SUPPLEMENTAL MATERIAL

Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames

Peggy Janich, Alaaddin Bulak Arpat, Violeta Castelo-Szekely, Maykel Lopes, David Gatfield

Supplementary Figure S1. Ribosome profiling and RNA-seq in mouse liver.

A) Insert size distribution of RNA-seq reads across all replicates and timepoints. Box-and-whisker plots: midline, median; box, 25^{th} and 75^{th} percentiles; whiskers, extend to the minimum and maximum values within 1.5 times the interquartile range from the box. While most RNA-seq reads are in the range of around 30 nt (i.e., similar to RPF-seq reads), the distribution is broader, as expected for chemically fragmented RNA. This graph is complementary to the RPF-seq distribution in main Fig. 1E.

B) Intra- and inter-group relationships between normalised CDS counts of RPF-seq libraries are visualised with a heat map of Spearman correlation coefficients. Size and dark shading of disks increase with increasing correlation coefficients; colour correspondence is given in the legend. Average Spearman correlation for biological replicates: 0.987 ± 0.004 (average ± SD).

C) Same as B), for RNA-seq counts. Average Spearman correlation for biological replicates: 0.985 ± 0.004 (average ± SD).

D) Scree plot of the PCA shown in main Fig. 1H. Components beyond the first three PCs shown in Fig. 1H only explain a small percentage of additional variation in the data.

Supplementary Figure S2: Transcript length features and uORF usage as predictors of translation efficiency.

A-C) Partial regression plots for the modelling of TE on the three variables, 5' UTR (A), CDS (B) and 3' UTR (C) lengths. Such partial regression plots are useful to investigate the relationship between a predictor and the response variable taking into account the effect of the other predictors. Both 5' UTR and CDS lengths show significant predictive value for the efficiency at which the transcripts' main ORFs are translated. The predictive value of 3' UTR length for TE is, however, not significant in this linear regression analysis. The statistically significant correlation between 3' UTR length and TE that was originally observed in main Fig. 2B therefore likely results from the correlation between 3' UTR with CDS length. Analysis was performed on transcripts from single protein isoform genes with an RPKM value >5 and 5' and 3' UTR lengths ≥10 nt (N=4277). In red, the slope of the red regression line and the significance level (t-test) are reported.

D) uORF translation correlates with low TE independently of 5' UTR length. The scatter plot of translation efficiency vs. 5' UTR length shows that transcripts without translated uORF (violet) have globally shorter 5' UTRs than those with translated uORFs (green). Nevertheless, the locally weighted fits (Lowess) indicate that even for transcripts with UTRs of similar length, uORF-containing transcripts have lower TEs (i.e., the green line is below the violet line throughout). The effect of low TEs is thus an effect of uORF translation and not merely of 5' UTR length. Same transcript set as in (A-C) was used.

E) Same analysis as D) for CDS length. The lower TEs of uORF-containing transcripts is seen independently of CDS length (i.e., the green line is below the violet line throughout).

F) Same analysis as D) for 3’ UTR length. The lower TEs of uORF-containing transcripts is seen independently of 3’ UTR length (i.e., the green line is below the violet line throughout).

G-K) Partial regression plots for the modelling of TE on the four variables, 5’ UTR (G), CDS (H), 3’ UTR (I) lengths and uORF presence (K). uORF-containing transcripts are depicted in green (N=920) and other in black (N=3357). uORF presence is thus a strong predictor of reduced TE (K). When compared to the three predictor model (A-C) it becomes clear that with the inclusion of uORF translation as a forth predictor, the 5’ UTR length lost much of its predictive value. This indicates that to a large extent the 5’ UTR effect seen in A) is in reality a uORF effect, whereby longer 5’ UTRs will be more likely uORF-containing than short 5’ UTRs. Given that relatively conservative uORF detection criteria were applied (e.g. only AUG-initiated uORFs with a certain minimal coverage) it is possible that 5’ UTR length *per se* has even lower (or no) predictive value for TE.

Supplementary Figure S3: Read distributions to 5’ UTR, CDS and 3’ UTR per transcript.

A) Read distribution within 5’ UTRs, CDS and 3’ UTRs of RPF-seq (blue) and RNA-seq data (orange) compared to the distribution expected by chance, which is determined by the feature sizes (grey). Box-and-whisker plots: midline, median; box, 25^{th} and 75^{th} percentiles; whiskers, extend to the minimum and maximum values within 1.5 times the interquartile range from the box. Note that 5’ UTR-mapping RPF-seq reads make up on average 6% of all reads (Fig. 1D), but that the actual percentage *per transcript* is variable (N=10829).

Supplementary Figure S4. Pause site occurrence is not correlated with TE.

A) Distribution of per-codon RPF-seq and RNA-seq counts. The plot shows the cumulative distribution of footprint (blue) counts at each codon, relative to the median density across the gene. In analogy to the method previously applied and published by (Ingolia et al. 2011) (see Fig. 2A in the cited publication), stall sites were defined at an arbitrary threshold for individual codons that exceeded read counts that are 25 x the median on the transcript (in our analysis we used the trimean instead of the median; see Supplemental Material). RNA-seq counts (red) control effects of library generation. The twelve blue and red curves represent the means of the 12 timepoints. For B), only pause sites were further considered when they appeared in the RPF-seq counts of both replicates of a timepoint. The dotted lines show the data from (Ingolia et al. 2011) for comparison and indicated that the overall pause site distribution is similar in liver and mESCs.

B) Pause sites are not associated with TE changes. Transcripts that did (red) or did not (blue) contain pause sites detected in A) were plotted according to their mRNA abundance (abscissa) and their TE (ordinate). Density curves for highlighted data points are plotted on the margins with the same colour code. Red numbers on density curves reflect the location shift (log2 values) of the transcripts with pause sites relative to those without pause sites; neither shift was significant at the 0.05 level (Wilcoxon rank sum test). A cut-off on RPF-seq RPKMs was applied – shown by the dotted diagonal – because for very low footprint density the probability to detect 25 x median levels is very low and would lead to false-negative pause site detection due to coverage effects.

91 The applied cut-off corresponds to the minimal RPKM at which pause sites were still
92 detectable.

Supplementary Figure S5. Constant, but widely differing TEs in clock transcripts.

A) Time-resolved TE plots for the core clock transcripts. This graph is complementary to main Fig. 3B and shows the ratios of RPF-seq/RNA-seq RPKMs around-the-clock. Means are plotted with error bars connecting the two biological replicates per timepoint. Grey dotted lines show a harmonic fit to the data points, and grey shading the 95% confidence interval of the fit. A fit was performed irrespective of whether rhythmicity was significant or not (when no fit was possible, e.g. for *Npas2* due to extremely low expression at ZT18, a straight line was plotted). Dark grey curves (e.g. for *Nr1d1*, *Per3*) indicated that the fit itself was significant, whereas light grey curves indicated non-significant harmonic fits or a fit with a <1.5 x peak-to-trough amplitude. Importantly, the meaning of a "significant fit" to TEs should not be over-interpreted as it does *not* consider if the TEs were significant between timepoints from the start – this important prerequisite was tested using the more sophisticated *Babel* analysis (Olshen et al. 2013). Indeed, none of the core clock transcripts (including *Nr1d1* and *Per3*, for which the harmonic fits themselves were significant) showed significant TE changes according to *Babel* analysis (Supplemental Table S3). It can therefore be concluded that core clock TEs were not subject to time-of-day-dependent control.

B) Bar graph in the top panel shows the mean ± SD of translation efficiencies over the day. Middle and lower panel show the sum of RNA abundance and ribosome footprint RPKMs integrated over-the-day, respectively, for the circadian core clock genes listed below. Ribosome footprints (as RPKMs) are a direct measure of protein biosynthesis rate and can thus be used to calculate the relative daily protein production of the core

clock components. Large differences that are seen at the level of RNA abundance (e.g. *Nr1d1* and *Nr1d2*) are thus compensated by TE, leading to protein output that explains some of the core clock biology better than the RNA levels do. See text for details.

C) RNA abundance and ribosome footprint plots as in B), but using the peak expression timepoints only.

Figure S6. Relationships between RPF and RNA rhythmicity parameters.

A) Evaluation of phase differences (left panel) and amplitude differences (right panel) between ribosome occupancy (RPF) and RNA abundance (RNA) of 'mRNA and footprints rhythmic' genes. The deviations of the means from 0 are not significant, indicating synchrony in RNA abundance and footprint rhythms for the vast majority of transcripts in the overlapping set.

B) Distinct out-of-phase translation was confined to few exceptional cases, which were initially identified by overlaying the 'mRNA and footprints rhythmic' set with the list of transcripts for which the *Babel* analysis had revealed significant TE changes between timepoints (Supplemental Table S3). The resulting 93 genes – many of which showed temporal changes in TE, but similar phase in RNA-seq and RPF-seq profiles – were further examined by visual inspection to identify out-of-phase translation events. Shown are the two most striking cases, Biogenesis of lysosomal organelles complex-1 subunit 1 *(Bloc1s1*) and Histone cluster 1, H4h *(Hist1h4h),* which both showed remarkable out-of-phase peaks in their mRNA and footprint profiles. Means per timepoint are plotted with error bars depicting the replicates.

C) Phase distribution (top panels) and heat maps (bottom panels) of genes identified as rhythmic only at the mRNA level (left) or only at the footprint level (right). mRNA abundances and footprint levels are standardised within each gene. Note that many of these supposedly non-rhythmic transcripts still had underlying, yet noisier rhythms that apparently escaped the detection algorithm – a problem that prompted us to use the

141 *Babel* method with the aim of obtaining a high-confidence set of “mRNA flat – footprints

142 rhythmic” transcripts (see main text and Fig. 5A).

Figure S7. Rhythmic control of ribosomal protein biosynthesis.

A) Table summarising the genes encoding ribosomal proteins of the 40S and 60S ribosomal subunits. Highlighted are those genes that were either detected as "mRNA flat – footprints rhythmic" in main Fig. 4A (bold print), or that had a similar profile as those in Fig. 4A (as judged by visual inspection) but were below the amplitude cut-off of >1.5 fold (normal print), or were not rhythmic, or not detected in our datasets altogether (grey print). The vast majority of *Rps* and *Rpl* transcripts thus undergoes similar regulation at the translational level.

B) Scatter plot of translation efficiencies and transcript abundances for *Rpl* and *Rps* genes of the cytoplasmic (pink) or mitochondrial (green; used as control) ribosome for timepoints ZT2 and ZT10 (small subunit proteins are shown in lighter shadings; large subunit proteins in darker shadings). Density curves of translation efficiencies and transcript abundances for highlighted data points are plotted on the margins with the same colour code. Numbers on density curves reflect the location shift (log2 values) relative to all genes (grey); a (*) denotes significance at the 0.05 level (Wilcoxon rank sum test). The two plots visualise the shift in translation efficiency of mRNAs of cytoplasmic ribosomal proteins between ZT2 and ZT10. Wilcoxon rank sum test indicated that only RPs (RPLs+ RPSs) changed their TEs (location shift=0.864, p=3.016e-10), whereas MRPs (MRPLs+MRPSs) remained unchanged (location shift=0.009, p=0.916) between ZT2 and ZT10.

See also Supplemental Movie M1.

Figure S8. Rhythmicity in iron metabolic genes and validation of other cases of rhythmic translation.

A) Mitochondria encode 13 protein-coding genes, of which the expression of 12 (exception: *mt-Nd4l*) was detectable in our datasets and all showed a characteristic translational spike at ZT12 (Supplemental Dataset 1). As an example, this graph shows RPKM-normalised mRNA abundance (orange) and footprint (blue) profiles for *mt-Co1* around the 24-hour daily cycle.

B) Western blot analysis for MT-CO1 and NDUFB8 proteins (loading control) in liver mitochondrial extracts. Numbers show relative MT-CO1 protein levels across different timepoints after normalisation to NDUFB8 protein levels. Data from one representative time series are shown (N=3). Note that despite the translational peak at ZT12, no protein rhythmicity was seen. An explanation may be the reported high stability of mitochondrial proteins (Kim et al. 2012). Nevertheless, one may speculate that the rhythmic translation has a role in assuring the coordinated biosynthesis of the individual proteins, thus enabling an efficient assembly of newly synthesised mitochondrial complexes.

C) RPKM-normalised mRNA abundance (orange) and footprints (blue) profiles for *Ggps1* around the 24-hour daily cycle.

D) Western blot analysis of GGPS1 was performed in total liver extracts (N=2). The sum of the signals from both bands was used for quantification. Beta-Tubulin served as a loading control and for normalisation. Note that rhythmic translation led to rhythmic protein accumulation.

E) Profiles for *Mxi1* similar as in A).

F) Rhythmic regulation of transferrin receptor (*Tfrc*) mRNA stability around the clock. The graph shows the temporal profile of pre-mRNA (a proxy for transcription rate; green) and mRNA abundance (orange) profiles. Note that the mRNA is rhythmic, but not the pre-mRNA, indicating that oscillations are engendered at the level of mRNA stability. Extracted from the RNA-seq data in (Du et al. 2014); error bars represent the two replicates of the time series. Of note, post-transcriptional rhythmicity of *Tfrc* mRNA was independently confirmed by (Le Martelot et al. 2012).

G) Schematic illustrating the various steps involved in the maintenance of cellular iron homeostasis. The small gene expression graphs show mRNA abundance (orange) and footprint profiles (blue) of rhythmically expressed genes associated with key regulatory processes of iron metabolism. Gene names highlighted in blue mark transcripts with rhythmic footprints shown in Fig. 5G. Moreover, marked rhythms in mRNA abundance were thus seen at the level of regulators of hepcidin expression, which is a liver-produced hormone that is crucial for the maintenance of systemic iron homeostasis and that was recently found to oscillate in human blood (Schaap et al. 2013). Moreover, iron uptake *(Tfrc, Slc25a37/Mfrn)*, export *(Slc11a2/Dmt1, Slc39a14/Zip14)*, storage *(Ftl1, Fth1)* and transport *(Pcbp4),* as well as several biosynthetic processes that use iron *(Alas1, Alas2, Glrx5)* all showed pronounced rhythmicity.

Figure S9. Analysis of uORF usage in circadian core clock genes.

A) Normalised footprint read counts for uORF1 (left panel), uORF2 (right panel) and the main ORF (red) for the *Nr1d1* transcript. uORF and main ORF translation thus occur in sync.

B) Sequence alignment of the hypothetical 63 amino acid Nr1d1-uORF1 polypeptide across 6 mammalian species (mouse, rat, human, chimpanzee, macaque, horse). The uORF1 sequence is well conserved even at the amino acid level.

C) Schematic representation of the *Nr1d1* 5' UTR mutants that were tested for influence on main ORF translation. Nr1d1-Wt represents the annotated, full-length *Nr1d1* 5' UTR. In Del-uORF1, the first uORF was deleted as a whole. In Del-uORF1+2 a region encompassing both uORFs was removed. While these deletion mutants can be considered relatively invasive as they not only remove the uORF but also shorten the UTR as a whole, the following point mutants are likely more specific and suitable to measure uORF translation effects. M1A-uORF1 thus carries a point mutation at the predicted AUG start codon of uORF1 (converting it to an alanine codon). M1A-uORF2 is mutated to an alanine codon at the AUG of uORF2, whereas M1A-uORF1+2 is a combination of the two point mutations.

D) Raw (not detrended) data for the Nr1d1-Luc reporter traces shown in main Fig. 6C.

E) Summary of luciferase bioluminescence signal at the first peak of rhythmicity recordings as that in D) from several experiments (relative to Nr1d1-Wt which was set to 1 within each experiment). The signals of the mutants thus showed variability which conceivably originated from several sources: On the one hand, it may represent a true

effect of impaired uORF translation, but quantitatively a likely stronger influence are technical parameters that are not controlled for in this assay. These are in particular the efficiency of lentiviral production and resulting lentiviral titres, the transduction efficiency of the NIH3T3 cells, the cellular growth rate and cell density in the transduced cells in the dishes subject to real-time recordings, etc. For these reasons, the UTRs shown in C) were recloned in a different assaying vector containing *Renilla* luciferase as an internal control gene, shown in main Fig. 6D.

F) shRNA-mediated *Denr* knockdown causes a short period phenotype of free-running circadian rhythms in NIH3T3 cells also using the *Bmal1-Luciferase* reporter (Nagoshi et al. 2004). Left panel: Summary of period change engendered by *Denr* shRNAs 1-3 relative to Scr (control) shRNA. Data are from 3 individual assayed dishes per condition from a single experimental series and show period length differences relative to the average of the Scr cells (error bars represent standard deviations). Within the experiment, the period shortening was significant for all three shRNAs at the p=0.05 level (t-test). Right panel: Representative bioluminescence tracks of Scr (control) and *Denr* shRNA-transduced *Bmal1-Luc* cells.

Figure S10. Comparison of published rhythmic proteome datasets.

A) Euler diagram showing the overlap in detected rhythmic and non-rhythmic proteins in the two publications (Mauvoisin et al. 2014; Robles et al. 2014). Robles et al. analysed a total of 2883 detected proteins, and Mauvoisin et al. 4197 detected proteins. 2482 proteins are present in both datasets, corresponding to 86% (2482/2883) of the Robles datasets and 59% (2482/4197) of the Mauvoisin dataset. 179 proteins are rhythmic according to Robles et al., and 191 proteins according to Mauvoisin et al., but here the overlapping set only consists of 32 proteins (18% and 17% of Robles and Mauvoisin rhythmic proteins, respectively). Only in 13 (Robles) and 52 (Mauvoisin) cases the underlying reason was absence from the other dataset, indicating that the relatively poor overlap is not only the result of proteome coverage differences. A likely cause are differences in rhythmicity algorithms, cut-offs, FDRs, etc., that were applied.

B) Area-proportional Venn diagram of the subset of proteins that Robles et al. (left) and Mauvoisin et al. (right) classified as “mRNA flat – protein rhythmic”. Only 3 proteins are in common, namely Albumin (*Alb*), Apolipoprotein A-I (*Apoa1*) and Progesterone receptor membrane component 1 (*Pgrmc1*).

Figure S11. Known regulators of translation of IRE-containing transcripts do not show high-amplitude rhythmicity.

A) The *Aco1* gene encodes iron regulatory protein 1 (IRP1), one of the main actors of IRE-controlled translational regulation. Neither at the mRNA and footprint level (left panel) nor at the protein level (right panel) does ACO1 show high-amplitude rhythmicity. The fluctuations in ACO1 protein shown in this blot were not consistently found in a second protein series from independent animals and should therefore likely not be considered oscillating protein expression. The western blot shows one out of N=2 series that were analysed.

B) As in A), but for the *Ireb2* gene/IREB2 protein (also known as iron regulatory protein 2, IRP2). A slight rhythm – that did not pass the 1.5x amplitude cut-off applied in our study – can be seen at the level of mRNA and footprint abundance. At the protein level, fluctuations were not consistent between the two western blots that were performed on independent extract series, suggesting that they did not represent robustly regulated rhythmicity.

C) mRNA and footprint quantification for *Fbxl5*, a known regulator of IREB2 protein levels. No rhythm is detectable. Due to the lack in a functioning antibody (several commercial sources tested), confirmation by western blot as in A) and B) was not possible.

SUPPLEMENTAL INFORMATION

Supplemental Table S1

Table reports the total number of sequenced reads for all RPF-seq and RNA-seq libraries reported in this study, including their mapping summary. Related to Fig. 1.

Supplemental Table S2

Table lists genes that were identified as rhythmic in Fig. 4A. Genes are thus either categorised as rhythmic mRNA only (set B), rhythmic footprints only (set C) or 'mRNA and footprints rhythmic' (set D). Calculated amplitude and phase values are reported for RPF-seq and RNA-seq in separate columns. Cyclic events with an amplitude >1.5 were considered in the analyses of this study; column 'cyclic fit' indicates the type of fit, sigmoid or sinusoidal, that the model selection algorithm assigned to the temporal profile (see Supplemental Experimental Procedures). Related to Fig. 4.

Supplemental Table S3

Gene list of transcripts with significant changes in ribosome occupancy after *Babel* analysis (sub-classified according to whether the change was observed "between" or "within" timepoints). In addition, genes are grouped according to their rhythmicity category (A: mRNA non-rhythmic – footprints non-rhythmic; B: mRNA rhythmic – footprints non-rhythmic; C: mRNA non-rhythmic – footprints rhythmic; D: mRNA rhythmic – footprints rhythmic). Related to Fig. 5 and Supplemental Fig. S6.

Supplemental Movie M1

Animated scatter plots showing translation efficiencies (ordinate) and transcript abundances (abscissa) of mitochondrial (green, as control) and cytoplasmic (pink) ribosomal protein transcripts across all timepoints (genes encoding ribosomal proteins of small subunits are shown in lighter shadings; ribosomal proteins of large subunits, darker shadings; timepoint replicates were averaged). Note the shift in translation efficiency of cytoplasmic ribosomal proteins starting around the day-night transition (ZT10). Related to Fig. 5.

Supplemental Dataset 1

This .zip file contains four folders with transcriptome-wide gene expression plots according to the rhythmicity categorisation in Fig. 4A (A-mRNA non-rhythmic, footprints non-rhythmic; B-mRNA rhythmic, footprints non-rhythmic; C-mRNA non-rhythmic, footprints rhythmic; D-mRNA rhythmic and footprints rhythmic). Left panels show RPF-seq (blue) and RNA-seq (orange) values, with error bars connecting the two biological replicates for each timepoint. The rhythmicity fits are plotted as well. Right panels show translation efficiencies around-the-clock (error bars connecting the two biological replicates). Dark grey curve fits indicate rhythmicity, light grey curves indicate that the harmonic rhythmicity fit was either non-significant or did not pass the 1.5-fold amplitude cut-off. The grey shaded area shows the 95% confidence interval of the fits. Please note that the rhythmic fits on the TE data points are not used for further analyses because they do not consider if the differences in the TEs themselves were significant

318 (they only consider the significance of the fit itself). In order to obtain the group of
319 genes with significantly changing TEs, the *Babel* analysis was performed.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Cell culture

NIH3T3, NIH3T3-Dbp-Luc (Stratmann et al. 2012), NIH3T3-Bmal1-Luc (Nagoshi et al. 2004) and HEK 293FT cells were cultured under standard conditions (DMEM; 10% FCS, 1% penicillin/streptomycin; all from Invitrogen; 37°C; 5% CO_2). Lentiviral particle production (from plasmids prLV1-*Nr1d1* or pLKO.1-shRNA, using envelope and packaging plasmids pMD2.G, psPAX2) and viral transduction of NIH3T3 cells followed published protocols (Salmon and Trono 2007). Puromycin selection in shRNA experiments occurred at 5 μg/ml for 4 days. Dexamethasone synchronisation and recording of circadian bioluminescence rhythms has been described (Stratmann et al. 2012). Detrending of raw bioluminescence data was performed using a 24-hour moving average. For dual luciferase assays, cells were lysed in passive lysis buffer and luciferase activities were determined using the DualGlo Luciferase assay system and a GloMax 96 Microplate luminometer (all Promega). For protein quantification, cells were lysed in RIPA-buffer and analysed by SDS-PAGE and western blotting.

Ribosome profiling

Freshly harvested mouse livers were homogenised in 3 volumes of lysis buffer (150 mM NaCl, 20 mM Tris-HCl pH7.4, 5 mM $MgCl_2$, 5 mM DTT, 100 μ g ml^{-1} cycloheximide, 1% Triton X-100, 0.5% Sodium deoxycholate, complete EDTA-free protease inhibitors (Roche) and 40 U ml^{-1} RNasin plus (Promega)) using a Teflon homogeniser. Lysates were

incubated for 10 min on ice and cleared by centrifugation at 1000 x g, 4°C for 3 min in a tabletop centrifuge. Supernatants were flash-frozen and stored in liquid nitrogen. Lysates were thawed on ice and the OD_{260} was determined using a Nanodrop spectrophotometer. For each replicate and timepoint equal amounts of OD_{260} lysate from two mice were pooled. From the lysate pool, 15 OD_{260} were incubated with 650 U RNase I (Ambion) and 5 U Turbo DNase (Ambion) for 45 min at room temperature and gentle agitation. Nuclease digestion was stopped through addition of 8.7 μl Superasin (Ambion). Subsequently, lysates were applied to sephacryl S-400 HR spin columns (GE Healthcare Life Sciences), pre-washed 3 times with 700 μl polysome buffer for 1 min at 600 x g, and centrifuged for 2 min at 600 x g and 4°C. The flow-through was immediately mixed with 1 ml Qiazol and ribosome-protected mRNA fragments were purified using miRNeasy RNA extraction kit (Qiagen) according to the manufacturer's instructions. For each sample 25 μg RNA were separated on a 15% urea-polyacrylamide gel and stained with SYBR-Gold (Invitrogen). Gel slices between 26-34 nucleotides were excised and RNA was extracted using 600 μl gel extraction buffer (0.5 M Ammonium acetate, 1 mM EDTA and 0.1% SDS) for 2 hours at room temperature and gentle agitation. Gel pieces were removed by centrifugation over spin filter tubes for 3 min at 5000 rpm. RNA was precipitated over night at -20°C in the presence of 900 μl isopropanol and 2 μl glycogen. RNA was pelleted and washed with 70% ethanol in a tabletop centrifuge at maximum speed and 4°C. Ribosomal RNA was depleted using the Ribo-Zero magnetic kit (Epicentre) according to the manufacturer's instructions. Sequencing libraries were

generated using the ARTseq ribosome profiling kit (Epicentre) and sequenced on an Illumina HiSeq 2500.

RNA sequencing

For each timepoint and replicate the same samples as for RPF-seq (pool of liver lysates from two mice) was used to extract total RNA using miRNeasy RNA extraction kit (Qiagen). Ribosomal RNA was depleted using the Ribo-Zero magnetic kit (Epicentre) according to the manufacturer's instructions. Sequencing libraries were generated using the total RNA preparation protocol of the ARTseq ribosome profiling kit (Epicentre). Libraries were sequenced on an Illumina HiSeq 2500.

RT-qPCR

Reverse transcription was performed with random hexamer primers and the following gene-specific primers were used for real-time PCR. *Denr* forward/reverse: AAAGGCGATACGAAGAACAGTG, CATCCGGCATGTATTCACAGT*; Nudt4* (control gene) forward/reverse: AAGTTCAAGCCCAACCAGACG, TCCTGGGACAATCCATTGGTC*; Firefly luciferase 1* forward/reverse: CCGCCTGAAGTCTCTGATTAAGT, ACACCTGCGTCGAAGATGTTG*; Firefly luciferase 2* forward/reverse: TGCAAAAGATCCTCAACGTG, AATGGGAAGTCACGAAGGTG*; Renilla luciferase* forward/reverse: GGAATTATAATGCTTATCTACGTGC, CTTGCGAAAAATGAAGACCTTTTAC.

Preparation of protein extracts

For the preparation of nuclear extracts, freshly harvested livers were homogenised each in 4 ml nuclear homogenisation buffer (10 mM HEPES pH 7.9, 10 mM KCl, 0.3 M sucrose, 0.1 mM EDTA, 0.74 mM spermidine, 1 mM DTT, supplemented with complete protease inhibitor tablets from Roche) for 1 min using a Teflon homogeniser and subsequently mixed with 8 ml cushion buffer (10 mM HEPES pH 7.9, 2.2 M sucrose, 0.1 mM EDTA, 0.74 mM spermidine, 1 mM DTT, complete protease inhibitors). The homogenate was split in two and each part was layered on top of 3 ml cushion buffer and centrifuged at 4°C for 1 h at 25000 rpm in a SW40 rotor. The nuclear pellets were re-suspended in 400 μl RIPA-buffer (150 mM NaCl, 50 mM Tris pH 8.0, 1% NP40, 0.5% DOC, 0.1% SDS and complete protease inhibitors) and incubated for 20 min on ice. The lysates were cleared by centrifugation at 16000 x g for 5 min at 4°C and stored at -80°C.

For the preparation of total cell extracts, livers were lysed in 3 volumes (w/v) RIPA-buffer using a Teflon homogeniser, incubated for 30 min on ice and centrifuged for 10 min at 16000 x g and 4°C. Supernatants were stored at -80°C. Protein concentrations were quantified using a BCA protein assay kit (Pierce).

For the isolation of mitochondria, livers were homogenised in 5 volumes (w/v) Mt-homogenisation buffer (10 mM Tris-MOPS pH 7.4, 1 mM EGTA-Tris pH 7.4, 200 mM sucrose) with 5 strokes in a Teflon homogeniser. The homogenate was centrifuged at 600 x g for 10 min at 4°C. The supernatant was transferred to a new tube and centrifuged at 7000 x g for 10 min at 4°C. Pellets were washed in 1 ml Mt-homogenisation buffer, re-centrifuged at 7000 x g for 10 min and re-suspended in SDS-PAGE sample buffer.

Protein extracts (25 μ g) were analysed by SDS-PAGE using standard conditions.

Western blot analysis

Immunoblotting to PVDF or nitrocellulose membranes was performed using a Tris/CAPS discontinuous buffer system with semi-dry transfer (Trans-Blot semi-dry electrophoretic transfer cell, Bio-Rad) according to the manufacturer's instructions. Membranes were blocked and incubated with primary and secondary antibodies in 5% milk/TBS-T or 5% BSA/TBS-T. Primary and secondary antibodies were used at the following dilutions: anti-FTL1 (1:100, Santa Cruz Biotechnology, sc-14420), anti-FTH1 (1:1000, Cell Signaling, #4393), anti-DENR (1:5000, Abcam, ab108221) anti-ALAS2 (1:1000, Abcam, ab184964), anti-NDUFB8 or anti-mt-CO1 (OxPhos rodent antibody cocktail, 1:250, Abcam, ab110413), anti-GGPS1 (1:1000, Abcam, ab167168), anti-DEAF1 (1:1000, LSBio, LS-C80262), anti-U2AF2 (also known as U2AF65) (1:5000, Sigma, U4758), anti-beta-Tubulin (1:5000, Sigma, T5201), anti-ACO1 (1:2000, Abcam, ab126595), anti-IREB2 (1:2000, Abcam, ab181153), anti-rabbit-HRP or anti-mouse HRP (1:10000, Promega) and anti-goat-HRP (1:5000, Santa Cruz Biotechnology). Signals were visualised using LumiGLO or LumiGLO-Reserve chemiluminescent substrate kits (KPL). For each timepoint 2-3 animals were analysed; shown are the results from one representative time series.

Cloning, plasmids

A 8.2 kb fragment containing 3.5 kb of the *Nr1d1* promoter region, exon 1, intron 1, the first 14 bp of exon 2 and the luciferase coding region was amplified by PCR from the

adenoviral plasmid pCV100-*Nr1d1*-Luc described previously (Saini et al. 2013) using the following primers: forward: ACGTTCGCGACGCGTGTGTGTATGTGTGTG, reverse: TGCACTCGAGTTACACGGCGATCTTTCCGCC. Fragments were cloned between *Xho*I and *Nru*I sites to replace the dual luciferase reporter cassette in the prLV1 lentiviral vector previously described (Du et al. 2014) to generate prLV1-*Nr1d1*-Wt. Subsequently, a 1.1 kb fragment of the *Nr1d1* promoter region was deleted by site-directed mutagenesis using forward: GAAGCCAGAAGCTGGGAGCACACACGCGTCGCGATT, reverse: AATCGCGACGCGTGTGTGCTCCCAGCTTCTGGCTTC to reduce plasmid size. The size-reduced plasmid was then used to generate *Nr1d1*-mutants by means of site-directed mutagenesis using the following primer combinations: M1A-uORF1, mutation of the uORF1 start codon to alanine, (F: CTCTCTGCTCTTCCCGCGCAAATCAGATCTCAGG, R: CCTGAGATCTGATTTGCGCGGGAAGAGCAGAGAG); M1A-uORF2, mutation of the uORF2 start codon to alanine, (F: CAAGGTCCAGTTTGAGCGACCGCTTTCAGCTGG, R: CCAGCTGAAAGCGGTCGCTCAAACTGGACCTTG); Del-uORF1, deletion of uORF1 (F: ACACTCTCTGCTCTTCCCATGACCGCTTTCAGCTGG, R: CCAGCTGAAAGCGGTCATGGGAAGAGCAGAGAGTGT); Del-uORF1+2, deletion of uORF1 and the first 4 codons of uORF2, (F: GGAGCTCCAGATTCATTACCAGCTGGTGAAGACATGACG, R: CGTCATGTCTTCACCAGCTGGTAATGAATCTGGAGCTCC).

For the generation of dual luciferase (Firefly/*Renilla*) reporter plasmids, fragments containing the 5' UTR and the first 10 amino acids of the *Nr1d1* coding sequence were amplified by PCR from prLV1-*Nr1d1*-Wt and mutant plasmids using the forward: AAAAGGATCCACTAGTGGGAAAGGCTCGGGCAAAAGGCGG and reverse primer:

TTTGGATCCTGTGTTGTTATTGGAGTCCAGGGT and cloned into the *BamH*I site of the prLV1 dual luciferase reporter plasmid (Du et al. 2014).

For shRNA vectors, the following hairpins targeting *Denr* were cloned: shRNA1: GTACCACAGAAGGTCACGATA, corresponding to clone TRCN0000308443; shRNA2: GTGCCAAGTTAGATGCGGATT, corresponding to clone TRCN0000098826; shRNA3: GTACCACAGAAGGTCACGATA, corresponding to clone TRCN0000098827); scramble shRNA (Addgene #1864) served as control.

Initial processing and quality assessment of sequencing data

Initial quality assessment of the sequencing reads was conducted based on the preliminary quality values produced by the Illumina pipeline Casava 1.82. Quality related statistics used were the percentage of clusters passed filtering (%PF clusters) and the mean quality score (PF clusters), as other quality related statistics, such as percentage of reads aligned, mean alignment score and percentage of alignment error were not very useful due to large contributions of adapter sequences. Adapter sequences were removed using cutadapt utility (Martin 2011) with following options: -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC –match-read-wildcards -m 6. The size distribution of insertions was employed to assess the quality of sequencing libraries, especially for ribosome footprints; for samples that did not produce a distinct peak around 29-30 nucleotides, either the sequencing or the library preparation was repeated. Next, trimmed read sequences were filtered by their size using an in-house

464 Python script with following inclusive ranges: [26,35] for footprints, [21,60] for total
465 RNA. Smaller or larger fragments were kept separately and not used in further analyses.

466 **Alignment to the mouse genome**

467 Trimmed and filtered insert sequences were mapped sequentially to following
468 databases: mouse rRNA, human rRNA, mt-tRNA, mouse tRNA, mouse cDNA from
469 Ensembl mouse database release 75 (Flicek et al. 2013) and, finally, mouse genomic
470 sequences (Genome Reference Consortium GRCm38.p2). All but last one were mapped
471 using bowtie version 2.2.1 (Langmead and Salzberg 2012) using the following
472 parameters: -p 2 -L 15 -k 20 –no-unal. Mapping against genomic sequence was
473 performed using tophat v2.0.11 (Trapnell et al. 2009) with the following parameters: –
474 transcriptome-index=Mmusculus.GRCm38.75.dna.ensembl_data -p 2. After each
475 alignment, only reads that were not aligned were used in the following mapping. For
476 further analysis, only alignments against mouse cDNA were used, unless specifically
477 stated otherwise. For each query sequence, only alignments with maximum alignment
478 score (AS) were kept.
479 Separately from this sequential alignment strategy, trimmed and filtered insert
480 sequences from each sample were also directly aligned against the mouse genome using
481 tophat with similar parameters. The output of this alignment was used to estimate
482 expressed transcript models out of all models contained in Ensembl mouse database
483 release 75. To this end, we used cufflinks v2.2.1 (Trapnell et al. 2010) to estimate the
484 number of fragments per kilo base of exon per million fragments mapped (FPKM) for

each transcript, with the following parameters: –GTF Mus_musculus.GRCm38.75.gtf –frag-len-mean 37 –frag-len-std-dev 8 –compatible-hits-norm –multi-read-correct –upper-quartile-norm –frag-bias-correct Mmusculus.GRCm38.75.dna.ensembl.fa -p 3. Transcript FPKM estimates from all total RNA samples were merged using cuffcompare with following parameters: -r Mus_musculus.GRCm38.75.gtf -R -V. Resulting FPKM tracking information was parsed with an in-house Python script to filter out transcripts which were not found to have an FPKM > 0.1, a LOW95 > 0.05 and a FMI > 2.0 in at least 3 samples. A database of expressed transcripts based on this filtering was used in further analysis.

Quantification of mRNA and ribosome footprint abundance

Abundance of mRNA and RPF (ribosome protected fragment) was estimated per gene. For this quantification, only reads that were uniquely mapped to a single gene and only transcripts that were identified as expressed (see Alignment to the Mouse Genome) were used. An in-house Python script was used to count reads within each annotation feature (5' UTR, CDS, 3' UTR) per gene. For genes that were associated with multiple expressed transcripts (multi-isoform genes), reads that were not mapped unambiguously within a single feature were assigned to one with the following preference order: CDS, 5' UTR, 3' UTR.

Prior to further analysis, variation captured by the CDS counts was inspected with the principal component analysis (PCA) tools provided in the R package DESeq v1.14.0 (Anders and Huber 2010). The aim of this type of analysis is to visualise the overall effect

of experimental covariates and to detect if batch effects or outlying samples are present in the experiment. For this, we followed the steps described in the reference manuals of DESeq. Briefly, a combined matrix of CDS counts for RPF and mRNA was built, where each row contained time-series data with biological replicates per gene. Normalisation factors for library sizes were calculated using upper quartile method and dispersions were blindly estimated using the "pooled" method (ignoring conditions) of DESeq. Prior to PCA, the normalised (division by size factors) count data were transformed using the getVarianceStabilizedData function from DESeq package, which calculates a variance stabilising transformation (VST) from the fitted dispersion-mean relation, yielding a matrix of values which are approximately homoskedastic. The variance-stabilised data matrix was then used for PCA using the plotPCA function of DESeq. Briefly, rows were ranked in decreasing order by their row-wise variances and the top 4000 genes were then used in PCA. The number of genes included was identified empirically by increasing the initial default number (500) iteratively until the PCA results reached a stable state. The plotPCA function, by default, performs a PCA without centring or further scaling.

Mappable and countable feature lengths (in nucleotides) for each locus were calculated by means of generating all possible 21-, 30-, 45-, and 60-nt long reads *in silico* (faux reads) for each expressed transcript in the database and counting the faux reads through identical mapping and counting work flow used for real experimental reads (Du et al. 2014). Weighted averages of these mappable and countable lengths were later used in RPKM calculations. For weights, proportions of 21-, 30-, 45-, and 60-nt inserts within averaged shifted histograms of insert length distributions were used.

Read counts of total RNA and RPF were normalised with upper quantile method of R package edgeR v3.4.2 (Robinson et al. 2010). Prior to normalisation, transcripts, which did not have at least 10 counts in at least one fourth of the samples, were removed from the datasets. For better comparability between datasets, RPKM values were calculated as the number of counted reads per 1000 mappable and countable bases per geometric mean of normalised read counts per million. The geometric mean of normalised read counts were 11,025,834 and 18,520,239 for total RNA and RPF datasets, respectively. Translational efficiencies (TE) were then calculated as the ratio of RPF-RPKM / mRNA-RPKM for each gene per sample. For most analysis downstream, TEs were log2 transformed and means or replicates per timepoint or grand-means over all timepoints were used as necessary.

Only for comparison of the translational efficiencies obtained from this study to those reported in (Ingolia et al. 2011), we first corrected the differences in location by adjusting the means trimmed 10%. Then, 13 data points at the extremities were removed from Ingolia's dataset to achieve similar ranges in both datasets.

Analysis of mRNA abundances and translation efficiencies

Changes in the mRNA abundance or TE distributions between different sets of genes were compared by estimating the amount of location shift and its significance using Wilcoxon rank sum test (equivalent to the Mann-Whitney test). For all cases, a two-sided test was performed and if necessary resulting p-values were corrected for multiple testing via BH method (Benjamini and Hochberg 1995). The median of all possible

differences between a sample from the first set and a sample from the second set was then reported as an estimate for the difference in location parameters (location shift). For all other statistical calculations and figure production, the R environment (R Core Team, 2013) with various packages (ggplot2 (Wickham 2009); MASS (Venables 2002)) was used. For production of plots of RNA and RPF densities along transcripts, an in-house Python script was used.

Analysis of significant changes in ribosome occupancy

Significance of changes in translational regulation within and between timepoints were assessed with the *Babel* computational framework (Olshen et al. 2013). To this end, unnormalised count data for both total RNA and RPF were used following the authors' instructions. 10,000,000 permutations were performed within comparisons to achieve precise estimates. The false discovery rate-adjusted p-values were then used to create lists of genes, whose ribosome occupancies were significantly altered (fdr-adjusted p-value < 0.05) within a timepoint or between any two timepoints. A union of these two gene lists was also created to include all genes that are predicted to have alterations in their ribosome occupancies (e.g. translational efficiencies).

Rhythmicity detection in mRNA and RPF profiles

Rhythmic parameter estimation methods that depend on harmonic regression typically fail to produce accurate estimates from abundance profiles of genes that show pronounced deviations from sinusoidal curves. An example is *Per1* (Fig. 3B), where the

569 increasing and decreasing portions of the profile are not symmetrical. To overcome this
570 shortcoming, we have devised an algorithm to detect rhythmicity and estimate rhythmic
571 parameters, which is based on Akaike information criterion (AIC) based model selection
572 (Wagenmakers and Farrell 2004) between three linear models: mean model with 1
573 parameter, harmonic model with 2 parameters (Symul 2013), and a sigmoid model with
574 5 parameters (Symul 2013). The sigmoid fit is composed of two sigmoid curves,
575 synthesis-up, *SU*, which lasts *l* hours, and synthesis-down, *SD*, which lasts *m* hours. The
576 two curves, together with the phase *p* describe a cyclic curve that runs from the basal
577 level *B* to max level, *B* times the fold-change *fc*. The equation for the sigmoid function
578 $f(B,fc,p,m,l)$ of time t is

579 $B(1+(fc-1)(SU+SD)))$,

580 where,

581 $$SU(m,p,t)=\frac{1}{1+\exp(\frac{8}{m}(t-p-\frac{m}{2}))},$$

582 $$SD(l,p,t)=\frac{1}{1+\exp(-\frac{8}{l}(t-p+\frac{l}{2}))}$$

583 For each gene, normalised CDS counts of total RNA and RPF reads were then fit to each
584 model using the method of least squares. For all regressions, the two biological
585 replicates per timepoint were treated as independent replicates and were not combined
586 into a false 48-hour data-series. For sigmoid curve fitting, the minpack.lm package in R
587 environment was used. Minimum and maximum values allowed for *m* and *l* were 3 and

16 hours. First, for each model, we computed the differences in AICc (AIC corrected for finite sample size) with respect to the AICc of the best candidate model, ΔAICc. Then, ΔAICc were used to calculate the Akaike weight for each model. A dataset was considered to be rhythmic if the evidence for mean-model (constant expression) was less than 0.05. For such genes, the sigmoid model was chosen over the sinusoidal (harmonic regression) only when the evidence for sigmoid model was at least 3.5 times higher than that for sinusoidal to prevent over-fitting. The rhythmic parameters, phase and amplitude-ratio, were then estimated from the selected model. Finally, genes that did not have an amplitude-ratio of at least 1.5 were flagged and not categorised as cyclic.

To enable a comparison of our results to those from previously published circadian proteomics studies (Mauvoisin et al. 2014; Robles et al. 2014), the UniProt and gene name identifiers in these reports were converted into EnsemblIDs using the external references maintained in the Ensembl database.

The Rayleigh test for uniformity and the Watson-Wheeler test for homogeneity were performed using the R package 'circular'.

Analysis of RPF positions on transcripts

For all total RNA and RPF reads that were counted towards genes, we have also tracked the position of the 5’ end of the read relative to the 5’ end of its corresponding transcript and the trimmed size of the read. The putative A-site of ribosomes was then calculated as 5’ position of a RPF read plus 15 if the read was ≤30 nt long or 16

609 otherwise. For analysis of the 3 nucleotide rhythmicity of RPFs, only a subset of genes
610 that i) had a single transcript isoform with a CDS length of at least 400 nt, ii) were
611 expressed at both mRNA and footprint levels, iii) had a RPF-RPKM > 5, iv) had at least 20
612 nt in UTRs preceding and following CDS region, and v) had at least 10 reads within both
613 the first and the last 200 nt, were considered (N=3237). For calculation of frame
614 preference in CDS, the last two criteria were not used (N=3793). Upstream ORFs (uORFs)
615 were identified within the 5' UTRs of a restricted set of genes, for which a single protein-
616 coding isoform was found to be expressed (N=6774), with the following criteria: i) start
617 with AUG, ii) are at least 18 nucleotides long, iii) can overlap with main CDS. To identify
618 uORFs that are actively translated we used a similar strategy described elsewhere
619 (Bazzini et al. 2014). Briefly, A-site counts for each reading frame within a predicted
620 uORF were tested for significant deviations from a uniform distribution with a chi-
621 square test. Resulting p-values were adjusted for FDR using BH (Benjamini and Hochberg
622 1995), and a uORF was considered as actively translated only then if it had an adjusted
623 p-value <0.05, the preferred reading frame was the first one (relative to uORF's 5'), and
624 it had a RPF coverage >10%. Same analysis was applied to total RNA reads, to validate its
625 specificity and also filter out unspecific cases (likely false positives) which suggested
626 translation events from total RNA reads (N=34).
627 The effects of feature (5' UTR, CDS, 3' UTR) size and presence of translated uORFs on the
628 translation efficiency (TE) from the main CDS were analysed using linear least squares
629 regression. While log-transformation of TE is widely used elsewhere in the manuscript
630 and in almost all other publications reporting ribosome footprinting, for this analysis we

found that cube root transformation was more optimum using the powerTransform tools from the R package 'car' v2.0.25 (Fox and Weisberg 2010). Feature sizes were log10 transformed, and the presence or absence of translated uORFs in a transcript was modelled as a binary predictor. The effects of the predictors were then visualised via partial regression plots using the avPlots function from package 'car'.

Identification of pause-sites

Analysis of ribosome pausing was carried out following previous report (Ingolia et al. 2011). The same restricted set of genes (N=6774) that was used for uORF analysis was used for pausing analysis. Mainly, two modifications to the original method were revised. Firstly, as an estimate of central tendency we devised to employ the trimean, which is the average of the median and the midhinge, instead of the median. The trimean is a more resistant measure and copes better where median can have a value of zero. Secondly, we have observed a small but significant positive correlation (Pearson r = 0.18, p-value < 2.2e-16) in normalised codon counts between RPF and mRNA, suggesting local RPF read densities along transcripts could also be influenced by other factors than ribosome pausing. To overcome this effect we devised a method where RPF codon values were normalised by mRNA codon values before identification of pause-sites. Specifically, we first calculated the pileup profile of mRNA reads along each transcript. At positions (nt), where pileup values were smaller than the first quartile of all pileup values, the values were set equal to the first quartile value. The resulting profile was smoothened with a correlation filter and subsequently used to calculate the

652 trimean normalised codon counts for mRNA reads. Finally, at each codon the RPF count
653 was divided by the mRNA count. This approach provided a conservative means of
654 normalisation where over-representation of pause-sites that coincide with prominent
655 peaks in the mRNA pileup profile were was lessened, without extensively promoting
656 non-pause-sites that coincide with troughs in the mRNA pileup into novel pause-sites.
657 The correlation in codon counts between RPF and mRNA dropped drastically after this
658 normalisation procedure ($r = 0.08$, p-value $< 2.2e-16$). A site is labelled as pause-site only
659 if both biological replicates had a score > 25 at that site.

SUPPLEMENTAL REFERENCES

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome biology* **11**: R106.

Bazzini AA, Johnstone TG, Christiano R, Mackowiak SD, Obermayer B, Fleming ES, Vejnar CE, Lee MT, Rajewsky N, Walther TC et al. 2014. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *The EMBO journal* **33**: 981-993.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* **57**: 289-300.

Du NH, Arpat AB, De Matos M, Gatfield D. 2014. MicroRNAs shape circadian hepatic gene expression on a transcriptome-wide scale. *eLife* **3**: e02510.

Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S et al. 2013. Ensembl 2013. *Nucleic acids research* **41**: D48-55.

Fox J, Weisberg S. 2011. *An R Companion to Applied Regression*. SAGE Publications.

Ingolia NT, Lareau LF, Weissman JS. 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**: 789-802.

Kim TY, Wang D, Kim AK, Lau E, Lin AJ, Liem DA, Zhang J, Zong NC, Lam MP, Ping P. 2012. Metabolic labeling reveals proteome dynamics of mouse mitochondria. *Molecular & cellular proteomics : MCP* **11**: 1586-1594.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: 357-359.

683 Le Martelot G, Canella D, Symul L, Migliavacca E, Gilardi F, Liechti R, Martin O,
684 Harshman K, Delorenzi M, Desvergne B et al. 2012. Genome-wide RNA
685 polymerase II profiles and RNA accumulation reveal kinetics of transcription
686 and associated epigenetic changes during diurnal cycles. *PLoS biology* **10**:
687 e1001442.
688 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput
689 sequencing reads. *EMBnetjournal* **17**: 10-12.
690 Mauvoisin D, Wang J, Jouffe C, Martin E, Atger F, Waridel P, Quadroni M, Gachon F,
691 Naef F. 2014. Circadian clock-dependent and -independent rhythmic
692 proteomes implement distinct diurnal functions in mouse liver. *Proceedings*
693 *of the National Academy of Sciences of the United States of America* **111**: 167-
694 172.
695 Nagoshi E, Saini C, Bauer C, Laroche T, Naef F, Schibler U. 2004. Circadian gene
696 expression in individual fibroblasts: cell-autonomous and self-sustained
697 oscillators pass time to daughter cells. *Cell* **119**: 693-705.
698 Olshen AB, Hsieh AC, Stumpf CR, Olshen RA, Ruggero D, Taylor BS. 2013. Assessing
699 gene-level translational control from ribosome profiling. *Bioinformatics* **29**:
700 2995-3002.
701 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for
702 differential expression analysis of digital gene expression data.
703 *Bioinformatics* **26**: 139-140.

Robles MS, Cox J, Mann M. 2014. In-vivo quantitative proteomics reveals a key contribution of post-transcriptional mechanisms to the circadian regulation of liver metabolism. *PLoS genetics* **10**: e1004047.

Saini C, Liani A, Curie T, Gos P, Kreppel F, Emmenegger Y, Bonacina L, Wolf JP, Poget YA, Franken P et al. 2013. Real-time recording of circadian liver gene expression in freely moving mice reveals the phase-setting behavior of hepatocyte clocks. *Genes & development* **27**: 1526-1536.

Salmon P, Trono D. 2007. Production and titration of lentiviral vectors. *Current protocols in human genetics / editorial board, Jonathan L Haines [et al]* **Chapter 12**: Unit 12 10.

Schaap CC, Hendriks JC, Kortman GA, Klaver SM, Kroot JJ, Laarakkers CM, Wiegerinck ET, Tjalsma H, Janssen MC, Swinkels DW. 2013. Diurnal rhythm rather than dietary iron mediates daily hepcidin variations. *Clinical chemistry* **59**: 527-535.

Stratmann M, Suter DM, Molina N, Naef F, Schibler U. 2012. Circadian Dbp transcription relies on highly dynamic BMAL1-CLOCK interaction with E boxes and requires the proteasome. *Molecular cell* **48**: 277-287.

Symul L. 2013. Kinetic Analysis of Transcriptional and Post-Transcriptional Processes During Circadian Cycles. in *Life Sciences*. EPFL, Lausanne.

Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105-1111.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq

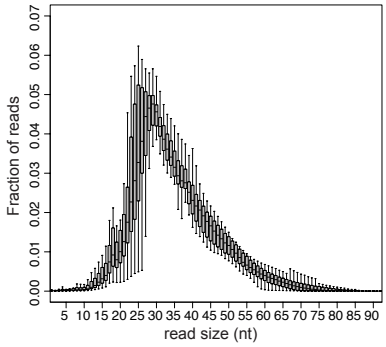
reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**: 511-515.

Venables WN. 2002. *Modern applied statistics with S / W.N. Venables, B.D. Ripley*. Springer, New York.

Wagenmakers EJ, Farrell S. 2004. AIC model selection using Akaike weights. *Psychonomic bulletin & review* **11**: 192-196.

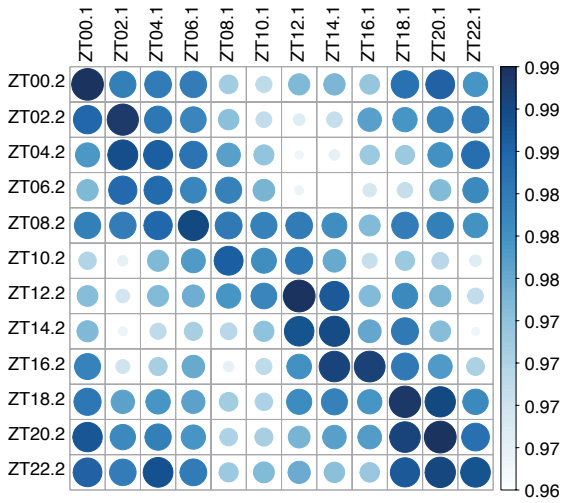
Wickham H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer Publishing Company, Incorporated.

A



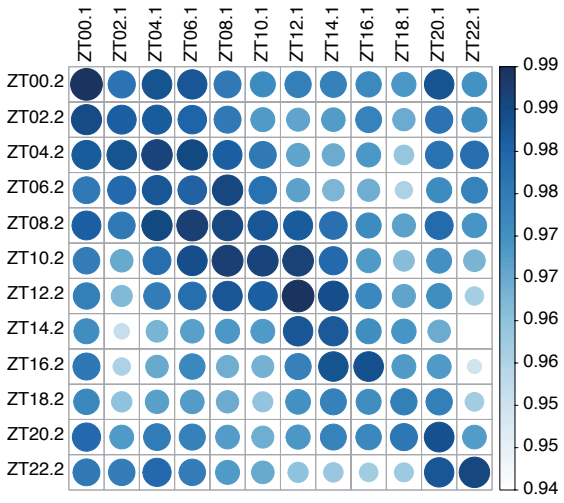
B

RPF-seq

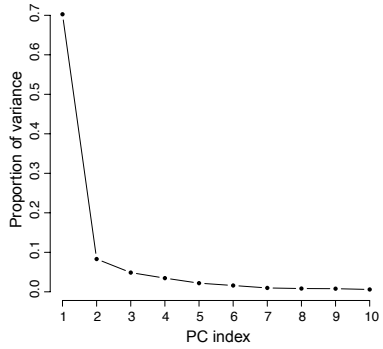


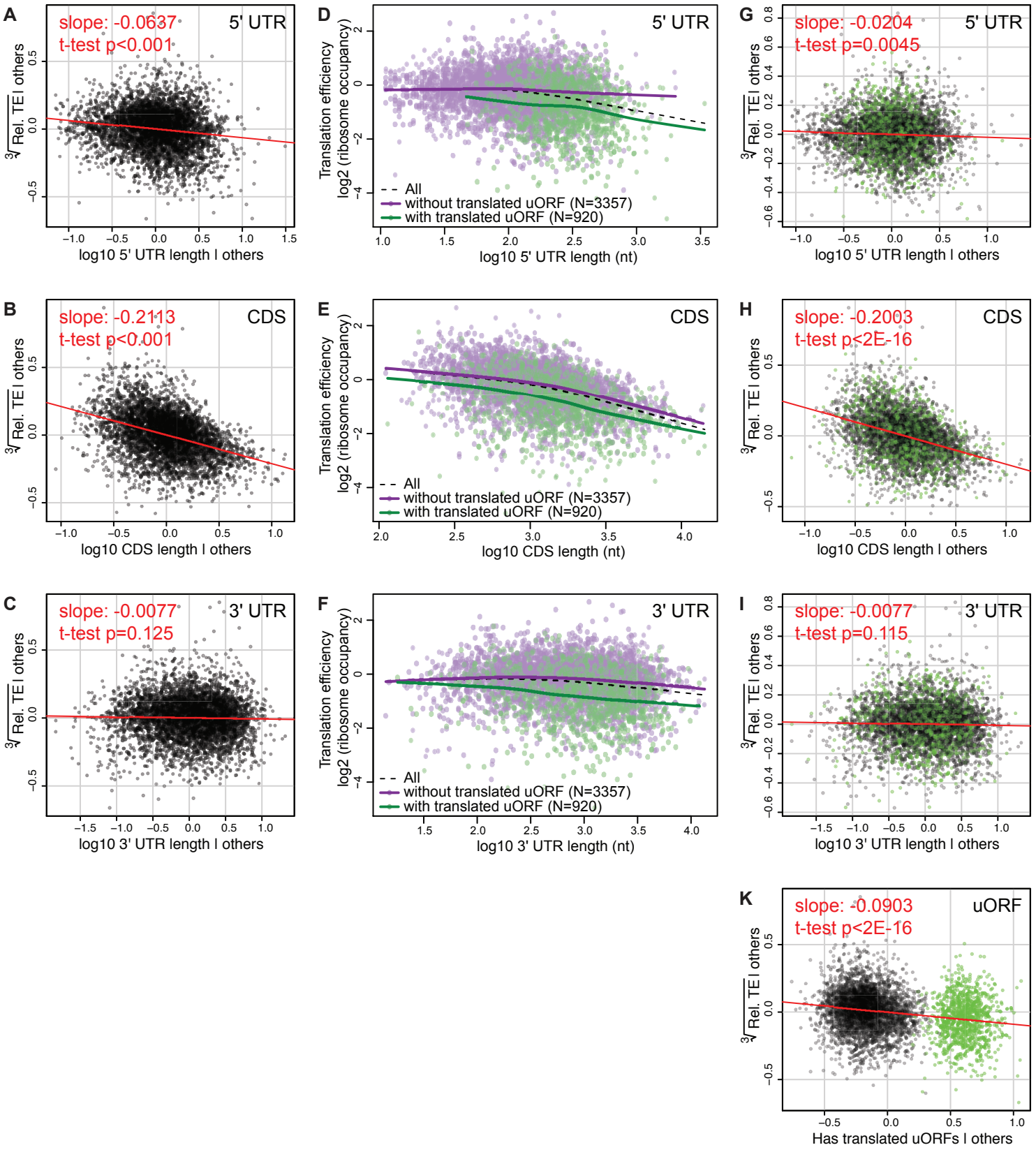
C

RNA-seq

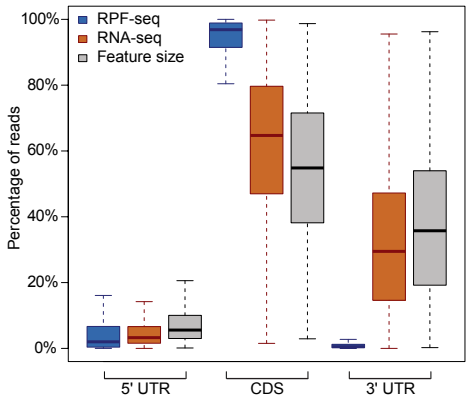


D

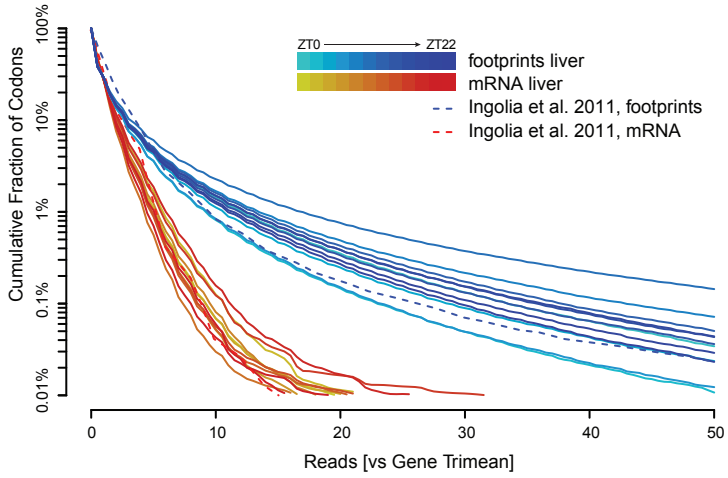



A
slope: -0.0637
t-test p<0.001
5' UTR
∛Rel. TE | others
log10 5' UTR length | others
D
5' UTR
Translation efficiency
log2 (ribosome occupancy)
All
without translated uORF (N=3357)
with translated uORF (N=920)
log10 5' UTR length (nt)
G
slope: -0.0204
t-test p=0.0045
5' UTR
∛Rel. TE | others
log10 5' UTR length | others
B
slope: -0.2113
t-test p<0.001
CDS
∛Rel. TE | others
log10 CDS length | others
E
CDS
Translation efficiency
log2 (ribosome occupancy)
All
without translated uORF (N=3357)
with translated uORF (N=920)
log10 CDS length (nt)
H
slope: -0.2003
t-test p<2E-16
CDS
∛Rel. TE | others
log10 CDS length | others
C
slope: -0.0077
t-test p=0.125
3' UTR
∛Rel. TE | others
log10 3' UTR length | others
F
3' UTR
Translation efficiency
log2 (ribosome occupancy)
All
without translated uORF (N=3357)
with translated uORF (N=920)
log10 3' UTR length (nt)
I
slope: -0.0077
t-test p=0.115
3' UTR
∛Rel. TE | others
log10 3' UTR length | others
K
slope: -0.0903
t-test p<2E-16
uORF
∛Rel. TE | others
Has translated uORFs | others

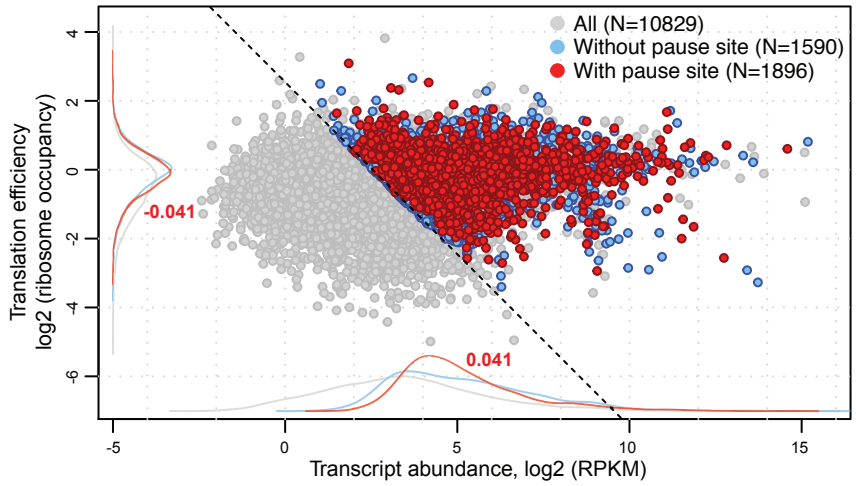
A

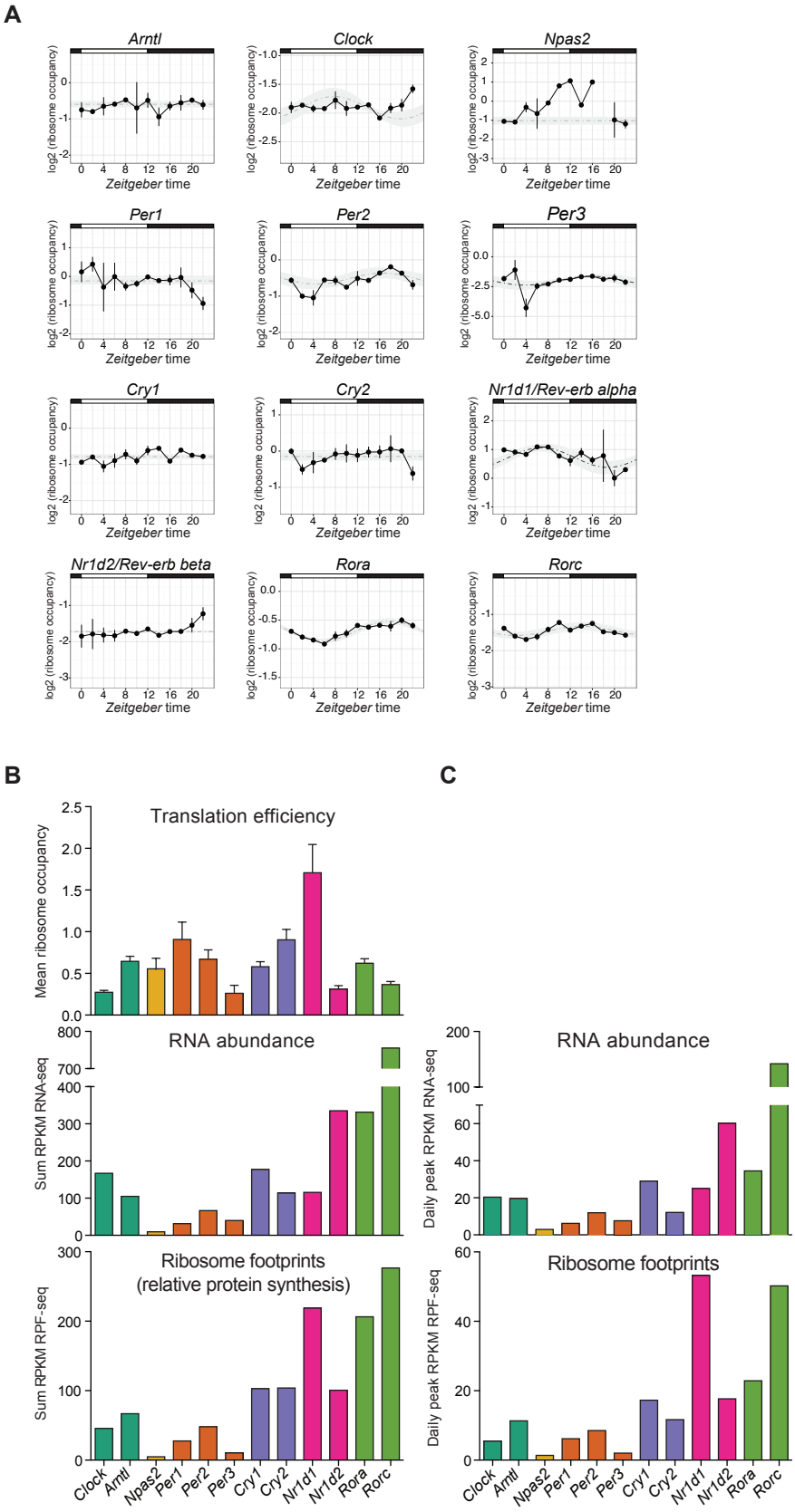

RPF-seq
RNA-seq
Feature size
100%
80%
60%
40%
20%
0%
Percentage of reads
5' UTR
CDS
3' UTR

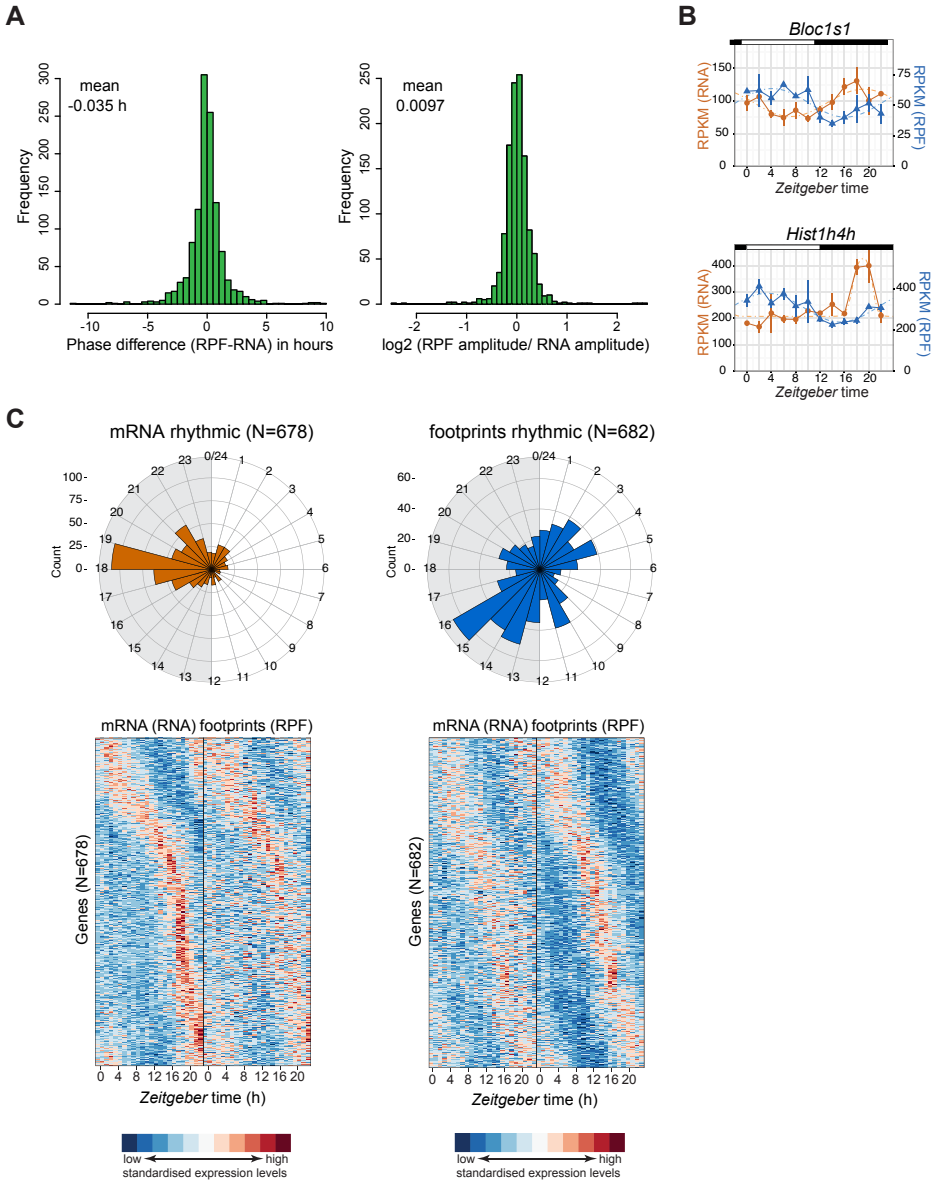
A


ZT0
ZT22
footprints liver
mRNA liver
Ingolia et al. 2011, footprints
Ingolia et al. 2011, mRNA
Cumulative Fraction of Codons
100%
10%
1%
0.1%
0.01%
0
10
20
30
40
50
Reads [vs Gene Trimean]

B


All (N=10829)
Without pause site (N=1590)
With pause site (N=1896)
Translation efficiency
log2 (ribosome occupancy)
-0.041
0.041
4
2
0
-2
-4
-6
-5
0
5
10
15
Transcript abundance, log2 (RPKM)

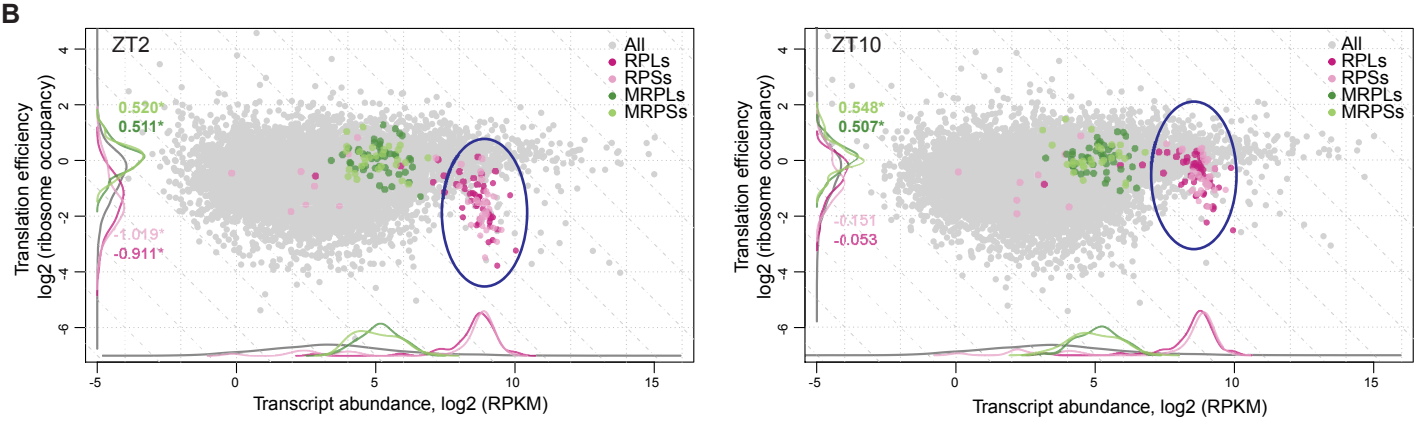

A
Arntl
Clock
Npas2
Per1
Per2
Per3
Cry1
Cry2
Nr1d1/Rev-erb alpha
Nr1d2/Rev-erb beta
Rora
Rorc
log2 (ribosome occupancy)
Zeitgeber time
B
Translation efficiency
Mean ribosome occupancy
RNA abundance
Sum RPKM RNA-seq
Ribosome footprints
(relative protein synthesis)
Sum RPKM RPF-seq
C
RNA abundance
Daily peak RPKM RNA-seq
Ribosome footprints
Daily peak RPKM RPF-seq
Clock
Arntl
Npas2
Per1
Per2
Per3
Cry1
Cry2
Nr1d1
Nr1d2
Rora
Rorc

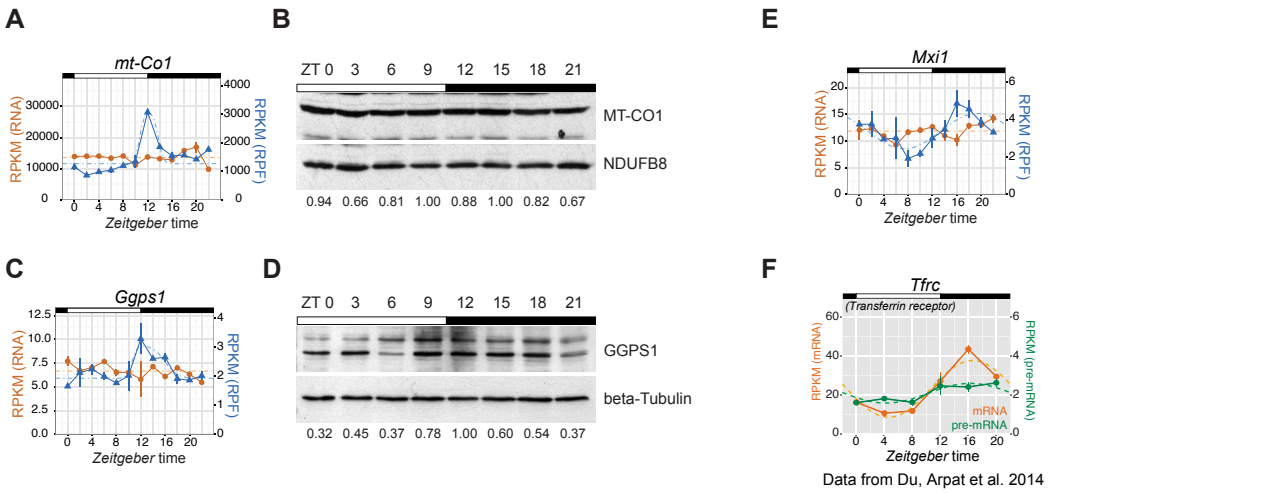

A
mean
-0.035 h
Frequency
Phase difference (RPF-RNA) in hours
mean
0.0097
log2 (RPF amplitude/ RNA amplitude)
B
Bloc1s1
RPKM (RNA)
RPKM (RPF)
Zeitgeber time
Hist1h4h
C
mRNA rhythmic (N=678)
footprints rhythmic (N=682)
Count
mRNA (RNA) footprints (RPF)
Genes (N=678)
Genes (N=682)
Zeitgeber time (h)
low
high
standardised expression levels

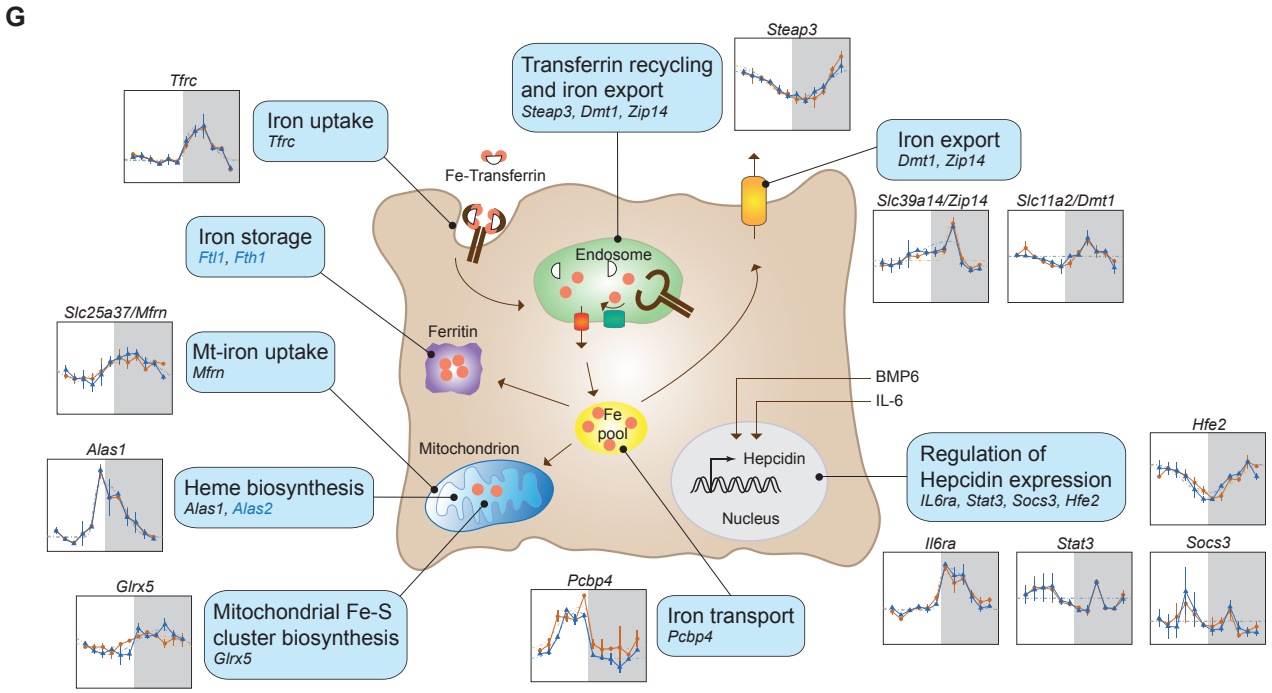
A

Ribosomal subunit	Gene name
40S subunit	*Rpsa, Rps2, Rps3,* ***Rps3a1****,* ***Rps4x****,* ***Rps5****,* ***Rps6****, Rps7,* ***Rps8****, Rps9,* ***Rps10****, Rps11,* ***Rps12****, Rps13, Rps14, Rps15, Rps15a,* ***Rps16****,* ***Rps17****,* ***Rps18****,* ***Rps19****,* ***Rps20****, Rps21, Rps23, Rps24,* ***Rps25****, Rps26, Rps27,* ***Rps27a****,* ***Rps28****, Rps29, Rps30*
60S subunit	*Rpl3, Rpl4,* ***Rpl5****,* ***Rpl6****,* ***Rpl7****,* ***Rpl7a****, Rpl8,* ***Rpl9****,* ***Rpl10****, Rpl10a,* ***Rpl11****, Rpl12, Rpl13, Rpl13a, Rpl14, Rpl15, Rpl17, Rpl18, Rpl18a, Rpl19, Rpl21,* ***Rpl22****,* ***Rpl23****,* ***Rpl23a****,* ***Rpl24****,* ***Rpl26****, Rpl27,* ***Rpl27a****,* ***Rpl28****, Rpl29,* ***Rpl30****,* ***Rpl31****,* ***Rpl32****, Rpl34, Rpl35, Rpl35a, Rpl36,* ***Rpl36a****,* ***Rpl37****, Rpl37a, Rpl38, Rpl39, Rpl40, Rpl41, Rplp0, Rplp1,* ***Rplp2***

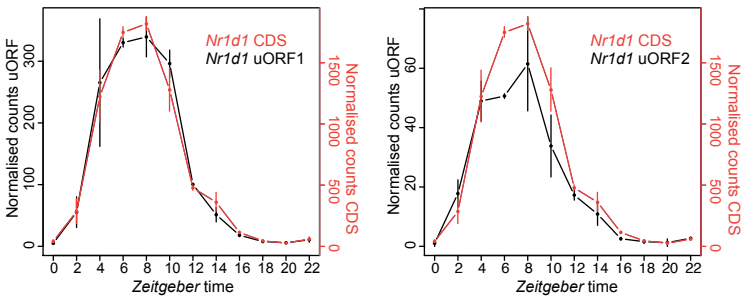
B




A
mt-Co1
RPKM (RNA)
RPKM (RPF)
Zeitgeber time
B
ZT 0 3 6 9 12 15 18 21
MT-CO1
NDUFB8
0.94 0.66 0.81 1.00 0.88 1.00 0.82 0.67
E
Mxi1
RPKM (RNA)
RPKM (RPF)
Zeitgeber time
C
Ggps1
RPKM (RNA)
RPKM (RPF)
Zeitgeber time
D
ZT 0 3 6 9 12 15 18 21
GGPS1
beta-Tubulin
0.32 0.45 0.37 0.78 1.00 0.60 0.54 0.37
F
Tfrc
(Transferrin receptor)
RPKM (mRNA)
RPKM (pre-mRNA)
mRNA
pre-mRNA
Zeitgeber time
Data from Du, Arpat et al. 2014


G
Tfrc
Iron uptake
Tfrc
Transferrin recycling
and iron export
Steap3, Dmt1, Zip14
Steap3
Iron export
Dmt1, Zip14
Fe-Transferrin
Slc39a14/Zip14
Slc11a2/Dmt1
Iron storage
Ftl1, Fth1
Endosome
Slc25a37/Mfrn
Mt-iron uptake
Mfrn
Ferritin
BMP6
IL-6
Fe
pool
Alas1
Mitochondrion
Heme biosynthesis
Alas1, Alas2
Hepcidin
Nucleus
Regulation of
Hepcidin expression
IL6ra, Stat3, Socs3, Hfe2
Hfe2
Il6ra
Stat3
Socs3
Glrx5
Mitochondrial Fe-S
cluster biosynthesis
Glrx5
Pcbp4
Iron transport
Pcbp4

A

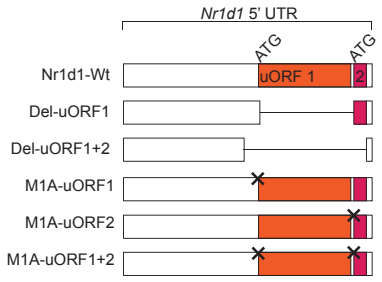


B

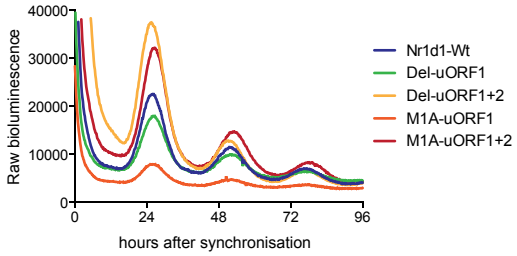
Nr1d1 uORF1

```
M. musculus       MQIRSQVPQRPTLLEGCSIATPRPYCSVPGILIANCGAHSSPSAIAQGAREAITTSSLCQGPV
R. norvegicus     MQIRSQVPQRPTLLEGCSIATPRPYCSVPGILIANCGAHSSPSAIAQGAREAITTSSLCQGPV
H. sapiens        MQSRTPLPQRPTLLQGCSPATPRPCCRVLRILIVSRGVHSPPLASAQGATAAIATSSLSQGPV
P. troglodytes    MQSRTPLPQRPTLLQGCSPATPRPCCRVLRILIVSRGVHSPPLASAQGATAAIATSSLSQGPV
M. fascicularis   MQSRTPLPQRPTLLQGCSPATPRPCCRVLRILIVSRGVLSPPLASAQGATAVIATSSLSQGPV
E. caballus       MQSRPPLPQRPTLLQGCSSATPRPCCRVPRIPIVSRGVHSPPSASAQGATAAIATSSLCQGPV
```

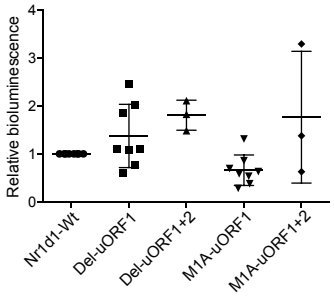
C



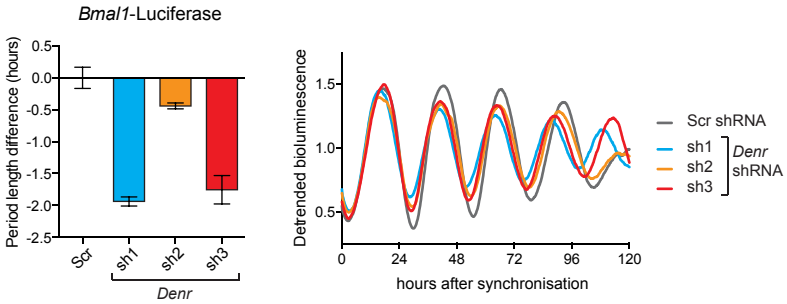
D



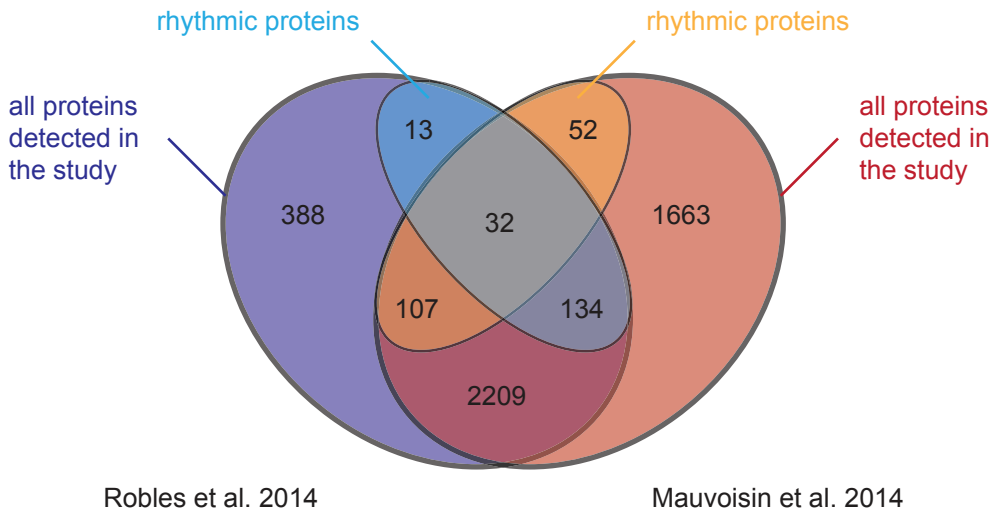
E



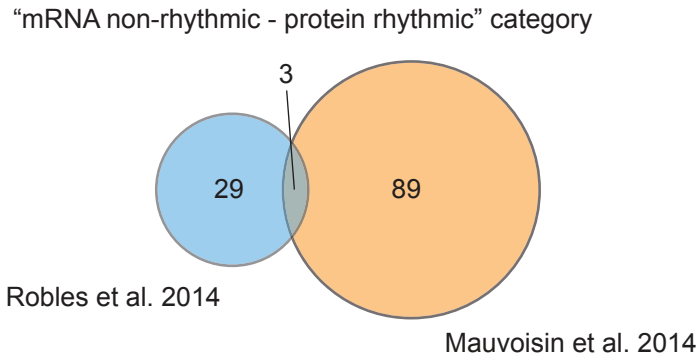
F

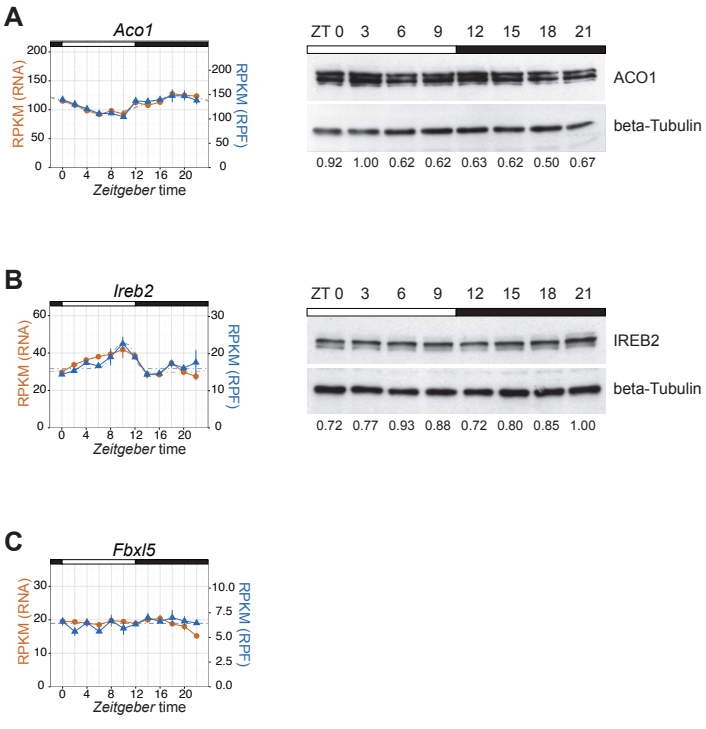


A


rhythmic proteins
rhythmic proteins
all proteins detected in the study
all proteins detected in the study
13
52
388
32
1663
107
134
2209
Robles et al. 2014
Mauvoisin et al. 2014

B


“mRNA non-rhythmic - protein rhythmic” category
3
29
89
Robles et al. 2014
Mauvoisin et al. 2014


A
Aco1
RPKM (RNA)
RPKM (RPF)
Zeitgeber time
ZT 0 3 6 9 12 15 18 21
ACO1
beta-Tubulin
0.92 1.00 0.62 0.62 0.63 0.62 0.50 0.67
B
Ireb2
RPKM (RNA)
RPKM (RPF)
Zeitgeber time
ZT 0 3 6 9 12 15 18 21
IREB2
beta-Tubulin
0.72 0.77 0.93 0.88 0.72 0.80 0.85 1.00
C
Fbxl5
RPKM (RNA)
RPKM (RPF)
Zeitgeber time