PNAS
www.pnas.org

Supplementary Information for

Systematic analysis of differential rhythmic liver gene expression mediated by the circadian clock and feeding rhythms

Benjamin D. Weger, Cédric Gobet, Fabrice P.A. David, Florian Atger, Eva Martin, Nicholas E. Phillips, Aline Charpagne, Meltem Weger, Felix Naef, and Frédéric Gachon

Frédéric Gachon, Félix Naef

Email: f.gachon@uq.edu.au, felix.naef@epfl.ch

This PDF file includes:

Supplementary Information Text
Figures S1 to S8
Legends for Datasets S1 to S3
SI References

Other supplementary materials for this manuscript include the following:

Datasets S1 to S3

Supplementary Information Text

SI Materials and Methods

RNA extraction and RNA-Seq. ~100 mg of frozen tissue was grinded with a Polytron PT 2500 E homogenizer in extraction buffer consisting of 3.9 M guanidium thiocyanate, 0.03 M sodium citrate, 0.2 M sodium acetate, and 1% (vol/vol) 2-mercaptoethanol. An equal volume of 0.5 volume of chloroform/isoamyl alcohol (49:1 [vol/vol]) and phenol (saturated in H_2O) was added to the homogenate. The mixture was vortexed and then centrifuged for 20 min at 12,000 × *g* at 4 °C for phase separation. Precipitation of RNA of the aqueous phase was performed in an equal volume of isopropanol at −20 °C for at least 20 min before precipitate was pelleted by centrifugation for 15 min at 12,000 × *g* at 4 °C. The pellet was resuspended in 4 M lithium chloride and then repelleted by centrifugation for 15 min at 12,000 × *g* at 4 °C. Subsequently, the pellet was washed with 75% ethanol for 15 min at 12,000 × *g* at 4 °C, dried at room temperature, and finally dissolved in RNase/DNase free water. Quantification of total RNA was done with Ribogreen (Life Technologies), and RNA quality was assessed with a Fragment Analyzer (Advances Analytical).

Libraries of the *Bmal1* KO and *Cry1*/*2* KO animal series were prepared using the TruSeq Stranded mRNA Library Prep protocol (Illumina) following the manufacturer's protocol. We used 250 ng of total RNA as starting material and performed 11 PCR cycles for library amplification. The optimal number of PCR cycles was assessed by qPCR (Kapa BioSystems). All 96 samples (Fig. 1B) were pooled, denatured, spiked with 3% of the PhiX Control v3 Library and loaded onto a paired-end read flow cell v3 at a final concentration of 9 pM for a paired-end 126 cycles run on a HiSeq 2500 (Illumina). Sequencing depth was equivalent to six samples per lane. For the data on *Hlf*/*Dbp*/*Tef* KO mice, sequencing libraries were prepared from 3 μg of total RNA using the TruSeq Stranded mRNA Library Prep Kit following the manufacturer's protocol. Purified libraries were quantified with Qubit DNA HS (Thermo Fisher Scientific), and profile analyses were retrieved on a TapeStation TS4200 (Agilent). Paired-end sequencing with 75 cycles were performed on a NextSeq 500 following Illumina's recommendations. We mapped the reads on the Ensembl *M. musculus* genome (GRCm38/mm10) using STAR-2.7.3a (1) with the annotation from Ensembl release 98 (2). To count uniquely mapped reads per gene, we used the quantMode GeneCounts option of STAR.

Modeling Temporal Gene Expression Profiles across Multiple Conditions. To assess rhythmicity and mean differences of gene expression in RNA-Seq count data, we developed a model selection framework based on generalized linear models (GLMs) (*SI Appendix*, Fig. S1*A*). As proposed (3), we modeled counts Y_{gs} that have been uniquely mapped to a gene g in a sample *s* as a negative binomial (NB) with a fitted mean μ_{gs} and a gene-specific but sample-independent dispersion parameter θ_g.

$$Y_{gs} \sim NB(\mu_{gs}, \theta_g)$$

The fitted mean is proportional to the quantity q_{gs} of fragments that correspond to a gene *g* in a sample s scaled by a sample-specific scaling factor λ_s (3). This scaling factor depends on the sampling depth of each library and was estimated using the median-of-ratios method of DESeq2 (3).

$$\mu_{gs} = \lambda_s q_{gs}$$

A gene-specific distribution θ_g was estimated using empirical Bayes shrinkage (4).

The fit was performed using a GLM with a logarithmic link function implemented in DESeq2. Sample specific size factor (λ_s) was defined as an offset. The full GLM was defined as follows:

$$\log_2(\mu_{gbcs}) = m^1_{gc} + m^2_{gb} + \alpha_{gc}\cos(\omega\, t(s)) + \beta_{gc}\sin(\omega\, t(s)) + \log_2(\lambda_s)$$

where μ_{gbcs} is the mean of the NB distribution for gene g and condition c in sample s at Zeitgeber/circadian time $t(s)$. The index b refers to a batch of samples. The parameters α_{gc} and β_{gc} are coefficients of the cosine and sine functions, respectively. m^1_{gc} is a coefficient to describe a condition specific mean expression level. We included a batch specific m^2_{gb} to account for technical batch effects when detected in the data. Specifically, we included m^2_{gb} to account for a batch effect observed for one series of *Cry1/2* KO mice under *ad libitum* feeding regimen (GSM4037515 – GSM4037520). To select an optimal gene-specific model, we proceeded in two steps. First, we assessed rhythmicity across the different conditions, where the parameters m^1_{gc} and m^2_{gb} were free but the parameters α_{gc} and β_{gc} had constraints depending on the particular model considered. To this end, we defined different models for four (*e.g., Bmal1* WT, *Bmal1* KO, *Cry1*/*2* WT, *Cry1*/*2* KO) or two conditions (*i.e., Hlf*/*Dbp*/*Tef* KO and *Hlf*/*Dbp*/*Tef* WT). Models were defined to have either zero (nonrhythmic pattern) or nonzero (rhythmic pattern) α and β coefficients for each analyzed condition. Moreover, for some models, the values of α and β can be also shared within any combination of the four conditions (*SI Appendix*, Fig. S1 *E* and *F*). For example, for four conditions, there are 52 such models (*SI Appendix*, Fig. S1*A*). The coefficients α and β were used to calculate the phase [atan2(β, α)] and amplitude ($\log_2$ fold-change peak-to-trough; $2\sqrt{\alpha^2+\beta^2}$) of a gene. Bayesian information criterion (BIC) based model selection was employed to account for model complexity (5) using the following formula:

$$\text{BIC}_j = k\log(n) - 2\log(\hat{L}_j).$$

$\hat{L}_j$ is defined as the log-likelihood of the model *j* from the regression, *n* is the number of data points, and *k* is the number of parameters. To assess the confidence of the selected model *j*, we calculated the Schwarz weight (BICW):

$$\text{BICW}_j = \frac{e^{\frac{1}{2}\Delta\text{BIC}_j}}{\sum_{m=1}^{M} e^{\frac{1}{2}\Delta\text{BIC}_m}},$$

where ΔBIC_j is defined as the difference in BIC between model *j* and the minimum BIC value in the entire model set BIC_{m^*}:

$$\Delta\text{BIC}_j = \text{BIC}_j - \text{BIC}_{m^*}.$$

BICW_j is interpreted as the probability for model j. The model with the highest BIC was selected as the optimal model within the set of all defined models. In the second step, we analyzed the mean levels (*SI Appendix*, Fig. S1*A*) and thus set the coefficient α and β to the values of the selected model in the first regression. A new regression was performed where the parameter m^1_{gc} was either free or subject to constraints between conditions based on several competing models. We considered all possible combinations for the mean coefficient m^1_{gc} with differing or shared means between conditions. For example, for four conditions, there are 15 such models (*SI Appendix*, Fig. S1 *A* and *D*). Each model was solved using generalized linear regression, and each gene was assigned to a preferred model based on the BICW as described above for the first iteration. Cook's distance is an estimate of how much a data point influences the fitted coefficients for a gene. A large value of Cook's distance is considered to be an outlier. Thus, genes that did not reach at least a BICW above 0.4 for the preferred model or genes that had at least a Cook's distance of more than 1 were categorized as "ambiguous" (model 0). In the case of two conditions (*i.e.,* PARbZip KO dataset) the threshold on BICW was set to 0.95. The difference in the two BICW thresholds reflected that the number of possible models is much larger for four compared with two conditions. Also, we considered only genes with a $\log_2$ amplitude of at least 0.25 to be rhythmic in the corresponding condition. To asses temporal variation of normally distributed measurements (*e.g.,* relative liver size, nuclear abundance of

PPARα), we modified model selection approach described above to deal with gaussian noise and therefore used linear models. A function handling normally distributed data is implemented in the *dryR* package. To assess differences in mean expression levels between *Bmal1* and *Cry1/2* KOs, we calculated a double-difference score that is the difference of the $\log_2$ fold-changes in mean levels of *Bmal1* KO vs. *Bmal1* WT and *Cry1/2* KO vs. *Cry1/2* WT:

$$\Delta\Delta(Bmal1_{\text{KO}}, Cry1/2_{\text{KO}}, Bmal1_{\text{WT}}, Cry1/2_{\text{WT}}) = (m_{Bmal1\,\text{KO}} - m_{Bmal1\,\text{WT}}) - (m_{Cry1/2\,\text{KO}} - m_{Cry1/2\,\text{WT}}),$$

where m is the condition specific $\log_2$ mean expression level retrieved from *dryR*.

Functional and Gene Set-Enrichment Analysis. We tested enrichment for annotated gene sets of several sources including Gene Ontology (GO) (6), Molecular Signatures Database (MSigDB) (7), Kyoto Encyclopedia of Genes and Genomes (KEGG) (8) and Reactome Pathway database (9). For phase-shifted rhythmic genes in *Cry1/2* and *Bmal1* KO animals, we performed an untargeted ChIP-enrichment analysis as described in ref. (10) to screen for potentially interesting transcription factors. Target genes for circadian clock regulators and PPARα were defined as outlined below. We used GOseq 3.1 (11) to assess overrepresentation of all gene sets for genes assigned to a statistical model. Temporal-resolved enrichments were calculated using a Fisher's exact test. We defined foreground genes as genes with a peak phase within in a sliding 4-h window. The gene universe was defined as all expressed genes. For each gene set, the window was moved by 0.1 h to calculate the *P* value.

Simulated Count Data and Comparing *dryR* with Other Methods. We simulated rhythmic count data from a NB distribution for four conditions and 5,000 features in an interval of 4 h using Simphony (12). The ratio of rhythmic genes in each condition was set to 25%. The period was 24 h with a minimum amplitude ($\log_2$ peak-to-trough) of 0.25 and a maximum of 6. The mean was between 10^0 to 10^6 counts. To test the aspect of rhythm detection across four conditions, we considered rhythm detection as true positive if all conditions were correctly detected as rhythmic or nonrhythmic. We compared *dryR* to the following alternative methods for rhythmicity detection algorithm: Jonckheere-Terpstra-Kendall (JTK) (13), Lomb-Scargle (14) and meta2d (15) on log-transformed count data. To compare *dryR* to a harmonic regression without using the BIC model selection, we fitted the full GLM model outlined in the previous paragraph. Subsequently, we employed a Wald test to test for significance of the α or β coefficients in a negative binomial GLM using DEseq2 (4). The retrieved *P* values were corrected for multiple testing according to Benjamini and Hochberg (16). The simulated data to test the performance for differential rhythmicity analysis was generated similarly as described above with minor modifications. We simulated data for 2 conditions with only rhythmic genes. Two datasets were generated with a phase shift of 0 or 4 h and an amplitude change of $\log_2$ fold-change of 0 or 1 between the two conditions. The standard datasets had a sampling interval of 4 h and consisted of two replicates. Differences in amplitude and phase were assessed with *dryR*. We compared the performance with CircaCompare (17), LimoRhyde (18) and DODR (19) on log-transformed data. CircaCompare was run with either default settings or an alpha threshold of 1 (test for presence of rhythmicity). *dryR* was run either with or without a threshold for a minimum phase difference of 2 h or minimum $\log_2$ peak-to-trough of 0.25. For the extended analysis using simulated datasets, we varied the sampling interval (i.e., 1, 2, and 6 h), or we changed the number of replicates (i.e., one, three, four, and eight replicates) while keeping the other parameter at the standard value. We also generated a standard dataset with one, two, or three missing samples in each condition.

Data resources and analysis. Target genes for clock genes and PPARα in mouse liver have been identified by published ChIP-Seq. ChIP-Seq data was retrieved from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (20) for BMAL1, CLOCK, NPAS2, CRY1, CRY2, PER1 and PER2, which are available under GEO Series accession no. GSE39977 (21); NR1D1 and NR1D2 are available under GEO Series accession no. GSE34020 (22). ChIP-

Seq data for NFIL3 and DBP were retrieved from the European Nucleotide Archive (ENA) (23) under accession no. PRJDB7796 (24). We mapped reads on the Ensembl *M. musculus* genome (GRCm38/mm10) using STAR-2.7.3a (1). Peaks were called using macs/2-2.1.1 (25) and peak annotation was performed with ChIPseeker 3.10 (26) using the genome annotation of Ensembl release 98 (2). Binding sites were visualized using the Integrative Genomics Viewer (27). Proteomics data were retrieved from the study of ref. (28), which are available in the ProteomeXchange Consortium (29) under accession number PXD003818. RNA-Seq data for *Nfil3* KO mice series in liver were retrieved from ENA (accession no. PRJDB7789) (24). Mapping and downstream analysis of the published reads followed the same procedures outlined in section RNA extraction and RNA-Seq.

Web application. We built a web application (https://clockprofile.epfl.ch), providing a user-friendly interface of the datasets and the statistical analysis. The application was built using the Ruby-on-Rails framework. The backend consists in a PostgreSQL relational database in which analysis results are stored. In the user interface, dynamic plots are produced with the Plotly.js library.

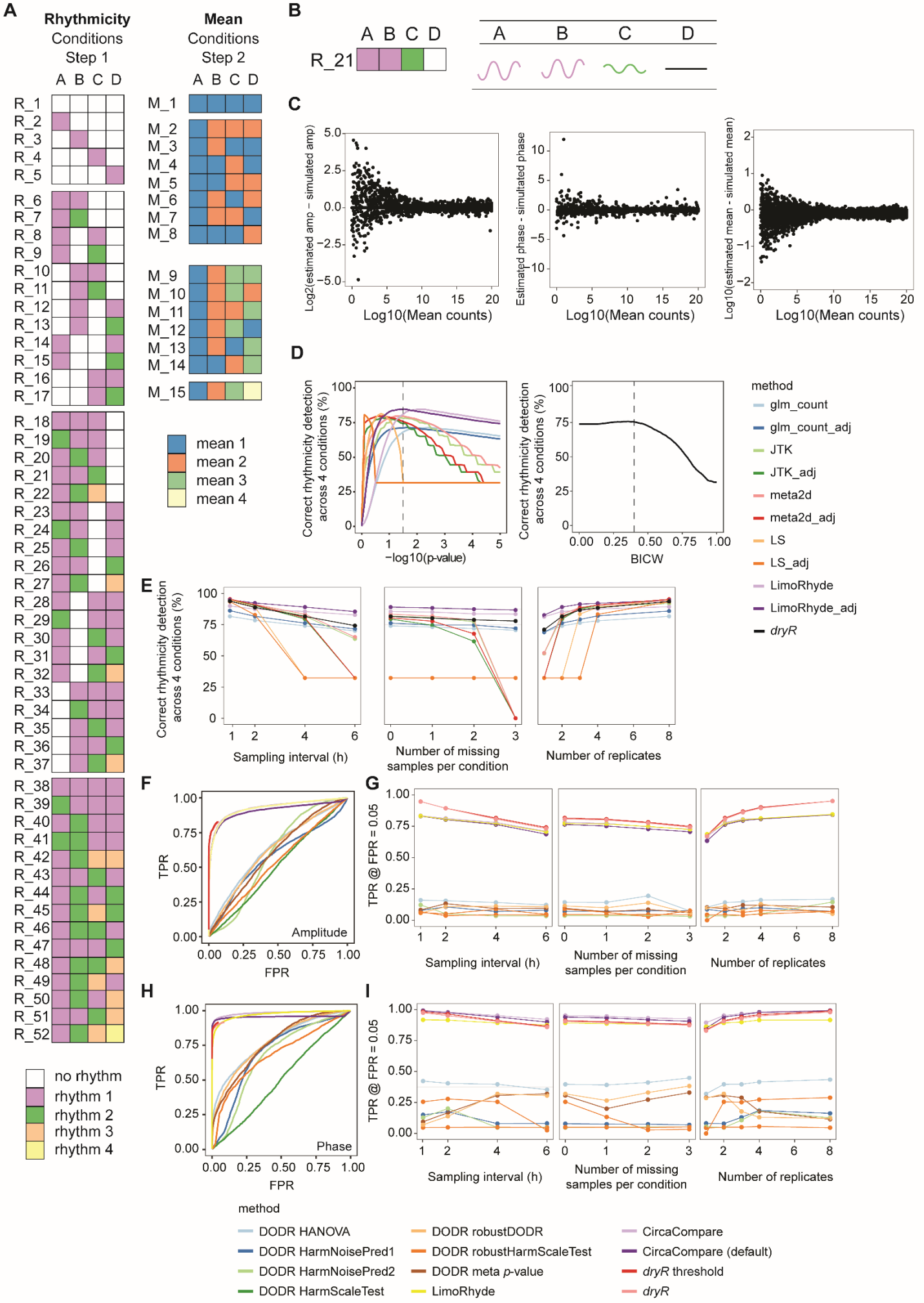

A
Rhythmicity
Conditions
Step 1
A B C D
Mean
Conditions
Step 2
A B C D
mean 1
mean 2
mean 3
mean 4
no rhythm
rhythm 1
rhythm 2
rhythm 3
rhythm 4
B
A B C D
R_21
C
Log2(estimated amp − simulated amp)
Log10(Mean counts)
Estimated phase - simulated phase
Log10(estimated mean - simulated mean)
D
Correct rhythmicity detection across 4 conditions (%)
−log10(p-value)
BICW
method
glm_count
glm_count_adj
JTK
JTK_adj
meta2d
meta2d_adj
LS
LS_adj
LimoRhyde
LimoRhyde_adj
dryR
E
Sampling interval (h)
Number of missing samples per condition
Number of replicates
F
TPR
FPR
Amplitude
G
TPR @ FPR = 0.05
H
Phase
I
method
DODR HANOVA
DODR HarmNoisePred1
DODR HarmNoisePred2
DODR HarmScaleTest
DODR robustDODR
DODR robustHarmScaleTest
DODR meta p-value
LimoRhyde
CircaCompare
CircaCompare (default)
dryR threshold
dryR

Fig. S1. Benchmarking of *dryR* to analyze rhythmic datasets that consist of multiple conditions.
A. Overview of all rhythmic and mean models that are fitted in case the data consists of four conditions. Same color indicates shared mean levels and rhythmic parameters between indicated conditions for mean and rhythmic models, respectively.
B. Schematic definition of an example rhythmic model. For model R_21 rhythmic parameters (*i.e.,* amplitude and phase) are shared between condition A and B. Condition C shows an altered rhythmicity pattern compared to condition A and B. No rhythm is detected for condition D.
C. Scatterplots to assess the differences of predicted and simulated values in function of mean expression for amplitude (left), phase (middle) and mean levels (right).
D. Rhythm detection across four conditions in a simulated dataset using *dryR* and other methods including JTK, Lomb-Scargle (LS), ARSER, meta2d and LimoRhyde on log transformed count data and a cosin fit on count data using generalized linear model (GLM_cos). *P* values were adjusted for multiple testing using Benjamini-Hochberg correction if indicated (_adj). The dashed lines represent a typical threshold used for rhythm detection (*P* value ≤ 0.05 and BICW ≥ 0.4).
E. Percentage of correctly detected rhythms across 4 condition at a typical threshold (*P* value ≤ 0.05 and BICW ≥ 0.4) for datasets with altered sampling intervals (left), number of missing samples in each of the 4 conditions (middle) and number of replicates (right).
F-G. ROC curves illustrating false positive rate (FPR, 1-specificity) and true positive rate (TPR, sensitivity) to detect a 2-fold change in amplitude (F) and a 4 hours phase difference (G) between two conditions. We compared *dryR* with CircaCompare (default settings and an alpha threshold of 1), LimoRhyde and DODR. For *dryR*_threshold, only features with differences in amplitude ≥ 0.25 and phase ≥ 2 h were considered, respectively. In contrast to DODR and CircaCompare which use *P* values to indicate differential rhythmicity of a certain feature, *dryR* provides likelihoods for a set of fitted models, which can then be converted to probabilities using the Bayesian Information Criterion weight (BICW, see methods). The model with the highest BICW is selected. ROC curves for *dryR* differ in shape from the other methods as loosening the decision threshold for the *p*-value based methods will ultimately result in a TPR/FPR ratio of 1, as true positives and false positives will reach their maximum number. A loosening of BICW threshold will still result in a selection of a model and therefore will not result in a TPR/FPR of 1 as the detection of false positive and true positive feature will reach a plateau.
H-I. Level of TPR at an FPR of 0.05 for differential rhythmicity analysis for amplitude (H) and phase (I) in synthetic rhythmic datasets with altered sampling intervals (left), number of missing samples in each of the four conditions (middle) and number of replicates (right).

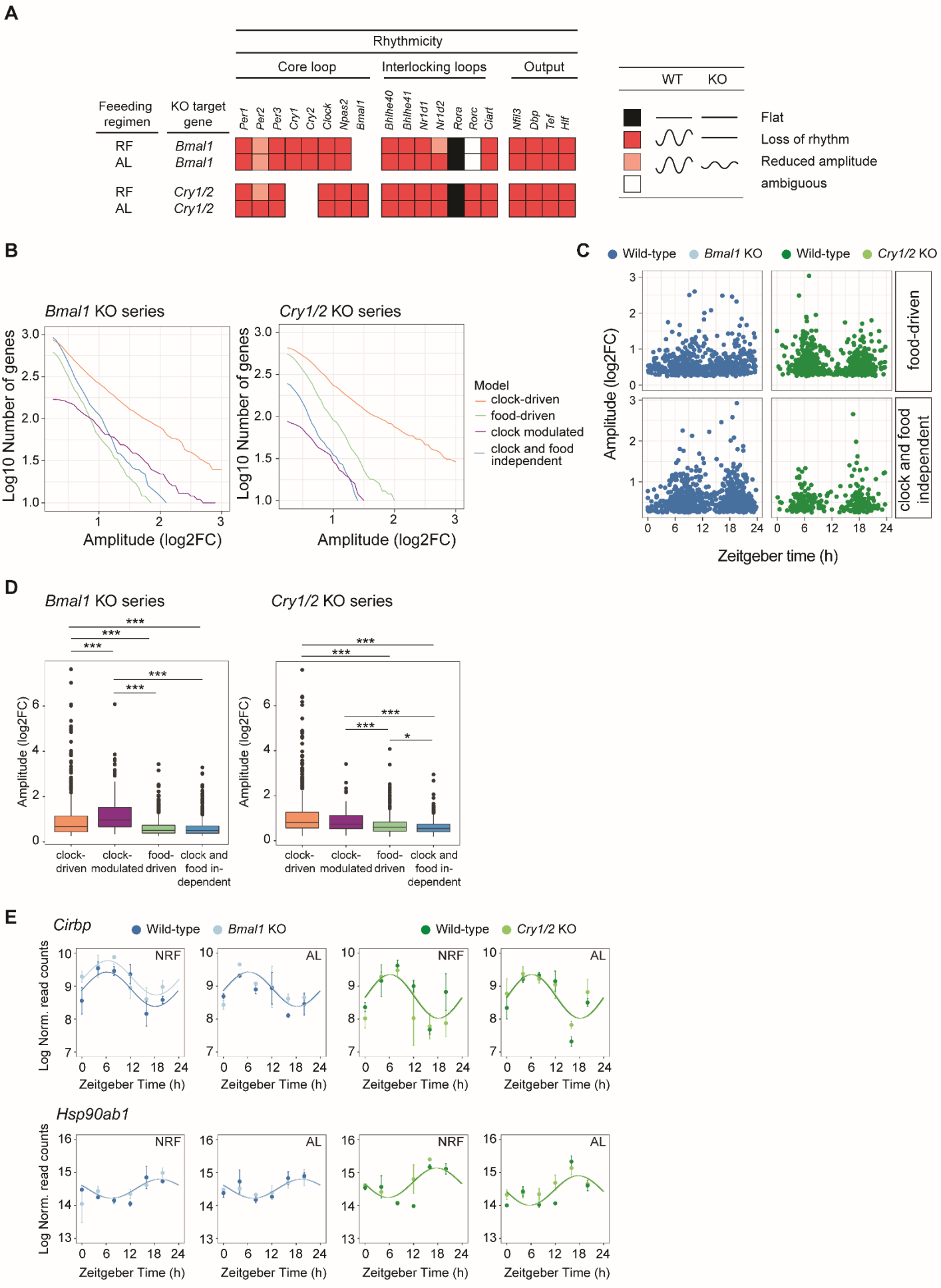

A
Rhythmicity
Core loop
Interlocking loops
Output
Feeeding regimen
KO target gene
Per1 Per2 Per3 Cry1 Cry2 Clock Npas2 Bmal1 Bhlhe40 Bhlhe41 Nr1d1 Nr1d2 Rora Rorc Ciart Nfil3 Dbp Tef Hlf
RF Bmal1
AL Bmal1
RF Cry1/2
AL Cry1/2
WT KO
Flat
Loss of rhythm
Reduced amplitude
ambiguous
B
Bmal1 KO series
Cry1/2 KO series
Log10 Number of genes
Amplitude (log2FC)
Model
clock-driven
food-driven
clock modulated
clock and food independent
C
Wild-type
Bmal1 KO
Wild-type
Cry1/2 KO
Amplitude (log2FC)
food-driven
clock and food independent
Zeitgeber time (h)
D
Bmal1 KO series
Cry1/2 KO series
Amplitude (log2FC)
clock-driven
clock-modulated
food-driven
clock and food independent
E
Cirbp
Wild-type
Bmal1 KO
Wild-type
Cry1/2 KO
NRF
AL
Log Norm. read counts
Zeitgeber Time (h)
Hsp90ab1

Fig. S2. Rhythmicity in *Bmal1* and *Cry1*/*2* KO mice in *ad libitum* (AL) or night restricted feeding (NRF) conditions.
A. Assignment of circadian clock genes to their corresponding temporal expression pattern (right).
B. Cumulative number of genes that show a larger amplitude than the value on the x-axis in the indicated model: circadian clock-driven genes (orange) have higher amplitudes than food-driven (green), clock and food independent rhythmic (blue), or clock-modulated genes (purple).
C. Amplitude *vs.* phase of rhythmic genes of indicated rhythmic model.
D. Amplitude represented as boxplot for the indicated model.
E. Temporal expression pattern of the cold inducible RNA-binding protein (*Cirbp)* and the heat shock protein *Hsp90ab1.* Rhythmic gene expression is food and clock independent for these genes as rhythmicity is kept under all conditions.

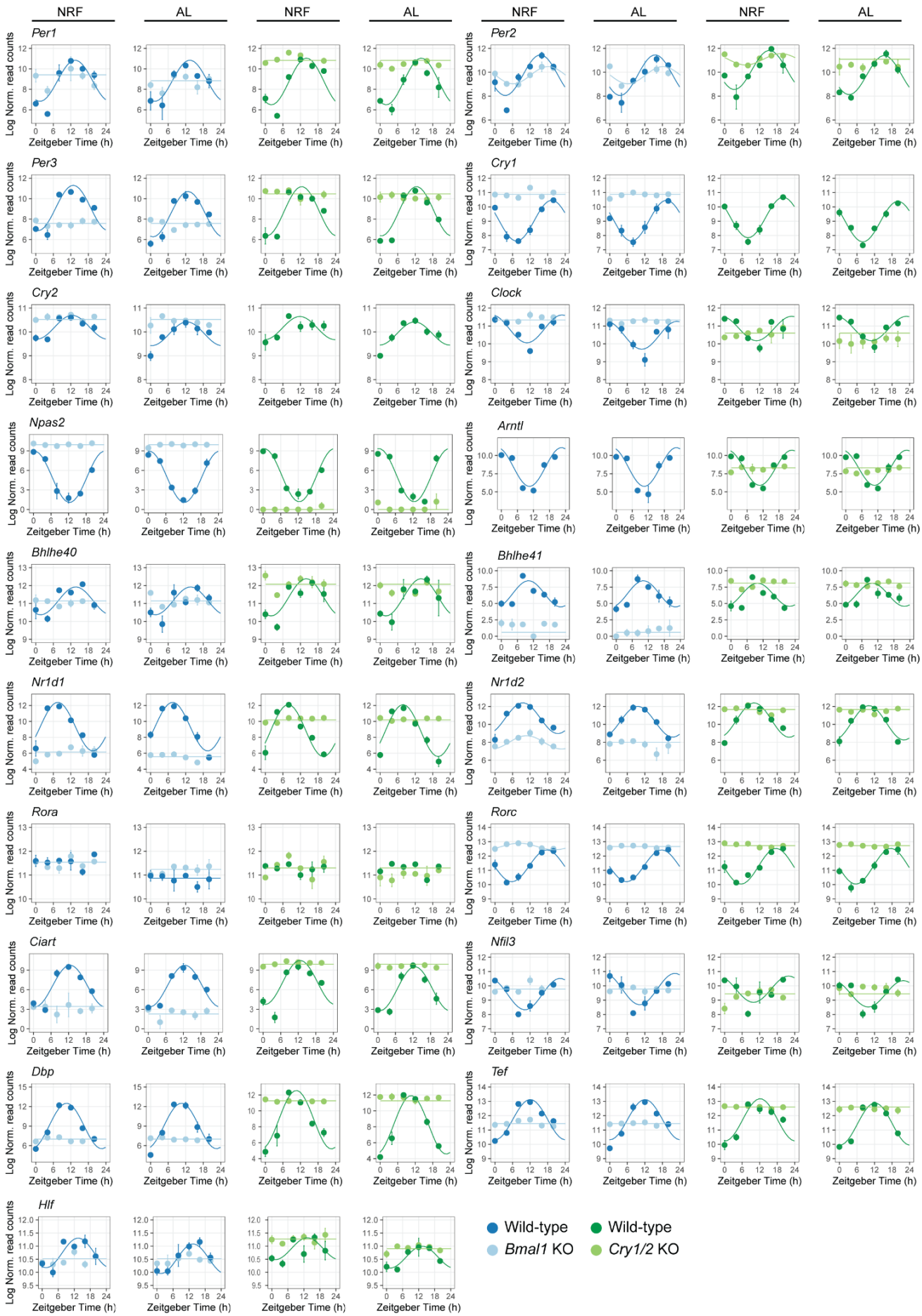

NRF
AL
NRF
AL
NRF
AL
NRF
AL
Per1
Per2
Per3
Cry1
Cry2
Clock
Npas2
Arntl
Bhlhe40
Bhlhe41
Nr1d1
Nr1d2
Rora
Rorc
Ciart
Nfil3
Dbp
Tef
Hlf
Log Norm. read counts
Zeitgeber Time (h)
Wild-type
Bmal1 KO
Wild-type
Cry1/2 KO

Fig. S3. Rhythmic expression of clock genes is blunted in the liver of *Bmal1* KO and *Cry1*/*2* KO mice, independent of feeding patterns. Circadian clock gene expression in *Bmal1* WT, *Bmal1* KO, *Cry1*/*2* WT, and *Cry1*/*2* KO under an *ad libitum* (AL) and night restricted feeding (NRF) regimen.

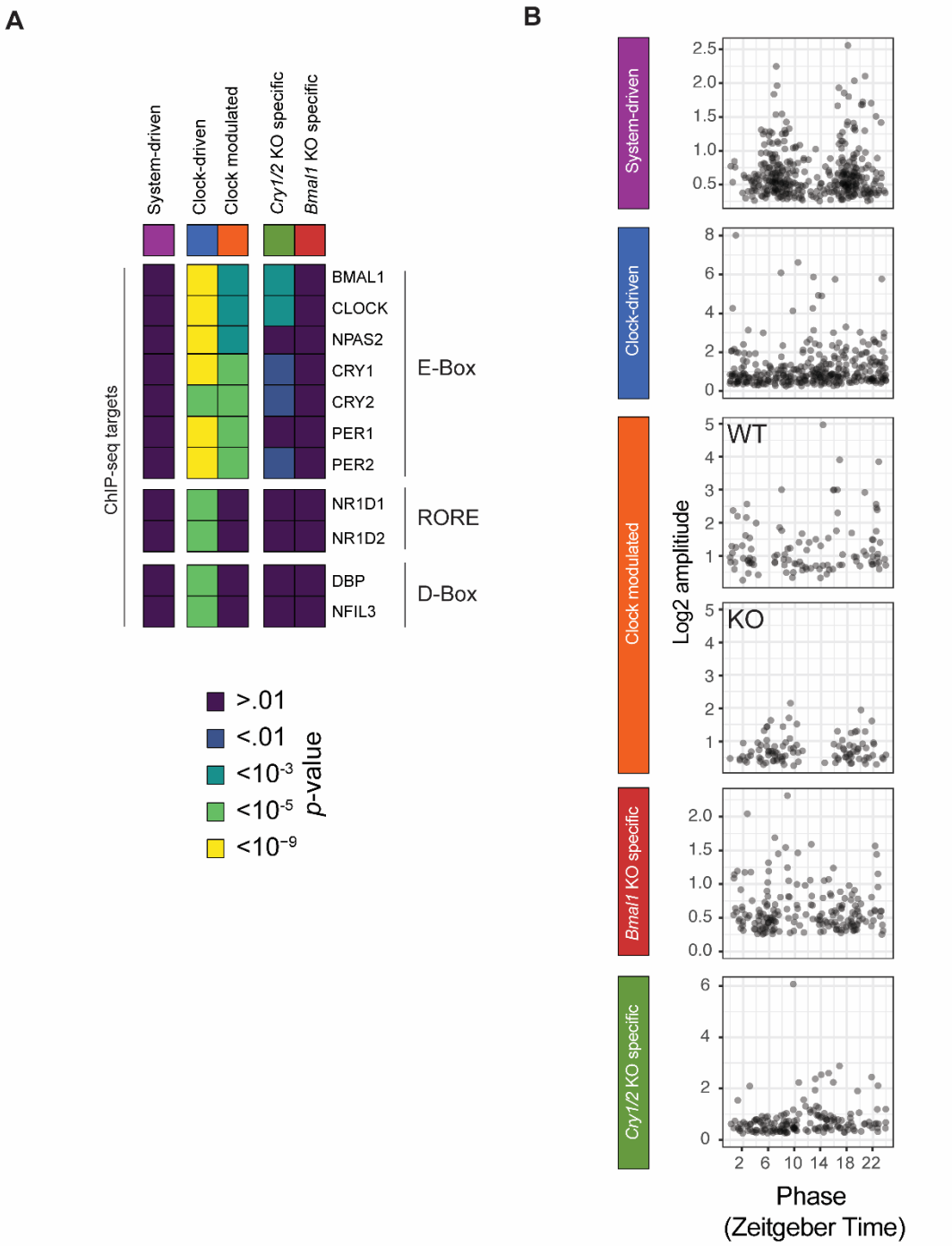

A
System-driven
Clock-driven
Clock modulated
Cry1/2 KO specific
Bmal1 KO specific
ChIP-seq targets
BMAL1
CLOCK
NPAS2
CRY1
CRY2
PER1
PER2
E-Box
NR1D1
NR1D2
RORE
DBP
NFIL3
D-Box
>.01
<.01
<10-3
<10-5
<10-9
p-value
B
System-driven
Clock-driven
Clock modulated
Bmal1 KO specific
Cry1/2 KO specific
WT
KO
Log2 amplitiude
Phase
(Zeitgeber Time)
C
***
**
Amplitude (log2FC)
clock-driven
clock-modulated
Cry1/2 KO specific
Bmal1 KO specific
System-driven
D
clock modulated
Amplitude KO (log2FC)
Amplitude WT (log2FC)

Fig. S4. System-driven and clock-driven hepatic gene expression under conditions of night restricted feeding.
A. Enrichment analysis of core circadian clock targets for genes classified into the indicated rhythmic model. Target genes have been identified using published chromatin immunoprecipitation sequencing (ChIP-Seq) data in mouse liver.
B. Amplitude *vs.* phase of rhythmic genes of indicated rhythmic model.
C. Amplitude represented as boxplots for indicated rhythmic model.
D. Most genes that show a consistent change in rhythmic expression in both KO model (clock modulated) show a declined amplitude in the absence of a functional circadian clock.

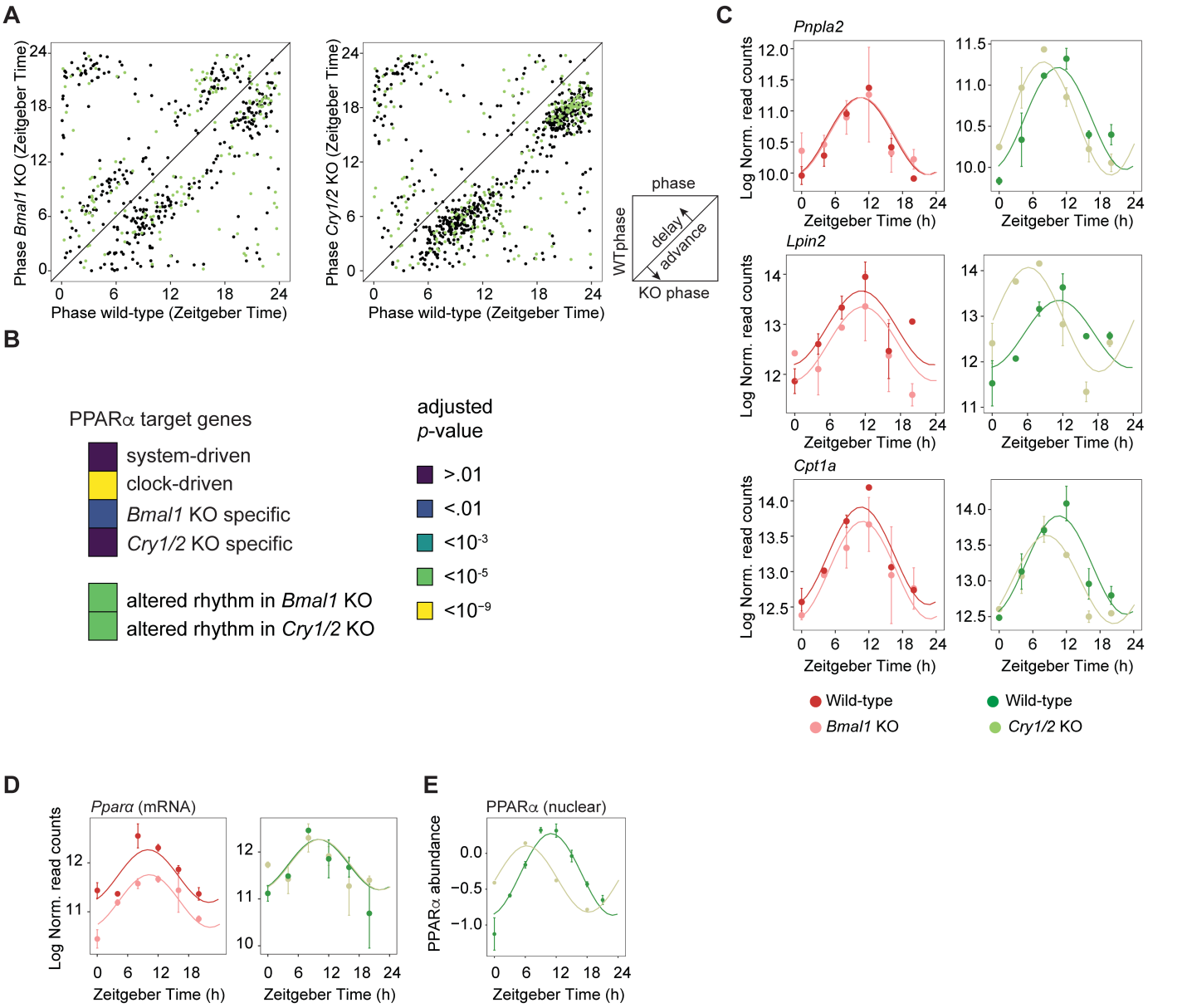
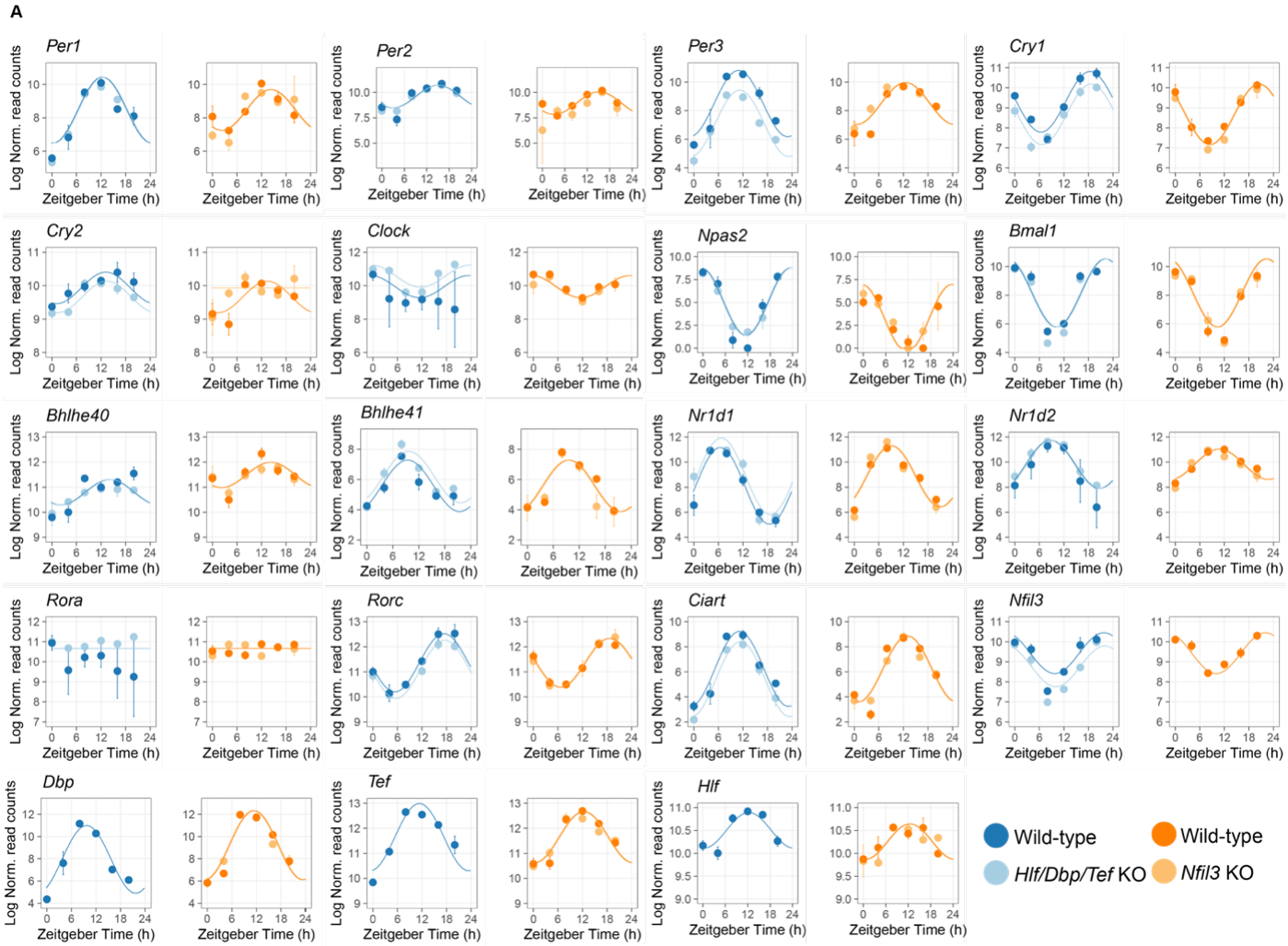
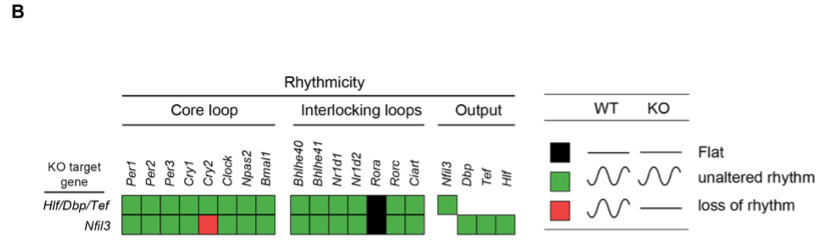
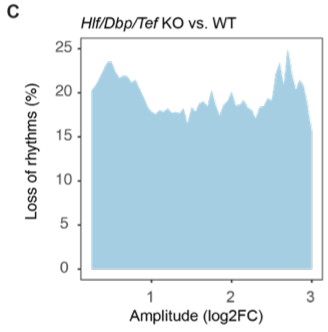
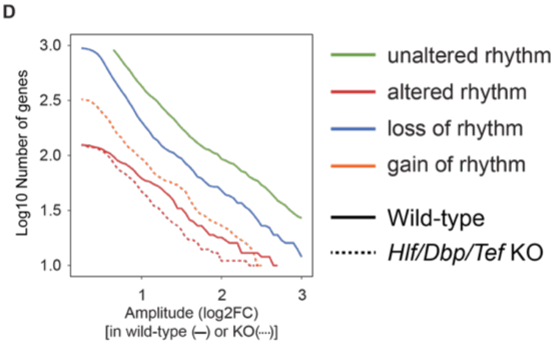


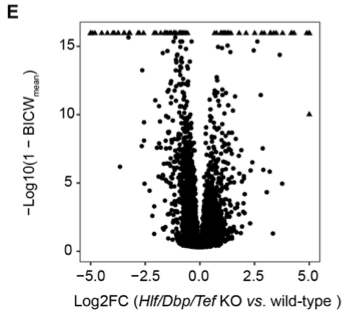
Fig. S5. *Cry1/2* KO but not *Bmal1* KO mice display a phase advance under night restricted feeding in liver gene expression
A. Phase distribution of differentially rhythmically expressed genes in *Bmal1* KO (left) and *Cry1/2* KO (right). *Bmal1* KO show a less pronounced phase advance than *Cry1/2* KO mice despite a matching NRF regimen.
B. Enrichment analysis of PPARα target genes that display differential rhythmicity in *Bmal1* KO and *Cry1/2* KO mice. Target genes have been identified by published chromatin immunoprecipitation sequencing (ChIP-Seq) data in mouse liver.
C. Representative examples of PPARα target genes that show a *Cry1/2* KO specific advance in phase in gene expression.
D. Hepatic *Pparα* expression exhibit lower mean levels in *Bmal1* KO animals.
E. *Cry1/2* KO mice show a specific phase advance in nuclear PPARα protein levels (data from ref. (28)).


A
Per1
Per2
Per3
Cry1
Cry2
Clock
Npas2
Bmal1
Bhlhe40
Bhlhe41
Nr1d1
Nr1d2
Rora
Rorc
Ciart
Nfil3
Dbp
Tef
Hlf
Log Norm. read counts
Zeitgeber Time (h)
Wild-type
Wild-type
Hlf/Dbp/Tef KO
Nfil3 KO


B
Rhythmicity
Core loop
Interlocking loops
Output
WT
KO
KO target gene
Per1
Per2
Per3
Cry1
Cry2
Clock
Npas2
Bmal1
Bhlhe40
Bhlhe41
Nr1d1
Nr1d2
Rora
Rorc
Ciart
Nfil3
Dbp
Tef
Hlf
Hlf/Dbp/Tef
Nfil3
Flat
unaltered rhythm
loss of rhythm


C
Hlf/Dbp/Tef KO vs. WT
Loss of rhythms (%)
Amplitude (log2FC)


D
Log10 Number of genes
Amplitude (log2FC)
[in wild-type (—) or KO(···)]
unaltered rhythm
altered rhythm
loss of rhythm
gain of rhythm
Wild-type
Hlf/Dbp/Tef KO


E
−Log10(1 − BICW mean)
Log2FC (Hlf/Dbp/Tef KO vs. wild-type)

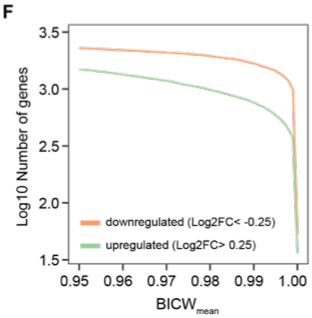

F
Log10 Number of genes
downregulated (Log2FC< -0.25)
upregulated (Log2FC> 0.25)
BICW mean

Fig. S6. Differential gene expression in the liver of *Hlf*/*Dbp*/*Tef* KO animals.
A. Clock gene expression in the liver of *Hlf*/*Dbp*/*Tef* WT, *Hlf*/*Dbp*/*Tef* KO, *Nfil3* WT, and *Nfil3* KO mice.
B. Assignment of circadian clock genes to their corresponding temporal expression pattern (right) in the indicated mouse model (row).
C. Relative number of genes that lose rhythmicity in *Hlf*/*Dbp*/*Tef* KO compared to wild-type in function of minimal amplitude.
D. Cumulative number of genes classified in indicated rhythmic model in function of minimal amplitude in wild-type or *Hlf*/*Dbp*/*Tef* KO animals.
E. Volcano plot of mean differences between *Hlf*/*Dbp*/*Tef* KO and WT mice.
F. Cumulative number of genes that are considered up or downregulated with a BICW equal or larger than the value indicated on the x-axis.

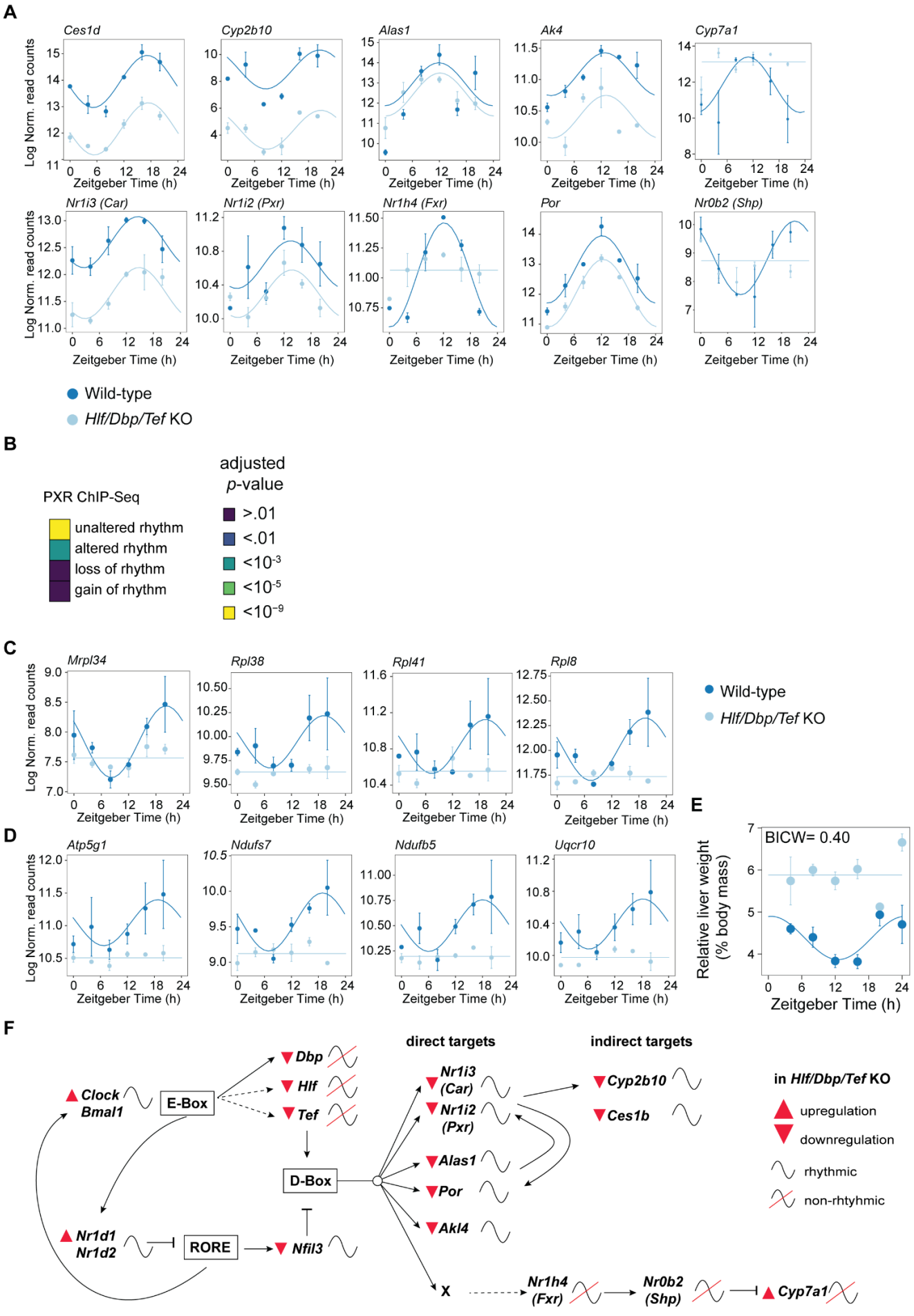
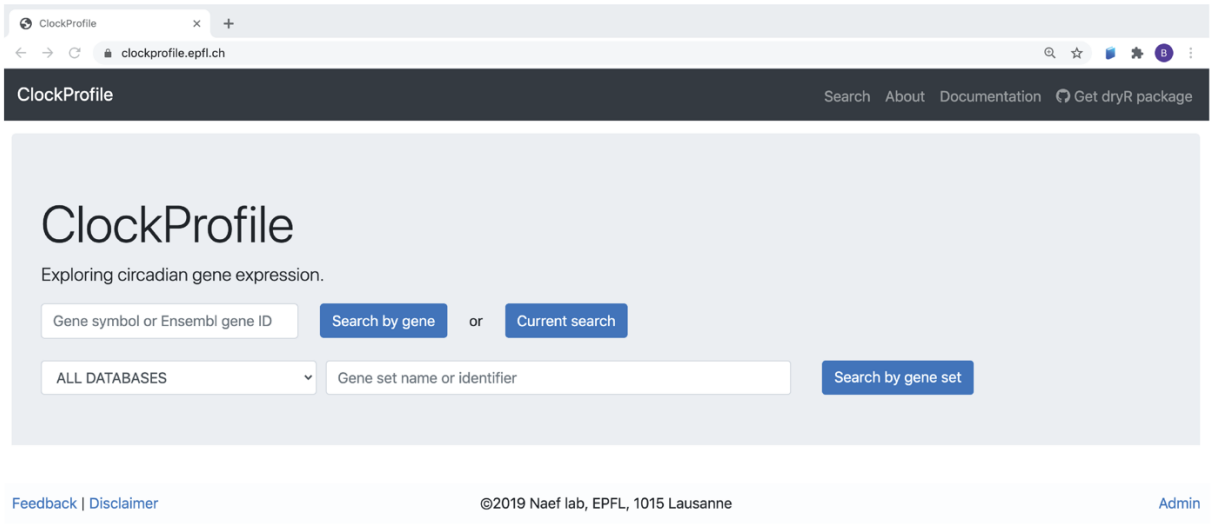
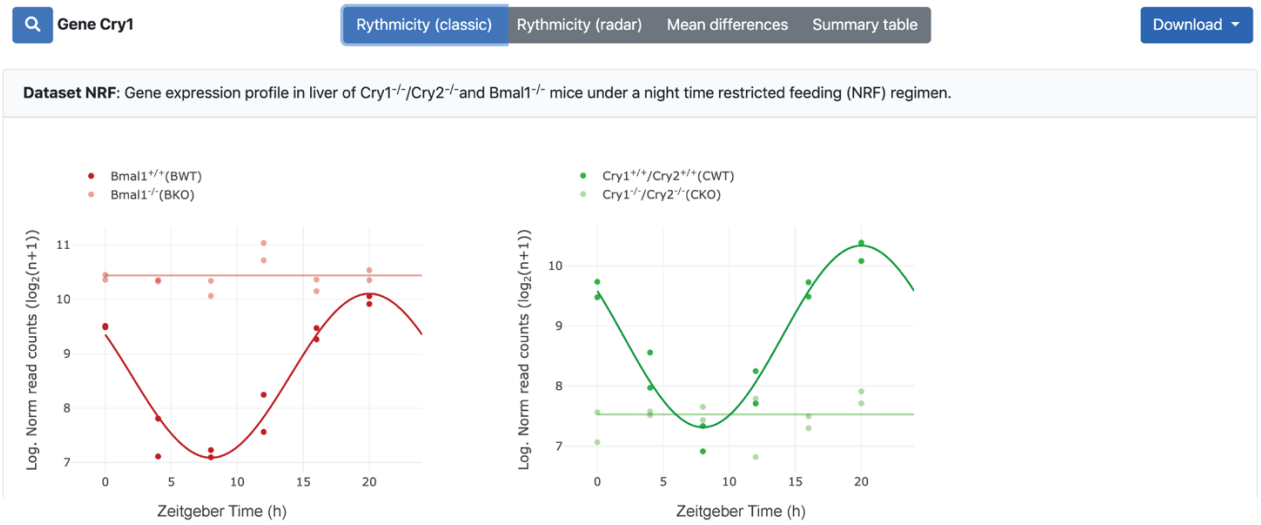

A
Ces1d
Cyp2b10
Alas1
Ak4
Cyp7a1
Nr1i3 (Car)
Nr1i2 (Pxr)
Nr1h4 (Fxr)
Por
Nr0b2 (Shp)
Log Norm. read counts
Zeitgeber Time (h)
Wild-type
Hlf/Dbp/Tef KO
B
PXR ChIP-Seq
unaltered rhythm
altered rhythm
loss of rhythm
gain of rhythm
adjusted p-value
>.01
<.01
<10-3
<10-5
<10-9
C
Mrpl34
Rpl38
Rpl41
Rpl8
D
Atp5g1
Ndufs7
Ndufb5
Uqcr10
E
Relative liver weight (% body mass)
BICW= 0.40
F
direct targets
indirect targets
Clock
Bmal1
E-Box
Dbp
Hlf
Tef
D-Box
Nr1d1
Nr1d2
RORE
Nfil3
Nr1i3 (Car)
Nr1i2 (Pxr)
Alas1
Por
Akl4
X
Nr1h4 (Fxr)
Nr0b2 (Shp)
Cyp7a1
Cyp2b10
Ces1b
in Hlf/Dbp/Tef KO
upregulation
downregulation
rhythmic
non-rhtyhmic

Fig. S7. Liver expression in *Hlf*/*Dbp*/*Tef* KO and WT mice.
A. Temporal profiles of genes associated with xenobiotic and bile acids metabolic network with altered expression in *Hlf*/*Dbp*/*Tef* KO mice.
B. Enrichment analysis of PXR targets for differentially rhythmically expressed genes in *Hlf*/*Dbp*/*Tef* KO and WT mice. Target genes have been identified by published chromatin immunoprecipitation sequencing (ChIP-Seq) data in mouse liver.
C-D. Gene expression of example genes associated with ribosomes (C) and oxidative phosphorylation (D).
E. The rhythmic fluctuation in relative liver weight (% body mass) is blunted in *Hlf*/*Dbp*/*Tef* KO mice.
F. Schematic representation of gene expression of xenobiotic metabolic gene network in the liver of *Hlf*/*Dbp*/*Tef* KO mice. Direction of triangles indicate if a gene is up (▲) or downregulated (▼) in *Hlf*/*Dbp*/*Tef* KOs.

A



B



C

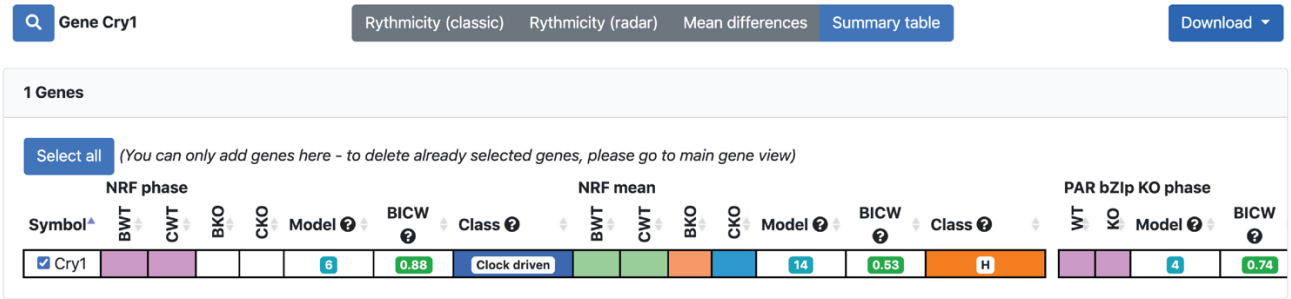


Fig. S8. The web app "ClockProfile" to assess temporal gene profiles in *Bmal1* KO, *Cry1/2* KO and PARbZip deficient mice. Screenshots from ClockProfile (https://clockprofile.epfl.ch/) illustrating features such as visualizing gene expression in WT and KO models, assessing statistical results from *dryR*, and browsing gene set enrichment analysis.
A. Homepage of ClockProfile with two options: search for 1. gene sets (search for gene set) or 2. individual genes (search by genes).
B. Gene expression visualization of *Cry1* as an example.
C. Summary table of *Cry1* as an example.

Dataset S1 (separate file). RNA-Seq analysis results of animal models with a genetically disrupted circadian clock under *ad libitum* or night restricted feeding regimen.
Sheet 1. Detailed table legend.
Sheet 2. Results of the rhythmicity analysis using multiple generalized linear regression and model selection in liver of WT and *Bmal1* KO series.
Sheet 3. Results of the rhythmicity analysis using multiple generalized linear regression and model selection in liver of WT and *Cry1*/*2* KO.

Dataset S2 (separate file). RNA-Seq analysis results of animal models with a genetically disrupted circadian clock under night restricted feeding regimen.
Sheet 1. Detailed table legend.
Sheet 2. Results of the rhythmicity analysis using multiple generalized linear regression and model selection in liver of WT*, Bmal1* KO and *Cry1*/*2* KO series.
Sheet 3. Functional enrichment analysis in system-driven genes.
Sheet 4. Functional enrichment analysis in clock-driven genes.
Sheet 5. Functional enrichment analysis in clock-modulated genes.
Sheet 6. Functional enrichment analysis of rhythmic genes that lose rhythmicity specifically in *Bmal1* KO.
Sheet 7. Functional enrichment analysis of rhythmic genes that lose rhythmicity specifically in *Cry1*/*2* KO.

Dataset S3 (separate file). RNA-Seq analysis results of *Hlf*/*Dbp*/*Tef* KO and WT under *ad libitum* feeding regimen.
Sheet 1. Detailed table legend.
Sheet 2. Results of the rhythmicity analysis using multiple generalized linear regression and model selection in liver of WT and *Hlf*/*Dbp*/*Tef* KO series.
Sheet 3. Functional enrichment analysis of genes that show an unaltered rhythmicity in *Hlf*/*Dbp*/*Tef* KO.
Sheet 4. Functional enrichment analysis of genes that lose rhythmicity in *Hlf*/*Dbp*/*Tef* KO.
Sheet 5. Functional enrichment analysis of genes that gain rhythmicity in *Hlf*/*Dbp*/*Tef* KO.
Sheet 6. Functional enrichment analysis of rhythmic genes with altered phase and/or amplitude in *Hlf*/*Dbp*/*Tef* KO.

SI References

1. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
2. T. Hubbard *et al.*, The Ensembl genome database project. *Nucleic Acids Res* **30**, 38-41 (2002).
3. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol* **11**, R106 (2010).
4. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
5. R. E. Kass, A. E. Raftery, Bayes Factors. *Journal of the American Statistical Association* **90**, 773-795 (1995).
6. M. Ashburner *et al.*, Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29 (2000).
7. A. Liberzon *et al.*, Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739-1740 (2011).
8. M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).
9. G. Joshi-Tope *et al.*, Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res* **33**, D428-432 (2005).
10. A. Lachmann *et al.*, ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* **26**, 2438-2444 (2010).
11. M. D. Young, M. J. Wakefield, G. K. Smyth, A. Oshlack, Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* **11**, R14 (2010).
12. J. M. Singer, D. Y. Fu, J. J. Hughey, Simphony: simulating large-scale, rhythmic data. *PeerJ* **7**, e6985 (2019).
13. M. E. Hughes, J. B. Hogenesch, K. Kornacker, JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms* **25**, 372-380 (2010).
14. E. F. Glynn, J. Chen, A. R. Mushegian, Detecting periodic patterns in unevenly spaced gene expression time series using Lomb-Scargle periodograms. *Bioinformatics* **22**, 310-316 (2006).
15. G. Wu, R. C. Anafi, M. E. Hughes, K. Kornacker, J. B. Hogenesch, MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics* **32**, 3351-3353 (2016).
16. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).
17. R. Parsons, R. Parsons, N. Garner, H. Oster, O. Rawashdeh, CircaCompare: a method to estimate and statistically support differences in mesor, amplitude and phase, between circadian rhythms. *Bioinformatics* **36**, 1208-1212 (2020).
18. J. M. Singer, J. J. Hughey, LimoRhyde: A Flexible Approach for Differential Analysis of Rhythmic Transcriptome Data. *J Biol Rhythms* **34**, 5-18 (2019).
19. P. F. Thaben, P. O. Westermark, Differential rhythmicity: detecting altered rhythmicity in biological data. *Bioinformatics* **32**, 2800-2808 (2016).
20. R. Edgar, M. Domrachev, A. E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**, 207-210 (2002).
21. N. Koike *et al.*, Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338**, 349-354 (2012).
22. H. Cho *et al.*, Regulation of circadian behaviour and metabolism by REV-ERB-alpha and REV-ERB-beta. *Nature* **485**, 123-127 (2012).
23. R. Leinonen *et al.*, The European Nucleotide Archive. *Nucleic Acids Res* **39**, D28-31 (2011).
24. H. Yoshitane *et al.*, Functional D-box sequences reset the circadian clock and drive mRNA rhythms. *Commun Biol* **2**, 300 (2019).
25. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
26. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).
27. J. T. Robinson *et al.*, Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).
28. J. Wang *et al.*, Nuclear Proteomics Uncovers Diurnal Regulatory Landscapes in Mouse Liver. *Cell Metab* **25**, 102-117 (2017).
29. J. A. Vizcaino *et al.*, ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* **32**, 223-226 (2014).