PNAS
www.pnas.org

Main Manuscript for

Circadian ribosome profiling reveals a role for the *Period2* upstream opening reading frame in sleep

Arthur Millius[1,2,3], Rikuhiro Yamada[1], Hiroshi Fujishima[1], Kazuhiko Maeda[2], Daron M. Standley[3], Kenta Sumiyama[4], Dimitri Perrin[5,6], and Hiroki R. Ueda*[1,7]

Affiliations

1) Laboratory for Synthetic Biology, RIKEN Quantitative Biology Center, Suita, Osaka, Japan
2) Laboratory for Host Defense, iFReC, Suita, Osaka, Japan
3) Laboratory for Systems Immunology, iFReC, Suita, Osaka, Japan
4) Laboratory for Mouse Genetic Engineering, RIKEN Center for Biodynamics Research, Suita, Osaka, Japan
5) School of Computer Science, Queensland University of Technology, Brisbane, Australia
6) Centre for Data Science, Queensland University of Technology, Brisbane, Australia
7) Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

*corresponding author: Hiroki R. Ueda
Email: uedah-tky@umin.ac.jp

Arthur Millius 0000-0002-5280-6730
Kazuhiko Maeda 0000-0003-0305-8163
Daron M. Standley 0000-0003-4078-0817
Kenta Sumiyama 0000-0001-8785-5439
Dimitri Perrin 0000-0002-4007-5256
Hiroki R. Ueda 0000-0001-8342-9176

Classification
Biological Sciences/Neuroscience
Keywords
Circadian rhythms, Ribosome profiling, uORF, Mass spectrometry, RNA

Author Contributions
A.M., D.P., R.Y., K.S., and H.R.U. designed research; A.M., R.Y., K.S., and H.F. performed research; A.M. and D.P. analyzed data; K.M. and D.M.S. contributed new reagents or computational tools; and A.M. wrote the paper with input from D.P., D.M.S., and H.R.U.

This PDF file includes:

Main Text
Figures 1 to 4

Abstract

Many mammalian proteins have circadian cycles of production and degradation, and many of these rhythms are altered post-transcriptionally. We used ribosome profiling to examine post-transcriptional control of circadian rhythms by quantifying RNA translation in the liver over a 24-h period from circadian-entrained mice transferred to constant darkness conditions and by comparing ribosome binding levels to protein levels for 16 circadian proteins. We observed large differences in ribosome binding levels compared to protein levels, and we observed delays between peak ribosome binding and peak protein abundance. We found extensive binding of ribosomes to upstream open reading frames (uORFs) in circadian mRNAs, including the core clock gene *Period2 (Per2)*. An increase in the number of uORFs in the 5'UTR was associated with a decrease in ribosome binding in the main coding sequence and a reduction in expression of synthetic reporter constructs. Mutation of the *Per2* uORF increased luciferase and fluorescence reporter expression in 3T3 cells without altering the phase or period and increased luciferase expression in PER2:LUC MEF cells. Mutation of the *Per2* uORF in mice increased PER2 protein levels and reduced total sleep time compared to that in wild-type mice. These results suggest that post-transcriptional processes shape translation of mRNA transcripts, which can impact physiological rhythms and sleep.

Significance Statement

Period (Per) is an iconic gene in the field of circadian rhythms since its discovery in 1971 by Seymour Benzer and Ronald Konopka in fruit flies. The inhibitory feedback loop of PER protein drives circadian rhythms, and here, we add to our understanding of PER protein production by analyzing a previously unknown upstream open reading frame (uORF) in the 5' untranslated region of *Period2* mRNA. Mutation of the *Per2* uORF altered the amplitude, but not the phase or period, of *Per2* expression in well-characterized cell culture models, and we observed that *Per2* uORF mutant mice exhibited sleep loss, particularly during light-to-dark and dark-to-light transitions, which suggests a role for uORFs in modulating molecular and physiological circadian rhythms.

Main Text

Introduction

Life is remarkably adapted to the 24-hour rotational movement of the earth. In mammals, the molecular time-keeping mechanism for circadian rhythms relies primarily on a hierarchical network of transcription activators and repressors in cells and tissues (1). In the past, circadian clocks have been measured using systems approaches to quantify genome-wide changes in RNA levels (2), which has resulted in an understanding of the transcriptional regulatory network; however, less is known about how translation and post-transcriptional regulation influence biological rhythms.

Although 10% of genes are rhythmic in the liver (3), de novo transcription is responsible for only some this rhythmicity (4). For example, the timing of when a circadian mRNA is

expressed does not necessarily correspond to that of mRNA translation or peak protein abundance (5). Some proteins have 24-h rhythms in abundance in the absence of rhythmic RNA expression (6, 7), which may suggest a role for rhythmic translation in regulating the clock (8).

In mouse liver, detection of low-abundant components of the core circadian circuit using systems proteomics is difficult (8, 9), unless special care is taken to examine a particular protein on a case-by-case basis (10) or by using advanced mass-spectrometry techniques (11, 12). Researchers have used next-generation sequencing of ribosome-bound mRNA protected from RNAse degradation to understand how translation regulation affects protein output (13, 14). Previous studies using ribosome profiling to measure daily rhythms focused on a cell culture model (15) or mouse tissues in light-dark conditions (16–18), which reflects diurnal gene expression, which may be influenced by non-circadian time-keeping systems. These studies also examined the timing between transcription and translation, but it remains unclear how circadian translation relates to peak protein abundance in terms of protein turnover and timing. For example, the peak abundance of PER and CRY proteins are delayed relative to expression of their mRNA in the liver (19) and suprachiasmatic nuclei (SCN) (20). This difference in timing may result from a delay in RNA processing before translation, a delay during translation, a delay in protein transport from one cellular location to another, or a delay in protein turnover (21, 22).

Previously, we developed a mass spectrometry method called MS-based Quantification By isotope-labeled Cell-free products (MS-QBiC) to determine the absolute protein levels of 16 selected circadian proteins in mice liver over a 24-h period (10). The method takes advantage of the PURE system (23) for reconstituting cell-free protein expression of optimal peptide standards for detection and quantification by selected reaction monitoring (SRM)-based targeted proteomics analysis. We found delays between the peak levels of RNA expression, as measured by quantitative PCR (qPCR), and the abundance of the corresponding protein, suggesting either a delay in post-transcriptional RNA processing or in protein turnover. Here, we investigated the same liver samples by ribosome profiling in order to understand the timing of RNA transcription, translation and protein production and the relative protein turnover (as measured by the relationship between ribosome binding and protein abundance). We found that upstream open reading frames (uORFs) modulated translation globally, repressed reporter expression in a combinatorial manner, and suppressed production of PER2 protein. Mutation of the *Per2* uORF in mice reduced total sleep duration, particularly during the early morning and early evening, without disrupting the circadian period, which suggests that translational repression of uORFs may impact physiological behaviors.

Results

Ribosome profiling of liver from mice in constant darkness conditions

An experimental workflow was designed to analyze ribosome-protected mRNA fragments from liver samples previously examined by MS-QBiC (10) (*SI Appendix*, Fig. S1). Briefly, mice were entrained to a 12-h light/12-h dark (LD) cycle for 14 days, transferred to constant darkness (DD) for 24 h, and sacrificed at circadian times (CT0, CT4, CT8, CT12, CT16, CT20, and CT24). Liver samples from two mice were collected and analyzed following established ribosome profiling protocols (24). We prepared

ribosome profiling libraries and sequenced ~70 million reads per sample; this resulted in 25-45 million reads that could be mapped to mRNA (*SI Appendix*, Table S1). Ribosome-protected fragments primarily aligned to the coding regions and 5' untranslated regions (5'UTRs) of the mRNA, with few reads mapping to the 3' untranslated region (3'UTR) (*SI Appendix*, Fig. S2A). Alignment of coding sequence (CDS)-mapped reads based on the footprint length revealed reading frame periodicity (*SI Appendix*, Fig. S2B). Reads were of the expected size, mapped with a high percentage to mRNA, and were correlated between samples (*SI Appendix*, Fig. S2C–E).
From ~14,000 well-translated transcripts (defined by a median of at least two reads per five-codon mRNA window), we identified rhythms in ribosome-protected read fragments using the JTK_CYCLE algorithm (25), yielding 2952 rhythmic transcripts with an adjusted p-values of 0.05 or less, including well-known circadian transcripts such as *Bmal1, Per1, Per2, Clock,* and *Cry1* (*SI Appendix*, Dataset S1, Table S2, Fig S6A).

Relationship between ribosome profiling reads and protein levels
We compared the timing of protein production output, as measured by ribosome profiling reads, to the absolute number of protein molecules per cell for 16 previously reported core circadian proteins (10). There was broad agreement in the timing of peak protein production compared to that of peak protein abundance (Fig. 1A). However, for several circadian proteins, such as BMAL1 and CLOCK, there was a delay of approximately 6 hours between peak protein biosynthesis and peak protein abundance, clearly outside the 4-h range of our sampling intervals (Fig. 1B), which suggested that there is a post-translational delay in protein turnover for these molecules rather than a post-transcriptional delay in the timing between mRNA transcription and translation. We compared the average number of ribosome profiling reads to the number of protein molecules over a 24-h period, but found that the two values were poorly correlated (Fig. 1C). We measured this protein production efficiency (as defined by the ratio of our mass-spectrometry protein levels to ribosome profiling reads) at each time point. For some proteins, such as BHLHE40, a large amount of translation resulted in a moderate amount of protein (Fig. 1D); whereas for other proteins, such as PER2, a much smaller amount of translation resulted in the same amount of protein as BHLHE40 (Fig. 1E), which suggests that BHLHE40 protein was more unstable than that of PER2. Thus, ribosome profiling reads can provide an approximate estimate for when a protein is produced, but protein abundance reflects both protein production and post-translational mechanisms to control overall protein levels (26).

When we compared mRNA and protein abundance phases in our dataset in DD to the reported phases in previously published studies of mouse liver in LD (12, 15, 16, 18), we observed a correlation with a 1-2 h delay in DD (*SI Appendix*, Fig. S10A, C). The transcripts with the largest differences in delay in the mRNA-to-peak-protein abundance between LD and DD conditions included *Nfil3, Per2, Per3,* and *Bmal1* (*SI Appendix*, Fig. S10B). The differences resulted from a slightly earlier peak in RNA transcription and translation in LD rather than a delay in the protein abundance peak; there was little difference between LD and DD in the delay between transcription and translation (*SI Appendix*, Fig. S10D). Although these small differences may reflect a true biological effect from light-dependent changes in transcription, they could also simply result from

differences in the algorithms used for phase determination, differences in experimental conditions, or both. Besides these 16 core circadian transcripts, most other transcripts with rhythmic translation in our dataset had a phase of around CT0 (*SI Appendix*, Fig. S6).

Upstream open reading frames (uORFs) suppress translation

We also noticed a relationship between ribosome occupancy and the presence of upstream open reading frames (uORFs), similar to previous reports (15, 18). Roughly half of all mouse transcripts contained at least one uORF (*SI Appendix*, Fig. S4B), and we found rhythmic ribosome binding in 602 uORFs with an adjust p value < 0.05 using JTK_cycle (*SI Appendix*, Dataset S2). In particular, for circadian transcripts, such as *Cry1* and *Bmal1*, there appeared to be increased ribosome binding in uORF regions, as measured by increased ribosome profiling reads (Fig. 2A), although ribosome binding in the 5'UTR was not always associated with uORFs in circadian transcripts (*SI Appendix*, Fig. S4A). mRNAs with increased numbers of uORFs had lower levels of ribosome occupancy in the downstream CDS (Fig. 2B), whereas the length of the uORF and the distance of the uORF to the start codon did not have a significant impact on ribosome binding in the CDS (*SI Appendix*, Fig. S4C, D). To investigate whether uORFs were sufficient to suppress translation in a combinatorial manner, we created a luciferase reporter vector with multiple synthetic uORFs. Predictably, increasing the number of uORFs reduced luminescence from the reporter (Fig. 2C). Using this synthetic reporter, we varied the position of a single uORF relative to the start codon or the uORF length and found a higher degree of uORF-mediated repression the closer the uORF was to the start codon; however uORF length had no effect on repression (*SI Appendix*, Fig. S4E, F). Both the position of the uORF and the number of uORFs altered the amplitude and mesor (mean luminescence) without altering the period using two different promoters in cell-based circadian luminescence assays (*SI Appendix*, Fig. S15). Other factors, such as the Kozak consensus of the uORF, overlap with the CDS, or uORF conservation may also impact repression strength (27, 28).

The Per2 uORF suppresses reporter expression without altering period

To further explore the impact of uORFs on circadian translation, we focused on the uORF in the circadian transcript *Per2*. This uORF is evolutionarily conserved and consists of only a start and stop codon (Fig. 3A and *SI Appendix*, Fig. S13), which eliminates any potential effect of a translated uORF peptide on the regulation of *Per2*. Ribosomes bound the *Per2* uORF rhythmically and slightly before peak ribosome binding on the *Per2* transcript (Fig. 3B). Mutation of the uORF in *Per2* increased the amplitude of expression without affecting the phase or period in 3T3 cells transfected with a luminescence reporter (Fig. 3C, D). This increase in amplitude was not affected by the amount of transfected plasmid, inclusion of the full-length *Per2* 5'UTR, or addition of PER2 protein (*SI Appendix*, Fig. S5). CRISPR/Cas9-mediated mutation of the *Per2* uORF in PER::LUC MEFs increased overall luciferase expression levels compared to that of wild-type MEFs; however, we were unable to generate circadian rhythms in these cells lines to determine the impact of the *Per2* uORF on other circadian parameters such as period and phase (*SI Appendix*, Fig. S12).

To understand uORF-mediated repression in individual cells, we created a GFP fluorescent reporter plasmid driven by the *Per2* promoter with or without a mutation in the *Per2* uORF. We transfected 3T3 cells with these plasmids and an mCherry normalization plasmid and analyzed the cells by FACS (*SI Appendix*, Fig. S11). Mutation of the uORF increased GFP brightness without affecting mCherry brightness in individual cells (Fig. 3E), and increased the GFP-to-mCherry ratio of the population (Fig. 3F) regardless of whether the cells were first gated on mCherry expression (Fig. 3E, F) or GFP expression (*SI Appendix*, Fig. S11C, D). Mutation of the uORF also increased the total number of GFP-expressing cells without affecting the total number of mCherry-expressing cells (Fig. 3G). Thus, the *Per2* uORF can repress translation of mRNA transcripts within individual cells to control circadian amplitude.

Mice with a mutation in the Per2 uORF have reduced sleep

We used CRISPR/Cas9 to generate a knock-in mouse harboring a mutation in the uORF of *Per2*, which removed both the start and stop codon of the uORF without disrupting a nearby E'-box (*SI Appendix*, Fig. S7A). Compared to wild-type mice, mutant mice had slightly higher levels of PER2 protein (*SI Appendix*, Fig. S7B). Wild-type and mutant mice were phenotyped over 13 days in LD and then for 12 days in DD using the Snappy Sleep Stager (29, 30), which is a respiration-based method to characterize sleep/wake parameters and circadian rhythms. Both male and female *Per2* uORF mutant mice had significantly reduced sleep per day (mean ± SEM: 717 ± 19 min and 642 ± 7 min, respectively) compared to their wild-type littermates (774 ± 11 min and 678 ± 4 min, respectively, $P < 0.001$ by two-way ANOVA), and a corresponding increase in wake duration per day (Fig. 4A, B and *SI Appendix*, Fig. S8A) in LD conditions. There was a decrease in sleep episode duration and a significant increase in the transition probability from sleep to awake termed P_{sw} in *Per2* uORF mutant mice (*SI Appendix*, Fig. S8B); however, there were no differences in other sleep parameters such as amplitude, P_{ws} (the transition probability from awake to sleep), or wake episode duration (*SI Appendix*, Fig. S8C). *Per2* uORF mutant mice exhibited reduced sleep duration particularly during the light-to-dark and dark-to-light transitions in the early morning and early evening (Fig. 4C, D). There were also differences in sleep duration later in the morning (*SI Appendix*, Fig. S8D), but not at other times during the day (*SI Appendix*, Fig. S8E). Next, we observed mice in constant darkness over 12 days and found no differences in activity rhythms or period (Fig. 4E); furthermore, there were less pronounced sleep differences in DD compared to LD (*SI Appendix*, Fig. S9). Cosinor analysis of the activity rhythms in LD and DD conditions revealed no differences in period and amplitude between wild-type and *Per2* uORF mutant mice (*SI Appendix*, Fig. S9). Taken together, these results suggest that *Per2* expression is regulated by the *Per2* uORF, which has a moderate impact on sleep duration.

Discussion

Global proteomics studies of mice livers revealed large numbers of diurnally rhythmic proteins. It has been estimated that 20–50% of these proteins are dependent on post-transcriptional and post-translational mechanisms (8, 9, 11, 12, 31). Previously, we used targeted proteomics to detect the rhythmicity and concentration of a select number of low abundance core circadian proteins in constant darkness conditions (10). Here, we used

ribosome profiling and quantitative mass spectrometry to understand the relationship between those circadian protein rhythms and mRNA translation.

We found a correlation between the timing of protein production, as measured by ribosome profiling, and peak protein abundance for most transcripts (Fig. 1). However, BMAL1 and CLOCK proteins had a delay between translation and peak protein levels (Fig. 1) caused by delayed protein turnover via post-translational modification pathways (32), rather than a delay in translation after transcription (*SI Appendix*, Fig. S3). The peak protein abundance of PER and CRY is also slightly delayed relative to their peak mRNA abundance (10, 19, 20), and we found this was also due to protein turnover rather than to a delay in translation after transcription (*SI Appendix*, Fig. S3), similar to previous ribosome profiling studies (15, 16, 18). There was a larger delay between the peak abundance of *Cry2* ribosome binding and CRY2 protein than that of *Cry1*, which may simply be a result of the lower amplitude of CRY2 rhythms.

Previous circadian ribosome profiling studies examined translation efficiency (TE), which is defined as the ratio of ribosome profiling reads to bulk RNA-seq reads. We lacked the RNA-seq data needed to compute TE in our samples, which limited our ability to determine how uORF features in 5'UTRs, such as number, length, and distance to CDS, alter TE by accounting for changes in transcript abundance. However, using the closest publicly available RNA-seq data (4) to our ribosome profiling data as a proxy for transcript abundance (*SI Appendix*, Dataset S3), we observed a decrease in TE in transcripts with one or more uORFs (*SI Appendix*, Fig. S16), but no relationship between TE and uORF length or uORF distance to CDS (*SI Appendix*, Fig. S17).

Most transcripts with rhythmic translation in our dataset had a phase around CT0, corresponding to what would be the dark-to-light transition (*SI Appendix*, Fig. S6). Less than 10% of the circadian translatome had significantly rhythmic mRNA transcription according to JTK cycle analysis ($P < 0.05$), and these transcripts were enriched for circadian and metabolic processes by Panther gene ontology (GO) analysis (Fig. S6B, C). Transcripts with a non-significant mRNA rhythmicity by JTK cycle analysis ($P > 0.05$) also were biased toward CT0, but there was no clear GO biological process underlying these rhythms (*SI Appendix*, Fig. S6D, E). The set of transcripts with phases between CT22 and CT02 (*SI Appendix*, Fig. S6F, G) included *Npas2*, *Bmal1*, *Cry1*, *Nfil3*, and *Clock* are all regulated by the nuclear receptor NR1D1, also known as Rev-erbα, which transcriptionally represses many genes involved in metabolism in a tissue- and circadian-dependent manner (33). In the liver, NR1D1 predominantly binds genomic locations at ZT08-ZT10 with little-to-no binding at ZT22 (34–36). Thus, NR1D1 de-repression may be driving the transcriptional, and subsequent translational, peak observed in our data between CT22 and CT02, which may indirectly account for the translation peaks observed in non-rhythmic mRNA transcripts.

We observed widespread binding of ribosomes to uORFs in the 5'UTRs of many core circadian transcripts and found an inverse relationship between uORF number and ribosome occupancy or luciferase expression in transiently transfected reporters or CRISPR/Cas9-generated MEF cells (Fig. 2, *SI Appendix*, Fig. S4, Fig. S12, and Fig. S15). Approximately 50% of mouse and human mRNAs contain uORFs, which are associated with widespread translational repression (28). Thus, uORF-mediated translation suppression may constitutively suppress the abundance of circadian proteins

or provide a post-transcriptional foothold to adjust protein abundance by altering the activities of ribosome reinitiation factors (37–39). Moreover, mutagenesis of uORFs have been shown to alter rhythmicity in other circadian systems (40, 41). Not all ribosome binding in 5'UTRs was associated with uORFs, and we observed extensive ribosome binding in areas without an apparent uORF for *Cry1*, *Nr1d1*, *Clock*, *Dbp*, and *Per3* (Fig. 2 and *SI Appendix*, Fig. S4). For the *Nr1d1* 5'UTR, we observed close overlap in the nucleotides necessary for internal ribosome entry site (IRES)-mediated translation (42) and ribosome binding (*SI Appendix*, Fig. S14A), and confirmed IRES-mediated translation from the *Nr1d1* but not the *Cry1* 5'UTR (*SI Appendix*, Fig. S14B). Near-cognate uORF translation (from a non-AUG start codon) may also drive ribosome binding (14). In silico mapping of near-cognate start codons with ribosome binding data revealed numerous potential near-cognate uORFs (*SI Appendix*, Fig. S14C) but experimental validation is still needed.

The uORF in the *Per2* 5'UTR is an attractive target to understand the role of uORFs in circadian biology, not only because cis-elements within the *Per2* promoter are well-understood (43–46), but also because the *Per2* uORF is too short to encode a peptide. PER2 production is post-transcriptionally controlled by miRNAs (47, 48), antisense transcription (49), and hnRNP1-mediated mRNA degradation (50), which suggests that *Per2* post-transcriptional control is important for producing the optimal amount of PER2 protein. Mutation of the *Per2* uORF increased the amplitude of a reporter plasmid without affecting the period (Fig. 3), similar to effects observed by mutating the *Per2* antisense transcript (51).

Both the abundance and timing of *Per2* expression are critical for maintaining circadian rhythmicity in mice because constitutive expression of PER2, unlike CRY1, disrupts behavioral rhythms (52). Several post-transcriptional mechanisms such as *Per2* anti-sense transcription (49) and miRNAs (47, 48, 53) in the *Per2* genomic locus alter the timing and abundance of PER2 protein expression. For example, replacing the *Per2* 3'UTR with a SV40 late poly(A) signal greatly amplifies bioluminescence rhythms in PER2:LUC mice and increases free-running periods (48). Mutation of the *Per2* uORF increased amplitude in reporter cells (Fig. 3), but did not increase the free-running period in mice (Fig. 4). *Per2* uORF mutant mice did have reduced total sleep in LD conditions (Fig. 4), as observed in *Per1/Per2* double mutant mice under similar sleep phenotyping conditions (29), but no significant change in sleep in DD conditions (*SI Appendix*, Fig. S9). Importantly, although our *Per2* mutation abolished the canonical ATGTGA uORF in the *Per2* 5'UTR, it also introduced a near-cognate uORF (CTGTAG) just two nucleotides upstream of the original uORF. Although we observed increased PER2 protein in *Per2* uORF mutant mice (*SI Appendix*, Fig. S7B) and increased luciferase expression in mutant PER2:LUC MEF cells (*SI Appendix*, Fig. S12), suggesting that the mutation effected protein levels, an alternative mutation that does not have the possibility of introducing a near-cognate uORF would further clarify the role of the *Per2* uORF on PER2 expression. In LD conditions, sleep in mutant mice was particularly reduced during light-to-dark and dark-to-light transitions (Fig. 4), which may result from a lower sleep episode duration and increase transition probability from sleep to wake (*SI Appendix*, Fig. S8B). Disruption of the E'box cis-element, which is only a few base pairs upstream of the *Per2* uORF, abolished molecular oscillations in mutant cells without disrupting the free running period in mice (43). These mice re-entrain quicker under an artificial jetlag

experiment than wild-type mice. Thus, it will be interesting to observe how *Per2* uORF mutant mice behave under similar conditions, especially considering that sleep was disrupted in *Per2* uORF mutant mice particularly at light-to-dark transitions. At a molecular level, de-repressed *Per2* translation in the absence of the *Per2* uORF may alter the timing and abundance of PER2 protein, leading to a reduction in sleep. However, further studies, particularly at a neurological level, are needed to understand the precise mechanism by which sleep is reduced in these animals.

Materials and Methods

Methods

Animals. All animal experiments were approved by the Institutional Animal Care and Use Committee of the RIKEN Kobe branch. Eight-to-ten week-old wild-type male mice (C57BL/6N, Japan SLC) were entrained under 12-h light (400 lx) 12-h dark (LD) for two weeks. Twenty-four hours after transferring to constant darkness (DD), wild-type mice were sacrificed every 4 h over one day (CT0, CT4, CT8, CT12, CT16, CT20, and CT24) for ribosome profiling analysis as in (10). Livers were excised, snap frozen in liquid nitrogen, and stored at -80 °C until use. For details regarding *Per2* uORF mutant mouse construction, see the *SI Appendix*.

Plasmids. For details regarding plasmid construction, see the *SI Appendix*.

Ribosome Profiling. Ribosome profiling was performed essentially as described in (24). Frozen liver samples (~50 mg) were pulverized and then homogenized in 400 µl polysome lysis buffer (150 mM NaCl, 20 mM Tris-HCl pH 7.4, 5 mM $MgCl_2$, 5 mM DTT, 100 µg/ml cycloheximide, 1% Triton X-100, 25 U/ml Turbo DNAse I). Lysates were incubated on ice for 5-10 min, triturated through a 26-G needle 10 times, clarified by centrifugation (20,000 × *g* for 10 min at 4 °C), and 300 µl of supernatant was transferred to a new tube. Unbound RNA was digested by the addition of 7.5 µl RNAse I (ThermoFisher) for 45 min and then stopped by 10 µl SUPERase In RNAse Inhibitor (ThermoFisher). The digestion was transferred to 13 mm × 51 mm polycarbonate ultracentrifuge tubes, layered on top of 0.9 ml sucrose cushion (150 mM NaCl, 20 mM Tris-HCl pH 7.4, 5 mM $MgCl_2$, 5 mM DTT, 100 µg/ml cycloheximide, 1 M sucrose, 10 U/ml SUPERase In), centrifuged in a TLA100.3 rotor at 70,000 rpm at 4 °C for 4 h. Ribosome pellets were resuspended in 0.7 ml Qiazol and mRNA was recovered using the miRNeasy RNA extraction kit (Qiagen) according to the manufacturer's instructions. For library construction and sequencing, see the *SI Appendix*.

Bioinformatic analysis of ribosome profiling. For details regarding sequence processing and alignment, see the *SI Appendix*. For each transcript, non-overlapping five-codon windows were tiled across the coding region, and the transcript was considered well-translated if it had a median value of at least 2 reads per window (excluding the excluding the first fifteen and the last five codons). Using all well-translated transcripts, we created file where each transcript is represented across all samples by its RPKM value. We used this file in JTK_CYCLE (25) to identify all rhythmic transcripts. We used a threshold of 0.05 on the adjusted p-value to assess significance.

To detect uORFs, we processed the 5'UTR of all transcripts, extracted their sequence, and identified all pairs of a start codon (AUG) and a stop codon (UGA, UAA, or UAG) in phase with each other. Using the same read assignment method as in the *SI Appendix*, we also allocated reads aligned to the UTR to their corresponding uORF (where appropriate).

Cells. NIH3T3 and MEFs from PER2::LUCIFERASE (PER2::LUC) knock-in reporter mice (54) were cultured in DMEM (ThermoFisher), 10% FBS (JRH Biosciences), and 1% penicillin/streptomycin (ThermoFisher) at 37 °C in 5% CO_2. For additional details, see the *SI Appendix*.

Real-time Circadian Luciferase Assay. Real-time circadian luciferase assays were performed as previously described (55). Briefly, the day before transfection, NIH3T3 cells or PER2:LUC MEFs were plated onto 35-mm dishes at a density 4×10^5 per well. The following day NIH3T3 cells were co-transfected using FuGene6 (Roche) with 1 µg of the indicated luciferase reporter plasmid according to the manufacturer's instructions (n = 3) and cultured at 37 °C. After 72 h, the media in well was replaced with 2 ml of DMEM containing 10% FBS supplemented with 10 mM HEPES (pH 7.2, ThermoFisher), 0.1 mM luciferin (Promega), antibiotics, and 10 µM forskolin (Fermentek, NIH3T3) or 100 µM dexamethasone (ThermoFisher, PER2::LUC MEFs). Luminescence was measured by a photomultiplier tube (LM2400R, Hamamatsu Photonics) for 1 min at 12 min intervals in a darkroom at 30 °C.

Dual-Luciferase Assay. NIH3T3 cells were plated on 6-well plates at a density of 2×10^5 per well. The following day cells were co-transfected with 0.95 µg of a Firefly *luciferase* reporter plasmid and 50 ng phRL-SV40 plasmid (Renilla luciferase, Promega) as an internal control for transfection efficiency using FuGene6 (Roche). Cells were harvested and assayed by the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions 48 h after transfection.

Preparation of mouse nuclear lysate and immunoblot analysis. Mice were sacrificed by cervical dislocation, and livers were dissected at CT2-4, snap frozen in liquid nitrogen, and stored at -80 °C. Liver extracts were prepared according to (56) with minor modifications. See the *SI Appendix* for further details.

FACS. NIH3T3 cells were plated onto plastic 35-mm dishes or 35-mm imaging dishes (Ibidi) at a density 4×10^5 per well. The following day cells were co-transfected using FuGene6 (Roche) with 0.5 µg of the indicated fluorescence reporter plasmids (1 µg total) according to the manufacturer's instructions (n = 3) and cultured at 37 °C. After 72 h, cells were trypsinized, sorted using FACSAria I (BD Biosciences), and analyzed by FlowJo version 10.8.1.

Sleep Phenotyping. Sleep phenotyping was conducted in in 12-week-old (LD) and 14-week- old (DD) mice in a Snappy Sleep Stager (SSS) using *Per2* uORF mutant mice and wild-type littermates as a control (29). SSS is a non-invasive, respiration-based sleeping staging system in which mice are placed in a chamber connected to a respiration sensor

that detects pressure differences between the outside and inside of the chamber. Detailed methods have been described previously (29). For details regarding sleep and wake parameters, see the *SI Appendix*.

Statistical Analyses. Statistical analyses were performed in R version 3.4.3, Prism 7.0, and custom Jupyter notebooks version 6.3. Two-way ANOVA and a Student's t-test was used to test differences in sleep parameters between wild-type and *Per2* uORF mutant mice. For the Student's t-test, data were confirmed gaussian-distributed by a Shapiro-Wilk normality test and an F test was used to check equality of variances. Lomb-Scargle periodograms were implemented using SciPy version 1.7, and cosinor analysis was implemented using CosinorPy (57). Statistical significance was defined *P < 0.05, **P < 0.01, ***P < 0.001, and n.s. for not significant.

Acknowledgments

We thank Ryohei Narumi for the circadian liver samples and for the use of the mass spectrometry and qPCR data in (10) and Dr. Shigehiro Kuraku for sequencing of the ribosome profiling libraries. We thank Dr. Shihoko Kojima and Dr. Seung-Hee Yoo for PER2::LUC MEFs. We thank all lab members at RIKEN BDR, in particular, Natsumi Hori and Yumika Sugihara for the production of mutant mice; and Masako Kunimi and Ruriko Inoue for breeding and helping phenotype measurements of the mutant mice. We thank all lab members at iFReC Systems Immunology lab and iFReC Laboratory for Host Defense, in particular, Kitiya Piboonprai for help with genotyping and Shizuo Akira for his kind support. This work was also supported by grants from Exploratory Research for Advanced Technology (ERATO) grant (JPMJER2001, H.R.U.) from the Japan Science and Technology Agency (JST), KAKENHI Grant-in-Aid from JSPS (Scientific Research S 18H05270, H.R.U.), MEXT Quantum Leap Flagship Program (MEXT QLEAP) Grant Number JPMXS0120330644, AMED under Grant Number JP20am0401011, and the intramural Grant-in-Aid from the RIKEN Center for Biosystems Dynamics Research to H.R.U. This research was supported by Research Support Project for Life Science and Drug Discovery (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP22ama121025. This work was supported by the Human Frontiers Science Program and a JSPS Grant-in-Aid for Early Career Scientists (18K14755 to A.M.). Two children were born to A.M. during the course of this project, and I thank them and Mari Nishino for their care, love, and support.

References

1. J. A. Mohawk, C. B. Green, J. S. Takahashi, Central and peripheral circadian clocks in mammals. *Annu. Rev. Neurosci.* **35**, 445–462 (2012).

2. A. Millius, H. R. Ueda, Systems Biology-Derived Discoveries of Intrinsic Clocks. *Front. Neurol.* **8**, 25 (2017).

3. R. A. Akhtar, *et al.*, Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. *Curr. Biol. CB* **12**, 540–550 (2002).

4. N. Koike, *et al.*, Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338**, 349–354 (2012).

5. S. P. Gygi, Y. Rochon, B. R. Franza, R. Aebersold, Correlation between Protein and mRNA Abundance in Yeast. *Mol. Cell. Biol.* **19**, 1720–1730 (1999).

6. C. Lim, R. Allada, Emerging roles for post-transcriptional regulation in circadian clocks. *Nat. Neurosci.* **16**, 1544–1550 (2013).

7. A. B. Reddy, *et al.*, Circadian orchestration of the hepatic proteome. *Curr. Biol. CB* **16**, 1107–1115 (2006).

8. D. Mauvoisin, *et al.*, Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal functions in mouse liver. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 167–172 (2014).

9. M. S. Robles, J. Cox, M. Mann, In-vivo quantitative proteomics reveals a key contribution of post-transcriptional mechanisms to the circadian regulation of liver metabolism. *PLoS Genet.* **10**, e1004047 (2014).

10. R. Narumi, *et al.*, Mass spectrometry-based absolute quantification reveals rhythmic variation of mouse circadian clock proteins. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E3461-3467 (2016).

11. M. S. Robles, S. J. Humphrey, M. Mann, Phosphorylation Is a Central Mechanism for Circadian Control of Metabolism and Physiology. *Cell Metab.* **25**, 118–127 (2017).

12. J. Wang, *et al.*, Nuclear Proteomics Uncovers Diurnal Regulatory Landscapes in Mouse Liver. *Cell Metab.* **25**, 102–117 (2017).

13. N. T. Ingolia, S. Ghaemmaghami, J. R. S. Newman, J. S. Weissman, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).

14. N. T. Ingolia, L. F. Lareau, J. S. Weissman, Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**, 789–802 (2011).

15. C. Jang, N. F. Lahens, J. B. Hogenesch, A. Sehgal, Ribosome profiling reveals an important role for translational control in circadian gene expression. *Genome Res.* **25**, 1836–1847 (2015).

16. F. Atger, *et al.*, Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E6579-6588 (2015).

17. V. Castelo-Szekely, A. B. Arpat, P. Janich, D. Gatfield, Translational contributions to tissue specificity in rhythmic and constitutive gene expression. *Genome Biol.* **18**, 116 (2017).

18. P. Janich, A. B. Arpat, V. Castelo-Szekely, M. Lopes, D. Gatfield, Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames. *Genome Res.* **25**, 1848–1859 (2015).

19. C. Lee, J. P. Etchegaray, F. R. Cagampang, A. S. Loudon, S. M. Reppert, Posttranslational mechanisms regulate the mammalian circadian clock. *Cell* **107**, 855–867 (2001).

20. M. D. Field, *et al.*, Analysis of clock proteins in mouse SCN demonstrates phylogenetic divergence of the circadian clockwork and resetting mechanisms. *Neuron* **25**, 437–447 (2000).

21. A. Honkela, *et al.*, Genome-wide modeling of transcription kinetics reveals patterns of RNA production delays. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 13115–13120 (2015).

22. Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–550 (2016).

23. Y. Shimizu, *et al.*, Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19**, 751–755 (2001).

24. N. T. Ingolia, G. A. Brar, S. Rouskin, A. M. McGeachy, J. S. Weissman, The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.* **7**, 1534–1550 (2012).

25. M. E. Hughes, J. B. Hogenesch, K. Kornacker, JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J. Biol. Rhythms* **25**, 372–380 (2010).

26. S. Kojima, D. L. Shingle, C. B. Green, Post-transcriptional control of circadian rhythms. *J. Cell Sci.* **124**, 311–320 (2011).

27. G.-L. Chew, A. Pauli, A. F. Schier, Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nat. Commun.* **7**, 11663 (2016).

28. T. G. Johnstone, A. A. Bazzini, A. J. Giraldez, Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J.* **35**, 706–723 (2016).

29. G. A. Sunagawa, *et al.*, Mammalian Reverse Genetics without Crossing Reveals Nr3a as a Short-Sleeper Gene. *Cell Rep.* **14**, 662–677 (2016).

30. F. Tatsuki, *et al.*, Involvement of Ca(2+)-Dependent Hyperpolarization in Sleep Duration in Mammals. *Neuron* **90**, 70–85 (2016).

31. Y. Wang, *et al.*, A proteomics landscape of circadian clock in mouse liver. *Nat. Commun.* **9**, 1553 (2018).

32. Y. Okamoto-Uchida, *et al.*, Post-translational Modifications are Required for Circadian Clock Regulation in Vertebrates. *Curr. Genomics* **20**, 332–339 (2019).

33. L. J. Everett, M. A. Lazar, Nuclear Receptor Rev-erbα: Up, Down, and All Around. *Trends Endocrinol. Metab. TEM* **25**, 586–592 (2014).

34. D. Feng, *et al.*, A circadian rhythm orchestrated by histone deacetylase 3 controls hepatic lipid metabolism. *Science* **331**, 1315–1319 (2011).

35. A. L. Hunter, *et al.*, Nuclear receptor REVERBα is a state-dependent regulator of liver energy metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 25869–25879 (2020).

36. Y. Zhang, *et al.*, Discrete Functions of Nuclear Receptor Rev-erbα Couple Metabolism to the Clock. *Science* **348**, 1488–1492 (2015).

37. J. Bohlen, *et al.*, DENR promotes translation reinitiation via ribosome recycling to drive expression of oncogenes including ATF4. *Nat. Commun.* **11**, 4676 (2020).

38. D. R. Morris, A. P. Geballe, Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.* **20**, 8635–8642 (2000).

39. S. Schleich, *et al.*, DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. *Nature* **512**, 208–212 (2014).

40. A. C. R. Diernfellner, T. Schafmeier, M. W. Merrow, M. Brunner, Molecular mechanism of temperature sensing by the circadian clock of Neurospora crassa. *Genes Dev.* **19**, 1968–1973 (2005).

41. H.-W. Wu, *et al.*, Noise reduction by upstream open reading frames. *Nat. Plants* **8**, 474–480 (2022).

42. D.-Y. Kim, K.-C. Woo, K.-H. Lee, T.-D. Kim, K.-T. Kim, hnRNP Q and PTB modulate the circadian oscillation of mouse Rev-erb alpha via IRES-mediated translation. *Nucleic Acids Res.* **38**, 7068–7078 (2010).

43. M. Doi, *et al.*, Non-coding cis -element of Period2 is essential for maintaining organismal circadian behaviour and body temperature rhythmicity. *Nat. Commun.* **10**, 2563 (2019).

44. N. Gekakis, *et al.*, Role of the CLOCK Protein in the Mammalian Circadian Mechanism. *Science* **280**, 1564–1569 (1998).

45. H. Hao, D. L. Allen, P. E. Hardin, A circadian enhancer mediates PER-dependent mRNA cycling in Drosophila melanogaster. *Mol. Cell. Biol.* **17**, 3687–3693 (1997).

46. H. R. Ueda, *et al.*, System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nat. Genet.* **37**, 187–192 (2005).

47. I. Park, *et al.*, microRNA-25 as a novel modulator of circadian Period2 gene oscillation. *Exp. Mol. Med.* **52**, 1614–1626 (2020).

48. S.-H. Yoo, *et al.*, Period2 3'-UTR and microRNA-24 regulate circadian rhythms by repressing PERIOD2 protein accumulation. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E8855–E8864 (2017).

49. R. A. Mosig, *et al.*, Natural antisense transcript of Period2, Per2AS, regulates the amplitude of the mouse circadian clock. *Genes Dev.* **35**, 899–913 (2021).

50. K.-C. Woo, *et al.*, Mouse period 2 mRNA circadian oscillation is modulated by PTB-mediated rhythmic mRNA degradation. *Nucleic Acids Res.* **37**, 26–37 (2009).

51. R. A. Mosig, *et al.*, Natural Antisense Transcript of Period2, Per2AS, regulates the amplitude of the mouse circadian clock. *bioRxiv*, 2020.07.27.222760 (2020).

52. R. Chen, *et al.*, Rhythmic PER abundance defines a critical nodal point for negative feedback within the circadian clock mechanism. *Mol. Cell* **36**, 417–430 (2009).

53. R. Nagel, L. Clijsters, R. Agami, The miRNA-192/194 cluster regulates the Period gene family and the circadian clock. *FEBS J.* **276**, 5447–5455 (2009).

54. S.-H. Yoo, *et al.*, PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 5339–5346 (2004).

55. M. Ukai-Tadenuma, *et al.*, Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell* **144**, 268–281 (2011).

56. S. Masuda, *et al.*, Mutation of a PER2 phosphodegron perturbs the circadian phosphoswitch. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 10888–10896 (2020).

57. M. Moškon, CosinorPy: a python package for cosinor-based rhythmometry. *BMC Bioinformatics* **21**, 485 (2020).

Figures

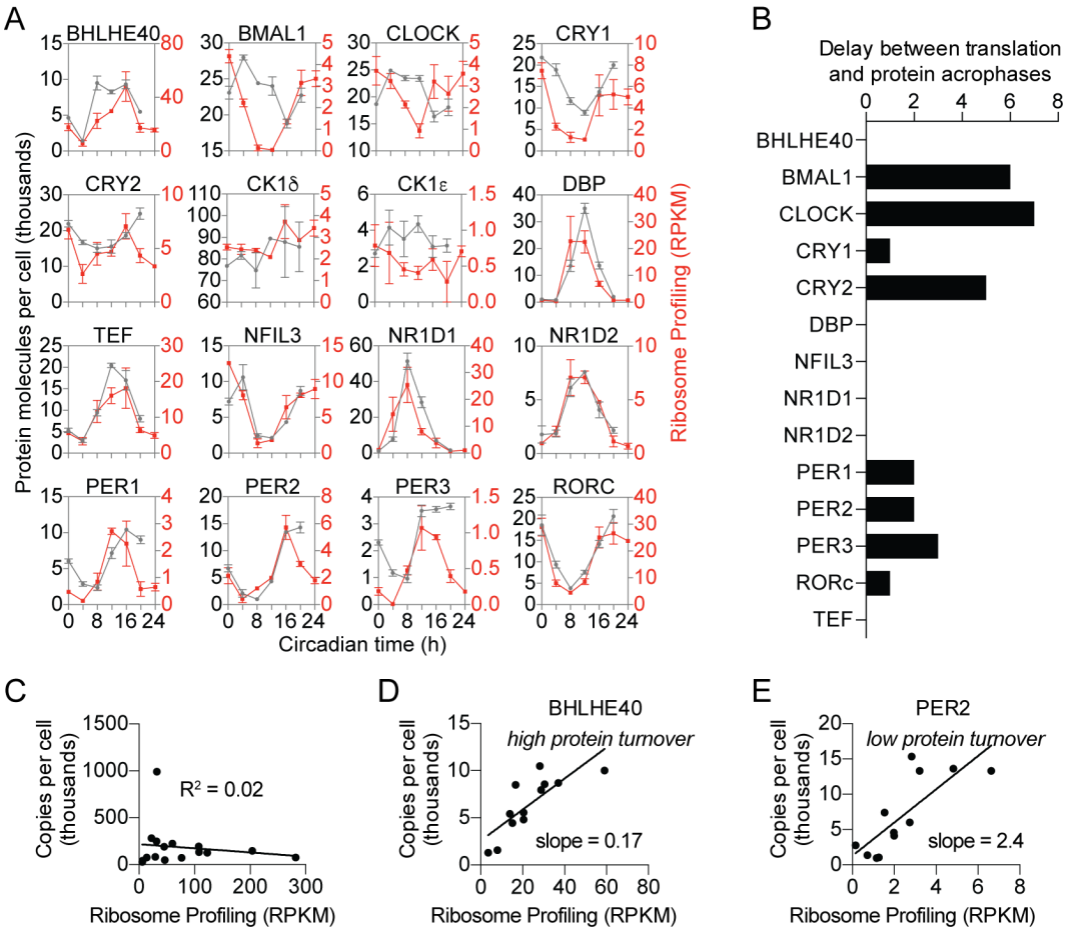

Figure 1 | Measuring translation and protein from the same liver samples for 16 circadian mRNAs. (A) Time course of the quantified proteins and the corresponding translation level. Protein concentration (gray) is expressed as copies per cell assuming the total amount of protein per cell is 0.5 ng for the average for all quantified peptides (10). Ribosome binding (red) is expressed as reads per kilobase of transcript per million mapped reads (RPKM). Data are mean ± SEM. (B) Phase delay between ribosome binding and protein abundance. Copies per cell were averaged from all peptides for each protein. JTK analysis was used to estimate the phases for these averaged protein levels and for ribosome binding RPKM levels. (C) Linear regression analysis comparing ribosome binding to protein amount averaged over 24 hours showed that there was no correlation between ribosome binding (protein biosynthesis output) and protein abundance (D and E) Linear regression analysis comparing ribosome binding to protein amount at each sampled time point showed that some proteins like BHLHE40 (D) had a high protein turnover whereas other proteins like PER2 (E) had a lower protein turnover.

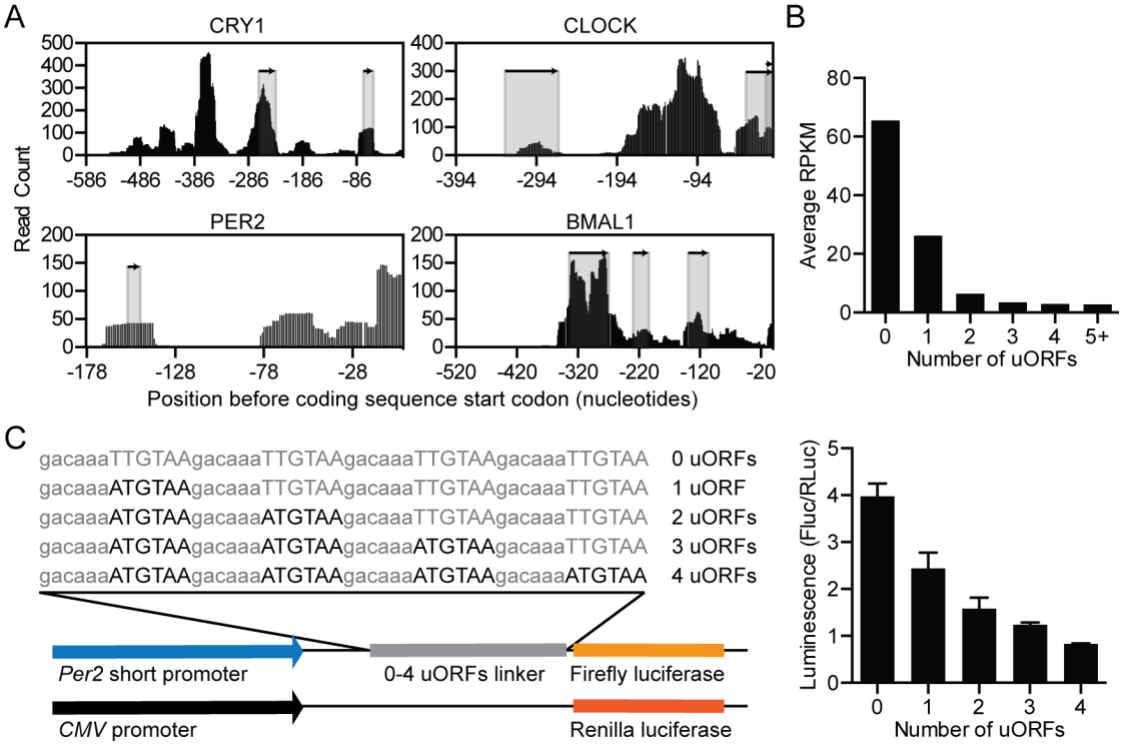

Figure 2 | Upstream open reading frames (uORFs) suppress protein expression. (A) The 5'UTRs of *Cry1*, *Clock*, *Per2*, and *Bmal1* transcripts contain at least one uORF (shaded regions). Distribution of raw read counts from ribosome profiling data (black bars) show ribosomes bounds in regions corresponding to uORFs. (B) Global distribution of uORFs compared to ribosome profiling translation level (RPKM) shows that mRNAs with more uORFs are translated less compared to mRNAs with fewer uORFs. (C) Introduction of a variable number of synthetic uORFs represses relative luminescence from a *Per2* short promoter in a dose-dependent manner.

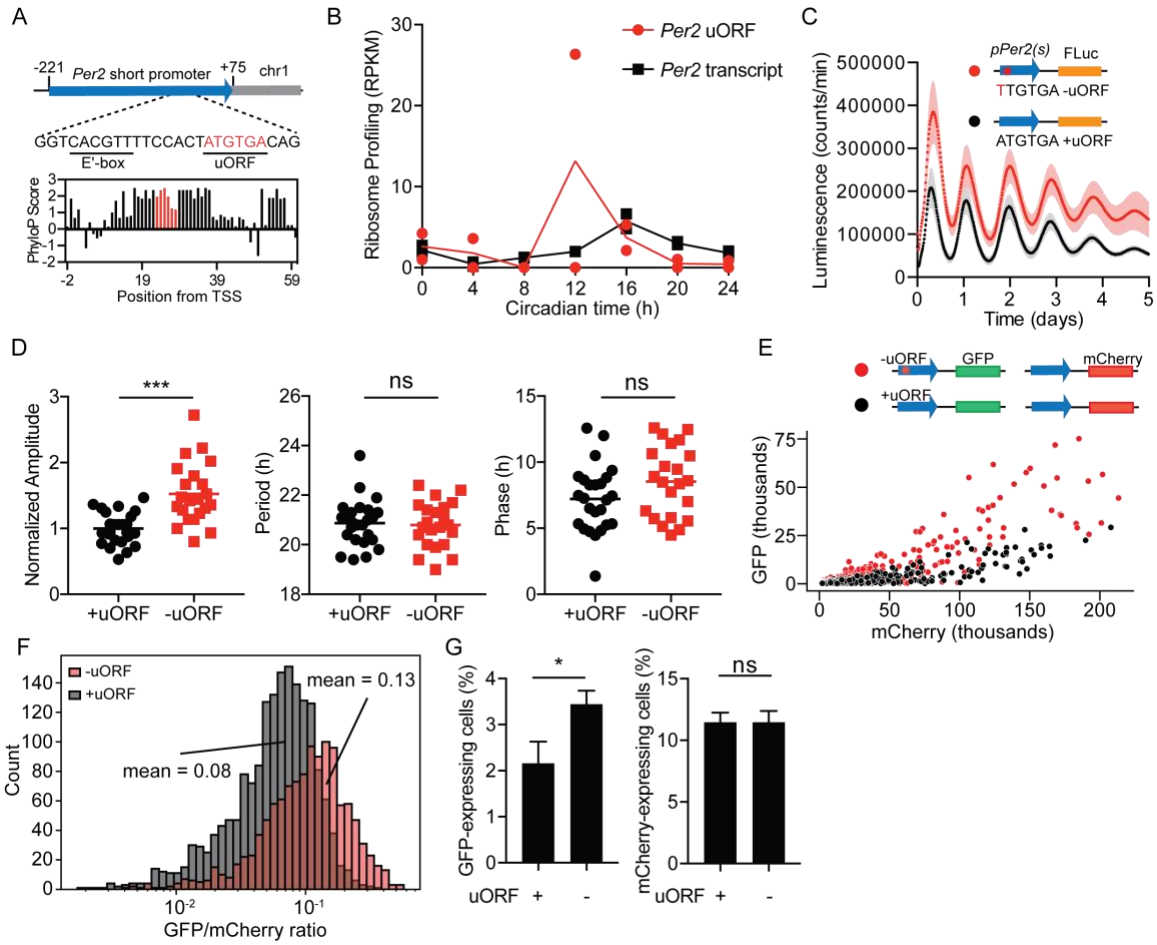

Figure 3 | The *Per2* uORF is evolutionarily conserved and suppresses PER2 expression. (A) Genomic position of the Per2 short promoter, which contains an E'box and uORF. Evolutionary conservation scores according to UCSC genome browser among 22 mammalian species in the *Per2* 5'UTR (black bars) with the uORF highlighted (red bars). (B) Ribosome binding (RPKM) on the *Per2* uORF (red) compared to the *Per2* transcript (5'UTR, CDS, and 3'UTR, black). (C) Representative bioluminescent traces showing that mutation of the *Per2* uORF (ATGTAA to TTGTGA) increases the expression level of a pGL3-P(*Per2*)-d*Luc* luminescent reporter. Reporter containing uORF (black), reporter with a mutant uORF (red). Shaded region is SD. (D) Cosinor analysis showing the normalized amplitude (*left*), period (*middle*), and phase (*right*) for cells transfected with the *Per2* uORF (black) or a mutant *Per2* uORF (red). Data from each group comprise at least 22 traces from 8 different experiments (see *SI Appendix*, Fig. S5 for the plots of individual traces for each experiment). (E) Representative scatterplot of GFP versus mCherry expression for 3T3 cells transfected with a GFP-fluorescence reporter with a *Per2* uORF (+uORF, black) or a mutant *Per2* uORF (-uORF, red) and a mCherry-normalization plasmid. The analyzed cells were gated on mCherry+ expression. (F) Histogram of the GFP/mCherry ratio for cells in (E). (G) The percentage of GFP+ cells (*left*) was significantly higher in the -uORF population compared to the +uORF population, whereas the percentage of mCherry+ cells (*right*) was unchanged (n = 3 independent experiments). See *SI Appendix*, Fig. S11 for FACS gating strategy and analysis of cells from the GFP+ gate.

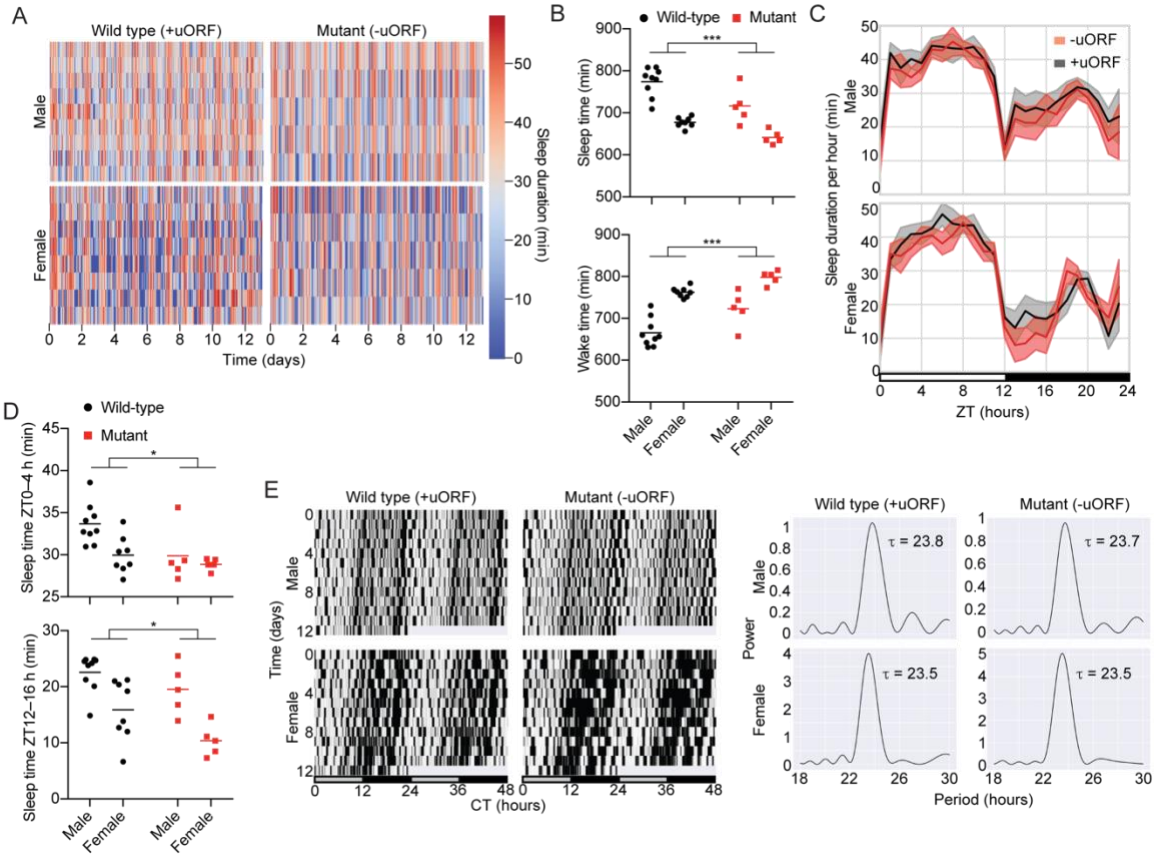

Figure 4 | Disruption of the *Per2* uORF reduces sleep in mice. (A) Sleep duration per hour over 13 d in 12-h light/12-h dark (LD) conditions for *Per2* uORF mutant and wild-type male and female mice. Each row indicates data from one mouse. (B) Mean sleep (*top*) and wake (*bottom*) duration over 24 h, averaged over 13 d. Red, *Per2* uORF mutant mice. Black, wild-type. (C) Sleep duration per hour over 24 h in LD, averaged over 13 d for *Per2* uORF mutant (red) and wild-type (black) male (*top*) and female (*bottom*) mice. Lines indicate mean sleep duration at each time of day for each strain. Shaded area, SD at each time point. (D) Mean sleep duration per hour averaged over 13 d during the early morning (*top*) and early evening (*bottom*). Per2 uORF mutant (red) and wild-type (black) dots are individual mice; lines indicate mean. (E) Representative double-plotted actograms of sleep duration per 6 min bins of wild-type and *Per2* uORF mutant mice in constant darkness (DD) over 12 d (*left*) and corresponding periodograms (*right*).