Translational contributions to tissue-specificity in rhythmic and constitutive gene expression

Violeta Castelo-Szekely[1], Alaaddin Bulak Arpat[1,2], Peggy Janich[1], David Gatfield[1,*]

[1] Center for Integrative Genomics, Génopode, University of Lausanne, and [2] Vital-IT, Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

[*] corresponding author: david.gatfield@unil.ch

Tel. +41 21 6923994

Fax. +41 21 6923905

Running title: Cross-organ analysis of translation

Keywords: circadian clocks, translation, ribosome profiling, kidney, liver

Abstract

Gene expression oscillations constitute the molecular basis of circadian rhythms in mammalian physiology and behavior. Two molecularly poorly understood aspects of rhythmic gene expression are (1) its striking tissue-specificity and (2) the contribution of post-transcriptional mechanisms to the generation of mRNA and protein abundance oscillations. We have used ribosome profiling in mouse kidney to quantify the translation of mRNAs intro protein transcriptome-wide and around-the-clock, and we have compared the protein biosynthesis rates in this organ with those from the livers of the same cohort of animals that we reported on recently. Our analyses revealed a small set of constantly abundant transcripts in kidney (<100 genes) that underwent daily rhythms in translation. Interestingly, these translational oscillations were almost exclusively tissue-specific, as neither the identity of genes showing this phenomenon, nor the global phase distribution of translational rhythms were shared between the two organs. The analysis of the rhythmically abundant transcripts in both tissues further revealed organ specificity in the relative timing of translational vs. RNA abundance oscillations, as well as in the protein biosynthetic output of core clock genes. Finally, our transcriptome-wide data uncovered marked tissue-specificity in translation rates across constitutively expressed genes. Although in magnitude less divergent than mRNA abundances, translation efficiency was identified as an important contributor to tissue-specificity in gene expression levels. Notably, we detected a clear signature of compensation of RNA expression differences at the translational level, leading globally to higher concordance across organs at the level of ribosome footprints than of RNA abundances.

Reviewer link to deposited data: [available on request]

Introduction

Circadian clocks serve organisms to anticipate daily recurring changes of the environment and to synchronize behavior, physiology and gene expression according to time of day. In mammals, the circadian system consists of a master clock in the brain's suprachiasmatic nuclei (SCN) that receives photic inputs from the retina and synchronizes peripheral clocks present in most cells throughout the body. The molecular timekeeping mechanism – the core clock – relies on a network of transcriptional activators and repressors interacting in negative feedback loops (reviewed in (Dibner et al. 2010; Partch et al. 2014)). In the core loop, the heterodimeric transcription factor ARNTL:CLOCK drives the expression of its own repressors, encoded by the *Period (Per1, Per2, Per3)* and *Cryptochrome (Cry1, Cry2)* genes – a configuration also known as the positive and negative limbs of the oscillator. Additional feedback – in particular an interconnecting limb involving nuclear receptors of the REV-ERB (encoded by genes *Nr1d1, Nr1d2*) and ROR *(Rora, Rorb, Rorc)* family – intersects with the core loop, and numerous post-translational modifications of clock proteins further add to the complexity of the circuitry. The final outcome is a set of robustly cycling transcriptional activities peaking at different phases of the day that drive the rhythmic expression of hundreds to thousands of other genes, known as the clock output or clock-controlled genes (CCGs). It is noteworthy that, despite the probably (near-)identical molecular makeup of the core clock across cell types, CCGs show considerable tissue-specificity (Zhang et al. 2014). The co-regulation by core clock and tissue-specific (non-rhythmic) transcription factors may engender such cell type-specific rhythmic expression patterns, as shown to occur in *Drosophila* (Meireles-Filho et al. 2014). Overall, however, the origins of tissue-specificity in rhythmic gene output (and even in certain core clock parameters (Yoo et al. 2004)) are poorly understood and an active field of research. Mechanisms that act at the post-transcriptional level and that impact daily mRNA and protein accumulation kinetics are plausible players in the generation of cell type differences as well.

So far, oscillatory gene expression has been mainly investigated at the transcriptome level i.e., using mRNA abundances as a primary readout. However, comparison of mRNA levels with genome-wide transcriptional activity and protein abundance data that has become available more recently, has suggested that a surprisingly large fraction of gene expression oscillations may have post-transcriptional

origins (reviewed in (Luck and Westermark 2016)). The many cases of protein rhythms that are independent of an underlying oscillating transcript – an observation initially reported in a low-throughput mass-spectrometric study from mouse liver already 10 years ago (Reddy et al. 2006) and recently confirmed at a comprehensive scale (Mauvoisin et al. 2014; Robles et al. 2014) – point to important roles for translation, protein degradation and protein secretion in shaping time of day-dependent proteomes. We (Janich et al. 2015) and others (Atger et al. 2015) have recently used ribosome profiling, a genome-wide method that assesses translation efficiency through the deep sequencing of ribosome-protected mRNA fragments, to chart the contribution of translational control to daily protein biosynthesis in mouse liver. From the identified cases of translationally generated protein oscillations it could be concluded that circadian clock activity and feeding rhythms both contribute to the regulation of rhythmic gene expression output (Atger et al. 2015; Janich et al. 2015). Notably, the most abundant group of transcripts subject to rhythmic translation, i.e. mRNAs encoding ribosomal proteins and other components of the translation machinery, which all contain 5'-terminal oligopyrimidine tract (5'-TOP) sequences responsive to the mammalian target of rapamycin (mTOR) regulation (Jouffe et al. 2013), appear to be under the dominant control of feeding (Atger et al. 2015).

Analogous to our liver datasets (Janich et al. 2015), we have now generated ribosome profiling data from a second organ from the same cohort of animals, the kidney, which is an emerging circadian model organ with distinct rhythmic functions (Bonny et al. 2013). By comparing kidney and liver datasets we comprehensively assessed commonalities and differences in the translatomes, and we evaluated in how far the regulation of translation efficiency contributed to tissue-specificity in rhythmic and constitutive protein biosynthesis.

Results

Around-the-clock ribosome profiling datasets from two organs

For our recent study of the mouse liver translatome around-the-clock (Janich et al. 2015) we had used the method ribosome profiling (RPF-seq) on a time series of organs collected from animals sacrificed every 2 hours over the 24-hour day (12 timepoints in duplicate; Fig. 1A). In order to generate a complementary

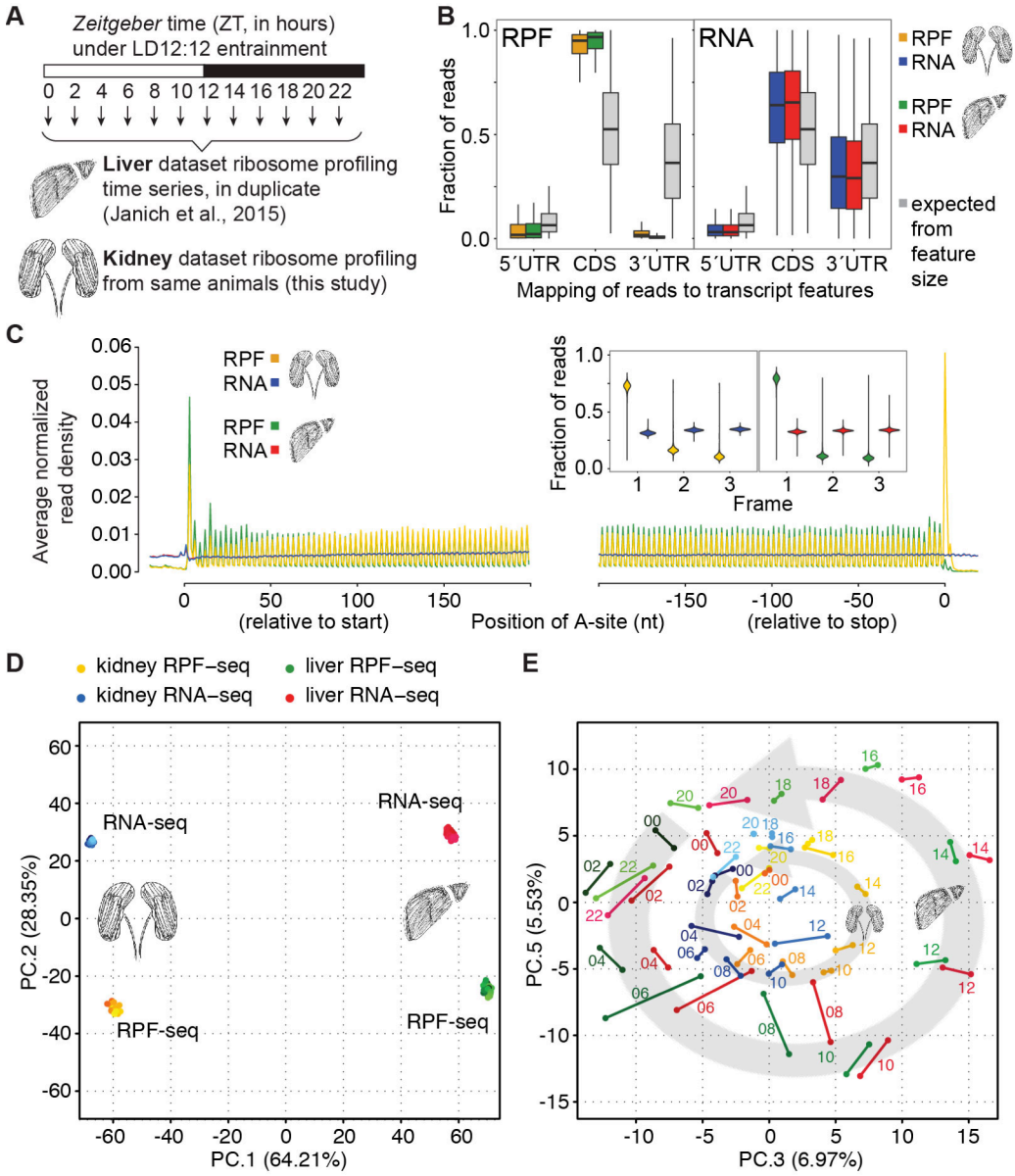


Figure 1. Ribosome profiling around-the-clock in mouse liver and kidney.
(A) Overview of the experimental design: livers and kidneys from the same mice were collected every 2 hours for two daily cycles and ribosome profiling was performed. Each timepoint sample was a pool of two mice livers or kidneys. Animals were kept under 12 hour:12 hour light-dark conditions, with *Zeitgeber* times ZT00 corresponding to lights-on and ZT12 to lights-off. **(B)** Read distribution within the transcripts' 5′ UTRs, CDS and 3′ UTRs for RPF-seq (left; yellow and green for kidney and liver, respectively) and RNA-seq (right; blue and red for kidney and liver, respectively) compared to a distribution expected from the relative feature sizes (grey). As the distributions based on the feature sizes were highly similar for both organs, only that for one organ (kidney) is shown. Note that RPF-seq footprints were enriched on the CDS and depleted from UTRs, whereas RNA-seq reads distributed more homogeneously along the transcript, according to feature size. **(C)** Predicted position of the ribosome's aminoacyl tRNA-site (A-site) of reads relative to the CDS start and stop codons. Read density at each position was averaged across single protein isoform genes (i.e. genes with one expressed transcript isoform) that had an average RPF RPKM > 5, a CDS > 400 nt in length and were expressed in both organs (n=3037 genes). This analysis revealed the trinucleotide periodicity of RPF-seq (but not RNA-seq) reads in both organs. *Inset:* Frame analysis of CDS reads showed preference for the annotated reading frame (i.e. frame 1) in RPF but not RNA-seq reads. Violin plots extend to the range of the data (n=3694 genes for liver, n=4602 genes for kidney). **(D)** Principal component (PC) analysis of kidney and liver RPF-seq and RNA-seq datasets, using the top-4000 most variable genes. The first two components reflected the variability coming from organ (64.21%) and from RPF/RNA origin of datasets (28.35%). **(E)** PC3 vs. PC5 (together 12.5% of variation) sequentially resolved the factor time within each dataset, resembling the face of a clock. Each dot represents a timepoint sample, replicates are joined by a line and timepoints within each dataset are sequentially colored. The circular arrangement of the liver data was larger than that of kidney, indicating a higher contribution of hepatic rhythmic genes to overall variability. A scree plot of the ten first PCs and a representation of PC4 can be found in Supplemental Fig. S3.

dataset from a second organ we chose the kidneys from the same cohort of animals. Liver and kidney express thousands of genes in common (Brawand et al. 2011; Zhang et al. 2014), thus providing a particularly suitable setting for a cross-organ comparison of translation rates.

Applying the same experimental and computational methods as for liver RPF-seq (Janich et al. 2015; Janich et al. 2016), we obtained comparable high-quality data for kidney (see Supplemental Table S1 and Supplemental Fig. S1A-C for details on sequencing and mapping outcomes). Briefly, ribosome footprints from both organs were similarly enriched for protein coding sequences (CDS) of mRNAs and depleted from untranslated regions (UTRs) (Fig. 1B). Like the footprints from liver, also those from kidney exhibited excellent reading frame preference, which allowed resolving the 3-nt periodicity of coding sequences transcriptome-wide (Fig. 1C). Moreover, the high correlation coefficients seen across replicates of the kidney time series for both RNA- and RPF-seq data indicated excellent biological and technical reproducibility (Supplemental Fig. S2A-B). Finally, principal component (PC) analysis on all available datasets (96 libraries, i.e. RPF-seq and RNA-seq from 2 organs, over 12 timepoints, and in duplicate) segregated the data according to the main experimental and biological covariates. PC1 (explaining 64.2% of variation) thus separated libraries according to organ, indicating that tissue origin represented the major source of variability, followed by PC2 (28.4%) that separated RNA-seq (mRNA abundance) and RPF-seq (footprints/translation) (Fig. 1D). The cyclic nature of the data was resolved in the PC3 (6.97% of variation) vs. PC5 (5.53% of variation) representation, in which timepoints assembled to a near-perfect clock (Fig. 1E). The larger circular arrangement of the liver vs. kidney time series suggested that rhythmic gene expression in the liver contributed more strongly to overall variation than did kidney rhythms. This observation is in line with the notion that there are more rhythmic transcripts in liver than in kidney and that hepatic oscillations are overall of higher amplitude (Zhang et al. 2014). Of the further components of the PCA (Supplemental Fig. S3A), PC4 (5.94% of variation) was remarkable as it grouped RNA-seq from one organ with RPF-seq from the other organ (Supplemental Fig. S3B). A plausible interpretation of this observation was the occurrence of translational buffering, which has recently been described in other systems (McManus et al. 2014; Schafer et al. 2015) and which compensates divergent RNA expression to lead to higher similarity at the level of protein biosynthesis (i.e., ribosome footprints). Taken together, we

112 concluded that the kidney data were of similar high quality as our previous liver datasets (Janich et al.
113 2015) and would be suitable for comparative analyses of time of day-dependent and constitutive
114 translation across two tissues.

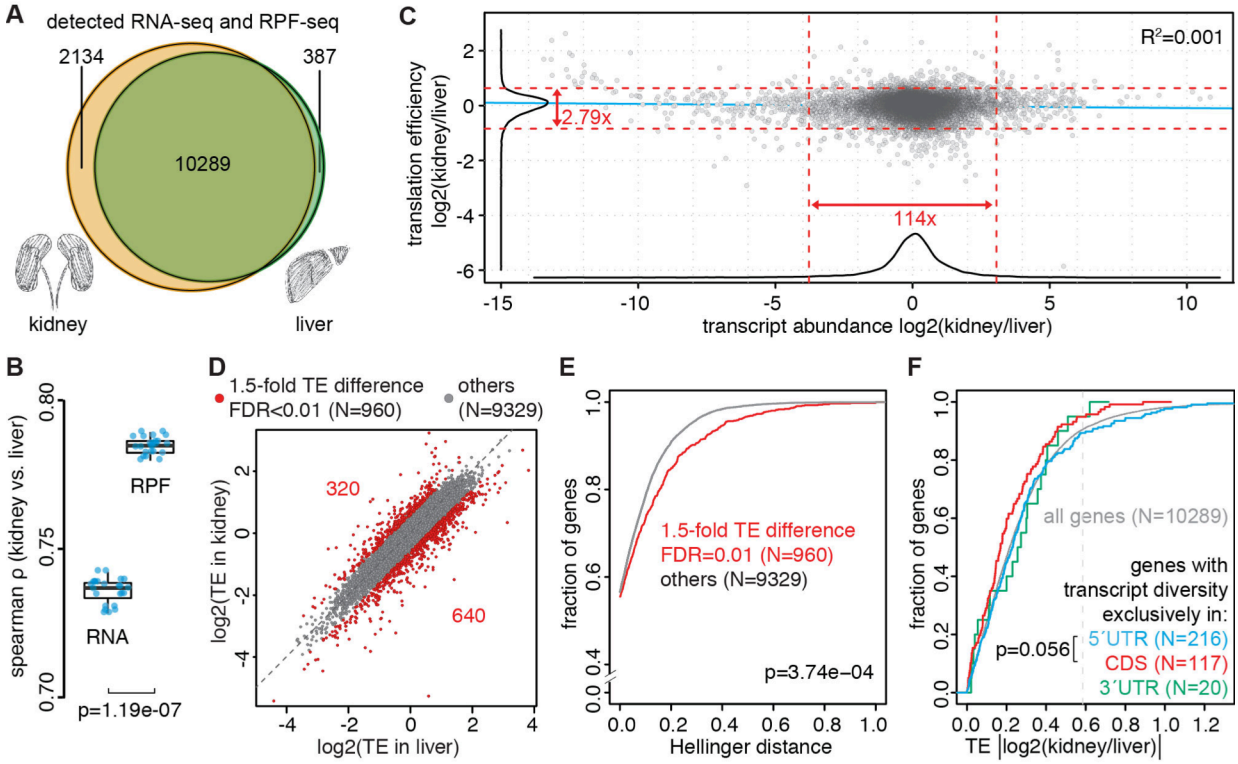


Figure 2. Contribution of translation efficiency to tissue-specific protein biosynthesis.
(A) Venn diagram showing the overlap in the expressed genome (i.e. detected at both RPF-seq and RNA-seq levels) between kidney (yellow, n=12423 genes) and liver (green, n=10676 genes). **(B)** Inter-organ Spearman correlation for RNA-seq and RPF-seq samples. Each dot represents the correlation coefficient for a timepoint and replicate sample. Note that RPF-seq samples consistently correlated significantly better than RNA-seq samples did (p=1.19e-07; Wilcoxon signed rank test). **(C)** Scatterplot of kidney-to-liver ratio of mRNA abundance vs. translation efficiency (TE) for all expressed genes (n = 10289), averaged over all timepoints. Corresponding density curves are plotted on the margins. Dashed red lines represent the 2.5 and 97.5 percentiles of each variable, and the corresponding fold-change is indicated. Linear regression line is depicted in blue (R^2=0.0009, p=0.0009). While 95% of genes spanned a 114-fold range in mRNA abundance differences across organs, the same number of genes changed less than 3-fold in TE, indicating that transcript abundance was the main contributor to divergent gene expression output. **(D)** Relative TE in liver vs. kidney, centered and averaged over all timepoints for all expressed genes (n=10289), showed an overall inter-organ correlation (grey, n=9329). However, differential TE was detected for ~9% genes (red, n=960). Differential TE genes are defined as having FDR-corrected p-value < 0.01 (Wilcoxon signed rank test on TE) and >1.5 difference in TE across organs. **(E)** Cumulative distribution of Hellinger distances for genes showing differential TE (red, n=960), or not (grey, n=9329), as detected in (D). Hellinger distance was used as a measure to quantify how divergent relative transcript isoform usage was across organs (see main text and methods section); the analysis shows that divergent TE correlated with larger diversity in transcript isoform expression (D=0.0702, p=3.74e-04, two-sample Kolmogorov-Smirnov [KS] test). **(F)** Cumulative distribution of the absolute kidney-to-liver TE ratio for genes whose transcript diversity originated only from the 5′ UTR (blue, N=216), only from the CDS (red, N=117) or only from the 3′ UTR (green, N=20). The vertical dotted grey line marks the 1.5-fold difference used to define differential TE (as in (D)). Although the difference between the "5′ UTR diversity only" and the "CDS diversity only" distributions did not reach statistical significance (D=0.15349, p=0.056, two-sample KS-test), these results suggested that tissue specificity in TE was partially achieved by expressing transcript isoforms that differ in their 5′ UTR.

115

Quantifying the contribution of translation efficiency to tissue-specific gene expression levels

Applying identical cut-offs on RPKM (reads per kilobase of transcript per million mapped reads) for both organs, we identified 10289 genes whose expression was shared and which were detectable at both levels, RNA and RPF (Fig. 2A). Of note, the number of genes expressed uniquely in one organ was greater for kidney than for liver (2134 vs. 387 cases of tissue-specific expression), likely owing to the more heterogeneous cellular composition of this organ.

Interestingly, we noted that between organs, footprint abundances correlated consistently better than did RNA abundances (Spearman ρ [RPF]: mean 0.784 vs. ρ [RNA]: mean 0.736; p=1.19e-07; Wilcoxon signed rank test) (Fig. 2B; Supplemental Fig. S2C-D). This observation lends further support to the idea that differences in mRNA expression across tissues are partially compensated by translation, leading to convergence at the level of protein biosynthetic output. Of note, this finding is conceptually similar to the reported higher evolutionary conservation of the proteome vs. the transcriptome that has been described (e.g. (Schrimpf et al. 2009; Khan et al. 2013)).

From the ratio of CDS-mapping RPF-seq to RNA-seq reads we next calculated relative translation efficiencies (TEs) per transcript and for each organ. We first quantified the degree to which TE differences contributed to organ-specific gene expression output. These analyses revealed that TEs were overall rather similar between organs, as 95% of genes fell into a less than 3-fold range for the kidney/liver relative TE ratio (Fig. 2C). By contrast, this range was greater than 100-fold for the transcript abundance ratio. Given that already within each organ, mRNA abundances showed a considerably broader spread than TEs (several hundredfold vs. barely greater than tenfold, respectively; Supplemental Fig. S4A-B; see also (Ingolia et al. 2011; Janich et al. 2015)), large differences in TEs between organs were not to be expected. Together with other recent analyses (discussed in (Li and Biggin 2015; Liu et al. 2016)), our observations make earlier suggestions that TE could be particularly good predictors of protein abundances seem improbable. They are rather in line with a dominant role for the regulation of mRNA levels (i.e., transcription and mRNA decay) in controlling gene output.

Despite the overall comparatively narrow dynamic range, TE differences between the organs reached statistical significance for a large proportion of the transcriptome. Comparison of the 24 kidney and

matching 24 liver samples conferred high statistical power and yielded 5013 genes whose transcripts' TEs were significantly different between tissues (Wilcoxon signed rank test for paired samples; FDR<0.01); we further implemented a 1.5-fold cut-off on TE ratio in order to select the cases that showed the strongest regulation, resulting in 960 "TE different" genes (Fig. 2D).

We next examined whether any overt transcript characteristic would have predictive value for differential TE. Remarkably, however, we were unable to identify any single, dominant mRNA feature that would potentially indicate an underlying mechanism. For example, we had previously observed that in the liver the presence of a translated upstream open reading frame (uORF) in the 5' UTR had strong predictive value for low TE on the transcripts' main ORF (Janich et al. 2015). An analogous analysis on the kidney datasets revealed a comparable relationship between uORF usage and low TE in this organ as well (Supplemental Fig. 4C). Nevertheless, we were unable to detect a significant correlation between differential uORF usage and TE differences of transcripts across organs (data not shown; see also below, Fig. 5F). Similarly, the 960 "TE different" transcripts were not enriched for any predicted miRNA binding sites, making it unlikely that this class of post-transcriptional regulators is a major player in establishing tissue-specific TEs (data not shown).

The only feature that we were able to identify as significantly associated with differential TE was transcript isoform diversity between the two organs i.e., the occurrence of tissue-specific mRNA variants as generated by alternative transcriptional start sites, by alternative splicing, and by alternative 3' cleavage/polyadenylation sites (Fig. 2E). In short, in this analysis we first used the RNA-seq data to compile an inventory of the annotated, protein-coding transcript isoforms and their estimated relative expression levels for each gene and separately for both tissues. We then used the Hellinger distance (Gonzalez-Porta et al. 2012) as a measure of dissimilarity of isoform expression levels between kidney and liver. A value of 0 for this metric signifies that a gene has identical distribution of isoform expression levels between the tissues and a value of 1 the lack of overlap in expressed isoforms. Globally, "TE different" genes showed significantly higher Hellinger distances than the remainder of the expressed genes (p=3.74e-04; Kolmogorov-Smirnov-test) (Fig. 2E), indicating that TE differences between tissues may in part have their origin in tissue-specific transcript variants. However, it is important to note that this

mechanism can potentially account only for some of the observed TE divergence, as slightly over half of all expressed genes, including those within the "TE different" set, showed Hellinger distance of 0 (i.e., the same, single protein-coding transcript isoform was expressed in kidney and liver) (Fig. 2E). Molecularly, the term "transcript isoform" comprises variations in mRNA structure that can affect 5′ UTR, CDS, 3′ UTR, or frequently combinations thereof. To evaluate whether any particular such variation would be more predictive of TE differences than another, we selected the genes for which the expressed variants affected only one of the features (5′ UTR, CDS, or 3′ UTR), omitting the more common, but potentially complicated combinatorial cases from our analysis. We found that transcript diversity affecting the 5′ UTR was more highly associated with differential TE than was CDS diversity (Fig. 2F). Of note, the association of the 5′ UTR with TE is in line with the view that initiation is rate-limiting for translation and that structure and sequence at and upstream of the translational start determine the efficiency with which scanning ribosomes commit to a productive engagement. Unfortunately, the low number of transcripts that showed exclusive 3′ UTR diversity (20 genes) precluded a reasonable interpretation, and although the generally low number of genes available for all three groups limited the statistical power of our analyses, it is remarkable that the difference between 5′ UTR and CDS only marginally failed the commonly used threshold of statistical significance (p=0.056; Kolmogorov-Smirnov-test) (Fig. 2F).

Finally, we were interested in whether cross-organ differences in TE were associated with functional classes of transcripts. For the 640 "TE different" genes that showed increased translation rate in liver (Fig. 2D), gene ontology (GO) analyses revealed significant enrichment for categories related to transcription (Supplemental Table S2). Conceivably, tissue-specific translational control of transcriptional regulators may have wide-ranging implications for the organs' transcriptomes. The 320 "TE different" genes that were translated better in kidney (Fig. 2D) did not show any enrichment for specific functional classes.

Translational modulation of phase of oscillation in kidney

We next turned to the analysis of factor time across the datasets. In order to annotate rhythmic events for kidney we used the same methodology, including a 1.5-fold cut-off on peak-to-trough amplitudes, as previously for the liver time series (Janich et al. 2015). A list of the detected RNA and RPF rhythms and

196 genome-wide gene expression plots can be found in Supplemental Table S3 and in Supplemental Dataset
197 S1, respectively. Our analyses yielded 1338 genes whose RNA abundance oscillated and 977 that cycled at
198 the footprint level, corresponding to 10.8% and 7.9% of expressed genes in kidney (Fig. 3A). The overlap of
199 542 genes corresponded to 41% and 55% of the "transcript rhythmic" and "footprint rhythmic" cases,
200 respectively (Fig. 3A). However, it is important to note that this rather modest concordance between the
201 rhythmic RNA and RPF gene sets certainly underestimates the true extent of rhythmicity shared between
202 RNA abundance and ribosome footprints, and overestimates the extent of "RNA only" and "translation
203 only" oscillations. The reasons lie in the propensity of rhythmicity detection algorithms to generate false-
204 negatives, and the lack of a canonical method to reliably determine true absence of rhythms (a common
205 problem in the field, see (Luck and Westermark 2016) for discussion). As in our previous study (Janich et
206 al. 2015) the only sector from the Venn diagram that we analyzed further was that of the 542 common
207 rhythmic gene, and we implemented more sophisticated methods to identify the true-positive
208 "translation only" cycling transcripts (see later).
209 Interestingly, the comparison of RNA abundance and footprint rhythmicity parameters across the 542
210 genes revealed that the timing of the RPF peaks relative to the RNA peaks had a significantly different and
211 broader distribution than the analogous set from liver ($p<1.0e-04$; permutation test) (Fig. 3B). This finding
212 suggested that the phase of protein biosynthesis rhythms undergoes stronger translational modulation in
213 kidney than it does in liver, where RPF peaks are more tightly gated by RNA abundance peaks. Moreover,
214 it was striking that the kidney data showed a tendency for maximal translation to *precede* maximal mRNA
215 abundance (Fig. 3C). Although the mean RPF peak time advance (-0.123 hours) did not reach statistical
216 significance ($p=0.16$, Wilcoxon rank sum test), the large number of transcripts (282) for which the rhythm
217 of translation was phase-advanced to its RNA accumulation was intriguing.
218 A shortcoming of the above phase analysis is that different rhythmic gene sets were compared between
219 the two organs. The alleged tissue differences in the RPF-RNA phase relationships could therefore have
220 simply arisen from transcript-specific rather than tissue-specific differences in the timing of translation.
221 We therefore determined the common rhythmic transcript set by overlapping kidney and liver with regard
222 to rhythmic events (Fig. 3D). A group of 178 genes (that included most core clock components;

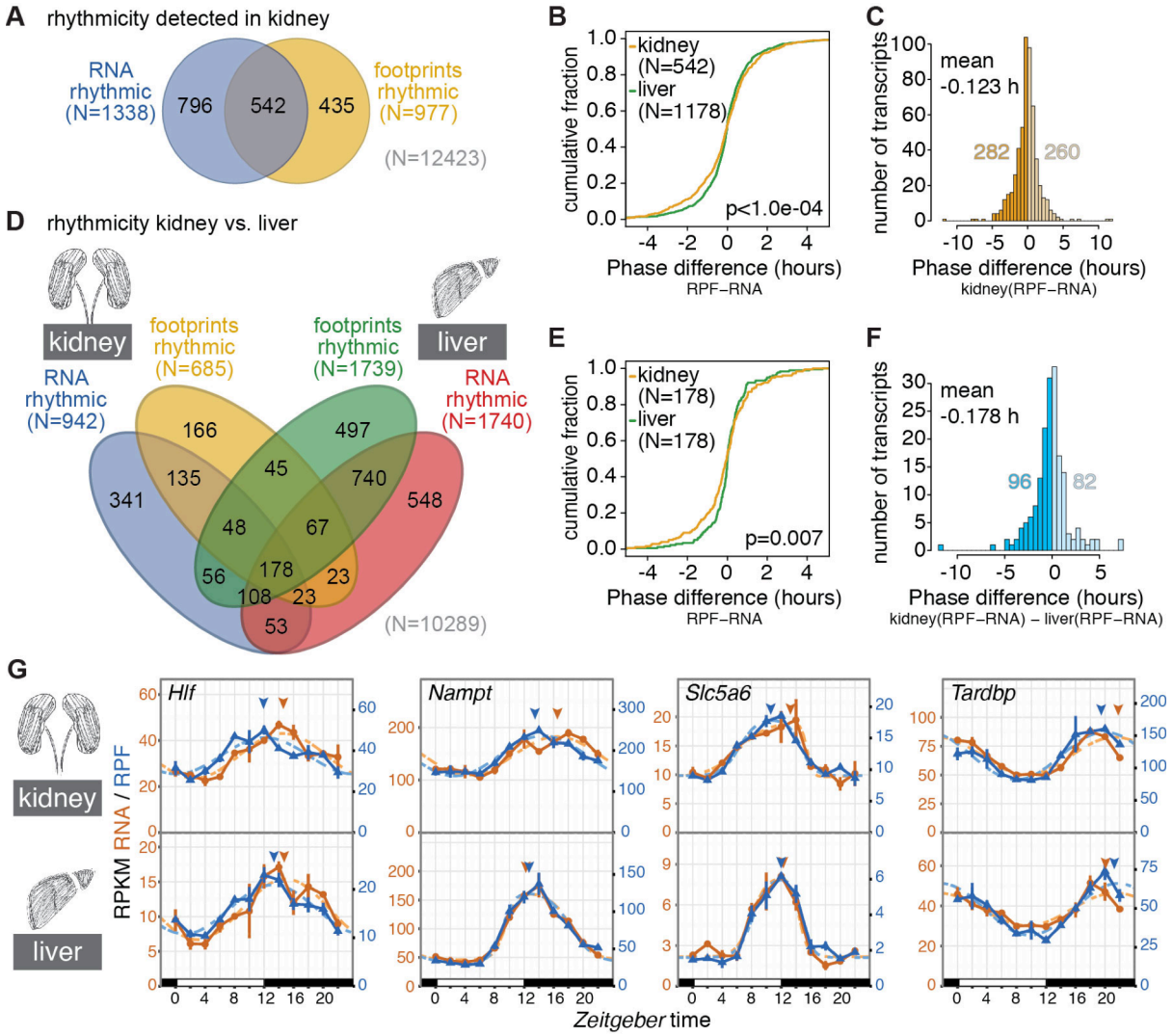


Figure 3. Rhythmicity analysis across organs and oscillation phase regulation.
(A) Venn diagram of the rhythmicity analysis in kidney. Of the 12423 genes detected, 1338 showed 24-hour oscillations of >1.5-fold amplitude in mRNA abundance (RNA-seq, 10.7%), 977 in footprints abundance (RPF-seq, 7.9%), and 542 (4.3%) were detected as rhythmic at both levels. 10650 (85.7%) genes were detected as non-rhythmic in our analysis. **(B)** Cumulative distribution of phase differences (RPF peak - RNA peak, in hours) for genes rhythmic at both RNA-seq and RPF-seq in liver (green, N=1178) and kidney (yellow, N=542). The two distributions were significantly different (p < 1e-04, permutation test), and reflected that maximal footprint abundance frequently preceded mRNA abundance peaks in kidney (note that the two distributions differed mostly in their negative tail). **(C)** Histogram of phase differences (RPF-RNA, in hours) for genes rhythmic at both levels in kidney (N=542). Although the distribution mean was not significantly different from 0, more genes had their footprint abundance peaks advanced (N=282) than delayed (N=260) with respect to their mRNA abundance peak. **(D)** 4-way Venn diagram of rhythmicity sets for genes expressed in both tissues (n=10289). 364 and 238 genes were detected as rhythmic in both organs at the RNA-seq and RPF-seq levels, respectively, and 178 genes were detected as rhythmic throughout (i.e. at RNA-seq and RPF-seq, in kidney and in liver). **(E)** Cumulative phase difference distribution in liver (green) and kidney (yellow) for the 178 genes rhythmic throughout. As in (B), the distributions were significantly different (p=0.007, permutation test), and reinstated that even when comparing the same set of genes, footprint peaks frequently preceded mRNA abundance maxima in kidney. **(F)** Histogram of the differential (kidney - liver) phase delay (RPF - RNA) for the 178 genes rhythmic throughout. **(G)** Daily profiles of RPF-seq RPKM (blue) and RNA-seq RPKM (orange) for four representative genes in which footprint abundance peaks preceded by several hours maximal mRNA abundance in kidney (top) but not, or less so, in liver (bottom). Arrowheads indicate the peak in footprint and mRNA abundance in their respective colors as estimated by the rhythmic fits.

223

Supplemental Fig. S5A; Supplemental Table S4; Supplemental Dataset S2), showed rhythmicity throughout, i.e. in both organs at RNA and RPF level. For this transcript set, the distribution of RPF-RNA intervals was significantly broader in kidney than in liver (Fig. 3E; p=0.007, permutation test) with an RPF peak phase advance in kidney (mean -0.143 h) and a phase delay in liver (mean 0.036 h) (Supplemental Fig. S5B-C). We next calculated the gene-wise RPF-RNA phase difference in kidney relative to that in liver. More genes showed their RPF peaks advanced (96) than delayed (82) in kidney vs. liver, with a mean phase advance that amounted to -0.178 hours (Fig. 3F). Although the phase advance globally did not reach statistical significance (p=0.152, Wilcoxon rank sum test), visual inspection of the RNA and RPF profiles identified numerous compelling cases of transcripts where specifically in kidney translation peaked up to several hours ahead of maximal RNA abundance, as shown by the examples *Hlf* (Hepatic leukemia factor, a circadian PAR-domain basic leucin zipper transcription factor), *Nampt* (Nicotinamide phosphoribosyltransferase, an enzyme involved in NAD biosynthesis), *Slc5a6* (solute carrier family 5 member 6, a sodium-dependent transporter for biotin and other vitamins) and *Tardbp* (TAR DNA binding protein) (Fig. 3G).

Translation that is phase-advanced to mRNA abundance is counterintuitive at first sight. Conceivably, however, it may occur when the translation rate is not constant over the lifetime of the mRNA but decreases, e.g. as a result of its gradual deadenylation liver (Kojima et al. 2012). In keeping with this hypothesis, we have observed that most subunits of the major cytoplasmic deadenylase complex, CCR4-NOT, are significantly more highly expressed in kidney than in liver (Supplemental Fig. S6A-C). Higher deadenylase activity in kidney could provide an attractive molecular explanation for the observed tissue-specific differences in RPF-RNA phasing, and in particular for phase-advanced RPF rhythms.

High tissue divergence in translationally driven rhythms

The aforementioned predisposition of rhythmicity detection methods to yield false-negatives, which was confirmed in our liver study (Janich et al. 2015) and evident in the kidney datasets as well (Supplemental Fig. S7), reinforced the notion that Venn diagrams that simply overlap rhythmic gene sets need to be interpreted with caution. Specifically, it led to an overestimation of the number of “RNA only” and of

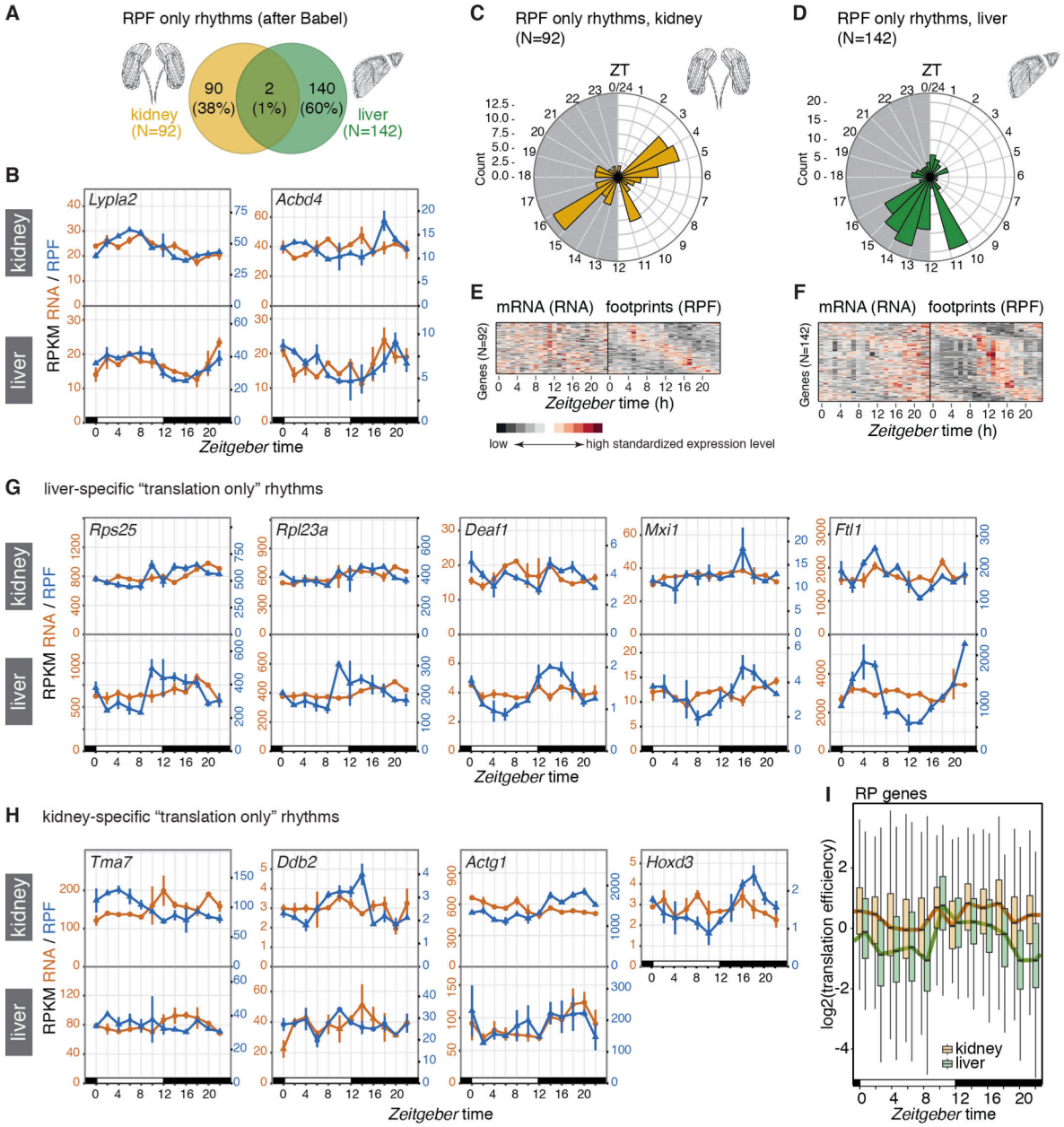


Figure 4. Tissue-specificity of translational rhythms.

(A) Venn diagram of rhythmic RPF-seq sets in liver (green, N=142) and kidney (yellow, N=92) after the *Babel* analysis showed the tissue-specificity of translational control. **(B)** Daily profiles of RPF-seq RPKM (blue) and RNA-seq RPKM (orange) for the two genes detected as translationally regulated in both tissues in (A). **(C)** Circular phase histogram for the 92 genes showing footprint rhythmicity in kidney. Maxima were clustered around ZT04 and ZT16, but not at the light-dark transition. **(D)** Circular phase histogram for the 142 genes showing footprint rhythmicity in liver. As reported before (Janich et al. 2015), translational rhythmicity in liver is enriched for maxima at the day-to-night transition. **(E)** Heatmap of RNA-seq (left) and RPF-seq (right) expression for the 92 genes translationally regulated in kidney. Genes are sorted by footprint phase and expression levels are standardized by row (gene). The heatmap reflects presence and absence of rhythmicity at the footprint and mRNA abundance levels, respectively. **(F)** As in (E) for the 142 rhythmically translated transcripts from liver. **(G)** Daily profiles of RPF-seq RPKM (blue) and RNA-seq RPKM (orange) for representative examples of genes with translational rhythms only in liver (lower panels) and not in kidney (upper panels). **(H)** Daily profiles of RPF-seq RPKM (blue) and RNA-seq RPKM (orange) for representative examples of genes with translational rhythms only in kidney (upper panels) but not in liver (lower panels). *Hoxd3* was not expressed in liver. **(I)** Translation efficiency (TE) along the day for ribosomal protein (RP) genes expressed in liver (green, n=86) and in kidney (yellow, n=89). For each timepoint (ZT) boxplots represent the interquartile range and whiskers extend to the minimum and maximum TE within 1.5 times the interquantile range. Lines connect the median of each boxplot to ease visualization. Notice the global TE upregulation at ZT10 in liver, whereas TEs in kidney remain high over the day.

250

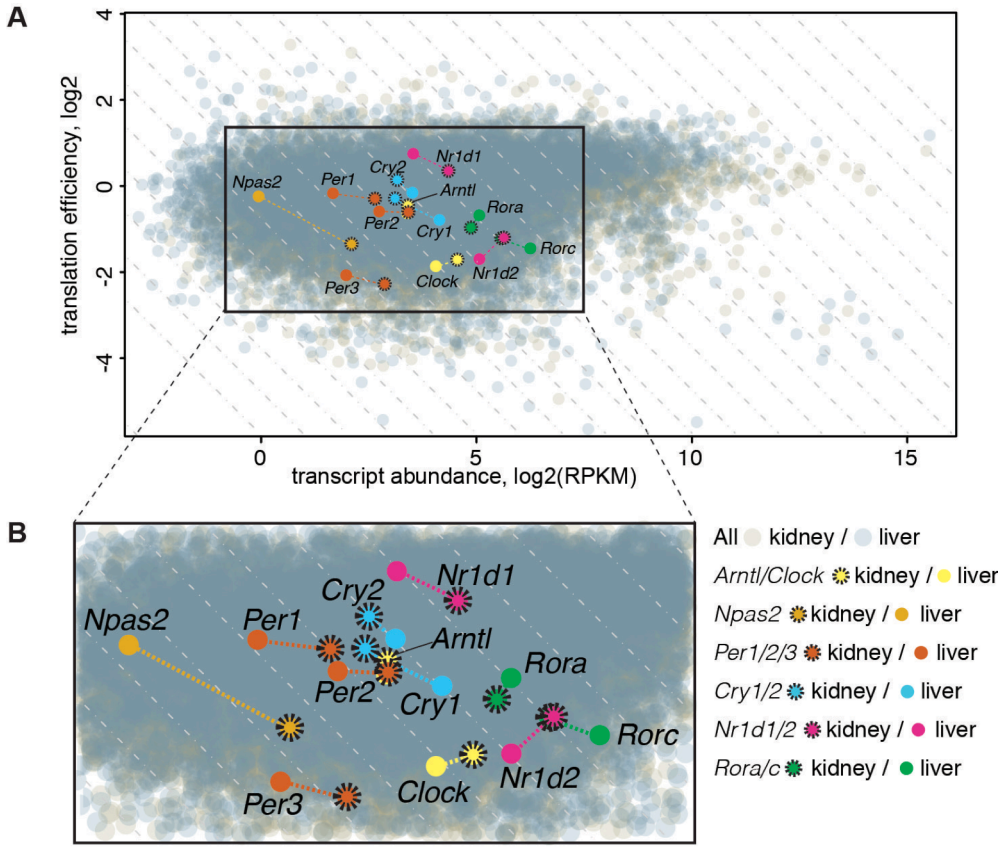
"footprints only" rhythmic genes (Supplemental Fig. S7B, D). In order to identify the rhythmically translated, constantly abundant transcripts in kidney with higher confidence, we implemented the analytical framework *Babel* (Olshen et al. 2013) that we had previously used for the liver data as well (Janich et al. 2015), to preselect the transcripts whose translation efficiency changed significantly over the day (and/or whose TEs deviated significantly from the global transcript population). Rhythmicity analyses that were then performed on this gene subset yielded 92 cases with the sought after temporal profiles of rhythmic translation on non-rhythmic mRNAs (Fig. 4A). Comparison with the 142 genes of the analogous set from liver revealed that translationally driven oscillations in protein biosynthesis showed near-perfect tissue specificity. Only two genes, *Abcd4* (ATP binding cassette subfamily D member 4) and *Lypla2* (lysophospholipase 2), were shared between the organs and, moreover, visual inspection indicated that they were among the least compelling cases of "translation only rhythms" that our method had identified (Fig. 4B).

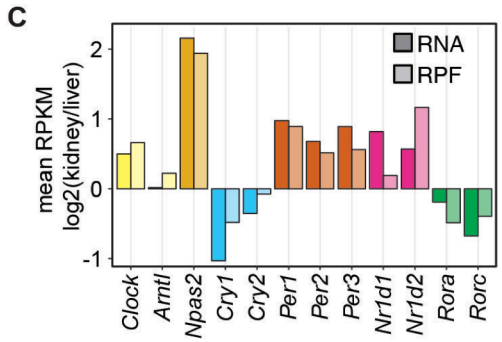
Marked tissue-specificity was further apparent for the daily timing of rhythmic translational events. The phase histograms thus showed striking differences in the peak time distribution between the organs (Fig. 4C-D; difference in distributions: p=1.66e-04; W = 17.403, df = 2; Watson-Wheeler test for homogeneity of angles). Of note, the enrichment for translational maxima at the light-dark transition (*Zeitgeber* time, ZT10-16; ZT00 corresponds to lights-on and ZT12 to lights-off) that dominated the distribution in liver (Fig. 4D, F) was virtually absent from kidney (Fig. 4C, E). Instead, kidney showed enrichment for transcripts with maximal translation occurring around ZT4 and ZT16. Visual inspection of individual examples confirmed liver- and kidney-specificity of RPF rhythms. The cases of robust translational oscillations that we (Janich et al. 2015) and others (Jouffe et al. 2013; Atger et al. 2015) had previously identified in liver were thus absent or severely blunted in kidney; of note, this included mRNAs encoding ribosomal proteins that make up the bulk of genes showing translational upsurge at the light-dark transition (e.g. *Rps25*, *Rpl23a*), as well as transcripts encoding the transcription factors *Deaf1* (deformed epidermal autoregulatory factor 1) and *Mxi1* (MAX interactor 1), and mRNAs containing iron-responsive elements in their 5' UTRs (e.g. Ferritin light chain 1, *Ftl1*) (Fig. 4G), all of which we had previously reported as translationally rhythmic in liver (Janich et al. 2015). Rhythmic translation exclusive to kidney was not significantly enriched for particular

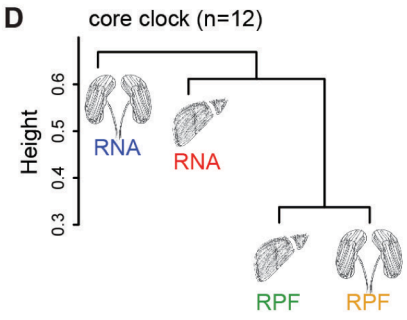
pathways (data not shown), and the temporal profiles were overall of lower amplitude than those seen for liver; *Tma7* (Translational machinery associated 7 homolog), *Ddb2* (Damage-specific DNA binding protein 2), *Actg1* (Actin, gamma, cytoplasmic 1) and *Hoxd3* (Homeobox D3; not expressed in liver) were among the most distinct examples (Fig. 4H). We concluded that temporal changes in TE were relatively rare in kidney and overall strikingly tissue-specific, possibly indicating differential sensitivity of the organs to the systemic signals or other mechanisms that drive such protein biosynthesis rhythms. Specifically for the most prominent group of genes subject to daily TE regulation in liver i.e., transcripts encoding ribosomal proteins (RPs) and other components of the translation machinery, it has been suggested that feeding-dependent mTOR-signaling underlies their translational upsurge at the light-dark transition via a mechanism that involves the 5'-terminal oligopyrimidine (5'-TOP) motifs that these transcripts carry (Jouffe et al. 2013; Atger et al. 2015), and our findings thus suggest that kidney is less sensitive to the responsible systemic cues. Moreover, the TE comparison between both tissues revealed that in kidney RPs were translated at a higher level throughout the day (Fig. 4I), indicating that the lack in rhythmicity in this organ resulted from an absence of translational repression during the light phase rather than an absence of activation in the dark phase.

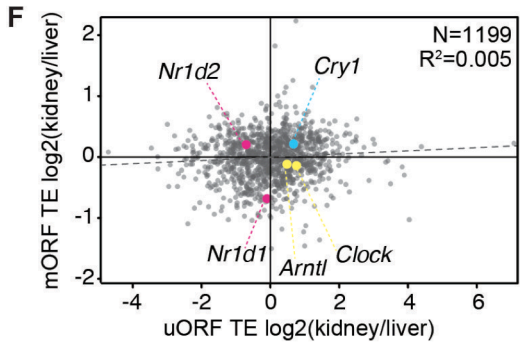
Tissue-specificity in core clock protein biosynthesis

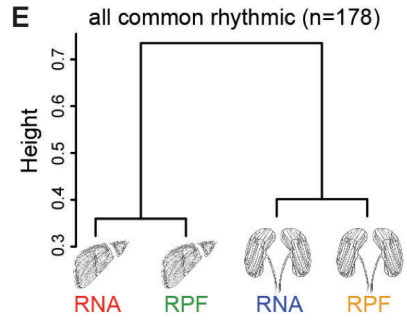
Observations such as the signature of translational compensation (Fig. 2B) or the phase modulation of CCGs (Fig. 3E), led us to conclude that the initial transcriptomal gene expression outputs underwent widespread refinement at the translational level. In particular within the core clock circuitry, such regulation could have important consequences. Conceivably, the rate and timing in the biosynthesis of individual clock proteins could underlie known cell type differences in core clock parameters (such as free-running period *ex vivo*/phase of oscillation *in vivo* (Yoo et al. 2004)), in clock output gene repertoires (Zhang et al. 2014), in oscillator strength and robustness (e.g. (Yagita et al. 2010; Lee et al. 2011)), or in clock gene loss-of-function phenotypes (e.g. (Landgraf et al. 2016)). We therefore aimed to gain quantitative insight into core clock protein biosynthesis in the two organs and at all three levels, RNA abundance, protein biosynthetic output (footprints), and translation efficiency. We first investigated RNA-

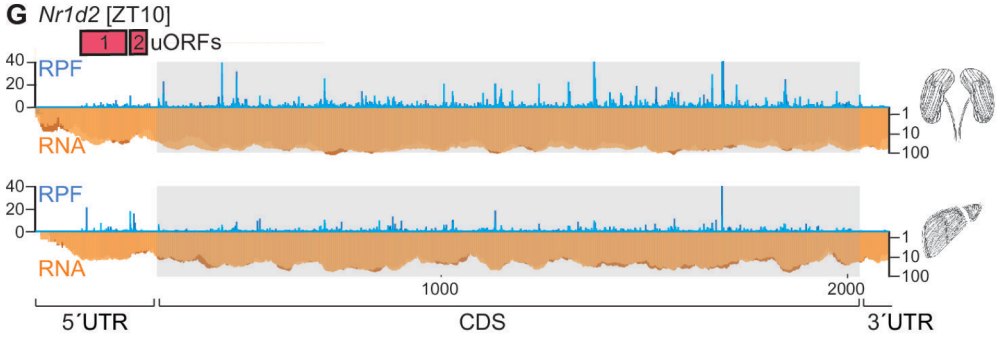












304

Figure 5. Translational buffering within the core clock circuitry.
(A) Scatterplot of transcript abundance vs. translation efficiency for liver (grey) and kidney (sepia), where core clock components are highlighted. Diagonal grey dashed lines indicate same amounts of absolute protein synthesis, where RNA abundance differences are compensated by TE differences. Colored dashed lines join the relative locations of each core clock genes between organs (kidney dots with dashed circles). Note that several core clock components (e.g. *Per* genes, *Cry1*, *Nr1d1*, *Rorc*, *Npas*2) are located along the descending diagonals, suggesting that translational buffering occurs for clock components and counteracts mRNA expression variations. **(B)** Magnification of the indicated area of (A). **(C)** Bar graph of the average RPKM ratio between kidney and liver for the main circadian core clock genes, at the level of mRNA abundance (dark shades) and ribosome footprints (light shades) showed that translational buffering led to a higher similarity at the level of protein biosynthesis (RPF) for several core clock genes. **(D)** Hierarchical clustering of the organs' transcriptomes and translatomes based on the similarities of the core clock genes expression patterns (n=12 genes shown in A-C). The height of the branches represents weighted average distances over the considered genes (see Methods). **(E)** Hierarchical clustering of the organs' transcriptomes and translatomes based on the similarities in expression patterns of genes detected as rhythmic throughout (i.e. at both RNA and RPF levels in both organs, N=178, see Fig. 3D). The height of the branches represents weighted average distances over the considered genes (see Methods). When compared to the clustering based on core clock gene expression patterns in (D) – for which the higher conservation of protein synthesis levels than mRNA levels was evident – this rhythmic gene set showed an organ-based clustering. **(F)** Scatterplot of upstream ORF (uORF) translation efficiency (TE) vs. main ORF (mORF) TE across organs for genes containing AUG-initiated translated uORFs in both organs (n=1199). uORF-containing core clock genes are highlighted. This analysis showed that differential uORF usage could not globally explain differences in mORF TE across organs (note the lack of negative correlation between the two variables, R^2 = 0.005, p=0.008). As an exception, the lower uORF TE of *Nr1d2* in kidney might explain its relatively higher mORF TE. **(G)** Raw read distribution for RPF (in blue) and RNA (in orange) along the 5' UTR and CDS of *Nr1d2* in kidney (top) and liver (bottom) for the timepoint of maximal CDS translation. Red boxes indicate AUG-initiated uORFs as predicted in our analyses.

seq RPKMs and TEs in a non-time-resolved fashion (averages over timepoints) in order to compare the cumulative daily production across all clock proteins. We noted that most core clock components showed a considerable degree of organ-specificity in their expression that was readily appreciable in the TE vs. RNA abundance representation with both organs plotted in a single graph (Fig. 5A-B). In this representation, identical amounts of biosynthesized protein (i.e., identical RPF RPKMs) locate along the descending diagonals on which differences in transcript abundance and TE cancel each other out. Interestingly, the majority of core clock genes (*Npas2, Cry1, Cry2, Per1, Per2, Per3, Nr1d1, Rorc*) showed, at least to some degree a diagonal vectorial component (Fig. 5B), which indicated compensatory TE changes that partially counteracted RNA abundance differences between the organs. For *Clock, Arntl, Nr1d2* and *Rora*, changes in TE exacerbated transcript abundance differences (Fig. 5B-C).

The inspection of quantitative relationships between clock components revealed that the main positive *(Clock, Arntl)* and negative *(Per1/2, Cry1/2)* limb members were all produced in roughly comparable amounts in kidney (i.e., they aligned along a relatively narrow diagonal zone in Fig. 5B), whereas interconnecting limb protein biosynthesis *(Nr1d1/2, Rora/c)* was 2-4-fold higher (i.e., they were shifted to the right in Fig. 5B). Overall, we had previously observed a similar pattern in the liver data (Janich et al.

2015). Despite the overall similarity between kidney and liver, however, we also noticed two striking manifestations of organ specificity in the amounts and relative ratios of clock protein biosynthesis. First, the interconnecting limb appeared to be subject to "reprogramming", with an increased biosynthesis of repressive *(Nr1d1/2)* and a decreased production of activating *(Rora/c)* elements in the kidney (Fig. 5C). Of note, beyond functioning in the rhythm-generating clock circuitry, interconnecting limb transcription factors also control an output branch of the oscillator. Consistent with the gene expression differences that we observed (i.e., more activators and less repressors in liver), major loss-of-function phenotypes of interconnecting limb components that have been reported are indeed associated with hepatic pathways, for example with lipid, cholesterol and bile acid metabolism (e.g. (Raspe et al. 2002; Le Martelot et al. 2009)). We deem it an attractive hypothesis that by controlling the relative levels of NR1D1/2 vs. RORs, clock output gene repertoires would be tailored in a tissue-specific fashion. Translational mechanisms are likely involved in the regulation, in particular to enhance NR1D2 production in kidney. Second, we observed that within the negative limb, the ratio between PER and CRY biosynthesis was shifted towards increased PERs in kidney (Fig. 5C). PERs (and in particular PER2) are considered stoichiometrically rate-limiting components of the inhibitory complex, and increased PER2 dose engenders long period (Gu et al. 2012; D'Alessandro et al. 2015). Consistent with our finding of increased PER production, the free-running kidney clock shows almost 1.5 hour longer period than that of the liver (Yoo et al. 2004). Conceivably, the modulation of the relative levels of core clock protein production could engender different stoichiometry of the circuitry components and lead to distinct oscillatory parameters of clocks across tissues. Of note, the increase in PER and decrease in CRY biosynthesis in kidney is already established at the transcriptomal level, with TE differences rather leading to partial compensation (Fig. 5B-C).

We next extended the core clock analysis to the time-resolved data. In the expression plots (Supplemental Fig. S8) we calculated for each gene individually the Euclidean distances between the four rhythmic traces (i.e., RNA, RPF in kidney, liver), which served as a measure of similarity between the temporal profiles. Hierarchical clustering of the similarities for the ensemble of the 12 core clock genes showed that RPF rhythms of the two organs grouped together (Fig. 5D), indicating higher similarity of clock protein biosynthesis rhythms between organs than of RNA and RPF rhythms within organs. As a control set, we

analyzed the 178 common rhythmic genes identified in Fig. 3D, which revealed within-organ clustering (Fig. 5E). These findings underscored that translational compensation was operative within the core clock, leading to more similar rhythms in clock protein biosynthesis than would have been predicted from the rhythmic RNA abundance profiles.

In our previous liver RPF-seq study we had identified uORF translation as a mechanism that is able to regulate the gene expression output for clock components, and we had annotated AUG-initiated uORFs in *Nr1d1, Nr1d2, Cry1, Clock* and *Arntl* (Janich et al. 2015). We therefore wished to evaluate whether differential uORF usage could potentially underlie any of the observed cross-organ TE differences for the core clock components. To this end, we calculated the translation efficiency specifically on uORFs (from the ratio of uORF-mapping RPF to RNA reads; see Methods) for all uORF-containing genes in the two organs, and then correlated uORF TE ratio (kidney/liver) with main ORF TE ratio (kidney/liver) (Fig. 5F). Importantly, given that uORF translation has been associated with decreased initiation at the main ORF ((Wethmar 2014; Janich et al. 2015) and Supplemental Fig. S4C), we would have expected uORF and main ORF TE ratios to negatively correlate if differential uORF usage were one of the main mechanisms to establish organ-specific TEs. This was, however, globally not the case, and we rather observed weak positive correlation (Spearman ρ=0.095). Among individual core clock genes, *Nr1d2* represented the clearest case for negative correlation, showing lower uORF TE and higher main ORF TE in kidney (Fig. 5F). Moreover, this negative correlation of uORF- and CDS-mapping footprints in the two organs could be confirmed by visual inspection of raw RPF reads mapping to the *Nr1d2* transcript (Fig. 5G). For *Nr1d2*, differential uORF usage could thus represent a plausible mechanism that contributes to organ-specific protein production, keeping its biosynthesis low in liver and high in kidney.

Discussion

Along the way from transcription to protein degradation, gene expression can be regulated at numerous levels. Certain steps and intermediates have been particularly well explored, including by genome-wide and quantitative approaches. This has led overall to the view that gene expression differences are typically generated transcriptionally and can be conveniently studied at the transcriptomal level. However, the

functionally relevant output of most gene expression is the protein rather than the transcript. Quantitative, genome-wide analyses of protein biosynthesis are thus of high interest to complement the wealth of available transcriptome data. Such studies are still scarce because it is only with the recent development of the ribosome profiling technique that a dedicated analysis of translational events in a high-throughput fashion has become possible (Ingolia 2014). Here, we report on a combined analysis of two paradigms of differential gene expression, namely its tissue dependence and its time of day-dependence, to evaluate the contribution of translation to the regulation of gene expression output. We have addressed several, rather fundamental, but still unanswered questions that are of interest to both chronobiology and the gene expression field at large: How does the dynamic range of translation efficiency compare to that of transcript abundance across two distinct organs? Is translation efficiency a default transcript property and comparable across two tissues, or do TEs become reinterpreted in different cellular environments? Does cross-tissue variability of TEs come with a direction i.e., is there a global tendency to either reinforce or to buffer transcriptomal differences? What is the extent and what are the properties of daily rhythms in translation efficiency, in an organ and between organs? Could differential TE underlie any of the known cross-tissue differences in core clock parameters?

To our knowledge, only one previous study has reported on RPF-seq datasets from two complementary mammalian tissues (Schafer et al. 2015). The authors recorded datasets from rat liver and heart but, of note, they also included animals with different genetic backgrounds as covariates in the experimental design. The study thus mainly focused on exploring strain differences and how genetic variation influenced translational levels, while tissue differences were not investigated in great detail. Our choice of liver and kidney (with more than 10'000 commonly expressed transcripts) from the same animals, and the high resolution of the time series (24 RPF-seq datasets per organ), provided high statistical power to our analyses.

Over the last years there has been some dispute regarding the contribution that differences in TE make to gene expression output (discussed in (Li and Biggin 2015; Liu et al. 2016)) – including suggestions that TE may actually represent the best predictor of protein abundances (Schwanhausser et al. 2011). Our data contribute to clarifying some of the disagreement and show that – across genes in a tissue and for

individual genes between tissues – the dynamic range of transcript abundances is about 30-50-fold broader than that of translation efficiencies. Gene expression differences are thus mainly set up by differences in transcription (and, possibly, RNA stability), whereas differences in translation rate have more of a modulatory role. It is noteworthy that this modulation is globally characterized by directionality – overall, TE differences thus help to buffer against mRNA abundance differences. Examples for such translational buffering of divergent gene expression (a phenomenon that was also covered in a recent review (Liu et al. 2016)) have been reported across yeast species (McManus et al. 2014) and across genetically different strains of rats (Schafer et al. 2015), and our study now extends this observation across organs. Conceptually, all these cases of buffering may reflect the fact that selective pressure on precise gene expression levels likely acts on protein abundances and that more tolerance may exist towards divergence in RNA levels. It will be exciting to further study the underpinnings of translational compensation, across tissues and across species.

At first site, it may appear unsatisfactory that our analyses did not identify specific, dominant transcript features and mechanisms that would explain TE differences between organs. Transcript isoform diversity (in particular at the 5' UTR) may play a role, but for more than half of the transcripts with differential TE, the same, single protein-coding isoform was expressed in both tissues under investigation. Tissue differences in the translation machinery and its regulators – including signaling pathways, the activity of trans-acting factors such as RNA binding proteins (RBPs), translation factors, and even ribosomal composition or tRNA repertoires – are likely involved. They may act in a combinatorial fashion and we expect that a complex translational regulatory universe thus awaits discovery. While ribosome profiling now allows us to record the consequences of such regulation, understanding all its causes represents an exciting challenge for the future.

Our study has allowed novel insights into rhythmic gene expression. The extent to which rhythmicity is generated by the temporal regulation of translation efficiency has been the subject of speculation ever since the first report that many rhythmic proteins in liver are encoded by non-rhythmic mRNAs (Reddy et al. 2006). Our new kidney datasets complement recent time-resolved ribosome profiling studies from liver (Atger et al. 2015; Janich et al. 2015) and from a circadian cell line, U2OS cells (Jang et al. 2015). Our

comparisons reveal that the number of transcripts subject to translational rhythms is slightly lower in kidney, but overall in a similar order of magnitude as in liver, affecting around 1% of the transcriptome. We were surprised to see that translational rhythms are essentially tissue-specific in terms of the identity of rhythmic translation events and in their phase distribution. A possible explanation is that these rhythms are not driven by local clocks, but by rhythmic systemic cues to which different tissues are not equally responsive. The effects of feeding and mTOR signaling, for example, may be more pronounced in liver than in kidney due to the dedicated role that this organ has in energy homeostasis and fasting responses. Nevertheless, the lack of rhythmicity of components of the translational machinery (ribosomal proteins) in kidney came as a surprise in the light of previous suggestions of conservation across tissues (Jouffe et al. 2013). In addition to the generation of rhythms by translation (which affected only a relatively small population of transcripts), our analyses have pointed to a rhythmicity-modulating role affecting the timing of protein biosynthesis oscillations relative to phase of mRNA abundance rhythms. Consistent with work by the Green lab that showed interactions between polyadenylation status of mRNAs and rhythmic protein expression in the liver (Kojima et al. 2012), it is tempting to speculate that such mechanisms are also operative across organs, with tissue-specific deadenylation kinetics tuning the timing of rhythmic protein biosynthesis.

Historically, the core clock mechanism has been referred to as a “transcript-translation feedback loop” (TTFL; see e.g. (Dunlap 1996) for an early mention of this term). The actual feedback, however, occurs at the transcriptional level, and a possible mechanistic role of translational regulation has not been much investigated. Our cross-organ comparison of core clock protein biosynthesis suggests that translational control – including through the activity of uORFs (Jang et al. 2015; Janich et al. 2015) – is of regulatory interest and represents a way by which the identical set of core clock genes could form circuitries with different stoichiometry of its main components. As a result, both clock parameters and output gene repertoires may be organ-specifically tuned.

Methods

Animals

12-week old male mice (C57BL/6J; Janvier Labs) were entrained for two weeks to light:dark 12:12 cycles with ad libitum access to food and water and were anesthetized (isoflurane) and sacrificed every two hours (ZT0 - ZT22, with ZT0 corresponding to "lights-on") for two daily cycles. Livers and kidneys were removed and processed either directly or flash-frozen in liquid N_2. All experimental procedures were approved by the Veterinary Office of the Canton Vaud (authorization VD2376).

Ribosome profiling

Generation of the ribosome profiling libraries (RPF-seq and RNA-seq) was described in our previous study (Janich et al. 2015). Kidney libraries were prepared in the same manner, with one modification. After RNA digestion and recovery of ribosome-protected fragments, 5µg of RNA were treated with Ribo-Zero magnetic kit (Epicentre) according to the manufacturer's protocol. Ribsomal RNA-depleted samples were then separated in a 15% PAGE gel. Gel was cut to obtain 26-35 nucleotides long fragments and library preparation was continued as done for liver samples and according to the ARTseq ribosome profiling kit instructions (Epicentre). These two steps (Ribo-Zero treatment and PAGE separation) had been inverted during the preparation of our liver samples in order to obtain sufficiently concentrated libraries for sequencing. However, applying this strategy for kidney samples recovered (and sequenced) the rRNA fragments of the Ribo-Zero kit intended for sample rRNA removal itself. We reasoned that this might be due to overall lower levels of translation (and therefore relatively less mRNA footprints) in kidney and decided to revert the steps and follow the original ribosome profiling protocol from Illumina. Total RNA libraries were prepared as previously described (Janich et al. 2015). RPF and RNA libraries were sequenced on an Illumina HiSeq 2500.

Sequencing data processing, alignment and quantification

Processing, quality assessment, alignment and quantification of sequencing data were performed as described in our previous study ((Janich et al. 2015; Janich et al. 2016)). Briefly, sequenced reads were trimmed of their adaptors using Cutadapt (Martin 2011) and the length distribution of trimmed reads was

used to assess the quality of nuclease digestion and size-selection steps, particularly important for RPF libraries (Supplemental Fig. S1B). Next trimmed reads were filtered by size (21-35 for RPF; 21-60 for RNA) using an in-house Python script, and sequentially mapped to mouse rRNA, human rRNA, mt-tRNA, mouse tRNA, mouse cDNA (Ensemble mouse database release 75) using Bowtie v2.2.1 (Langmead and Salzberg 2012) and mouse genome (GRCm38.p2) using Tophat v2.0.11 (Trapnell et al. 2009). Trimmed and filtered sequences were also directly mapped against the mouse genome (using Tophat) in order to estimate expressed transcript models in each organ (using Cufflinks v2.2.1 (Trapnell et al. 2010). Transcriptome-mapping reads in the sequential alignment were counted towards their location into the 5' UTR, CDS or 3' UTR of the transcript, based on feature annotation (Ensemble mouse release 75). Mappable and countable feature lengths were not calculated for this study (see "faux reads analysis" in the "Quantification of mRNA and ribosome footprint abundance" section of (Janich et al. 2015) Supplemental Experimental Procedures) as its contribution was negligible for further analyses. Therefore RPKM calculations in this study were not corrected with such factor. Read counts in RNA-seq and RPF-seq datasets were normalized with upper quantile method of edgeR (Robinson and Oshlack 2010) and RPKM values were calculated as the number of reads per 1000 bases per geometric mean of normalized read counts per million. Relative translation efficiencies (TE) were calculated as the ratio of RPF-RPKM to RNA-RPKM per gene per sample. Reading frame and nucleotide periodicity analyses were done using the same criteria as in (Janich et al. 2015). Principal Component Analysis (PCA) was done using a combined matrix of CDS counts for RPF and RNA from both liver and kidney and following the same approach as before (Janich et al. 2015).

Correlation of RNA-seq and RPF-seq across organs

Inter-organ correlation at the levels of RNA and RPF-seq (Fig. 2B) was done per timepoint and replicate. Significance of the difference in the spearman coefficient between both distributions was assessed by Wilcoxon rank sum test in R (stats package).

Analysis of differential translation efficiency

Significance of differences in translation efficiency (TE) between liver and kidney was assessed using the Wilcoxon-rank sum test in R (stats package). A two-sided, paired test was performed on centered TE

values per timepoint and replicate, and resulted p-values were FDR-corrected. A gene was defined as having differential TE when FDR < 0.01 and the inter-organ difference in TE was at least 1.5-fold (Fig. 2D).

Analysis of transcript usage diversity across organs

For each gene g, P(g) = (p1,...,pn) is the vector of the relative expression proportions of its n protein-coding transcripts, as estimated in our RNA-seq analysis (see Sequencing data processing, alignment and quantification). To quantify the dissimilarity in relative transcript isoform expression between liver L and kidney K, the Hellinger distance H (Gonzalez-Porta et al. 2012) is defined as:

$$H(P_L(g), P_K(g)) = 1/sqrt(2)\sqrt{\sum_{i=1}^{n}(\sqrt{p_L i} - \sqrt{p_K i})^2} \qquad (1)$$

Similarity between the distribution of the genes detected as differentially translated (Fig. 2D) and the overall distribution was tested by two-sided two-sample Kolmogorov-Smirnov (KS) test. In order to detect the transcript feature that mostly determines the tissue specificity in translation efficiency, we selected genes whose transcript diversity in both organs originated only from either the 5' UTR, the CDS, or the 3' UTR of the transcripts, based on the annotation information for the protein-coding transcripts detected. Similarity of the "5' UTR-only diversity" distribution to the "CDS-only diversity" distribution was tested with the two-sided two-sample Kolmogorov-Smirnov (KS) test, although the low and very different number of genes in each group (n=216 vs. n=117) might limit the power of the test to detect a significant difference between the distributions.

Rhythmicity analyses

Rhythmicity detection and rhythmic parameters estimation in each dataset (RNA-seq and RPF-seq, liver and kidney) were done based on Akaike information criterion (AIC) model selection as in our previous study (Janich et al. 2015). The Babel computational framework (Olshen et al. 2013) was used to detect rhythmically translated genes from constantly expressed mRNAs within each organ.

Hierarchical clustering of rhythmic genes

In order to study the similarity of rhythmic genes based on their expression profiles, a dissimilarity matrix was computed for each gene of interest, based on the Euclidean distance between the RNA-seq and RPF-

seq expression profiles within and across organs. A hierarchical clustering tree was constructed on the weighted average of the dissimilarities matrices under consideration (core clock genes in Fig. 5D or all rhythmic genes in Fig. 5E), using the "average" clustering method. The R functions{packages} dist{stats}, fuse{analogue} and hclust{stats} were used for computing the individual dissimilarity matrices, the weighted mean dissimilarity matrix, and the hierarchical clustering, respectively.

Upstream open reading frame (uORF) translation efficiency calculation

In order to assess the impact of differential upstream ORFs (uORFs) usage on translation efficiency differences across organs, uORFs were identified as in our previous study (Janich et al. 2015). Briefly, genes expressing a single protein-coding isoform in both organs were selected (n=5815) and uORFs starting with an AUG and being at least 18 nucleotides long were considered as translated if they showed significant frame bias towards the first reading frame (relative to the uORF 5') and had a coverage >10%. uORF translation efficiency was calculated from the ratio of RPF-seq to RNA-seq reads on the uORF regions. If several uORFs partially or completely overlapped on a given transcript 5' UTR, a non-overlapping composite uORF was considered for read counting.

Data access

The sequencing data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number [available on request].

Acknowledgments

We thank staff at the Lausanne Genomic Technologies Facility for high-throughput sequencing support and Ioannis Xenarios and staff at Vital-IT for computational support. D.G. was supported by Swiss National Science Foundation (professorship grants 128399, 157528); National Center for Competence in Research (NCCR) RNA & Disease; Fondation Pierre Mercier; Fondation Leenaards; Olga Mayenfisch Stiftung; SystemsX.ch StoNets consortium; and the University of Lausanne. P.J. was supported by Human Frontiers Science Program long-term fellowship LT000158/2013-L.

References

Atger F, Gobet C, Marquis J, Martin E, Wang J, Weger B, Lefebvre G, Descombes P, Naef F, Gachon F. 2015. Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proceedings of the National Academy of Sciences of the United States of America* **112**: E6579-6588.

Bonny O, Vinciguerra M, Gumz ML, Mazzoccoli G. 2013. Molecular bases of circadian rhythmicity in renal physiology and pathology. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association* **28**: 2421-2431.

Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M et al. 2011. The evolution of gene expression levels in mammalian organs. *Nature* **478**: 343-348.

D'Alessandro M, Beesley S, Kim JK, Chen R, Abich E, Cheng W, Yi P, Takahashi JS, Lee C. 2015. A tunable artificial circadian clock in clock-defective mice. *Nature communications* **6**: 8587.

Dibner C, Schibler U, Albrecht U. 2010. The mammalian circadian timing system: organization and coordination of central and peripheral clocks. *Annual review of physiology* **72**: 517-549.

Dunlap JC. 1996. Genetics and molecular analysis of circadian rhythms. *Annu Rev Genet* **30**: 579-601.

Gonzalez-Porta M, Calvo M, Sammeth M, Guigo R. 2012. Estimation of alternative splicing variability in human populations. *Genome research* **22**: 528-538.

Gu X, Xing L, Shi G, Liu Z, Wang X, Qu Z, Wu X, Dong Z, Gao X, Liu G et al. 2012. The circadian mutation PER2(S662G) is linked to cell cycle progression and tumorigenesis. *Cell death and differentiation* **19**: 397-405.

Ingolia NT. 2014. Ribosome profiling: new views of translation, from single codons to genome scale. *Nature reviews Genetics* **15**: 205-213.

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.

Jang C, Lahens NF, Hogenesch JB, Sehgal A. 2015. Ribosome profiling reveals an important role for translational control in circadian gene expression. *Genome research* **25**: 1836-1847.

Janich P, Arpat AB, Castelo-Szekely V, Gatfield D. 2016. Analyzing the temporal regulation of translation efficiency in mouse liver. *Genomics data* **8**: 41-44.

Janich P, Arpat AB, Castelo-Szekely V, Lopes M, Gatfield D. 2015. Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames. *Genome research* **25**: 1848-1859.

Jouffe C, Cretenet G, Symul L, Martin E, Atger F, Naef F, Gachon F. 2013. The circadian clock coordinates ribosome biogenesis. *PLoS biology* **11**: e1001455.

Khan Z, Ford MJ, Cusanovich DA, Mitrano A, Pritchard JK, Gilad Y. 2013. Primate transcript and protein expression levels evolve under compensatory selection pressures. *Science* **342**: 1100-1104.

Kojima S, Sher-Chen EL, Green CB. 2012. Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. *Genes & development* **26**: 2724-2736.

Landgraf D, Wang LL, Diemer T, Welsh DK. 2016. NPAS2 Compensates for Loss of CLOCK in Peripheral Circadian Oscillators. *PLoS genetics* **12**: e1005882.

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

Le Martelot G, Claudel T, Gatfield D, Schaad O, Kornmann B, Lo Sasso G, Moschetta A, Schibler U. 2009. REV-ERBalpha participates in circadian SREBP signaling and bile acid homeostasis. *PLoS biology* **7**: e1000181.

Lee Y, Chen R, Lee HM, Lee C. 2011. Stoichiometric relationship among clock proteins determines robustness of circadian rhythms. *The Journal of biological chemistry* **286**: 7033-7042.

Li JJ, Biggin MD. 2015. Gene expression. Statistics requantitates the central dogma. *Science* **347**: 1066-1067.

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

Luck S, Westermark PO. 2016. Circadian mRNA expression: insights from modeling and transcriptomics. *Cellular and molecular life sciences : CMLS* **73**: 497-521.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**: 10-12.

607 Mauvoisin D, Wang J, Jouffe C, Martin E, Atger F, Waridel P, Quadroni M, Gachon F, Naef F. 2014.
608 Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal
609 functions in mouse liver. *Proceedings of the National Academy of Sciences of the United States of*
610 *America* **111**: 167-172.
611 McManus CJ, May GE, Spealman P, Shteyman A. 2014. Ribosome profiling reveals post-transcriptional
612 buffering of divergent gene expression in yeast. *Genome research* **24**: 422-430.
613 Meireles-Filho AC, Bardet AF, Yanez-Cuna JO, Stampfel G, Stark A. 2014. cis-Regulatory Requirements for
614 Tissue-Specific Programs of the Circadian Clock. *Current biology : CB* **24**: 1-10.
615 Olshen AB, Hsieh AC, Stumpf CR, Olshen RA, Ruggero D, Taylor BS. 2013. Assessing gene-level translational
616 control from ribosome profiling. *Bioinformatics* **29**: 2995-3002.
617 Partch CL, Green CB, Takahashi JS. 2014. Molecular architecture of the mammalian circadian clock. *Trends*
618 *in cell biology* **24**: 90-99.
619 Raspe E, Duez H, Mansen A, Fontaine C, Fievet C, Fruchart JC, Vennstrom B, Staels B. 2002. Identification
620 of Rev-erbalpha as a physiological repressor of apoC-III gene transcription. *J Lipid Res* **43**: 2172-
621 2179.
622 Reddy AB, Karp NA, Maywood ES, Sage EA, Deery M, O'Neill JS, Wong GK, Chesham J, Odell M, Lilley KS et
623 al. 2006. Circadian orchestration of the hepatic proteome. *Current biology : CB* **16**: 1107-1115.
624 Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-
625 seq data. *Genome biology* **11**: R25.
626 Robles MS, Cox J, Mann M. 2014. In-vivo quantitative proteomics reveals a key contribution of post-
627 transcriptional mechanisms to the circadian regulation of liver metabolism. *PLoS genetics* **10**:
628 e1004047.
629 Schafer S, Adami E, Heinig M, Rodrigues KE, Kreuchwig F, Silhavy J, van Heesch S, Simaite D, Rajewsky N,
630 Cuppen E et al. 2015. Translational regulation shapes the molecular landscape of complex disease
631 phenotypes. *Nature communications* **6**: 7200.

632 Schrimpf SP, Weiss M, Reiter L, Ahrens CH, Jovanovic M, Malmstrom J, Brunner E, Mohanty S, Lercher MJ,
633 Hunziker PE et al. 2009. Comparative functional analysis of the Caenorhabditis elegans and
634 Drosophila melanogaster proteomes. *PLoS biology* **7**: e48.
635 Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global
636 quantification of mammalian gene expression control. *Nature* **473**: 337-342.
637 Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*
638 **25**: 1105-1111.
639 Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L.
640 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and
641 isoform switching during cell differentiation. *Nature biotechnology* **28**: 511-515.
642 Wethmar K. 2014. The regulatory potential of upstream open reading frames in eukaryotic gene
643 expression. *Wiley interdisciplinary reviews RNA* **5**: 765-778.
644 Yagita K, Horie K, Koinuma S, Nakamura W, Yamanaka I, Urasaki A, Shigeyoshi Y, Kawakami K, Shimada S,
645 Takeda J et al. 2010. Development of the circadian oscillator during differentiation of mouse
646 embryonic stem cells in vitro. *Proceedings of the National Academy of Sciences of the United*
647 *States of America* **107**: 3846-3851.
648 Yoo SH, Yamazaki S, Lowrey PL, Shimomura K, Ko CH, Buhr ED, Siepka SM, Hong HK, Oh WJ, Yoo OJ et al.
649 2004. PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian
650 oscillations in mouse peripheral tissues. *Proceedings of the National Academy of Sciences of the*
651 *United States of America* **101**: 5339-5346.
652 Zhang R, Lahens NF, Ballance HI, Hughes ME, Hogenesch JB. 2014. A circadian gene expression atlas in
653 mammals: Implications for biology and medicine. *Proceedings of the National Academy of*
654 *Sciences of the United States of America* **111**: 16219-16224.
655

SUPPLEMENTAL MATERIAL

Translational contributions to tissue-specificity in rhythmic and constitutive gene expression

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

Figure S1. Outcome of ribosome profiling around-the-clock in mouse liver and kidney.

A) Summary of the sequential mapping outcome, indicating the number (y-axis) and percentage (within bars) of reads mapping to each database, averaged over all timepoints. For each library of the four datasets an average of more than 20 million reads aligned to cDNA.

(B, C) RPF-seq (B) and RNA-seq (C) read lengths after trimming of adaptors showed that RPF-seq reads were mostly 29-30 nucleotides in both organs, whereas RNA-seq fragments showed a broader distribution as expected from chemical RNA fragmentation. Boxplots represent the interquartile range and whiskers extend to the minimum and maximum value within 1.5 times the interquartile range.

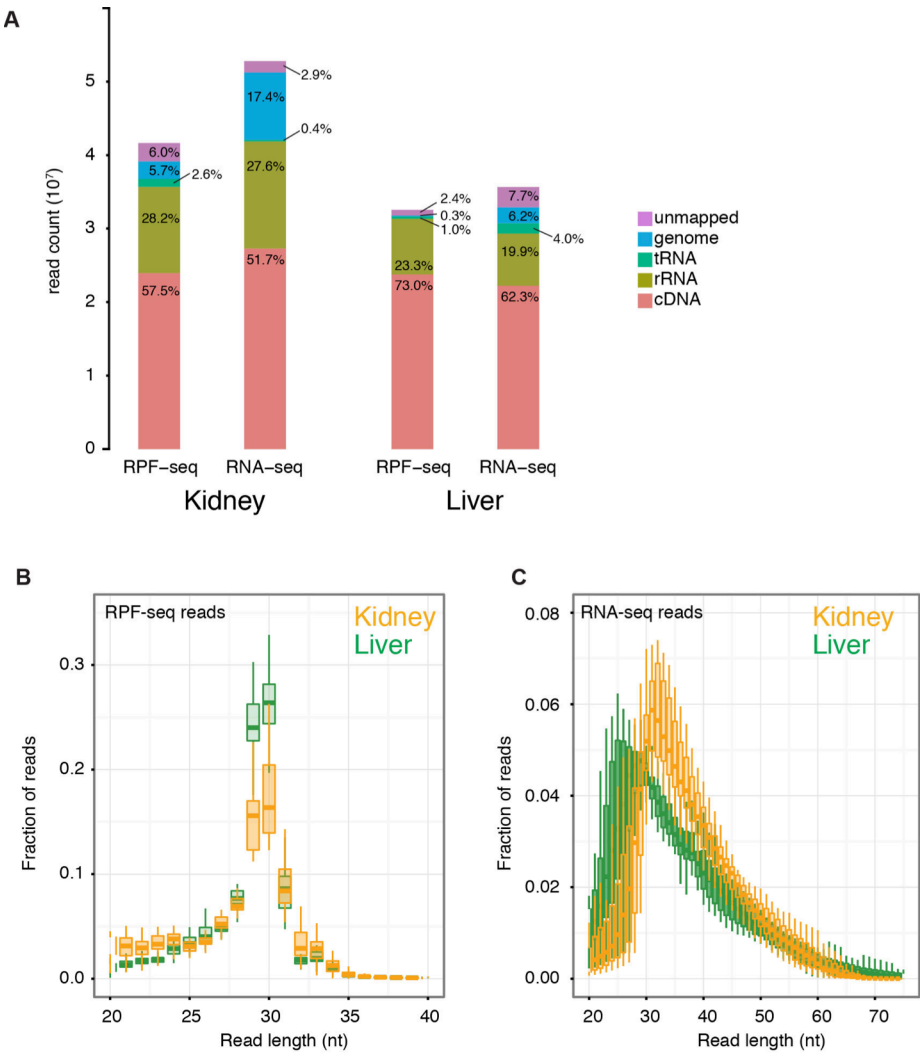
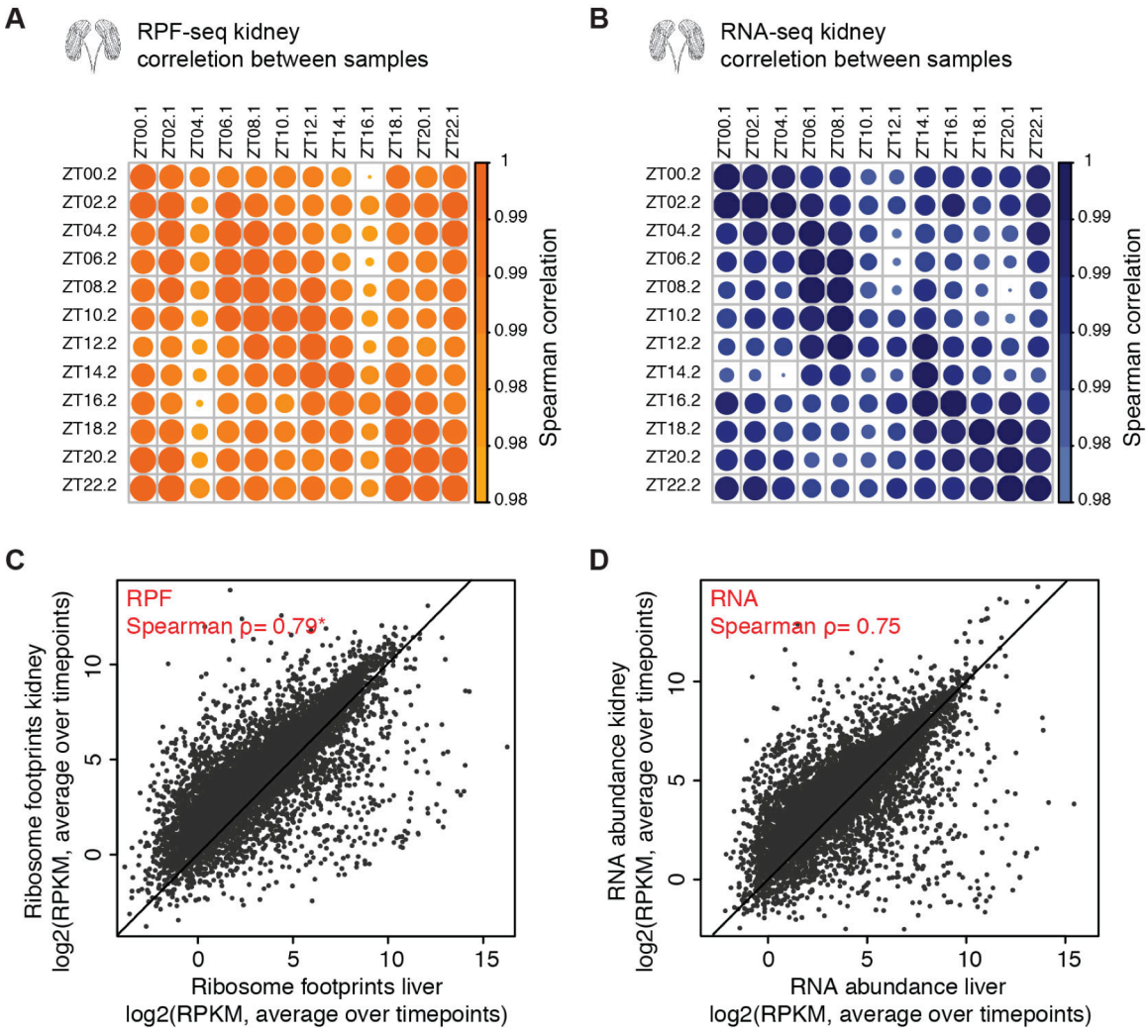


Figure S2. High reproducibility of the datasets.

(A, B) Spearman correlation of normalized CDS read counts between timepoints and between replicates for kidney RPF-seq (A) and kidney RNA-seq (B) datasets. Correlation coefficient is indicated by the size and intensity of the disks.

(C, D) Normalized CDS read counts (RPKM) in liver vs. kidney at the RNA-seq (C) and RPF-seq level (D), plotted as averages over all timepoints. Note the overall higher Spearman correlation at the footprint level than mRNA level.



* correlations are significantly different
Steiger test for difference between 2 dependent correlations, Z value=-22.23, p=8.7e-110 (n=10289)

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87, 245-251.

Figure S3. Principal component analysis across datasets.

(A) Scree plot of the first 10 principal components of the PCA in Fig. 1D. Components 4 and 5 explained a closely similar proportion of variation (between 5-6%) and a plateau was reached from component 6.

(B) Principal component (PC) 3 vs. PC4 of the PCA. PC4 clustered RPF-seq data of one organ with RNA-seq data of the other, hinting towards translational buffering as a relevant mechanism impinging on variability in gene expression across organs. Each dot represents a timepoint sample, replicates are joined by a line and timepoints within each dataset are sequentially colored.

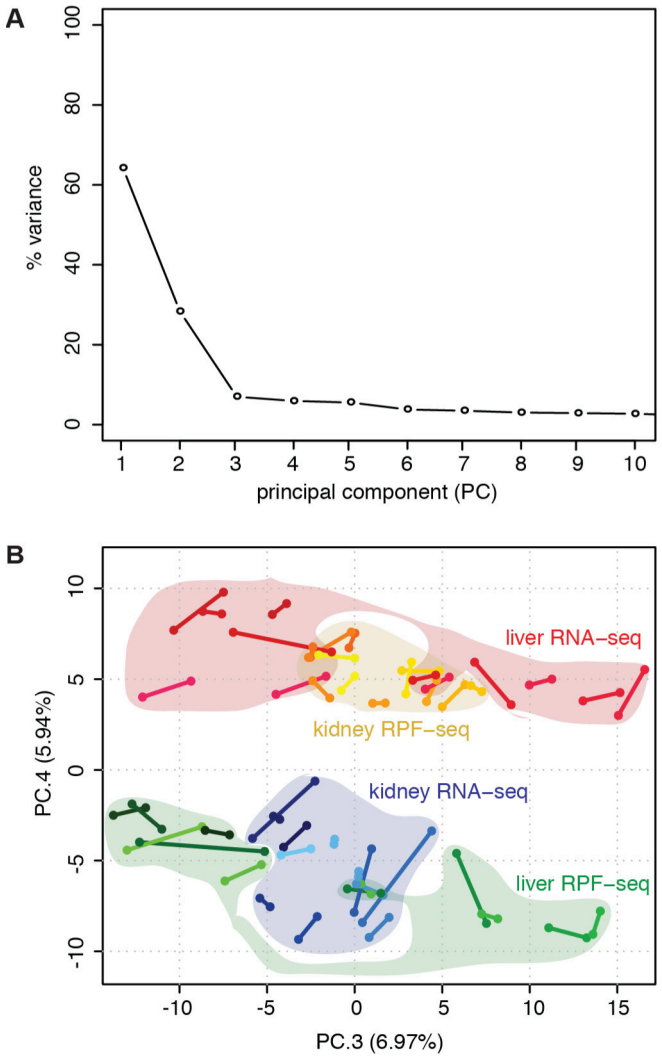
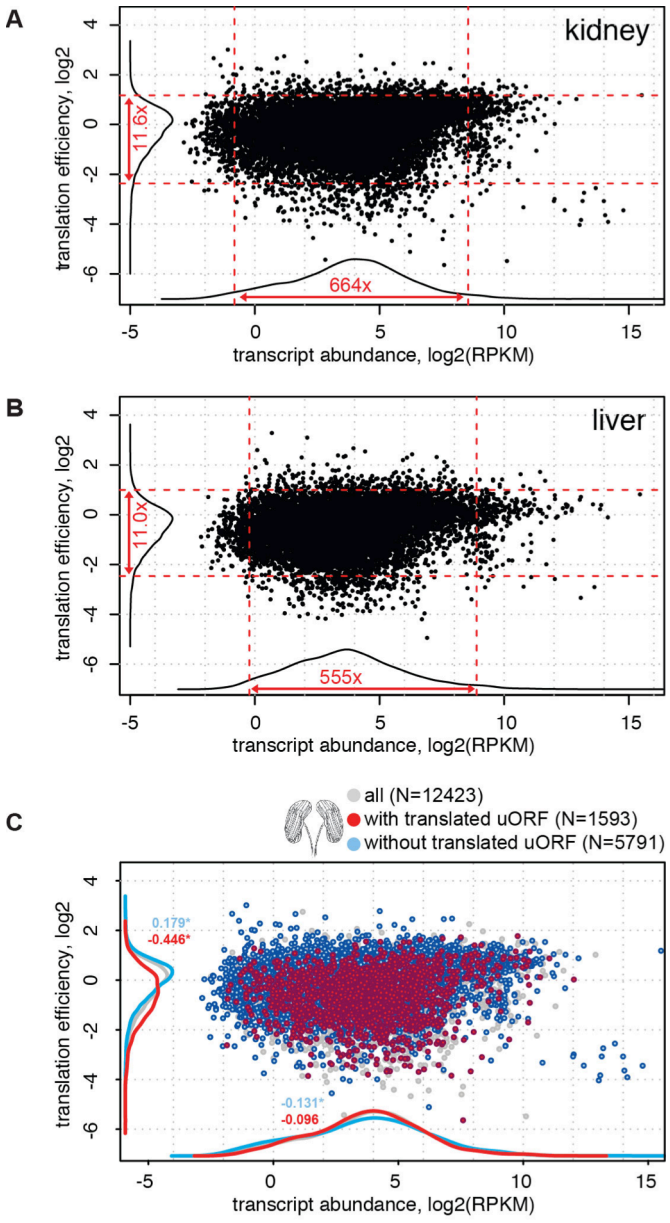


Figure S4. Contribution of translation efficiency to protein biosynthesis within organs.

(A, B) Scatterplot of mRNA abundance vs. translation efficiency (TE) in kidney (A) and liver (B) averaged over all timepoints (N=12423 and N=10676 genes, respectively). Corresponding density lines are plotted on the margins. Dotted red lines represent the 2.5 and 97.5 percentiles of each variable, and the corresponding fold-change is indicated. Transcript abundance range for 95% of genes spanned 2-3 orders of magnitude, whereas TE dynamic range was less than 12-fold in either organ, suggesting the former as the main contributor to gene expression variability within tissues.

(C) Scatterplot of mRNA abundance (TA) vs. translation efficiency (TE) in kidney, averaged over all timepoints. Highlighted are single protein-coding genes that do (red) or do not (blue) contain AUG-initiated translated uORFs. Corresponding density lines are plotted on the margins. uORF translation is associated with significantly reduced translation efficiency, but not with a strong effect on mRNA level. Numbers on the density curves indicate the location shift (difference of the distribution



54 medians) relative to all transcripts. Genes with translated uORFs: TA, p=0.16; TE, p<2.2e-16
55 (Wilcoxon rank sum test). Genes without translated uORFs: TA, p=8.7e-05; TE, p<2.2e-16
56 (Wilcoxon rank sum test).

Figure S5. Phase and amplitude relationships for RNA and RPF in both organs.

(A) Scatterplot of phase differences at the RNA (x-axis) vs. the RPF (y-axis) levels between kidney and liver for the genes rhythmic throughout (n=178). Clock genes are highlighted in red. The blue diagonal line represents the linear regression. Note that the phase concordance between organs is larger at the footprint level than at the mRNA level (i.e., the blue line is less steep than the grey dashed line that marks he 1:1 relationship).

(B, C) Histogram of the phase differences (footprints to mRNA abundance, in hours) in kidney (B) and liver (C) for the 178 genes rhythmic in both organs. Note the broader distribution of phase differences in kidney and an average phase advance of RPF with respect to RNA (0.143 hours) as opposed to an overall more synchronous phase of RPF and RNA in liver. See also Fig. 3B, E.

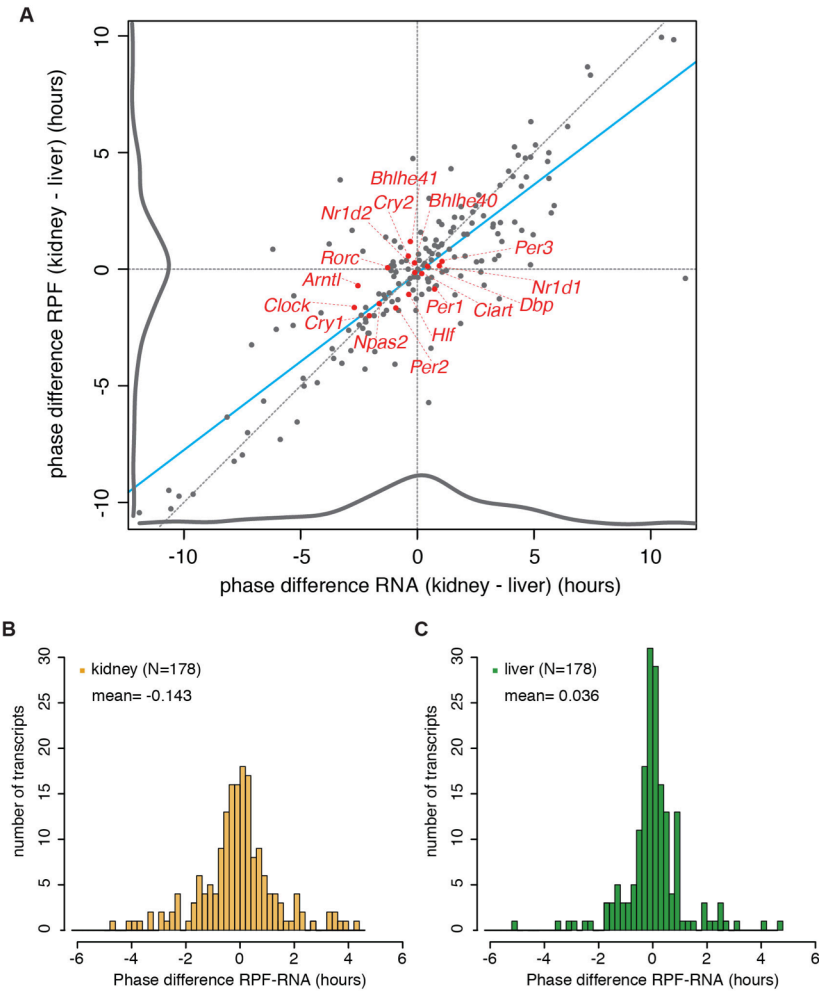
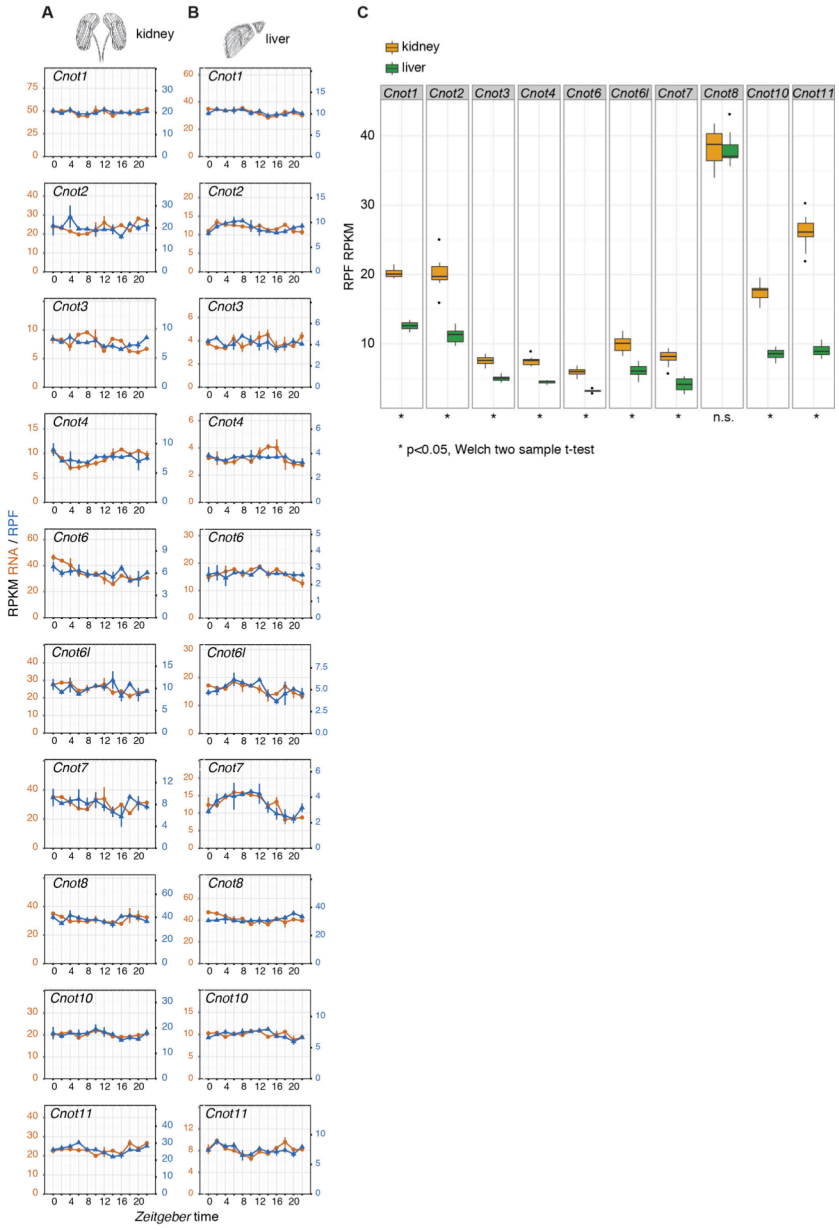


Figure S6. Higher expression of deadenylase complex subunits in kidney.



(A, B) Daily expression profiles of the CCR4-NOT complex components in kidney (A) and liver (B). Note that for all but one subunit, the daily expression is higher in kidney than in liver.

(C) RPF expression of the CCR4-NOT subunits averaged over the day. Boxplots represent the interquartile range and whiskers extend to the minimum and maximum TE within 1.5 times the interquartile range. Note that differences in protein biosynthesis are statistically significant for all but one of the subunits ($p<0.05$, two-sample t-test).

Figure S7. Heatmaps of all detected RPF and RNA rhythms indicates false-negatives of rhythmicity detection method.

(A) Same as Fig. 3A, but re-plotted here for ease of comparison with (B-D).

(B-D) Heatmap of RNA-seq (left panels) and RPF-seq (right panels) expression in kidney for genes detected as rhythmic only at the mRNA level (B, n=796), at both levels (C, n=542), and for the ribosome footprints only (D, n=435). Gene expression levels are standardized by row (gene). Even the panels that should represent "non-rhythmicity" (i.e. right panel in B and left panel in D) clearly show underlying rhythmicity, albeit with more noise and/or lower amplitude. Many of these cases are therefore probably not truly "non-rhythmic", indicating that the conventional rhythmicity detection s prone to false negatives, as described in the main text.

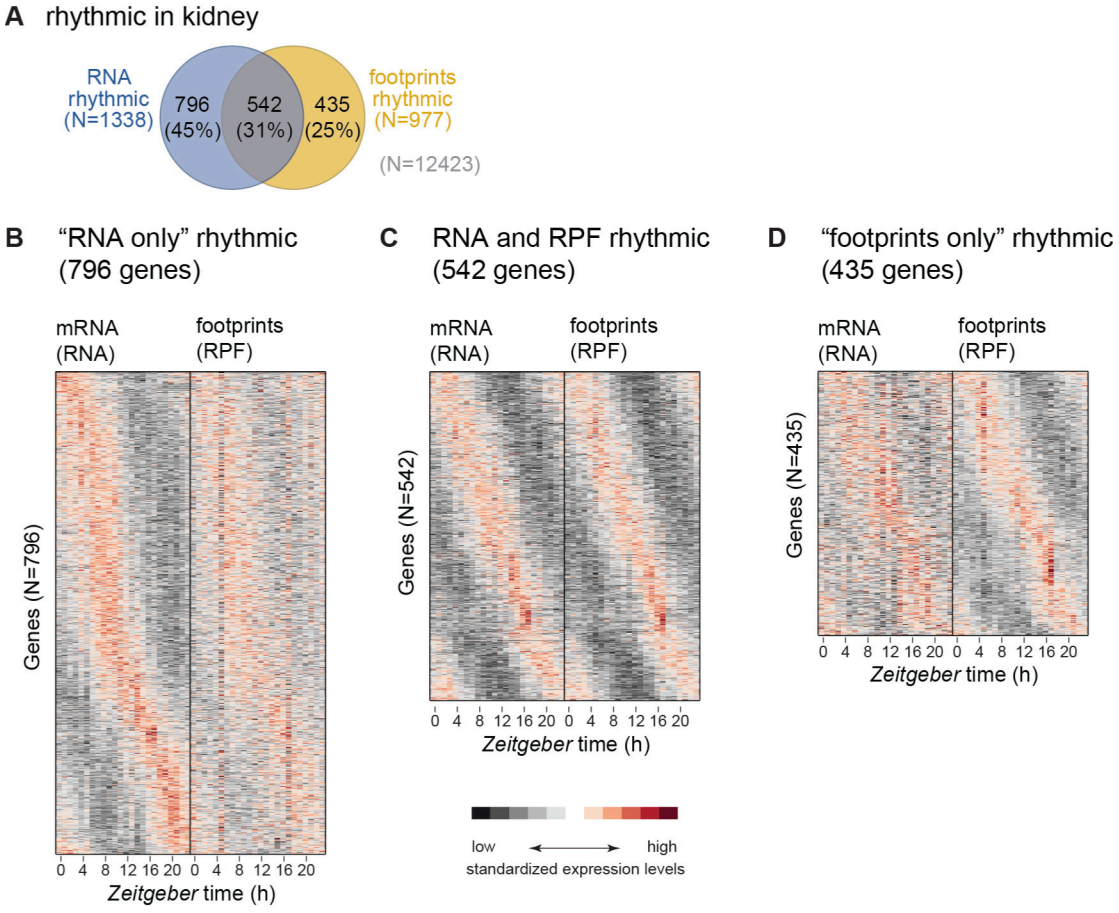


Figure S8. Core clock gene expression at RNA and RPF levels in both organs.

Left panels: Daily expression profiles of the main core clock genes shown in Fig. 5A-D.

Right panels: Hierarchical clustering of the organs' RNA and RPF profiles for each clock gene. Branch height represents the average Euclidean distance. Note that for 7 out of the 12 core clock genes, protein synthesis profiles are more similar across organs than mRNA abundance and than RPF-RNA within organs.