annotated CDS; and brown area to the alternative CDS, in frame with the annotated one. **D)** Quantification of relative translation ratio (*Denr* Kd-to-control) on both uORFs, and on the CDS, showing increased uORF translation and reduced CDS translation in *Denr*-deficient cells. **E)** Quantification of relative translation efficiency (*Denr* Kd-to-control) on both uORFs, and on the CDS showed higher TE on the overlapping uORF (uORF2). **F)** (Top) Schematic representation of the dual luciferase reporter construct, where *Clock* 5' UTR and first 14 codons of the CDS (containing both the annotated and alternative start codons) were cloned upstream of FL. (Bottom) Dual luciferase assay results for *Clock* WT and mutants. Empty vector contained only the vector-encoded 5' UTR. FL signal was internally normalised to RL signal; all constructs and conditions were first normalised to the value of condition "empty vector, scr shRNA"; all *Clock* mutant constructs were then additionally normalised to the value from condition "*Clock* WT, scr shRNA". Signal from all *Denr* shRNA-transduced *Clock* constructs were also normalised to the signal obtained for "empty vector, *Denr*-shRNA", to remove DENR effects that were not specific to the 5' UTR. **G)** Schemcatic of how uORFs and reinitiation regulate the translation of the transcript encoding the circadian core clock component *Clock*. *Clock* transcript contains two translated AUG uORFs, the second of which shows high ribosome occupancy and overlaps with the annotated CDS. An additional in-frame AUG is located 9 amino acids downstream of the annotated AUG; usage of the downstream AUG would give rise to a so far unnanotated CLOCK protein variant. Our data suggest that this second AUG is the true start codon for CLOCK biosynthesis in NIH3T3 cells, and that DENR is necessary for translation reinitiation on the alternative CDS after translation of the second uORF.


A
NIH3T3
Control
3x Scramble shRNA
2x Gfp shRNA
Denr knockdown
3x shRNA1
3x shRNA2
Ribosome profiling (RPF-seq)
5'UTR
uORF
CDS
B
5'UTR
CDS
3'UTR
Scramble shRNA
Gfp shRNA
Denr shRNA1
Denr shRNA2
Normalized RNA-seq read count
1000
2000
C
Denr
Mcts1
Eif2d
RPKM (RNA-seq)
Scr
Gfp sh
Denr sh1
Denr sh2
D
CDS Translation
All detected
N = 10516,
Median = 0.006
With 5'UTR coverage
N = 5148,
Median = -0.019
#5016
#2848
#5420
#2300
Number of genes
Log2 (Denr Kd / Ctrl)RPF
E
Relative 5'UTR translation
With 5'UTR coverage
N = 5148,
Median = 0.32
p < 2.2e-16
#336
#4812
Number of genes
Log2 (5'UTR/CDS)RPFDenr / (5'UTR/CDS)RPFCtrl


A
TE upregulated
n=29
TE downregulated
n=211
Diff.TE, N=240
Log2 CDS TE (Denr Kd/Ctrl)
Log2 CDS RNA (Denr Kd/Ctrl)
B
All
49.2% (4755)
50.8% (4919)
5.731e−04 (Fisher)
Diff. TE (Xtail)
72.7% (160)
27.3% (60)
% genes
uORF
Non-translated
Translated
C
Translation regulation
negative regulation of gene expression
peptide biosynthetic process
cellular amide metabolic process
peptide metabolic process
organonitrogen compound biosynthetic process
amide biosynthetic process
regulation of translation
translation
regulation of cellular amide metabolic process
RNA binding
RNA stability
regulation of mRNA stability
regulation of mRNA metabolic process
cellular response to stress
positive regulation of mRNA metabolic process
endonucleolytic cleavage-dependent decay
regulation of mRNA catabolic process
positive regulation of mRNA catabolic process
Protein catabolism
modification-dependent protein catabolic process
ubiquitin-dependent protein catabolic process
protein modification by small protein conjugation or removal
cellular catabolic process
cellular protein catabolic process
cellular macromolecule catabolic process
proteolysis involved in cellular protein catabolic process
organic substance catabolic process
modification-dependent macromolecule catabolic process
proteasome-mediated ubiquitin-dependent protein catabolic process
macromolecule catabolic process
positive regulation of cellular catabolic process
kinase binding
protein kinase binding
Kinase binding
regulation of cell cycle
cell cycle
Cell cycle
positive regulation of protein metabolic process
positive regulation of cellular protein metabolic process
protein phosphorylation
positive regulation of protein modification process
regulation of protein phosphorylation
Protein phosphorylation


A
Klhdc8a
Etaa1
Map2k5
Vegfd
Slc20a1
Ctrl
Denr.Kd
Log2 CDS TE
CDS Translation efficiency
in Control cells
in Denr Kd cells
B
Log2 RPF FC
5'UTR
CDS
Translation (RPF Denr Kd / RPF Ctrl)
on 5'UTR
on CDS
C
PGK1 promoter
ATG/mut
uORF
ATG
Renilla Luc.
5' UTR and first few
codons of gene of interest
Firefly Luc.
Scramble shRNA
Denr shRNA
Firefly/Renilla luciferase signal
5'UTR
-
Cry1
Fth1
Deaf1
Rpl30
wt
uORF
mut
"negative controls"
Klhdc8a
Etaa1
Map2k5
Lrrc28


A
Number of uORFs
#1
Start codon
#2
CDS
5'UTR
aacaguCUGg
gacgcaAUGg
gccgccAUGg
Distance from the 5' cap
Kozak context
uORF length
Distance to the CDS start
B
Log2 TE (DenrKd/Ctrl)
b= -0.0193; p< 2e-16
Log2 Number of uORFs
C
b= -0.05; p=1.15e-9
Proportion of AUG uORFs
D
b= -0.0026 ; p=0.013
uORF Kozak score
E
b= -0.0006 ; p=0.38
CDS Kozak score
F
b= 0.005 ; p=0.44
FALSE
TRUE
Overlapping uORFs
G
b= 0.0006; p=0.81
Log2 uORF length (nt)
H
b= -0.01; p=7.89e-7
Log2 Distance to CDS start (nt)
I
b= -0.0018; p=0.18
Log2 Distance from the cap (nt)
J
b= 0.058 ; p=0.005
uORF GC content (%)


A
Log2 (Denr Kd/Ctrl)
0.50
0.25
0.00
−0.25
−0.50
RNA
RPF
Clock
Bmal1
Per1
Per2
Per3
Cry1
Cry2
Nr1d1
Nr1d2
Rora
B
Log2 CDS TE (DenrKd/Ctrl)
0.2
0.1
0
−0.1
−0.2
C
Clock
Annot. CDS
uORF 1
uORF 2
Alter. CDS
Control
RPF-seq
RNA-seq
Denr Kd
Normalized read count
75
25
0
100
200
200
400
D
Relative translation
Log2 RPF (Denr Kd/Ctrl)
0.4
0.2
0
−0.2
−0.4
uORF1
uORF2
Annot. CDS
E
Translation efficiency
Log2 TE (Denr Kd/Ctrl)
0.3
0.15
0
0.15
F
Renilla Luc.
PGK1 promoter
uORF 1
uORF 2
Annot ATG
Alter ATG
Firefly Luc.
Normalized Firefly/Renilla luciferase signal
2.0
1.5
1.0
0.5
0.0
Scramble shRNA
Denr shRNA
=1
=1
=1
0.71
0.93
0.71
0.05
0.04
1.19
0.79
1.54
1.39
1.93
1.61
1.07
0.79
1.53
1.27
1.66
1.44
0.06
0.05
0.19
0.13
0.22
0.15
Empty
Clock Wt
Mut.Annot
Mut.Alter
Mut.uORF1
Mut.uORF2
Mut.uORF1+2
Mut.Annot+uORF1
Mut.Annot+uORF2
Mut.Annot+uORF1+2
Mut.Alter+uORF1
Mut.Alter+uORF2
Mut.Alter+uORF1+2
G
Clock
DENR
uORF 1
uORF 2
"Alternative" CDS
9 aa
MVFTVSCSKMSSIVDRD
Annotated AUG
Alternative AUG