scientific reports

Check for updates

OPEN

Genome-wide correlation analysis to identify amplitude regulators of circadian transcriptome output

Evan S. Littleton, Madison L. Childress, Michaela L. Gosting, Ayana N. Jackson & Shihoko Kojima✉

Cell-autonomous circadian system, consisting of core clock genes, generates near 24-h rhythms and regulates the downstream rhythmic gene expression. While it has become clear that the percentage of rhythmic genes varies among mouse tissues, it remains unclear how this variation can be generated, particularly when the clock machinery is nearly identical in all tissues. In this study, we sought to characterize circadian transcriptome datasets that are publicly available and identify the critical component(s) involved in creating this variation. We found that the relative amplitude of 13 genes and the average level of 197 genes correlated with the percentage of cycling genes. Of those, the correlation of *Rorc* in both relative amplitude and the average level was one of the strongest. In addition, the level of *Per2AS*, a novel non-coding transcript that is expressed at the *Period 2* locus, was also linearly correlated, although with a much lesser degree compared to *Rorc*. Overall, our study provides insight into how the variation in the percentage of clock-controlled genes can be generated in mouse tissues and suggests that *Rorc* and potentially *Per2AS* are involved in regulating the amplitude of circadian transcriptome output.

Circadian clocks regulate the daily fluctuations of biochemical, physiological, and behavioral rhythms[1]. In mammals, signals originating in the suprachiasmatic nucleus (SCN) of the hypothalamus synchronize independent oscillators in other peripheral tissues, such as the brain and even in fibroblasts[2,3]. The molecular circadian clock within each cell is comprised of interlocking transcriptional-translational feedback loops, whose coordinated action is essential to generating cell-autonomous circadian oscillation[4].

At its core mechanism, the BMAL1 (official gene name: *Arntl*) and CLOCK (or its paralogue NPAS2) form a heterodimer and activate the transcription of *Period (Per) 1–3* and *Cryptochrome* (*Cry)1–2*, whose promoters contain target DNA regulatory elements, called E-boxes. As the level of PER and CRY proteins increases, they form a heterodimer and translocate back to the nucleus to repress their own transcription. As repression of *Per* and *Cry* transcription progresses, the level of the PER/CRY protein decreases, thereby allowing ARNTL and CLOCK to begin a new cycle of transcription. As an auxiliary loop, ARNTL and CLOCK also activate the expression of *Rev-erba/b* (official gene name: *Nr1d1* and *Nr1d2*), and *Rora-c*, all of which are nuclear receptors. REV-ERB and ROR proteins, in turn, repress or activate the target mRNA expression including *Arntl*, *Clock,* and *Npas2*, respectively, by recognizing DNA elements termed REV-ERB/ROR binding motifs (ROREs) in their promoters. As an additional loop, ARNTL/CLOCK activates the expression of *Dbp,* which activates the transcription of target mRNAs that possess a DNA element, called a D-box, while REV/ROR proteins regulate the expression of *Nfil3,* which represses D-box containing genes. Targets include *Rev-erbs, Rors*, and *Pers*[4].

Cell-autonomous circadian clocks also drive thousands of rhythmic output genes (i.e., clock-controlled genes) that, ultimately, produce daily rhythms of many types of physiology and behavior[5–10]. Interestingly, the number of cycling genes is vastly different among mouse tissues. In some tissues, more than 10% of the entire transcriptome is rhythmic, while only a few percent are rhythmic in other tissues[10–12]. Nevertheless, it remains unclear how the core clock machinery drives different numbers of clock-controlled genes, even though the core clock mechanism is nearly identical in each tissue.

To gain mechanistic insights into how some tissues produce more cycling genes than others, we characterized the circadian transcriptome data from various mouse tissues and attempted to identify a parameter(s) that correlates with the percentage of clock-controlled genes. We found that the differences in the percentage of cycling genes are not due to the difference in the overall gene expression level in each tissue. Interestingly, however, the relative amplitude of 13 genes as well as the average level of 197 genes correlated with the percentage of cycling

Department of Biological Sciences, Fralin Life Sciences Institute, Virginia Tech, Blacksburg, VA, USA. ✉email: skojima@vt.edu

genes. Of our particular interest was *Rorc,* whose correlation in both the relative amplitude and the average level was one of the strongest. We also found that the level of *Per2AS*, a novel non-coding transcript that is expressed at the *Period 2* locus, also showed a correlation. Based on these data, we propose that *Rorc* is involved in regulating the amplitude of circadian transcriptome output.

Methods

Microarray data processing. Microarray data were downloaded through NCBI GEO from series GSE54650[10]. Data was originated from 12 different tissues, with 24 time points from each tissue in 2-h intervals over the course of 48 h. Extracted data was normalized by Robust Multichip Average (RMA) normalization[13] and annotated by the Affymetrix Transcriptome Analysis Software package (http://www.affymetrix.com/support/technical/byproduct.affx?product=tac). Unannotated probesets, as well as those that had values lower than the average of all negative probesets across all timepoints in the respective tissue, were eliminated from the downstream analysis. For multiple probesets annotated to the same gene, the probeset with the highest average value was selected.

RNA-seq data processing. Mouse RNA-seq data were downloaded as fastq files through the NCBI database from SRA ID SRP036186[10]. Data contained information from 12 different tissues, with 8 time points from each tissue in 6-h intervals over the course of 48 h. Reads were mapped to the Ensembl mouse genome release 95 using STAR 2.7.0a[14] with outFilterScoreMinOverLRead = 0.3 and outFilterMatchNMinOverLRead = 0.3 options. We added the option 'Condensegenes' to select the most abundant isoform as the representative of a gene, as well as the option 'count exons' to measure only mRNA. Baboon circadian transcriptomic datasets were downloaded from series GSE54650[12]. Reads were mapped to the Ensembl baboon genome (*Papio Anubis* 2.0) release 90 using STAR 2.7.2b[14] with outFilterScoreMinOverLRead = 0.3 and outFilterMatchNMinOverLRead = 0.3 options. The quantification of expression level was performed by HOMER[15] using the transcripts per million (TPM) option. Any transcript with an average TPM < 0.5 across all timepoints were eliminated from the downstream rhythmicity analysis. We also used TPM to normalize the expression levels of each transcript. We eliminated white adipose data from the downstream analysis because no transcripts were rhythmic with our statistical threshold (BH. Q-value < 0.05), even though more than 13,500 transcripts were detected after applying the filter of TPM > 0.5. The expression of *Per2AS* was measured with the "strand-" option in HOMER. We did not apply the filter (TPM > 0.5 to call 'expressed') in quantifying the level of *Per2AS*, because non-coding transcripts generally have low expression levels[16–18].

Rhythmicity analysis. We used MetaCycle[19] to determine the rhythmicity of each gene. MetaCycle integrates three different algorithms ARSER, JTK_CYCLE, and Lomb-Scargle and calculates the p-value, Benjamini–Hochberg q-value (BH.Q value), period, phase, baseline value, amplitude (AMP), and relative amplitude (rAMP), which is the ratio between amplitude and baseline expression level. We defined the expression rhythmic when meta2d BH.Q < 0.05.

Correlation analysis. Pearson and Spearman correlation tests were performed in R to determine the linear and non-linear correlation between the percentage of cycling transcripts in each tissue and the rhythmicity (using BH.Q value), phase, and relative amplitude of the 15 clock genes, as well as *Per2AS*, calculated by the MetaCycle package in R. A significant correlation was defined as a p-value < 0.05. For the transcriptome-wide correlation analysis of mouse and baboon data, we used the rcorr function from the Hmisc and tidyverse packages in R to perform Pearson or Spearman correlation tests and used the average gene expression of transcripts expressed in all 12 (microarray) or 11 tissues, excluding white adipose tissue (RNA-seq)[20,21]. Rat data expression values were Log2 normalized, and we then used JMP to perform Pearson or Spearman correlation tests. Fisher Z-scores were calculated from the Rho or R^2 with Fisher transformation. GO enrichment analysis of the significantly correlated genes was performed using the Gene Ontology Resource[22,23].

Promoter motif analysis. We first retrieved the promoter sequences (−1000 to +100 bp with respect to the transcription start site: TSS) of all the cycling transcripts from the UCSC Genome Browser, and performed a motif search using Find Individual Motif Occurrence (FIMO) with a p-value = 1×10^{-4} as the threshold[24]. Input motif matrices were downloaded from JASPER (RORA: MA0071.1, RORA(var.2): MA0072.1, RORB: MA1150.1, RORC: MA1151.1, NR1D1: MA1531.1, NR1D2: MA1532.1, ANRTL: MA0603.1, CLOCK: MA0819.1, NPAS2: MA0626.1, NFIL3: MA0025.1, and MA0025.2, DBP: MA0639.1)[25].

Single value decomposition analysis. Matrix of clock gene expression values in 12 tissues (11 tissues for RNA-seq) were input into the la.svd function in R. Returned eigengenes 1 and 2 were projected onto expression values. Returned eigentissues 1 and 2 were also projected onto expression values.

Cell culture. *Per2::LucSV* (a gift from Dr. Yoo at the University of Texas Health Science Center at Houston), NIH3T3/*Dbp*-luc, or NIH3T3/*Bmal1*-luc cells (a gift from Dr. Schibler, University of Geneva) were grown in Dulbecco's Modified Eagle Medium (DMEM) (Life Tech) with 10% fetal bovine serum (FBS) (ATLANTA biologicals) at 37 °C with 5% CO_2.

Real-time bioluminescence and gene expression measurements. Cells were grown in 35 mm dishes until confluent. Samples used for gene expression analysis were circadian synchronized by DMEM sup-

plemented with 50% horse serum for 2 h. Cells were then treated with either 45 μM nobiletin (Toronto Research Chemicals), 10 μM SR1001 (Sigma), or DMSO (Thermo Fisher Scientific). 24 h after treatment, samples were collected every 4 h, and total RNA was extracted with TRIZOL reagent (Life Tech) according to the manufacturer's instructions. RNAs were treated with TURBO DNaseI (Life Tech), then subjected to reverse transcription using High Capacity cDNA Reverse Transcription Kits (Applied Biosystems). qPCR was performed using QuantStudio 6 (Life Tech) with SYBR Power Green (Applied Biosystems). Primers used in this study are as follows: 36B4_Fwd 5′-CACTGGTCTAGGACCCGAGAAG-3′, 36B4_Rev 5′-GGTGCCTCTGAAGATTTTCG-3′, Rora_Fwd 5′-ACCGTGTCCATGGCAGAAC-3′, Rora_Rev 5′-TTTCCAGGTGGGATTTGGAT-3′, Rorc_Fwd 5′-TCTACACGGCCCTGGTTCT-3′, Rorc_Rev 5′-ATGTTCCACTCTCCTCTTCTCTTG-3′. 1 μM dexamethasone was added to media for 2 h in samples used for bioluminescence recordings. After 2-h dexamethasone treatment, media was changed to phenol red-free DMEM (Cellgro 90-013-PB) supplemented with 100 μM luciferrin, 10 mM HEPES pH 7.2, 1 mM sodium pyruvate, 0.035% sodium bicarbonate, 2% FBS, 1 × Penicillin/Streptomycin, and 2 mM L-glutamine. Samples for real-time bioluminescence recordings were treated with various concentrations of nobiletin (1.5, 5, 15, or 45 μM) or SR1001 (1, 5, or 10 μM), and measurements were performed for 7 days using a LumiCycle (Actimetrics, Inc. Wilmette, IL). The first 24 h were removed from measurements to quantify amplitude using JMP software[26].

Results

Characteristics of circadian transcriptomic output in various mouse tissues. To gain insight into what determines the number of clock-controlled genes in each tissue, we first retrieved existing circadian transcriptome datasets from various mouse tissues[10]. We found this particular dataset best-suited to our study, because it covered the highest number of tissues (12 total) and provided the highest time resolution (2 h intervals), compared to other studies[5,7–9].

Our in-house analysis was able to replicate the previous findings, in which the percentage of cycling genes was highest in liver, followed by kidney, lung, brown adipose, and heart, and lowest in brainstem (Fig. 1A, Supplemental Data Files 1a and 1b). The ranks are slightly different from the original study[10], which is most likely due to the differences in the analytical methods and statistical criteria used in our study (see "Methods"). Distribution of Benjamini–Hochberg q-values from the rhythmicity analysis was also widest in liver, followed by kidney, lung, brown adipose, and heart, and was particularly narrow in white adipose and brainstem (Fig. 1B). These data indicate that the expression levels of transcripts in liver are most variable across different times of the day compared to other tissues such as white adipose or brainstem, regardless of their rhythmicity.

The difference in the percentage of rhythmic transcripts in each tissue was not due to the difference in the number of genes expressed, because the total number of transcripts detected in each tissue was comparable, irrespective of the percentage of rhythmic transcripts (Fig. 1C). It was not due to the median microarray signals per transcript in each tissue either, as it did not correlate with the percentage of rhythmic transcripts (Pearson $r^2 = 0.115$, $p = 0.368$; Spearman $rho = -0.042$, $p = 0.904$) (Fig. 1D). When we focused only on the transcripts that were rhythmically expressed (Benjamini–Hochberg $q < 0.05$), the median of the relative amplitude (i.e., the ratio between amplitude and baseline expression level) (Fig. 1E) or the amplitude itself (Fig. 1F) of cycling gene expression did not correlate with the percentage of cycling genes in each tissue (relative amplitude: Pearson $r^2 = 0.060$, $p = 0.189$; Spearman $rho = 0.070$, $p = 0.834$, amplitude: Pearson $r^2 = 0.439$, $p = 0.154$; Spearman $rho = 0.175$, $p = 0.588$), indicating that the amplitude of gene expression is comparable in each tissue if they are rhythmic.

We also performed the same set of analyses using the RNA-seq data[10], which surveyed the same set of tissues but with a lower time resolution (Microarray: 2 h, RNA-seq: 6 h)[10]. Even though the order of the tissues was slightly different than the microarray datasets (Fig. S1A), which was likely due to the differences in time resolution and the threshold to eliminate low-expressed transcripts (see "Methods"), the results were essentially the same: distribution of Benjamini–Hochberg q-values from the rhythmicity analysis was wider in tissues with a high percentage of rhythmic transcripts (Fig. S1B), the number of genes expressed was comparable among tissues (Fig. S1C, D), and the median of the relative amplitude or the amplitude of cycling gene expression was comparable in each tissue (Fig. S1E,F, Supplemental Data File 2).

Characterization of cycling gene expression in various mouse tissues. Because we did not observe any characteristics that had a correlation with the percentage of cycling genes at a genome-wide scale, we shifted our focus on single gene level analyses. We first analyzed a total of 15 core clock genes (*Arntl*, *Clock*, *Npas2*, *Per1-3*, *Cry1-2 Rora-c, Nr1d1-2*, *Dbp*, *Nfil3*), and found that most of these genes were expressed ubiquitously across all tissues, except for *Rorb,* whose expression was restricted to brain and brown adipose tissue (Supplemental Data File 3). A majority of these genes were also rhythmically expressed, except for the *Rors*: *Rora* was rhythmic in four tissues (liver, lung, heart, and muscle) but not in the other eight tissues (kidney, BAT, adrenal, aorta, cerebellum, hypothalamus, WAT, and brainstem), *Rorb* was arrhythmic in all four tissues that it is expressed in (BAT, hypothalamus, cerebellum, and brainstem), and *Rorc* was rhythmic in most of the peripheral tissues but not in the hypothalamus, muscle, or brainstem (Supplemental Data File 3), which was consistent with previous reports[7,9,27,28]. *Clock* and *Cry1* were also arrhythmic in hypothalamus, making the hypothalamus the tissue with the lowest number of rhythmic core clock genes, even though hypothalamus ranked 9th out of 15 tissues in the percentage of cycling transcripts (Fig. 1A). Similar results were obtained from the RNA-seq data, although the number of rhythmic core clock genes were lower, most likely due to the lower time resolution of the RNA-seq data compared to the microarray data (Supplemental Data File 4).

As was previously reported, the phases of core clock gene expression were confined to a relatively narrow window[12,28–31], except for a few genes such as *Cry2, Rorc,* and *Nfil3* (Fig. 2A). On the other hand, the relative

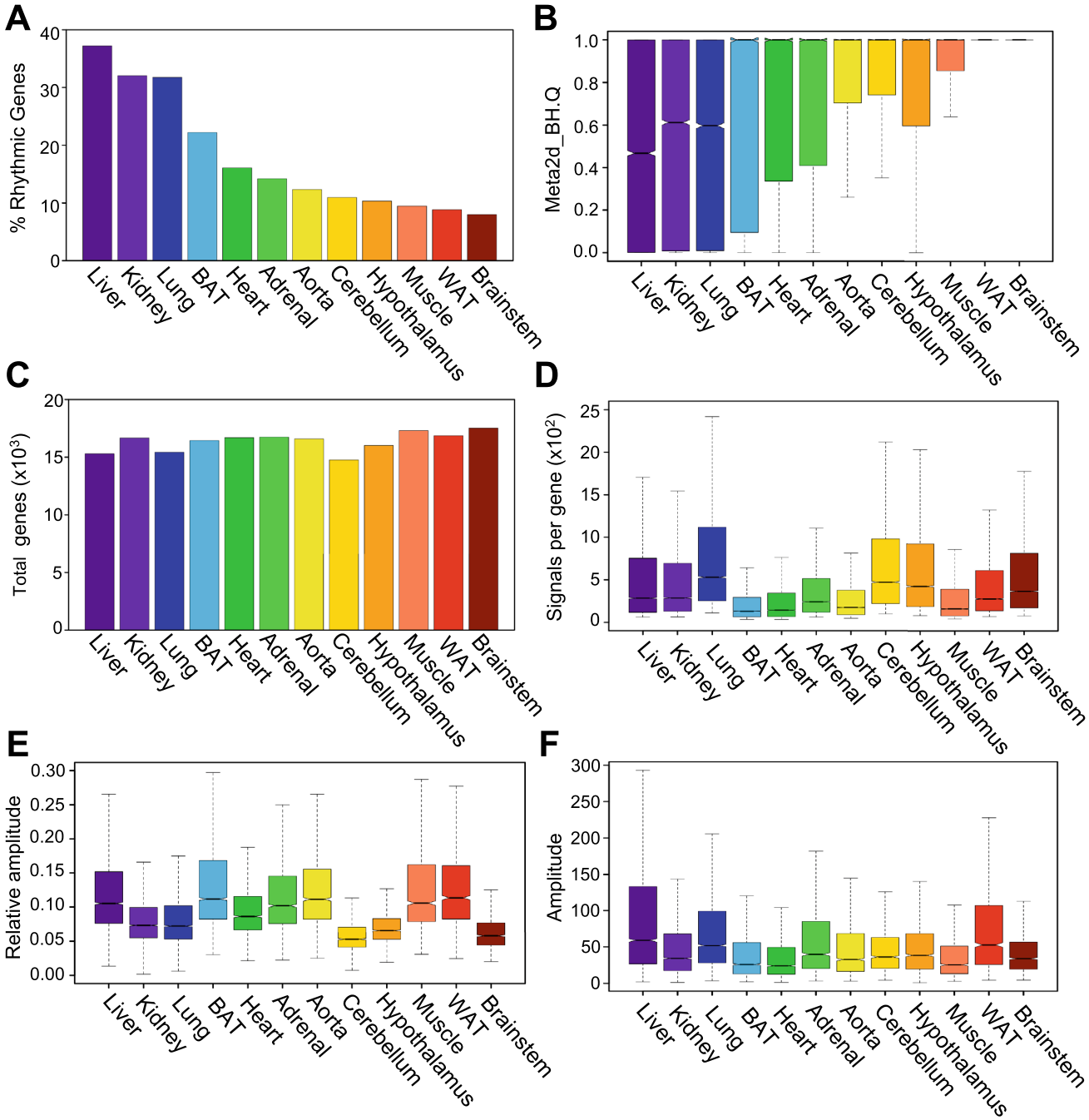

Figure 1. Characteristics of the mouse circadian transcriptome (microarray) in various mouse tissues. (**A**) Percentage of cycling genes in each tissue from highest (left) to lowest (right) % cycling. Rhythmicity of a gene was defined as Benjamini–Hochberg q values < 0.05 by MetaCycle. (**B**) Distribution of Benjamini–Hochberg q values of all expressed genes in each tissue. (**C**) Numbers of genes expressed in each tissue. (**D**) Average microarray signals per gene for all the probesets. (**E**) Distribution of relative amplitude of cycling genes in each tissue calculated by MetaCycle. (**F**) Distribution of the amplitude of cycling genes in each tissue calculated by MetaCycle. (**D**–**F**) The central line represents the median, and each box represents the 25th and 75th percentiles, respectively. The notch represents the 95% confidence interval around the median. Numbers of expressed genes or rhythmic genes in each tissue can be found in the Supplementary Data File 1. Each color corresponds to a tissue; liver (purple), kidney (light purple), lung (blue), brown adipose (BAT) (light blue), heart (green), adrenal (light green), aorta (yellow), cerebellum (gold), hypothalamus (orange), muscle (coral), white adipose (WAT) (red), and brainstem (dark red).

amplitude of core clock gene expression was more variable between tissues (Fig. 2B), and nine genes, *Dbp, Npas2, Nr1d1, Arntl, Per3, Per2, Rorc, Cry1, and Cry2* had their relative amplitude positively correlated with the

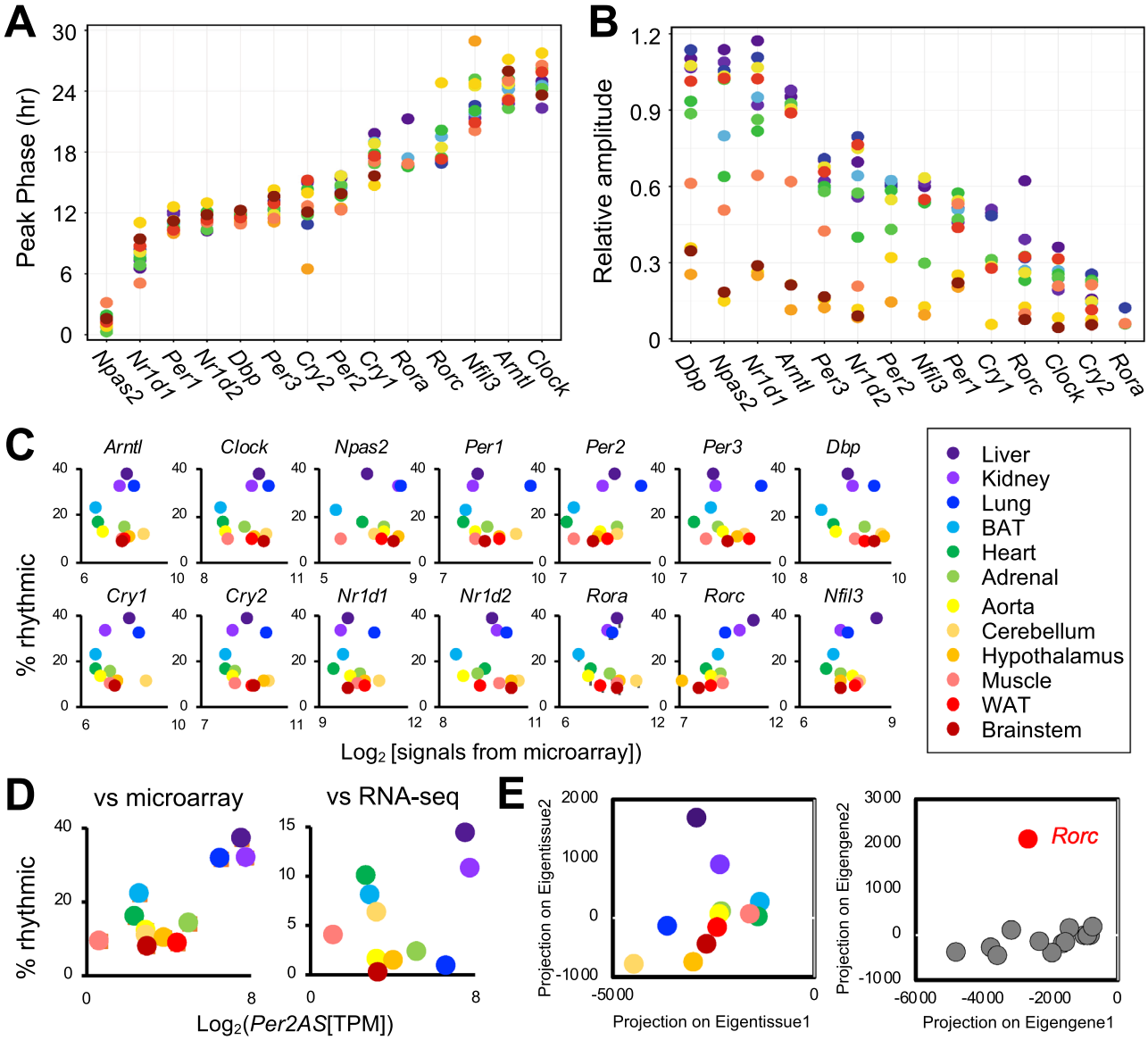

Figure 2. A positive correlation between the percentage of cycling transcripts and the mean level of *Rorc*. (**A**) The peak phase of core clock gene expression in CT (Circadian Time) determined by MetaCycle. (**B**) The relative amplitudes of core clock gene expression determined by MetaCycle. (**C**) Correlation between % rhythmic transcripts and the mean levels of each core clock gene determined by MetaCycle. (**D**) Correlation between % rhythmic transcripts and the mean levels of *Per2AS* in each tissue. Each color corresponds to a tissue; liver (purple), kidney (light purple), lung (blue), brown adipose (light blue), heart (green), adrenal (light green), aorta (yellow), cerebellum (gold), hypothalamus (orange), muscle (coral), white adipose (red), brainstem (dark red). Core clock gene expression that did not fulfill Benjamini–Hochberg $q < 0.05$ criteria for rhythmicity was not included. (**E**) Single value decomposition (SVD) of tissues (left) or core clock genes (right). Eigentissues and eigengenes 1 and 2 were projected onto clock gene expression values, *Rorc* indicated in red.

percentages of cycling transcripts in either Pearson and Spearman correlation analyses (Table 1). Additional ten clock-controlled genes were expressed and rhythmic in all tissues, for which we calculated the correlation between their relative amplitude and the percentage of cycling transcripts. Among those, the relative amplitude of three genes (*P4ha1*, *Tsc22d3*, and *Lonrf3*), or another set of three genes (*Tsc22d3*, *Lonrf3*, and *Usp2*) correlated significantly with the percentage of cycling transcripts in Spearman or Pearson analyses, respectively (Table 1). Notably, the strongest correlation was observed for *Rorc* (Pearson $r^2 = 0.852$, $p = 0.004$; Spearman *rho* = 0.917, $p = 0.001$). In addition, the number of rhythmic core clock genes also correlates with the percentages of cycling transcripts in Spearman analysis (*rho* = 0.728, $p = 0.007$), but not in Pearson analysis ($r^2 = 0.510$, $p = 0.0906$). The sum of relative amplitudes from rhythmic core clock genes also correlates with the percentages of cycling transcripts (Pearson $r^2 = 0.653$, $p = 0.0212$; Spearman *rho* = 0.734, $p = 0.0091$). It is unclear, however, whether the higher amplitude of core clock gene expression leads to a higher percentage of rhythmic transcripts in each tissue or vice versa. We also analyzed the data from RNA-seq and performed the same analyses. However, the lower

	Pearson (linear)		Spearman (non-linear)	
	R^2	p-value	*Rho*	p-value
Arntl	0.5775568	0.04923*	0.8741259	0.0003089*
Clock	0.3962219	0.2277	0.3181818	0.3414
Npas2	0.6054532	0.04839*	0.7181818	0.0168*
Per1	0.4538226	0.1384	0.4685315	0.2175
Per2	0.6617942	0.01907*	0.7972028	0.0031161*
Per3	0.5899605	0.04347*	0.5944056	0.04575*
Cry1	0.7952259	0.003433*	0.7090909	0.01873*
Cry2	0.5342694	0.07355	0.6853147	0.01731*
Nr1d1	0.6248975	0.02981*	0.5944056	0.04575*
Nr1d2	0.5542365	0.06149	0.4405594	0.1542
Rora	0.3681567	0.5421	0.2	0.7833
Rorc	0.8522986	0.003518*	0.9166667	0.001312*
Dbp	0.6452423	0.02346*	0.6853147	0.01731*
Nfil3	0.4639099	0.1506	0.3272727	0.327
P4ha1	0.5738972	0.05103	0.6923077	0.01588*
Tsc22d3	0.677488	0.01549*	0.6293706	0.03239*
Lonrf3	0.74423	0.005506*	0.7972028	0.003161*
Usp2	0.5768827	0.04956*	0.4685315	0.1275

Table 1. Correlations between the percentage of rhythmic transcripts and the relative amplitude of core clock genes in each tissue. *Asterisks denote $p < 0.05$. [a]*Rorb* was excluded from our correlation analyses due to its low expression in all tissues except brain.

number of rhythmic core clock genes detected in the RNA-seq dataset significantly compromised our ability to calculate the correlations between the percentage of rhythmic transcripts in each tissue and the phase and amplitude of core clock gene expression.

We also investigated the correlation between the mean levels of core clock gene expression across all time points and the percentage of cycling genes in each tissue, as non-cycling genes could contribute to the differences in the percentage of cycling genes. Interestingly, we again found that there was a positive correlation between the percentage of cycling transcripts and the mean level of *Rorc,* but not with any other core clock genes (Fig. 2C, Table 2). This correlation was also observed in the RNA-seq dataset (Fig. S2, Table 2). We also tested *Per2AS*, a newly identified non-coding RNA[32–34], because the expression of *Per2AS* is rhythmic and antiphasic to *Per2* in liver, adrenal gland, lung, and kidney[10], and it was suspected that *Per2AS* is involved in regulating the circadian system[35]. Indeed, we found a linear correlation between the mean levels of *Per2AS* and the percentages of rhythmic transcripts in both microarray and RNA-seq datasets (Fig. 2D). Single value decomposition (SVD) analysis did not detect any clear clustering patterns for tissues (Fig. 2E). In contrast, all the clock genes were clustered together except for *Rorc* (Fig. 2E, Fig. S2B), indicating that the correlation between *Rorc* and the percentage of rhythmic transcripts is not due to innate differences between the tissues. Rather, it is most likely due to the difference in expression patterns of *Rorc* between tissues that are distinct from other clock genes.

To test how robust the correlation of *Rorc* is, we extended the analysis to the genome-wide scale. We found that among 12,024 genes expressed in all 12 tissues from microarray datasets, the mean level of 1131 and 400 genes was correlated significantly with the percentage of rhythmic genes in each tissue from the Pearson (linear) and the Spearman (non-linear) correlation test, respectively (Supplemental Data File 5). Similarly, among 8269 genes expressed in all 11 tissues from the RNA-seq datasets, we found the mean level of 925 and 1664 genes correlated significantly with the percentage of rhythmic genes from the Pearson and Spearman correlation tests, respectively (Supplemental Data File 5). Of those, 135 (Pearson) or 77 (Spearman) genes were found correlated in both microarray and RNA-seq datasets (Supplemental Data File 5), and we therefore considered those as more robustly correlated. Gene ontology (GO) analyses were then performed to assess whether a specific process contributes to the high percentage of rhythmic transcripts. No pathways were detected as statistically significant ($FDR < 0.05$) among those that correlated robustly in the Spearman analysis. Whereas numerous metabolic processes were enriched among those that correlated robustly in the Pearson analysis (Supplemental Data File 6). We also calculated Fisher Z-scores from each test to evaluate the relative strength of *Rorc* correlation, compared to other genes. *Rorc* was ranked 5th (Pearson) or 14th (Spearman) when we used average Z-scores from both microarray and RNA-seq datasets. These data suggest that the correlation between the level of *Rorc* and the amplitude of the mouse circadian transcriptome is one of the strongest.

To further gain more mechanistic insights into how *Rorc* contributes to the increase in the number of cycling mRNAs without driving mRNA expression, we next tested the correlation between the mean levels of *Rorc* and other core clock genes. Not surprisingly, we found a linear correlation between the mean levels of *Rorc* and *Per2AS* in both the microarray and RNA-seq datasets (Table 3). The mean level of *Rorc* also linearly correlated with *Nfil3* (Microarray) or *Cry2* (RNA-seq) (Table 3); however, the biological significance of these correlations is

	Pearson (linear)				Spearman (non-linear)			
	Microarray		RNA-seq		Microarray		RNA-seq	
	R^2	p-value	R^2	p-value	*rho*	p-value	*rho*	p-value
Arntl	0.00025806	0.9993649	−0.2477924	0.46255259	−0.0629371	0.8459309	−0.3545455	0.28469274
Clock	0.20709174	0.5184007	−0.3077513	0.35722663	−0.020979	0.9484022	−0.3818182	0.24655958
Npas2	0.13268917	0.6810054	NA	NA	0.03496503	0.9140933	NA	NA
Per1	0.07460386	0.8177602	−0.4232656	0.19458479	−0.2797203	0.3785687	−0.4454545	0.1697326
Per2	0.4444646	0.1477156	−0.3037159	0.36389063	0.25174825	0.4299188	−0.1	0.769875
Per3	−0.004058	0.9900137	−0.5366997	0.08871493	−0.3006993	0.3422595	−0.6181818	0.04264557
Cry1	0.17502327	0.5863932	−0.1887732	0.57828653	−0.1258741	0.6966831	−0.4272727	0.18994372
Cry2	−0.0725671	0.8226649	−0.469578	0.14504423	−0.3636364	0.245265	−0.6454545	0.0319628
Nr1d1	−0.1931836	0.5474633	−0.4214839	0.19667007	−0.4055944	0.1908359	−0.5818182	0.0604199
Nr1d2	−0.1093403	0.7351637	−0.3684187	0.26490747	−0.3006993	0.3422595	−0.5090909	0.10973723
Rora	−0.154493	0.6316449	−0.0850269	0.8037002	−0.2097902	0.5128409	−0.4	0.22286835
Rorc	0.81551562	0.0012236*	0.80572369	0.0027537*	0.65734266	0.0201855*	0.8	0.0031104*
Dbp	−0.3043041	0.3362164	−0.5439689	0.08366	−0.4615385	0.1309481	−0.4454545	0.1697326
Nfil3	0.2775392	0.3824536	0.18797349	0.57992718	−0.1188811	0.7128842	0.4	0.22286835
Per2AS	0.8546421	0.0003983*	0.6079273	0.04723*	0.4545455	0.1404	0.07272727	0.8388

Table 2. Correlations between the percentage of rhythmic genes and the mean expression level of core clock genes in microarray and RNA-seq datasets. *Asterisks denote $p < 0.05$. [a]*Rorb* was excluded from our correlation analyses due to its low expression in all tissues except brain.

	Pearson (linear)				Spearman (non-linear)			
	Microarray		RNA-seq		Microarray		RNA-seq	
	R^2	p-value	R^2	p-value	*Rho*	p-value	*Rho*	p-value
Arntl	0.1424678	0.6587	−0.1930034	0.5696	0.2517483	0.4301	−0.3	0.3711
Clock	0.2017612	0.5295	−0.5247839	0.09743	0.2307692	0.4709	−0.5636364	0.07594
Npas2	−0.1087769	0.7341	−0.1029352	0.8262	0	1	−0.1071429	0.8397
Per1	−0.1271442	0.6938	−0.3823291	0.2459	0.02097902	0.9562	−0.5727273	0.0706
Per2	0.2802252	0.3777	0.03352763	0.922	0.5244755	0.08388	0	1
Per3	−1.93E−01	0.5478	−0.5847199	0.05884	−0.1328671	0.6834	−0.7818182	0.007012
Cry1	0.2039933	0.5248	−0.3352892	0.3135	0.2237762	0.4849	−0.5909091	0.06073
Cry2	−0.1406772	0.6628	−0.7071451	0.01495*	−0.020979	0.9562	−0.8	0.005202
Nr1d1	−0.1865557	0.5615	−0.2836711	0.3979	−0.013986	0.9737	−0.5636364	0.07594
Nr1d2	−0.075446	0.8157	−0.6073742	0.4749	−0.0489511	0.8863	−0.6818182	0.02548
Rora	0.08273569	0.7982	−0.2710108	0.4202	0.0979021	0.7663	−0.4545455	0.1634
Dbp	−0.2537079	0.4262	−0.2642677	0.4323	−0.1608392	0.6194	−0.4272727	0.1926
Nfil3	0.6783749	0.01531*	0.4541654	0.1605	0.4195804	0.1766	0.4727273	0.1456
Per2AS	0.858603	0.0003493*	0.6431022	0.0328*	0.5034965	0.09875	0.1454545	0.6734

Table 3. Correlations between the mean level of *Rorc* and the mean level of other clock genes in each tissue. *Asterisks denote $p < 0.05$. [a]*Rorb* was excluded from our correlation analyses due to its low expression in all tissues except brain.

unclear, as the correlations were not consistent between microarray and RNA-seq datasets. We did not detect any correlation between *Rorc* and *Rorc*-target genes such as *Arntl*, *Cry1*, *Nfil3*, and *Nr1d1* (Table 3), whose promoter regions contain RORE motifs and the amplitude of their rhythmic mRNA expression was dampened in most of the $Rorc^{-/-}$ tissues[36–39]. We did not find any correlations between the level of *Per2AS* and other core clock genes either, except for *Per2* and *Npas2* (Table 4). The significance of these correlations is also unclear, because they were found only in microarray but not in RNA-seq datasets.

The effect of RORC as a transcriptional activator in regulating the circadian transcriptome. Because *Rorc* directly activates the transcription of RORE-containing genes[27,29,40], we hypothesized that, if *Rorc* was directly driving rhythmic gene expression leading to a high number of cycling transcripts, then the number of rhythmic genes with RORE motifs in their promoter would be higher in tissues with a higher number of rhythmic genes. To test this, we first retrieved the promoter sequence of rhythmic genes in each tissue

	Pearson (linear)				Spearman (non-linear)			
	Microarray		RNA-seq		Microarray		RNA-seq	
	R^2	p-value	R^2	p-value	*Rho*	p-value	*Rho*	p-value
Arntl	0.1072332	0.7401	−0.1923029	0.5711	0.4405594	0.1542	0.4727273	0.1456
Clock	0.3049452	0.3351	−0.102975	0.7632	0.5804196	0.05209	0.5272727	0.1001
Npas2	0.308457	0.3292	−0.0209776	0.9644	0.6783217	0.01883*	0.5357143	0.2357
Per1	−0.0469417	0.8848	−0.3423037	0.3028	0.4125874	0.1845	0.0090909	0.9892
Per2	0.3363529	0.2851	0.03514764	0.9183	0.7202797	0.01102*	0.5636364	0.07594
Per3	−6.68E−02	0.8367	−0.2394287	0.4783	0.4265734	0.1689	−0.0454546	0.9029
Cry1	0.1160046	0.7196	−0.1523629	0.6547	0.4125874	0.1845	0.2090909	0.5391
Cry2	−0.0904382	0.7798	−0.360859	0.2756	0.2727273	0.3912	0.1090909	0.7549
Nr1d1	−0.2317755	0.4685	−0.4086626	0.2121	0.1258741	0.6997	−0.0545455	0.8815
Nr1d2	0.03833355	0.9058	−0.2832654	0.3986	0.1188811	0.7162	−0.0181818	0.9676
Rora	−0.027873	0.9315	−0.1883284	0.5792	0.2447552	0.4435	0.1363636	0.6935
Dbp	−0.1331303	0.68	−0.2910061	0.3853	0.2937063	0.3543	0.08181818	0.8177
Nfil3	0.357542	0.2539	−0.1930106	0.5696	0.1468531	0.6511	−0.3909091	0.2365

Table 4. Correlations between the mean *Per2AS* TPM and the mean level of other clock genes in each tissue. *Asterisks denote $p < 0.05$. [a]*Rorb* was excluded from our correlation analyses due to its low expression in all tissues except brain.

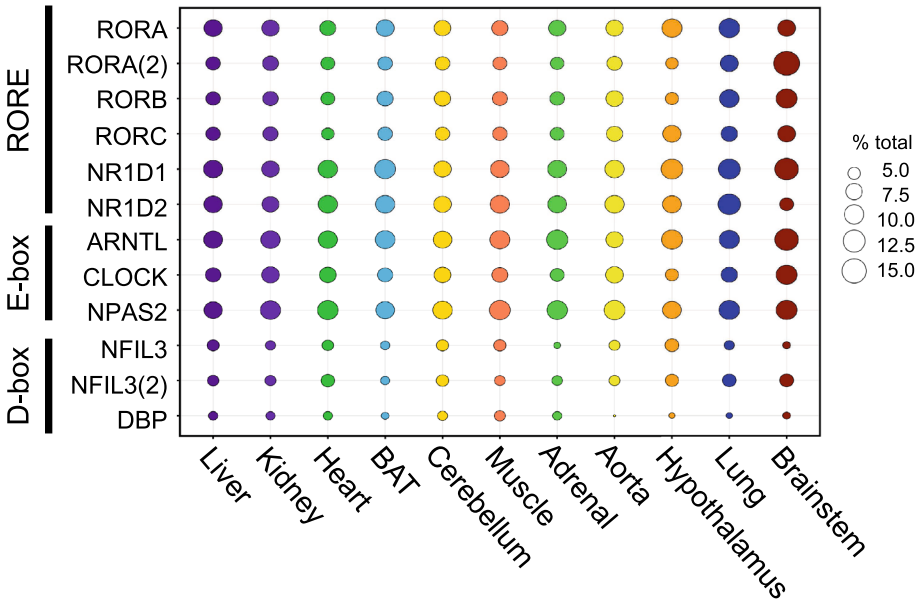

Figure 3. The number of RORC-binding motifs does not correlate with the percentage of rhythmic genes in each tissue. Weighted scatterplot representing the percentage of rhythmic genes containing binding motifs of circadian transcription factors listed on the left. The size of each circle represents the % and each color corresponds to a tissue.

using the RNA-seq dataset (−1000 to +100 bp with respect to TSS), and then determined the number of DNA motifs that can be recognized by RORC in silico. We surveyed the recognition sequences of not only RORC, but also RORA, RORB, NR1D1, NR1D2, ARNTL, CLOCK, NPAS2, NFIL3, and DBP, because these proteins are all considered important to drive rhythmic gene expression[39,41]. We found that the RORC binding motif was found in approximately 8% of the rhythmically expressed genes, and this was consistent in all 11 tissues we examined (Fig. 3). Of the genes with RORC motifs in their promoter region, we found 960 genes that were expressed in all 11 tissues and rhythmic in at least one tissue. Of these 960 genes, the mean level of 291 and 252 genes correlated with the mean level of *Rorc* by Spearman or Pearson analysis, respectively (Supplemental Data File 7). It is plausible that RORC regulates the expression and/or rhythmicity of these downstream RORC target genes. The binding sites for NPAS2, ARNTL, NR1D1, and NR1D2 were the most highly represented (~10–14%), and NFIL3 and DBP were the least represented (~3–6%) (Fig. 3).

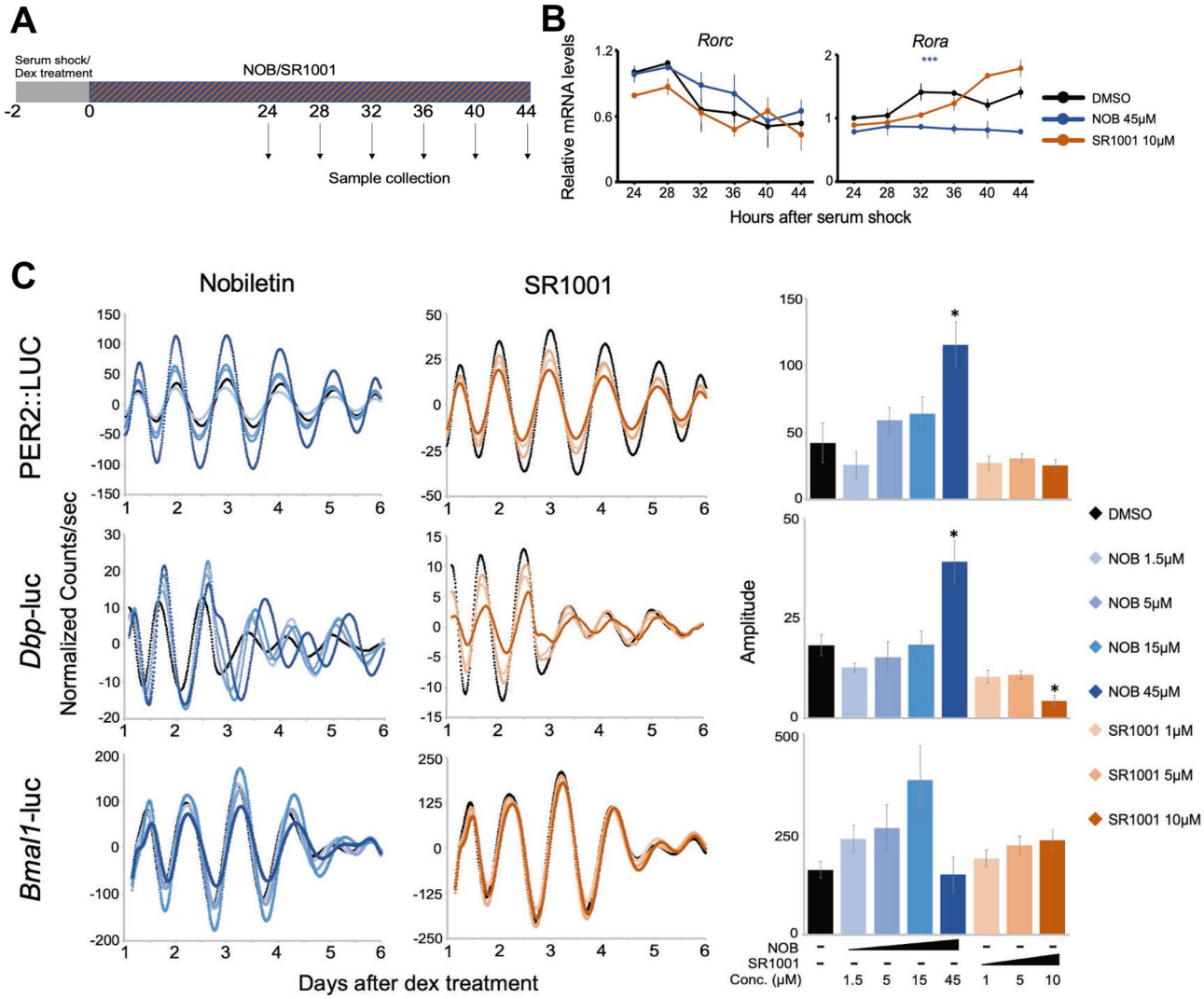

Figure 4. The effect of nobiletin (NOB) and SR1001 on the clock gene expression. (**A**) Schematic representation of our study design. (**B**) Relative mRNA levels of *Rorc* and *Rora* in *Per2::LucSV* cells treated with 45 μM nobiletin, 10 μM SR1001, or DMSO. (**C**) Normalized luciferase activity measurements and quantified amplitudes of PER2::LUCSV, *Dbp*-luc, or *Bmal1*-luc cells treated with varying concentrations of nobiletin, SR1001., or DMSO as control (n = 2–4). All the data represent Mean ± SEM. *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001.

To experimentally test the biological significance of *Rorc*, we utilized three independent luciferase reporter cell lines; *Per2::LucSV* cells (mouse embryonic fibroblasts derived from *Per2::LucSV* knockin mice)[42] as well as *Bmal1*-luc and *Dbp*-luc cells (derived from NIH3T3 cells that stably express luciferase genes under the control of either a *Dbp* or *Bmal1* promoter)[43]. We treated these cells with the *Rora/c* agonist nobiletin or inverse agonist SR1001 to determine whether it is the mRNA level or the transcriptional activity of RORC that is important for the amplitude of core clock genes expression patterns (Fig. 4A). We first analyzed the mRNA level of *Rora/c* in *Per2::LucSV* cells, as the level of *Rorc* cells was under the detection limit in *Dbp*-luc or *Bmal1*-luc cells, both of which derived from NIH3T3 cells[44]. We found that the mRNA levels of *Rorc* were unchanged in the presence of either nobiletin or SR1001, whereas the mRNA levels of *Rora* decreased when treated with nobiletin (Fig. 4B). We also found that nobiletin increased the amplitude of PER2::LUCSV and *Dbp*-luc reporter output compared to control cells (Fig. 4C). In contrast, SR1001 decreased the amplitude of *Dbp*-luc reporter output. Interestingly, neither nobiletin nor SR1001 altered the amplitude of *Bmal1*-luc despite *Bmal1* being under direct control of *Rora/c* (Fig. 4C)[27]. Because both nobiletin and SR1001 have a higher inhibition constant for RORC compared to RORA[45,46], the changes in the amplitude of reporter output are likely due to their effect on RORC.

Discussion

Among the three key parameters for cycling systems (period, phase, and amplitude), the regulatory mechanisms of period and phase have been relatively well-characterized, whereas that of amplitude have remained much more enigmatic. Forward genetics or screening approaches using pharmacological or genetic perturbation have not been successful, as the variance of amplitude is much higher than that of period, compromising the statistical

ability to distinguish true positives from false positives[47–51]. Furthermore, it is currently unclear whether it is a single gene, a combination of genes, one of the feedback loops, and/or topology of the network that is important for amplitude. To make things even more complicated, amplitude can be measured by various outputs, such as gene expression, firing patterns in neurons, body temperature, and locomotor activity, each of which can be under the regulation of both cell-autonomous (intracellular or local) and systemic (or extracellular) rhythms. It also remains unclear whether all the amplitude of various rhythms is regulated by the same mechanism.

In this study, we focused on the percentage of cycling genes in various mouse tissues and explored the possible mechanisms of amplitude regulation of circadian transcriptomic output. The circadian transcriptome can be influenced both by cell-autonomous and systemic cues in each tissue. However, the circadian gene expression is a common feature of the circadian clock system in various tissues and allows us to directly compare the difference in amplitude between tissues without relying on their respective physiology.

We found that 18 genes (eight core clock genes and ten clock-controlled genes) are rhythmically expressed in all tissues that we examined (Table 1, Supplemental Data File 3). This is consistent with previous observations that the rhythmicity of each gene is often tissue-specific, and while only a handful of genes are cycling in all or most tissues, others are rhythmic only in certain tissues[7–11]. The mechanism that drives tissue-specificity of rhythmic gene expression still remains largely unknown. Nonetheless, recent studies demonstrated that the BMAL1 DNA binding is largely tissue-specific, and the tissue-specific rhythmic gene expression can be driven by rhythmic BMAL1 binding in coordination with the activity of enhancers nearby that form chromatin looping[52,53]. It is possible that RORC drives tissue-specific rhythmic target gene expression using a similar mechanism. We also found that the relative amplitude of 13 genes (nine core clock genes and four clock-controlled genes) were correlated with the percentage of cycling genes, while the mean level of 197 genes correlated with the percentage of cycling genes. These genes are not necessarily expressed rhythmically, albeit about a half (100/197) are, and vast majority of these genes were involved in the metabolic processes (Supplemental Data Files 5, 6). Given that the energy cost for cycling genes are higher than non-cycling genes[54], it is reasonable that metabolism related genes are highly expressed in tissues that have higher number of cycling transcripts. Interestingly, metabolically active tissues, such as liver, brown fat, and skeletal muscle have rhythmic RNAs with higher amplitude (Fig. 1E), which is consistent with the previous report[53], suggesting that the metabolic activity in each tissue affects the amplitude of rhythmic RNA expression. It is also plausible that systemic cues, including metabolites, contributed to the differences in the percentage of rhythmic gene expression in each tissue, and these genes act as mediators to connect systemic cues and rhythmic gene expression both independent and dependent on the core clock machinery. However, it is more tempting to postulate that the role of *Rorc* and/or *Per2AS* in the core-clock circuit gives it a more promising function in potentially regulating the amplitude of the circadian transcriptome, at least in the tissues where *Rorc* is expressed.

We also attempted to verify our findings using independent datasets. To achieve this, datasets must include gene expression data taken at multiple circadian time points (to detect rhythmicity) in at least three different tissues (to perform correlation analyses). To date, no other datasets are available in mouse. One study using rats analyzed gene expression patterns in four tissues (liver, lung, muscle, and adipose) at multiple circadian time points[11] (Supplemental Data File 8). Unfortunately, however, some of the microarray platforms used in this study did not include probes for *Rorc* and we were unable to calculate the correlation between the percentage of rhythmic genes and the level of *Rorc.* A recent study in Baboon also surveyed diurnal gene expression patterns in 64 tissues[12]. We analyzed 14 tissues from this dataset (Aorta, Adrenal Cortex, Adrenal Medulla, Bone Marrow, Heart, Hippocampus, Kidney Cortex, Kidney Medulla, Liver, Lung, Pancreas, Prostate, Smooth Muscle, and SCN) that closely mirror 12 mouse tissues (or 11 for RNA-seq) that we analyzed in our study.

When we used the MetaCycle p-value < 0.05 as a rhythmicity threshold, we did observe a positive correlation between the level of *RORC* with the percentage of rhythmic transcripts (Pearson: $r^2 = 0.603$, $p = 0.029$, Fig. S3A, Supplemental Data Files 9a, 9b). However, this correlation was not statistically significant when we used the JTK_CYCLE algorithm with p-value < 0.05 (Hughes et al., 2010) (Pearson: $r^2 = 0.167$, $p = 0.587$, Fig. S3B) or when we used q-value < 0.05 as a rhythmicity threshold. In contrast, the correlation between the *Rorc* level and the percentage of rhythmic transcripts in mouse was statistically significant with both MetaCycle (Table 2, Fig. 2C) and JTK_CYCLE (Pearson: $r^2 = 0.738$, $p = 0.006$, Spearman: $rho = 0.678$, $p = 0.019$) (Fig. S3C) with B.H. Q value < 0.05. Of note, we were also unable to detect *PER2AS* in any of these tissues we examined (Fig. S3D), indicating that either *PER2AS* does not exist in baboon, or the baboon reference genome we used (*Papio Anubis* 2.0) does not have high enough resolution to annotate and detect *PER2AS*. Overall, we concluded that it is unclear whether the *RORC* level and the percentage of rhythmic transcriptome correlate in baboon.

We have also found that the mRNA levels of 584 and 473 genes also correlated with the percentage of the rhythmic mRNAs in each tissue in rat and baboon, respectively (Supplemental Data File 7)[11,12]. Among these, 40 and 4 genes also showed a correlation in mouse dataset[10] whereas none was commonly detected as correlated in all three datasets (Supplemental Data File 7). The significance of these genes in regulating the amplitude of circadian transcriptome output is unclear, however. Experimental parameters (sampling resolution, number of tissues examined, transcriptomic platform) are significantly different between studies, and these have a significant impact on detecting rhythmicity of each gene as well as calculating the percentage of rhythmic transcripts (Supplemental Data Files 1a, 1b, 2, 8, 9a, 9b). It is also possible that the regulatory mechanism of amplitude is species-specific and not conserved.

The positive loop (*Clock-Arntl-Rev-Ror*) was originally considered to confer additional robustness to the system and, therefore, stabilizes the system. However, it is not required for circadian rhythm generation[37,55]. Recent studies have also highlighted the role of the positive loop as the central axis of amplitude regulation[45,56]. Our study is consistent with these findings and suggest that the positive loop, particularly the level of *Rorc*, is important in setting the amplitude of the circadian transcriptome. In addition, our study also suggested that *Per2AS* is involved in the positive loop, because the level of *Per2AS* positively correlated with the level of *Rorc* as

well as the percentage of cycling genes in each tissue, even though it was originally assumed to only interact with *Per2*[35]. Interestingly, the mathematical model predicted the functional interaction between *Ror* and *Per2AS*, as *Per2AS* would restore stable circadian rhythms when they are disrupted by the overproduction of *Ror* or *Rev-erb* mRNAs[35]. It is possible that *Rorc* and *Per2AS* function synergistically in the circadian clock system. Because long non-coding RNAs, such as *Per2AS,* can function *in trans* and interacts with DNAs, other RNAs, or proteins to regulate target gene expression[57,58], it is possible that *Per2AS* RNA interacts with RORC proteins to modify its transcriptional activity. *Per2AS* could also alter the level of *Rorc* by interacting with the promoter or enhancer sequences of *Rorc* (i.e., DNA), transcriptional regulators and/or epigenetic modifiers of *Rorc* or vice versa. Alternatively, the correlation between *Rorc* and *Per2AS* may simply indicate that their expression is regulated by the same or similar mechanism.

It is still unclear from our study whether the relationship between *Rorc* and the percentage of rhythmic transcripts is simply a correlation or causation. Because the molecular clock system is quite complex, and the expression of each clock gene is dependent on the expression/activity of other clock genes directly or indirectly, we think it is critically important to keep the network intact to fully understand the function of each component. In contrast to experimental approaches in which perturbation of each genetic component (i.e., gene knock-out, knock-down, or overexpression) would often disturb the network itself, a computational approach excels in this area to shed light into the function of a component within a network. To circumvent these issues, we used a pharmacological approach and altered the transcriptional activity of RORC without changing its levels. We found that the transcriptional activity of RORC alters the amplitude of reporter bioluminescence output (Fig. 4), suggesting that the transcriptional activity of RORC is important for regulating the amplitude of the circadian clock machinery and potentially the circadian transcriptome output. Regardless, we think the level of *Rorc* is still biologically relevant, because the changes in the *Rorc* mRNA level can lead to the changes in the level of RORC and/or its activity.

It still remains unclear why *Rorc*, but not *Rora* and *Rorb*, correlates with the amplitude of the circadian transcriptome, as all the ROR proteins share significant sequence similarities[40,59]. Unfortunately, the physiological relevance of each ROR paralogue has never been clarified in the mammalian circadian system. One notable difference among *Ror* paralogues, however, is their unique expression patterns (Supplementary Data Files 1a, 1b, 2). It is possible that the systemic cues, which are, in theory, the same to all the tissues have a tissue-specific impact in regulating the level of *Rorc*. Understanding the difference in the regulatory mechanisms of *Ror* gene expression would provide insight into how their tissue-specific expression pattern is achieved and how *Rorc* specifically impacts the amplitude of the circadian transcriptomic output.

Overall, our study highlighted the potential role of *Rorc* in regulating the amplitude of the circadian transcriptome, although it is unclear whether the correlation between the *Rorc* and the percentage of rhythmic transcriptome is specific to mouse. Follow-up experimental studies would further complement our observations from the rich transcriptomic datasets that are publicly available and delineate the mechanisms of circadian amplitude regulation.

Data availability

The datasets generated during and/or analyzed during the current study are available in the NCBI GEO repository, from series GSE54650 (mouse), GSE98965 (baboon), or GSE8988, GSE8989, GSE20635, and GSE25612 (rat).

Received: 12 April 2020; Accepted: 20 November 2020
Published online: 14 December 2020

References

1. Mohawk, J. & Takahashi, J. Cell autonomy and synchrony of suprachiasmatic nucleus circadian oscillators. *Trends Neurosci.* **34**, 349–358. https://doi.org/10.1016/j.tins.2011.05.003 (2011).
2. Lowrey, P. & Takahashi, J. MAMMALIAN CIRCADIAN BIOLOGY: Elucidating genome-wide levels of temporal organization. *Annu. Rev. Genomics Hum. Genet.* **5**, 407–441. https://doi.org/10.1146/annurev.genom.5.061903.175925 (2004).
3. Takahashi, J. S., Hong, H.-K., Ko, C. H. & McDearmon, E. L. The genetics of mammalian circadian order and disorder: Implications for physiology and disease. *Nat. Rev. Genet.* **9**, 764–775. https://doi.org/10.1038/nrg2430 (2008).
4. Takahashi, J. S. Transcriptional architecture of the mammalian circadian clock. *Nat. Rev. Genet.* **18**, 164–179. https://doi.org/10.1038/nrg.2016.150 (2017).
5. Akhtar, R. *et al.* Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. *Curr. Biol.* **12**, 540–550. https://doi.org/10.1016/s0960-9822(02)00759-5 (2002).
6. Duffield, G. E. DNA microarray analyses of circadian timing: The genomic basis of biological time. *J. Neuroendocrinol.* **15**, 991–1002. https://doi.org/10.1046/j.1365-2826.2003.01082.x (2003).
7. Panda, S. *et al.* Coordinated transcription of key pathways in the mouse by circadian clock. *Cell* **109**, 307–320. https://doi.org/10.1016/s0092-8674(02)00722-5 (2002).
8. Storch, K.-F. *et al.* Extensive and divergent circadian gene expression in liver and heart. *Nature* **417**, 78–83. https://doi.org/10.1038/nature744 (2002).
9. Ueda, H. R. *et al.* A transcription factor response element for gene expression during circadian night. *Nature* **418**, 534–539. https://doi.org/10.1038/nature00906 (2002).
10. Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: Implications for biology and medicine. *Proc. Natl. Acad. Sci.* **111**, 16219–16224. https://doi.org/10.1073/pnas.1408886111 (2014).
11. Mavroudis, P., DuBois, D., Almon, R. & Jusko, W. Daily variation of gene expression in diverse rat tissues. *PLoS ONE* https://doi.org/10.1371/journal.pone.0197258 (2018).
12. Mure, L. S. *et al.* Diurnal transcriptome atlas of a primate across major neural and peripheral tissues. *Science* **359**, eaaao0318. https://doi.org/10.1126/science.aao0318 (2018).
13. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–264. https://doi.org/10.1093/biostatistics/4.2.249 (2003).

14. Dobin, A. & Gingeras, T. R. Mapping RNA-seq reads with STAR. *Curr. Protoc. Bioinform.* **51**, 11.14.11-11.14.19. https://doi.org/10.1002/0471250953.bi1114s51 (2015).
15. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589. https://doi.org/10.1016/j.molcel.2010.05.004 (2010).
16. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789. https://doi.org/10.1101/gr.132159.111 (2012).
17. Ravasi, T. *et al.* Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res.* **16**, 11–19. https://doi.org/10.1101/gr.4200206 (2006).
18. Ørom, U. A. *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143**, 46–58. https://doi.org/10.1016/j.cell.2010.09.001 (2010).
19. Wu, Y.-W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605–607. https://doi.org/10.1093/bioinformatics/btv638 (2015).
20. Wickam, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686. https://doi.org/10.21105/joss.01686 (2019).
21. Harrell, F. E. *et al.* Hmisc: Harrell Miscellaneous. R package version 4.3-1. https://CRAN.R-project.org/package=Hmisc (2020).
22. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29. https://doi.org/10.1038/75556 (2000).
23. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338. https://doi.org/10.1093/nar/gky1055 (2019).
24. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: Scanning for occurrences of a given motif. *Bioinformatics (Oxford, England)* **27**, 1017–1018. https://doi.org/10.1093/bioinformatics/btr064 (2011).
25. Fornes, O. *et al.* JASPAR 2020: Update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92. https://doi.org/10.1093/nar/gkz1001 (2019).
26. Oklejewicz, M. *et al.* Phase resetting of the mammalian circadian clock by DNA damage. *Curr. Biol.* **18**, 286–291. https://doi.org/10.1016/j.cub.2008.01.047 (2008).
27. Sato, T. K. *et al.* A functional genomics strategy reveals Rora as a component of the Mammalian Circadian Clock. *Neuron* **43**, 527–537. https://doi.org/10.1016/j.neuron.2004.07.018 (2004).
28. Yang, M.-Y. *et al.* Downregulation of circadian clock genes in chronic myeloid leukemia: Alternative methylation pattern of hPER3. *Cancer Sci.* **97**, 1298–1307. https://doi.org/10.1111/j.1349-7006.2006.00331.x (2006).
29. Akashi, M. & Takumi, T. The orphan nuclear receptor RORα regulates circadian transcription of the mammalian core-clock Bmal1. *Nat. Struct. Mol. Biol.* **12**, 441–448. https://doi.org/10.1038/nsmb925 (2005).
30. Matsumura, R., Okamoto, A., Node, K. & Akashi, M. Compensation for intracellular environment in expression levels of mammalian circadian clock genes. *Sci. Rep.* **4**, 4032–4032. https://doi.org/10.1038/srep04032 (2014).
31. Yamazaki, S. *et al.* Resetting central and peripheral circadian oscillators in transgenic rats. *Science* **288**, 682–685. https://doi.org/10.1126/science.288.5466.682 (2000).
32. Koike, N. *et al.* Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science (New York, N.Y.)* **338**, 349–354. https://doi.org/10.1126/science.1226339 (2012).
33. Menet, J. S., Rodriguez, J., Abruzzi, K. C. & Rosbash, M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *Elife* **1**, e00011. https://doi.org/10.7554/eLife.00011 (2012).
34. Vollmers, C. *et al.* Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell Metab.* **16**, 833–845. https://doi.org/10.1016/j.cmet.2012.11.004 (2012).
35. Battogtokh, D., Kojima, S. & Tyson, J. J. Modeling the interactions of sense and antisense Period transcripts in the mammalian circadian clock network. *PLoS Comput. Biol.* **14**, e1005957. https://doi.org/10.1371/journal.pcbi.1005957 (2018).
36. Etchegaray, J.-P., Lee, C., Wade, P. A. & Reppert, S. M. Rhythmic histone acetylation underlies transcription in the mammalian circadian clock. *Nature* **421**, 177–182. https://doi.org/10.1038/nature01314 (2003).
37. Liu, A. *et al.* Redundant function of REV-ERBα and β and non-essential role for Bmal1 cycling in transcriptional regulation of intracellular circadian rhythms. *PLoS Genet.* **4**, e1000023. https://doi.org/10.1371/journal.pgen.1000023 (2008).
38. Takeda, Y., Jothi, R., Birault, V. & Jetten, A. M. RORγ directly regulates the circadian expression of clock genes and downstream targets in vivo. *Nucleic Acids Res.* **40**, 8519–8535. https://doi.org/10.1093/nar/gks630 (2012).
39. Ueda, H. R. *et al.* System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nat. Genet.* **37**, 187–192. https://doi.org/10.1038/ng1504 (2005).
40. Kojetin, D. J. & Burris, T. P. REV-ERB and ROR nuclear receptors as drug targets. *Nat. Rev. Drug Discov.* **13**, 197–216. https://doi.org/10.1038/nrd4100 (2014).
41. Kumaki, Y. *et al.* Analysis and synthesis of high-amplitude *Cis*-elements in the mammalian circadian clock. *Proc. Natl. Acad. Sci.* **105**, 14946–14951. https://doi.org/10.1073/pnas.0802636105 (2008).
42. Yoo, S.-H. *et al. Period2* 3′-UTR and microRNA-24 regulate circadian rhythms by repressing PERIOD2 protein accumulation. *Proc. Natl. Acad. Sci.* **114**, E8855–E8864. https://doi.org/10.1073/pnas.1706611114 (2017).
43. Saini, C., Morf, J., Stratmann, M., Gos, P. & Schibler, U. Simulated body temperature rhythms reveal the phase-shifting behavior and plasticity of mammalian circadian oscillators. *Genes Dev.* **26**, 567–580. https://doi.org/10.1101/gad.183251.111 (2012).
44. Johnson, J. L. *et al.* Lineage-determining transcription factor TCF-1 initiates the epigenetic identity of T cells. *Immunity* **48**, 243-257.e210. https://doi.org/10.1016/j.immuni.2018.01.012 (2018).
45. He, B. *et al.* The small molecule nobiletin targets the molecular oscillator to enhance circadian rhythms and protect against metabolic syndrome. *Cell Metab.* **23**, 610–621. https://doi.org/10.1016/j.cmet.2016.03.007 (2016).
46. Solt, L. A. *et al.* Suppression of TH17 differentiation and autoimmunity by a synthetic ROR ligand. *Nature* **472**, 491–494. https://doi.org/10.1038/nature10075 (2011).
47. Hirota, T. *et al.* High-throughput chemical screen identifies a novel potent modulator of cellular circadian rhythms and reveals CKIα as a clock regulatory kinase. *PLoS Biol* **8**, e1000559. https://doi.org/10.1371/journal.pbio.1000559 (2010).
48. Hirota, T. *et al.* Identification of small molecule activators of cryptochrome. *Science (New York, N.Y.)* **337**, 1094–1097. https://doi.org/10.1126/science.1223710 (2012).
49. Hirota, T. *et al.* A chemical biology approach reveals period shortening of the mammalian circadian clock by specific inhibition of GSK-3β. *Proc. Natl. Acad. Sci.* **105**, 20746–20751. https://doi.org/10.1073/pnas.0811410106 (2008).
50. Takahashi, J. S., Shimomura, K. & Kumar, V. Searching for genes underlying behavior: Lessons from circadian rhythms. *Science (New York, N.Y.)* **322**, 909–912. https://doi.org/10.1126/science.1158822 (2008).
51. Zhang, E. E. *et al.* A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell* **139**, 199–210. https://doi.org/10.1016/j.cell.2009.08.031 (2009).
52. Beytebiere, J. R. *et al.* Tissue-specific BMAL1 cistromes reveal that rhythmic transcription is associated with rhythmic enhancer-enhancer interactions. *Genes Dev.* **33**, 294–309. https://doi.org/10.1101/gad.322198.118 (2019).
53. Yeung, J. *et al.* Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs. *Genome Res.* **28**, 182–191. https://doi.org/10.1101/gr.222430.117 (2018).
54. Wang, G. Z. *et al.* Cycling transcriptional networks optimize energy utilization on a genome scale. *Cell Rep.* **13**, 1868–1880. https://doi.org/10.1016/j.celrep.2015.10.043 (2015).

55. Hogenesch, J. B. & Herzog, E. D. Intracellular and intercellular processes determine robustness of the circadian clock. *FEBS Lett.* **585**, 1427–1434. https://doi.org/10.1016/j.febslet.2011.04.048 (2011).
56. Zhao, X. *et al.* Circadian amplitude regulation via FBXW7-targeted REV-ERBα degradation. *Cell* **165**, 1644–1657. https://doi.org/10.1016/j.cell.2016.05.012 (2016).
57. Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166. https://doi.org/10.1146/annurev-biochem-051410-092902 (2012).
58. Marchese, F. P., Raimondi, I. & Huarte, M. The multidimensional mechanisms of long noncoding RNA function. *Genome Biol.* **18**, 206. https://doi.org/10.1186/s13059-017-1348-2 (2017).
59. Solt, L. A., Griffin, P. R. & Burris, T. P. Ligand regulation of retinoic acid receptor-related orphan receptors: Implications for development of novel therapeutics. *Curr. Opin. Lipidol.* **21**, 204–211. https://doi.org/10.1097/MOL.0b013e328338ca18 (2010).

Acknowledgements

The authors thank Drs. Frank Aylward (Department of Biological Sciences, Virginia Tech), Natalia Toporikova (Department of Biology, Washington and Lee University), for technical assistance and Dr. Janet Webster (Fralin Life Sciences Institute, Virginia Tech), as well as the members of the Kojima laboratory, for critical reading of the manuscript. *Per2::LucSV* cells, as well as *Dbp*-luc and *Bmal1*-luc cell lines were a kind gift from Dr. Seung-Hee Yoo (University of Texas Health Science Center at Houston) and Dr. Ueli Schibler (University of Geneva), respectively. This work was supported by R01GM126223 from the NIH (to S.K.).

Author contributions

E.S.L., M.L.C., M.L.G., and A.N.J. performed data analysis, S.K. designed and conceived the study, and both E.S.L. and S.K. generated figures, wrote, and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-020-78851-9.

Correspondence and requests for materials should be addressed to S.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.