

Tissue-specific circadian transcriptional regulation

Thèse N° 8893

Présentée le 18 janvier 2019

à la Faculté des sciences de la vie

Unité du Prof. Naef

Programme doctoral en biotechnologie et génie biologique

pour l'obtention du grade de Docteur ès Sciences

par

JAKE YEUNG

Acceptée sur proposition du jury

Prof. B. Deplancke, président du jury

Prof. F. Naef, directeur de thèse

Prof. J. Hogenesch, rapporteur

Prof. E. van Nimwegen, rapporteur

Prof. J. Lingner, rapporteur

2019

"In the fields of observation luck favors the prepared mind."
— Louis Pasteur

"Beware of finding what you're looking for."
– Richard Hamming

To my parents...

Acknowledgements

During my PhD, I have been fortunate to have worked with the right people at the right time. First of all, this thesis would not have been possible without my advisor Felix. He has been more than I could have hoped for as a mentor. I am grateful for all of our discussions over the last four years on wide topics of science and also the deep but important details. Felix, you have taught me that treasure can be found by getting the details right. Your critical eye and sound judgement have guided me toward paths of exciting discoveries. You inspired me to maintain a high technical level while making the broadest impacts in science. I was always eager to soak up knowledge and seek wisdom from your discussions. If I could have gleaned even a tiny fraction of your wisdom and insights through osmosis over the last few years, I would be delighted. I would also like to thank Fred and Paul for having confidence in me to work on a variety of projects. Paul, thank you for showing me the world of sleep and for broadening my view of biology. Fred, I benefited from your deep insights on metabolism, physiology, and the circadian clock. Your encyclopedic knowledge of biology is inspirational.

I would also like to thank my colleagues over the years who have contributed to my growth over the last few years, Jingkui, Cédric, Daniel, Benjamin, Saeed, Nick, Damien, Clémence, and Eric. Our discussions have inspired many ideas. Special thanks to Sophie for helping me with administrative tasks, taking them off my plate so I can focus on my work. I would like to thank the facilities, both experimental and computational, for providing expertise and insights to technical problems.

I would like to give special thanks to Jérôme for a fruitful collaboration over the last few years. Your bold thinking and ability to sense the right experiments to do have benefited me greatly. I owe you a debt of gratitude for all of your work. Thank you also to Charlotte for letting me in on your experiments and inspiring new ways of thinking about dynamics with your experimental designs. Thank you Maxime for your work and analysis, which revealed treasure where I was unable to find on my own.

I would like to thank my family and friends both here and abroad for support over the years. Last but certainly not least, I would like to thank Bingqing for going on this journey with me these last several years. You have given me structure in my life and allowed me to focus on the important things. Thank you for interweaving your life with mine. Together this bond makes us stronger.

Lausanne, 23 August 2018

Jake Yeung

Abstract

Circadian rhythms in physiology and behavior evolved to resonate with daily cycles in the external environment. In mammals, organs orchestrate temporal physiology over the 24-hour day, which requires extensive gene expression rhythms targeted to the right tissue. Although a core set of gene products oscillate across virtually all cell-types, gene expression profiling across tissues over the 24-hour day showed that rhythmic gene expression programs are tissue-specific. We highlight recent progress in uncovering how the circadian clock interweaves with tissue-specific gene regulatory networks involving functions such as xenobiotic metabolism, glucose homeostasis, and sleep. This progress hinges on not only comprehensive experimental approaches but also computational methods for multivariate analysis of periodic functional genomics data. This thesis first explores how circadian gene expression is regulated across tissues. Second, we investigate how dynamic chromatin interactions underlie circadian gene transcription, core clock functions, and ultimately behavior. Third, we elucidate how the temporal transcriptome in mouse cortex responds to sleep deprivation. Finally, we discuss perspectives on extending the knowledge of the circadian clock in mice to human chronobiology.

Key words: systems chronobiology, chromatin interactions, gene regulation, feeding-fasting cycles, sleep-wake cycles

Résumé

Les rythmes circadiens de la physiologie ont évolué pour correspondre aux cycles quotidiens de l'environnement externe. Chez les mammifères, les organes orchestrent la physiologie temporelle sur une journée de 24 heures, ce qui nécessite des rythmes d'expression génique dans le bon tissu. Bien qu'un ensemble de base de produits génétiques oscillent dans pratiquement tous les types de cellules, le profil d'expression génétique dans différents tissus au cours des 24 heures a montré que certains programmes d'expression de gènes rythmiques sont spécifiques à des tissus. Nous mettons en évidence les progrès récents dans la découverte de l'imbrication de l'horloge circadienne avec les réseaux de régulation génique spécifiques aux tissus impliquant des fonctions telles que le métabolisme xénobiotique, l'homéostasie du glucose et le sommeil. Ce progrès repose non seulement sur des approches expérimentales complètes, mais aussi sur des méthodes informatiques pour l'analyse multivariée des données périodiques sur la génomique fonctionnelle. Cette thèse explore d'abord comment l'expression des gènes circadiens est régulée à travers les tissus. Deuxièmement, nous étudions comment les interactions de la chromatine dynamique sous-tendent la transcription des gènes circadiens, les fonctions de base de l'horloge et, en fin de compte, le comportement. Troisièmement, nous expliquons comment le transcriptome temporel dans le cortex de souris répond au traitement de privation de sommeil. Enfin, nous discutons des perspectives d'extension de la connaissance de l'horloge circadienne chez la souris à la chronobiologie humaine.

Mots clefs : systems chronobiology, chromatin interactions, gene regulation, feeding-fasting cycles, sleep-wake cycles

Contents

Acknowledgements	i
Abstract (English/Français)	iii
1 Introduction	1
1.1 Circadian rhythms at all scales of biological organization	1
1.1.1 Mammalian circadian timing is organized as a hierarchy of distributed clocks	2
1.1.2 The molecular components of the circadian clock	2
1.2 The function of the circadian clock is tissue-specific	3
1.3 Regulatory mechanisms underlying tissue-specific rhythmic gene expression .	4
1.3.1 Interactions between tissue-specific transcription factors and clock regu- lators	4
1.3.2 Rhythmic systemic signals can drive tissue-specific rhythms in gene ex- pression	6
1.3.3 Integrating temporal analysis of WT mice with clock-deficient mutants .	6
1.4 Statistical analysis of rhythms	7
1.4.1 Multivariate and modular analysis of rhythmic gene expression	7
1.4.2 Higher harmonics	7
1.4.3 Cell-type heterogeneity	9
1.4.4 Rhythmic gene expression in humans	9
1.4.5 Dynamic responses to acute perturbations	9
1.4.6 Mapping chromatin interactions	10
1.4.7 Chromatin interactions and the circadian clock	10
1.4.8 Dynamic chromatin interactions in cell culture models	10
1.4.9 Tissue-specific and dynamic chromatin interactions in tissues	11
1.4.10 Dynamic chromatin interactions and transcription	11
1.5 Specific questions addressed in thesis	13
2 Transcription factor activity rhythms and tissue-specific chromatin interactions ex- plain circadian gene expression across organs	15
2.1 Introduction	15
2.2 Extended methods: model selection with Zellner's g -priors	15

Contents

2.2.1	Enumerating harmonic regression models to identify possible combinations of rhythms across tissues	15
2.2.2	Selecting models by the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and Bayes Factors	17
2.3	Extended methods: complex-valued singular value decomposition	20
2.4	Extended methods: Gene ontology (GO) analysis around the clock	21
2.5	Contributions	22
2.6	Conclusion and perspectives	52
3	Clock-dependent chromatin topology modulates circadian transcription and behavior	53
3.1	Introduction	53
3.2	Contributions	53
3.3	Conclusion and perspectives	86
4	Dynamic gene expression and regulation from circadian and sleep-wake processes in the mouse cortex	87
4.1	Introduction	87
4.2	Contributions	87
4.3	Conclusion and perspectives	142
5	Conclusions and Perspectives	143
	Bibliography	161
	Curriculum Vitae	163

1 Introduction

Overview of the thesis

The thesis begins with an introduction that uses ideas and text that are discussed in a review paper in revision. References in this section to Yeung et al. and Mermet et al. highlight how the work in this thesis fits into the current challenges of chronobiology (Yeung et al., 2018; Mermet et al., 2018). This section concludes with three broad questions tackled in the thesis. After the introduction are three papers investigating regulation of gene regulatory dynamics. The first paper (Yeung et al., 2018) focuses on tissue-specific regulation. The second paper (Mermet et al., 2018) goes into details on the role of chromatin interactions in regulating the circadian clock. And the third paper, under preparation, explores how perturbing oscillatory gene expression by a short sleep deprivation treatment can reveal insights into how dynamic gene expression is regulated in the mouse cortex. The thesis concludes with the broad significance of my work and outstanding questions for the field.

1.1 Circadian rhythms at all scales of biological organization

Many organisms exhibit rhythms in behavior and physiology that are synchronized to the daily cycles in the environment (Bell-Pedersen et al., 2005). This internal circadian rhythm (period length of around one day) resonates with daily cycles, predicting and adapting to external cues such as light abundance and food availability (Schibler et al., 2015). Competition experiments in cyanobacteria suggest that this resonance enhances fitness (Ouyang et al., 1998). In mammals, a genetically encoded molecular clock oscillates in virtually every cell of the body (Takahashi, 2017). These clocks are synchronized to an internal body rhythm set by the master pacemaker, located in the suprachiasmatic nucleus (Hastings et al., 2018). Chronobiology researchers study how 24-hour periodicity in external signals is dynamically integrated at all scales of biological organization, from gene expression to behavior, promising to yield new insights on how temporally structured external signals such as artificial light, unconventional meal times, and drug therapies dynamically interact with circadian rhythms

in physiology (Roenneberg and Merrow, 2016).

1.1.1 Mammalian circadian timing is organized as a hierarchy of distributed clocks

In mammals, the circadian timing system is organized as a hierarchy of oscillators, containing a master clock and peripheral clocks. The master clock, located in the suprachiasmatic nucleus (SCN), establishes rhythms in behavior by synchronizing peripheral clocks, located outside of the SCN, through systemic signals such as hormones, body temperature, and direct innervation (Mohawk et al., 2012). As insinuated by their name, circadian rhythms (*'circa diem'* means 'approximately a day') do not have a period of exactly 24 hours but differ between species (Bell-Pedersen et al., 2005), individuals as well as age (Czeisler et al., 1999). For example, the free-running circadian period for humans is longer than 24 hours while for the mouse is shorter. Therefore, the circadian timing cannot measure 24 h with accuracy (although it is very precise) but is periodically synchronized to the environment by external cues. These rhythmic cues are called zeitgebers. A prominent one is external sunlight light, which synchronizes the central clock in the SCN to geophysical time. Other zeitgebers include external temperature and feeding-fasting rhythms, which contribute to entrainment of clocks in tissues outside of the SCN (Damiola et al., 2000; Saini et al., 2011).

1.1.2 The molecular components of the circadian clock

Single-cell analyses have shown that circadian oscillations rely on a cell-autonomous clock that is genetically encoded, suggesting that potentially every cell in the body resonates with environmental time (Nagoshi et al., 2004). A widespread, though not necessarily definitive, model of the mammalian circadian clock consists of interlocked transcriptional and translational feedback loops that regulate the expression of core clock genes (Dibner et al., 2010). The genetic components and regulation of the circadian clock have been discussed in several excellent reviews (Takahashi, 2017; Mohawk et al., 2012; Dibner et al., 2010). Briefly, activating transcription factors (TFs) drive expression of many genes, including its own repressors, which decreases the accumulation of activating TFs and subsequently its repressors. The low accumulation of negative TFs derepress the activating TFs and a new cycle begins. A simplified diagram of two interlocked negative feedback loops is shown in Figure 1.1. Core clock transcription factors BMAL1 and CLOCK heterodimerize (BMAL1-CLOCK) and activate transcription of genes containing E-box (enhancer box) *cis*-regulatory elements. Important targets of BMAL1-CLOCK are the core clock genes, *Period* (*Per1*, *Per2*, *Per3*) and *Cryptochrome* (*Cry1*, *Cry2*), whose protein products (CRY-PER complexes) directly abrogates the transcriptional activity of BMAL1-CLOCK (Lowrey and Takahashi, 2004). Additionally, BMAL1-CLOCK activates transcription of another clock gene *Rev-erba*, whose protein product binds to ROREs (ROR response elements) and thus inhibits transcription of the *Bmal1* gene (Partch et al., 2014). These transcriptional and translational negative feedback loops establish the circadian clock (Yoo et al., 2004).

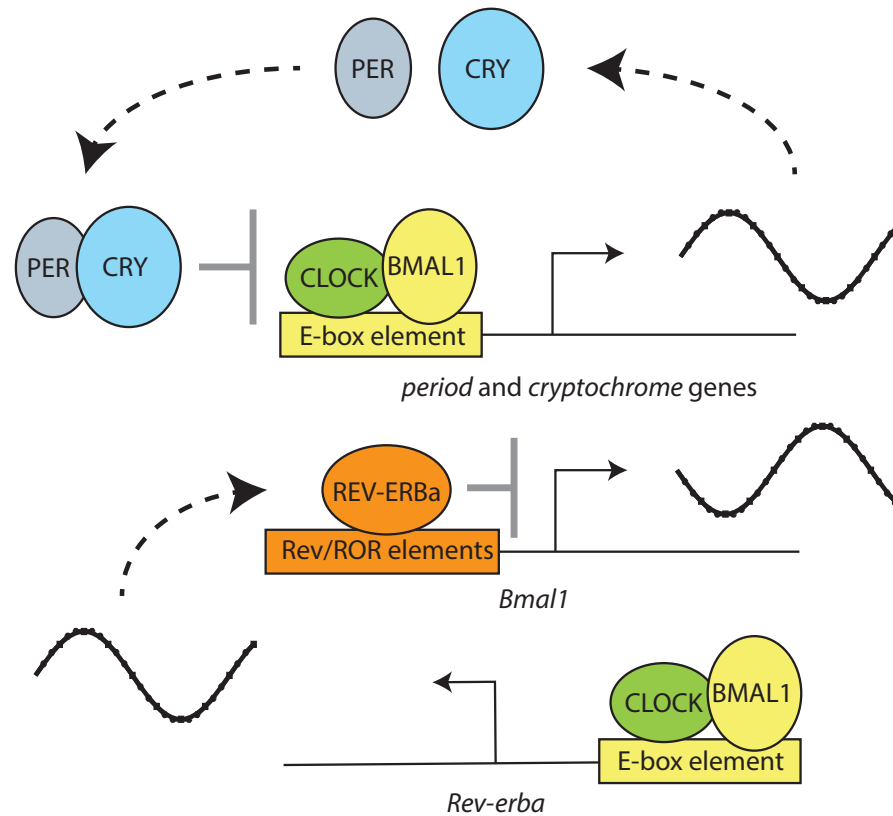


Figure 1.1 – Simplified diagram of the molecular clock. Two interlocking feedback loops robustly oscillate with a circadian period (~24 hours). Transcription factors (TF) CLOCK and BMAL1 bind to E-box elements (a TF binding motif) and activate the expression of many genes, including its own repressors. One set of repressors are period and cryptochrome genes, which are translated and subsequently abrogates the activity of the CLOCK:BMAL1 complex. A second loop, consisting of REV-ERBa, directly binds to RORE (also a TF binding motif) and represses the transcription of the gene, *Bmal1* (also known as *Arntl*), which codes for the positive TF.

Extensive studies mutating different components of the clock in mouse have identified a wide range of period phenotypes. A comprehensive table in a review by Ko and Takahashi summarizes clock mutants and their period phenotypes (Ko and Takahashi, 2006). For example, *Cry1* knockout (KO) mice show a 1-hour shorter period while *Cry2* KO mice show a 1-hour longer period. *Bmal1* KO mice are arrhythmic. These studies show that the circadian clock relies on a complex molecular architecture.

1.2 The function of the circadian clock is tissue-specific

Although this molecular clock ticks in virtually all cells of the body (Dibner et al., 2010), early transcriptome profiling studies around the clock in different mouse tissues have found that rhythmic gene expression is highly tissue-specific (Storch et al., 2002). How this clock is used

in different cell types and tissues to regulate diverse physiological processes is at the core of contemporary chronobiology. Daily rhythms in cell-type specific functions are pervasive. For example, recent work showed that fibroblasts exhibit circadian rhythm in actin dynamics, allowing wound healing to be more efficient during the active versus resting phase (Hoyle et al., 2017). Size of liver cells, its protein accumulation, and ribosome number oscillate over the 24-hour day (Gerber et al., 2013; Sinturel et al., 2017), making metabolism and xenobiotic detoxification more efficient when animals are active and feed. Macrophages exhibit circadian rhythm in cytokine production (Keller et al., 2009), which may improve response against bacterial infections.

Intriguingly, disrupting the clock in different tissues can lead to opposite phenotypes (Bass and Lazar, 2016). For example, pancreas-specific ablation of the clock leads to hyperglycaemia (Marcheva et al., 2010) while liver-specific ablation leads to hypoglycaemia (Lamia et al., 2008). Alternatively, tissue-specific rescue of the clock in whole-body knockouts can also reveal surprises. For example, whole-body *Bmal1* KO mice have increased non-rapid eye movement (NREM) sleep, and restoring *Bmal1* expression specifically in skeletal muscles can rescue the amount of NREM sleep, although not the timing (Ehlen et al., 2017). These tissue-specific phenomena are likely underpinned by tissue-specific rhythmic gene expression signatures. Indeed, a high-resolution circadian gene expression atlas across twelve organs has shown that mRNA abundances oscillate mostly in an organ-specific manner (Zhang et al., 2014), indicating that the circadian clock regulatory network interweaves with tissue-specific gene regulatory mechanisms.

1.3 Regulatory mechanisms underlying tissue-specific rhythmic gene expression

The molecular circadian clock and systemic signals likely contribute to tissue-specific rhythmic gene expression (Hughes et al., 2012; Kornmann et al., 2007) (Figure 1.2A). What are the regulatory mechanisms? Liver has been a productive model to study temporal gene regulation across multiple omics data types, from transcriptional (Le Martelot et al., 2012; Sobel et al., 2017; Trott and Menet, 2018), post-transcriptional (Luck et al., 2014), translational (Atger et al., 2015), to post-translational (Mauvoisin et al., 2017; Robles et al., 2017; Wang et al., 2017). Extending such analyses to multiple tissues will uncover general principles of how tissues physiologically interact with each other and with the circadian clock.

1.3.1 Interactions between tissue-specific transcription factors and clock regulators

Recently, temporal RNA-seq analysis of multiple mouse tissues suggested that interactions between tissue-specific and a clock transcription factor (TF) underlies tissue-specific rhythmic transcription regulation (Yeung et al., 2018) (Figure 1.2B). The binding of a tissue-specific

1.3. Regulatory mechanisms underlying tissue-specific rhythmic gene expression

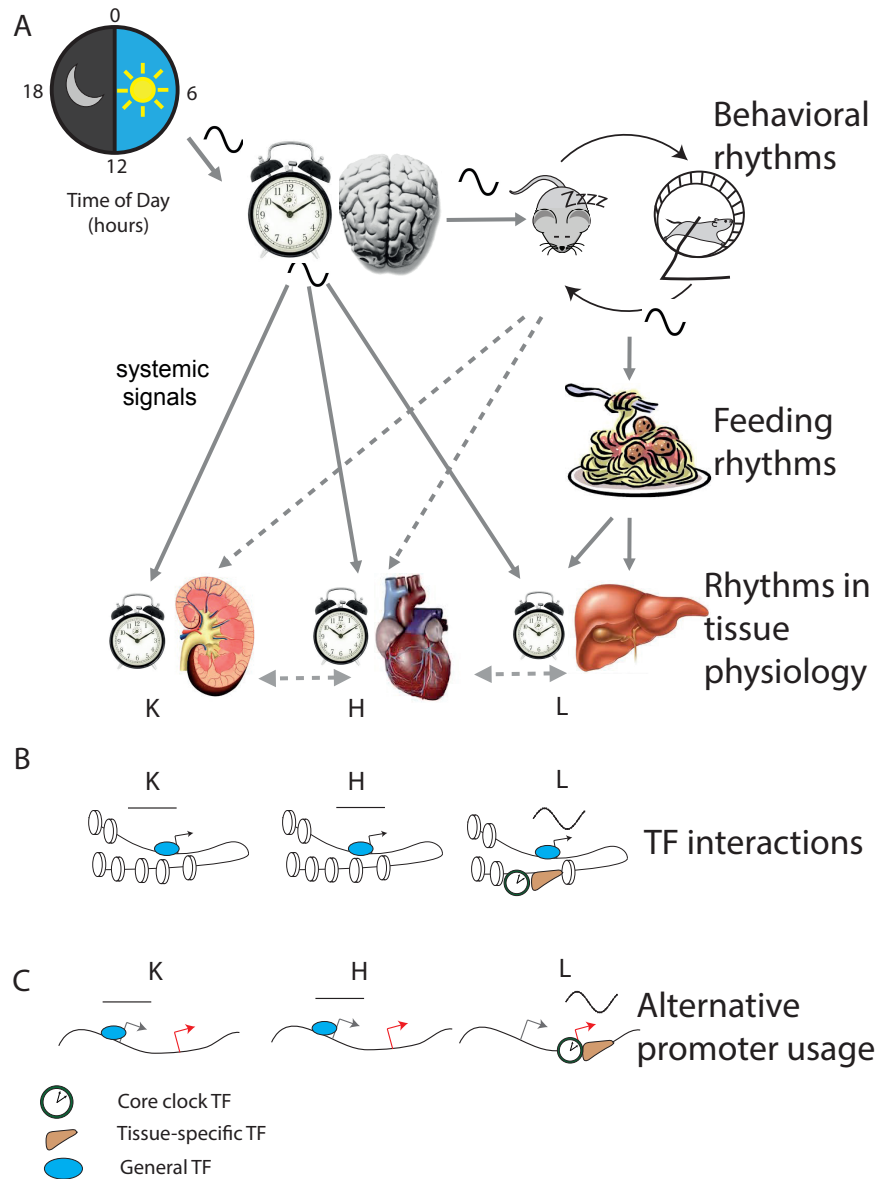


Figure 1.2 – Generation of tissue-specific rhythms. (A) Scheme of how external rhythmic cues entrain circadian rhythms across tissues. The central clock, located in the suprachiasmatic nucleus of the hypothalamus, takes light from the environment and synchronizes clocks in peripheral tissues through by regulating locomotor activity and feeding rhythms as well as through systemic signals such as hormones and metabolites. Peripheral tissues, such as kidney, heart, and liver, orchestrate rhythms in tissue physiology, such as sodium homeostasis, carbohydrate metabolism, and blood pressure, respectively. (B,C) Examples of how tissue-specific rhythmic gene expression can be generated transcriptionally. (B) Interactions between tissue-specific and clock transcription factors (TFs) can generate tissue-specific gene expression. For example, a gene can be rhythmically transcribed in one tissue but not in others by the presence or absence of a tissue-specific TF, which renders a nearby clock TF binding site accessible. (C) Different tissues can regulate rhythmic transcription of a gene by using different alternative promoters. In this example, the rhythmically transcribed promoter is used in liver but not in other tissues. A rhythmically transcribed gene is shown with a sinusoid, flat transcription is shown with a flat horizontal line.

TF may render the local chromatin region more accessible, increasing the affinity of clock TFs to their respective binding sites. For example, liver-specific rhythmic gene expression has been found to be enriched for liver-specific DNase-I hypersensitive sites containing binding sites for both liver-specific and clock TFs (Yeung et al., 2018). Direct protein-protein interactions between clock and other transcriptional regulators may also regulate rhythmic gene expression in a tissue-specific manner. For example, co-immunoprecipitation studies examining interactions of mouse nuclear receptors (NR) with CRY1 found that one-third of mouse NRs, some of which were expressed in a tissue-specific manner such as PXR and CAR (which regulates xenobiotic detoxification in liver) interact with CRY1 (Kriebs et al., 2017; Lamia et al., 2011). Another mechanism generating specificity in transcription dynamics is alternative promoter usage, where one transcript is rhythmically transcribed while the other is not, which explains a fraction of tissue-specific rhythmic transcript abundances (Figure 1.2C) (Yeung et al., 2018). Further analyses estimating the frequency of each mechanism that defines the rhythmic transcriptome will clarify which regulatory modes are most prevalent.

1.3.2 Rhythmic systemic signals can drive tissue-specific rhythms in gene expression

Generally, it is still unclear how rhythmic systemic signals such as hormones and metabolites can drive responses of varying magnitudes or amplitudes, and also activate different downstream pathways depending on the tissue context (Bass and Takahashi, 2010) (Figure 1.2A). One explanation is that signaling molecules such as hormones act in a tissue-specific manner; for example, fibroblast growth factor hormones regulate bile acid homeostasis in liver but contribute to thermogenesis in brown adipose tissue (Owen et al., 2015). Temporal profiles of serum concentrations could also differ across the body; for example, human IL-6 levels in cerebrospinal fluid versus plasma show distinct rhythms and peaked at different times of day (Agorastos et al., 2014), suggesting that dynamics in permeability of the blood-brain barrier can regulate distribution of cytokines across the body (Pan and Kastin, 2017). Thus tissue-specific decoding of systemic signals or local variation in ligand concentrations can induce transcription factor activities that oscillate in one tissue but not others.

1.3.3 Integrating temporal analysis of WT mice with clock-deficient mutants

Since rhythmic gene expression in organs can be generated through systemic cues as well as by the local clock (Hughes et al., 2012; Kornmann et al., 2007; Lamia et al., 2008), study designs incorporating wild-type (WT) and clock-deficient mutants allow clock-dependent and independent mechanisms to be distinguished (Atger et al., 2015), as demonstrated in a study of WT and *Bmal1* KO clock mutants in liver and kidney (Yeung et al., 2018).

1.4 Statistical analysis of rhythms

1.4.1 Multivariate and modular analysis of rhythmic gene expression

Many genome-wide descriptions of circadian or diurnal rhythmicity have considered genes one at a time, and typically in one or few conditions (Hutchison et al., 2015). To assess rhythmic gene expression in a given condition, parametric (Atger et al., 2015; Fisher, 1929) and non-parametric (Hughes et al., 2010; Hutchison et al., 2015; Thaben and Westermark, 2014) methods have been used to statistically test, for each gene, whether a measured temporal pattern shows evidence for rhythmicity compared to a null model (e.g. flat gene expression in time).

Today, an increasing number of datasets include multiple conditions, such as multiple tissues, genotypes, feeding conditions, or combinations thereof (Korenčič et al., 2014; Atger et al., 2015; Zhang et al., 2014). Analysis methods that integrate oscillatory patterns across multiple conditions could reveal novel patterns not easily identified from standard analyses. One approach that extends classic tests to two conditions is the Chow test (Chow, 1960). Beyond two conditions, there is the model selection method, where multi-condition data is fit to a set of models representing combinations of rhythmic and non-rhythmic outcomes across the conditions. The best model is identified by balancing goodness of fit and model complexity, such as the Bayesian Information Criterion (Figure 1.3A) (Atger et al., 2015).

Another area of interest is constructing low-dimensional representations of periodic data across tissues. One method, complex-valued singular value decomposition (cv-SVD), extends conventional real-valued SVD analysis to periodic data. cv-SVD can identify gene sets with large 24-hour amplitudes and phase shifts across many conditions (Figure 1.3B, details in Section 2.3) (Yeung et al., 2018). For example, modules of genes with coherent phase and amplitude relations, which occurs often in chronobiology datasets, can be adequately captured in low-dimensional complex-valued representations. Other methods, such as Zeitzeiger (Hughey et al., 2016), can identify sparse sets of genes that oscillate across conditions which has been successfully applied to a NanoString assay, called BodyTime, to predict internal circadian time from blood monocytes (Wittenbrink et al., 2018). Interestingly, new methods such as Oscope and CYCLOPs can now reconstruct cyclic dynamics from unlabeled data. This enables identifying oscillating genes in datasets without explicit time labels, such as in large-scale human tissue gene expression datasets (Anafi et al., 2017; Leng et al., 2015; Ruben et al., 2018).

1.4.2 Higher harmonics

Although 24-hour rhythms contribute the largest temporal variance in circadian datasets across tissues, 12-hour rhythms (also classified as ultradian rhythms) are often the second largest (Hughes et al., 2009; Yeung et al., 2018). Methods that include linear combinations of higher harmonics can systematically analyze ultradian dynamics, defined as a period length

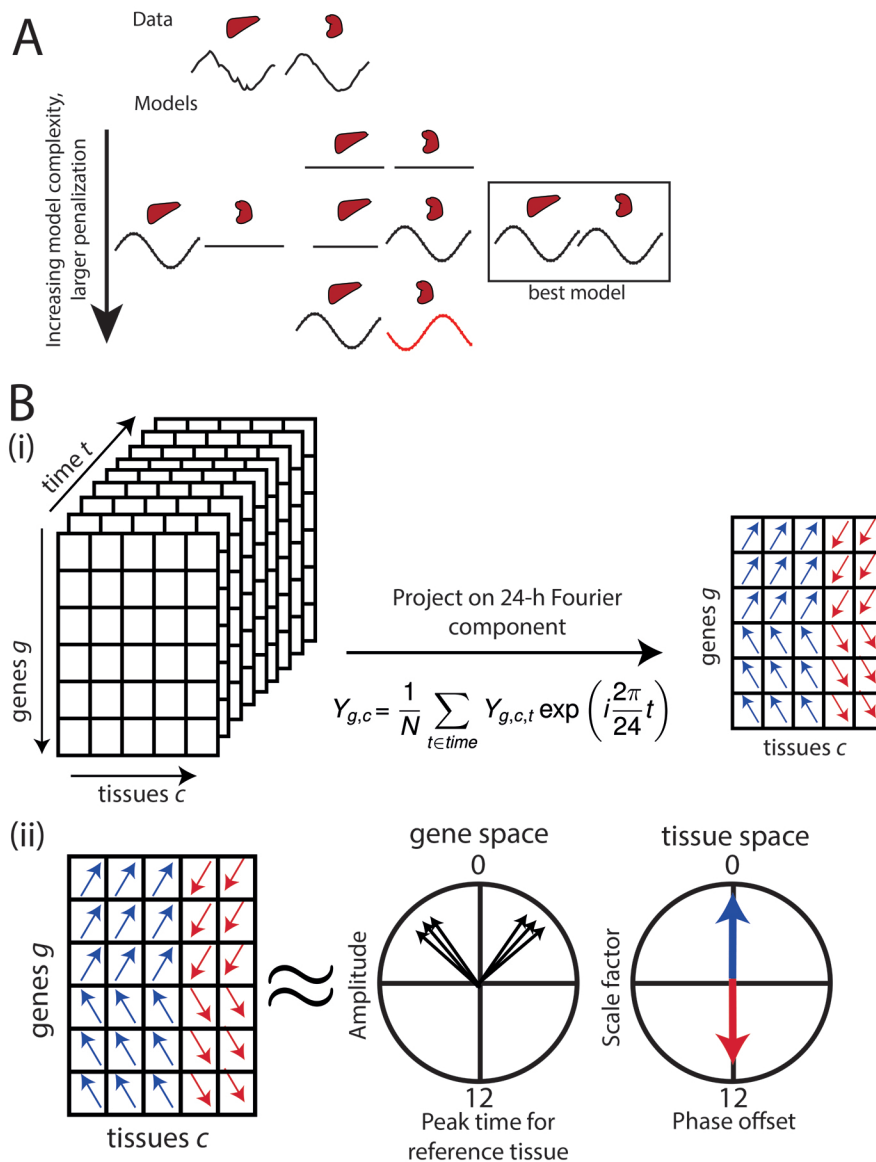


Figure 1.3 – Computational methods integrating multiple genes and conditions. (A) Model selection methods can integrate rhythms across multiple conditions (e.g. tissues). In this toy example, noisy data (shown as a roughened sinusoid) come from rhythmic gene expression from two tissues that oscillate in synchrony. The data is fit to each rhythmic model and the best model is selected. In this example, the best model is a model of intermediate complexity where rhythmic parameters are shared across two tissues. More complex models fit the data better, but are penalized for having more parameters. (B) Complex-valued singular-value decomposition (cv-SVD) allows rhythmic gene expression datasets to be factorized into gene space and tissue spaces. This representation identifies modules of genes with coherent phase and amplitude relationships across tissues. (i) First, the 3-dimensional matrix of real numbers are projected onto the 24-hour Fourier component, resulting in a 2-dimensional matrix of complex numbers (represented as arrows pointing to time of peak gene expression, length of arrow represent oscillation amplitudes). (ii) The 2D matrix is then decomposed into a gene space and a tissue space by SVD (conventional implementations, for example in R, can handle complex numbers). The gene space shows the amplitude and phase for a reference tissue. Amplitude and phase of genes in other tissues are multiplied and added, respectively, based on the tissue loading. In the reality, the tissue space is often in phase. The antiphase example highlights the method.

shorter than a day (Costa et al., 2013; Zhu et al., 2017). In liver, clock-dependent (Cretenet et al., 2010; Westermarck and Herzel, 2013) and independent (Zhu et al., 2017) mechanisms of 12-hour rhythms have been studied. Mechanisms include pairs of clock TFs (Westermarck and Herzel, 2013), or ultradian regulation of stress response pathways such as the unfolded protein response in the endoplasmic reticulum (Cretenet et al., 2010). Some of these ultradian rhythms consist of two peaks of mRNA expression per day, which interestingly, revert to 24-hour rhythms in clock mutants (Atger et al., 2015; Cretenet et al., 2010; Hughes et al., 2009), indicating that the two peak times are differentially controlled by the clock and feeding cycles. Extending the analysis of higher harmonics to multiple tissues may reveal interactions between tissue-specific TFs with regulators of ultradian rhythmicity.

1.4.3 Cell-type heterogeneity

The degree of cell type heterogeneity and structure varies from tissue to tissue (Han et al., 2018; Yang et al., 2014), which needs to be considered if this heterogeneity changes over time (Scheiermann et al., 2013). Certain immune cell types, such as T-cells and macrophages, are recruited to tissues in a circadian manner (Keller et al., 2009; Scheiermann et al., 2012), suggesting that cellular heterogeneity within tissues could fluctuate over the day. When studying circadian oscillations in cytosine modifications, such cellular heterogeneity were taken into account as potential confounding factors (Oh et al., 2018).

1.4.4 Rhythmic gene expression in humans

Recently, an increasing number of gene expression datasets of human tissue samples have become available (Ardlie and Guigó, 2017), most of which are sampled without explicit regard to the time of day. Interestingly, it is possible to predict internal circadian time in gene expression in humans by computationally assigning time labels to unlabeled samples using different statistical learning approaches (Anafi et al., 2017; Hughey et al., 2016; Leng et al., 2015; Ruben et al., 2018; Wittenbrink et al., 2018). For example, knowledge in mice about which genes are rhythmically transcribed and in which tissues can provide prior information for analyzing rhythms in human tissue gene expression data, under the assumption that rhythmicity in mRNA expression profile is sufficiently conserved across mammalian evolution (Anafi et al., 2017).

1.4.5 Dynamic responses to acute perturbations

Novel experimental designs may motivate analyses beyond fitting periodic functions. For example, studies have shown that sleep deprivation alters expression of clock genes (Franken et al., 2007). But what are the dynamics of gene expression during and after sleep deprivation? How quickly does the system adapt to a new environment? Is the response tissue-specific? How do these transient dynamics apply to other environmental inputs such as feeding patterns

(Acosta-Rodríguez et al., 2017)? Most circadian studies use periodic functions to fit gene expression data. However, the mentioned situations require us to think beyond periodic functions and investigate other dynamics that may occur during perturbations (we look at an example of this in Chapter 4). These temporal dynamics in gene expression may be accompanied by dynamics in other gene regulatory layers. We focus on chromatin interactions as a regulatory layer underlying circadian transcription and behavior.

1.4.6 Mapping chromatin interactions

The conformation and organization of chromatin in the nucleus is known to be important for the control of gene regulation (Pombo and Dillon, 2015), particularly gene transcription. Chromatin conformation capture (3C) technique (Dekker et al., 2002) and its high-throughput variants, such as 4C (Simonis et al., 2006), 5C (Nora et al., 2012), Hi-C (Lieberman-Aiden et al., 2009), and ChIA-PET (Zhang et al., 2013), as well as orthogonal techniques such as GAM (Beagrie et al., 2017) and SPRITE (Quinodoz et al., 2018), have revealed the spatial organization of the genome. An emerging picture is that interphase chromosomes are organized in a hierarchy of structural layers (Gibcus et al., 2018). At the largest scale, there are multi-megabase compartments A and B. A compartments are generally gene rich and transcriptionally active while B compartments are gene poor and transcriptionally repressed (van Steensel and Belmont, 2017). Generally, the compartment A occupies a central nuclear position while compartment B associates with the nuclear lamina (Bouwman and de Laat, 2015). Within these compartments, at sub-megabase scale, the chromatin is organized as topologically associating domains (TADs). At the TAD scale, genes show maximal enrichment of coexpression during differentiation from embryonic stem cells, compared to other genomic scales (Zhan et al., 2017). TAD boundaries are often demarcated by binding of cohesion and CTCF proteins oriented in a convergent manner (de Wit et al., 2015). At a fine scale, promoter-enhancer contacts allow transcription factors to regulate distal (often less than 100 kb away) target genes (Gibcus and Dekker, 2013). Generally, TAD interactions are fairly stable across cell-types (Dixon et al., 2012), whereas sub-TAD interactions can be cell-type specific as well as dynamic (Phillips-Cremins et al., 2013).

1.4.7 Chromatin interactions and the circadian clock

1.4.8 Dynamic chromatin interactions in cell culture models

What are the circadian dynamics in the spatial organization of the genome? The first studies applying 4C in the circadian context used cell line models. One study using mouse embryonic fibroblasts to evaluate chromatin contacts to the promoter of *Dbp*, a clock output gene, reported circadian fluctuations in inter-chromosomal contacts (Aguilar-Arnal et al., 2013). Another study found that the *Pard3* locus was rhythmically recruited in the nuclear lamina in a human colon cancer cell line (HCT116) (Zhao et al., 2015). However, these models have not studied the dynamics of promoter-enhancer looping and whether such interactions may have

a function in the core clock.

1.4.9 Tissue-specific and dynamic chromatin interactions in tissues

Recent *in vivo* work showed that sub-TAD interactions play a role in circadian gene expression by connecting gene promoters with enhancers (Beytebiere et al., 2018; Kim et al., 2018; Mermet et al., 2018; Xu et al., 2016; Yeung et al., 2018). For example, tissue-specific promoter-enhancer interactions enable the circadian clock to regulate gene expression in a tissue-specific manner (Yeung et al., 2018) (Figure 1.4A), and such regulation can also be dynamic over the course of the day (Figure 1.4B). Moreover, 4C-seq performed around the clock focusing on a core clock gene, *Cry1*, and a clock-output gene, *Gys2*, demonstrated dynamic rhythms in promoter-enhancer contacts coinciding with rhythms in the active enhancer mark, H3K27ac (Mermet et al., 2018). As revealed by 4C-seq experiments in *Bmal1* KO, these contacts lost rhythmicity in mice without a functioning clock. Moreover, CRISPR-Cas9 deletion of the contacted intronic *Cry1* enhancer in mice indicated that the contact rhythms shortens the period of the locomotor activity (Mermet et al., 2018), revealing the dynamic function of a noncoding DNA element that propagates from chromatin interactions, gene expression, to locomotor activity. A Hi-C study performed at two time points in the liver reported chromatin interactions at the sub-TAD scale, which could be dynamic or stable over time (Kim et al., 2018). Consistently, a study employing ChIA-PET argued that promoters of rhythmically active genes form stable contact with enhancers (Beytebiere et al., 2018). It will be interesting to investigate why dynamic transcription of some genes is accompanied by dynamic chromatin looping, while other genes exhibit static looping.

1.4.10 Dynamic chromatin interactions and transcription

Recent studies in cell culture showed that promoter-enhancer looping influence transcriptional parameters such as burst fraction (fraction of active transcription sites in each nucleus, which is related to burst frequency) or burst size (Bartman et al., 2016; Kalo et al., 2015; Senecal et al., 2014). Combined live imaging and smRNA FISH studies to investigate molecular mechanisms found that targeted acetylation of promoter histones increases burst frequency while not affecting burst size. (Nicolas et al., 2018). In liver tissue, bursts in gene transcription have been characterized using smRNA FISH (Bahar Halpern et al., 2015). Combining smRNA FISH with functional genomics assays such as 4C-seq and measurements of chromatin H3K27ac have uncovered how promoter-enhancer looping influences transcriptional bursting *in vivo* (Mermet et al., 2018). In particular, deleting an enhancer element using CRISPR-Cas9 demonstrated that disrupting promoter-enhancer looping decreases burst fraction and shortens the period of locomotor activity (Figure 1.4C).

How promoter-enhancer looping in single-cells leads to bursts of gene transcription remains unknown. Are promoter-enhancer contacts preceding initiation of transcriptional bursts, or what are the time delays between the two events? Current techniques such as sm-

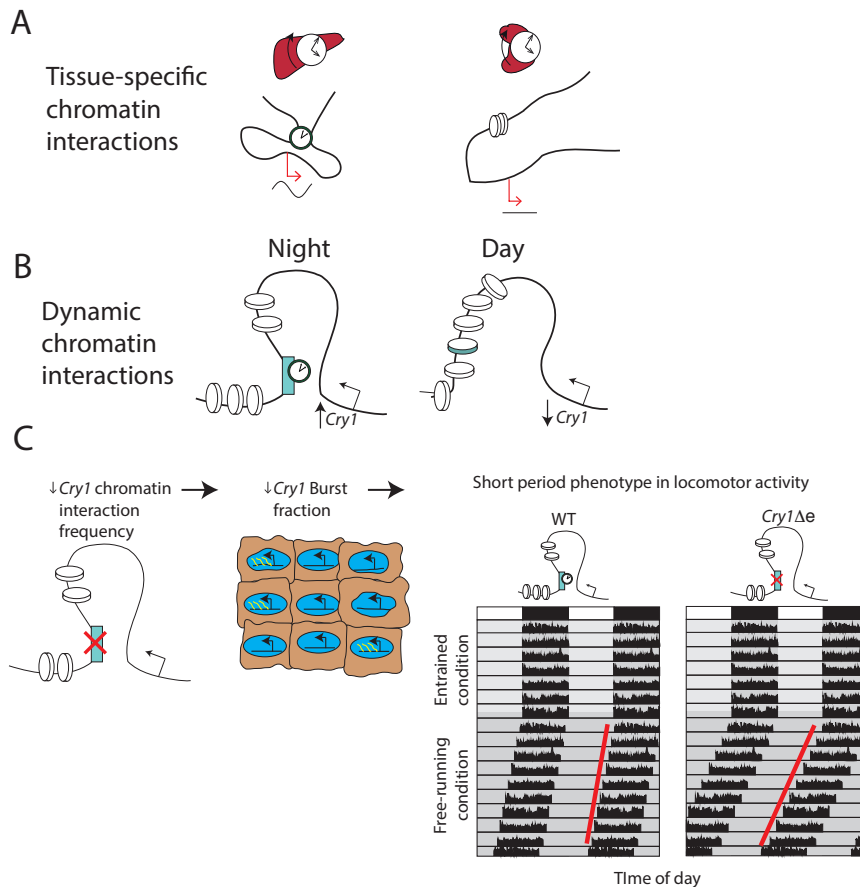


Figure 1.4 – Chromatin interactions as a regulatory layer underlying circadian gene expression and behavior. (A) Tissue-specific chromatin interactions can regulate clock TFs to bind in a tissue-specific manner. Here, the clock TF binding site is accessible in liver but not in kidney, shown as open and closed chromatin, respectively. (B) Dynamic chromatin interactions underlie circadian transcription of rhythmic genes such as the core clock gene *Cry1*. Clock object represents a clock TF. Rectangle represents an enhancer element. (C) Deleting a non-coding DNA element downstream of *Cry1* in mouse (mutant named *Cry1Δe*) decreases chromatin interactions between promoter and enhancer, reduces burst fraction (fraction of active transcription start sites per nucleus, which is related to burst frequency), and shortens the period in locomotor activity. Simplified cartoon of two actograms indicate a shorter period of locomotor activity in *Cry1Δe* versus WT under free-running conditions. During entrained conditions (12-hours light, 12-hours dark, shown as white and black rectangles at the top), the mice are active (activity level in black) during lights off and inactive during lights on. During free-running conditions (in complete darkness), the period of locomotor activity depends on the genotype, which can be calculated by the slope of the red line.

RNA FISH can detect transcriptional bursting *in vivo* by analyzing snapshots of RNA molecules across many cells (Bahar Halpern et al., 2015). But monitoring the same cells to determine the relative timing of promoter-enhancer contacts with transcriptional bursting remains challenging. In the context of molecular clocks, linking gene regulation and transcription at single cells will uncover how noisy gene expression is regulated to produce robust pacemakers at the tissue and organism level.

1.5 Specific questions addressed in thesis

This thesis tackles three questions related to the regulation of dynamic, and more specifically temporal gene expression. First two focus on circadian biology; the third relates to sleep.

1. How does the circadian clock regulate gene expression in a tissue-specific manner?
 - What fraction of variance in multi-tissue circadian gene expression data is of temporal or tissue-specific origin?
 - What transcriptional mechanisms underlie tissue-specific circadian gene expression?
 - How do tissue-specific chromatin landscapes regulate circadian gene expression?
2. What is the role of chromatin interactions in regulating circadian gene expression and circadian rhythms?
 - Are circadian dynamics in gene expression accompanied by dynamics in promoter-enhancer looping?
 - Do dynamics in promoter-enhancer looping require the circadian clock?
 - Does dynamics in promoter-enhancer looping regulate circadian gene transcription and period of locomotor activity?
3. How does diurnal gene expression respond to acute perturbations such as sleep deprivation?
 - What are the gene expression dynamics in response to acute sleep deprivation in mouse cortex?
 - How do the sleep homeostat and circadian processes contribute towards gene expression output?
 - What are possible transcriptional regulators underlying sleep-wake dynamics?

2 Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs

2.1 Introduction

This chapter investigates transcriptional mechanisms underlying diurnal gene expression rhythms that can be regulated in a tissue-specific manner. The work has been published in *Genome Research* 2018 under the Creative Commons license and is reproduced here. The main text as well as supplemental figures are below. Due to the size of supplemental tables, these are not attached but can be found on the open online version of the *Genome Research* article. No changes were made to the main text or supplemental figures, which was downloaded from doi:10.1101/gr.222430.117.

The pursuit to uncover transcriptional mechanisms underlying tissue-specific circadian gene expression involved developing and applying a variety of computational and statistical methods. Some of these methods are specific to the problem I tackled while others may be broadly applied to other problems. Here, I will go into more detail of some methods which I think may be of use in other gene expression analysis projects: model selection for identifying different combinations of rhythms across conditions, complex-valued singular value decomposition, and GO term analysis around the clock.

2.2 Extended methods: model selection with Zellner's g -priors

2.2.1 Enumerating harmonic regression models to identify possible combinations of rhythms across tissues

This chapter relies on identifying whether gene expression oscillates in no tissues, one tissue, or a combination of tissues. We use a model selection approach to select the combination of

Chapter 2. Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs

tissues that fit the data while penalizing for model complexity (i.e., the number of parameters).

To illustrate an example, we take the case of two tissues ($|c|=2$) (Figure 2.1). The most complex model (i.e. most parameters) corresponds to each tissue having its own rhythmic parameters $\beta_{1,c}$, $\beta_{2,c}$ in addition to the intercepts α_c .

$$M_f : \quad Y_{t,c_i} = \alpha_{c_i} + \beta_{1,c_i} \cos(\omega t) + \beta_{2,c_i} \sin(\omega t) + \epsilon, \quad (2.1)$$

where $i = 1, 2$ denote the index of tissues.

The restricted model, meaning both tissues are flat, would have only the intercept parameters $(\alpha_{c_1}, \alpha_{c_2})$:

$$M_r : \quad Y_{t,c_i} = \alpha_{c_i} + \epsilon. \quad (2.2)$$

For two tissues, there will be 3 intermediate models, each with 4 parameters. The first two models (M_{I_1}, M_{I_2}) capture the case where a gene is rhythmic in a single tissue (tissue-specific) and a shared model (M_{I_3}) captures the case where a gene is rhythmic in both tissues with the same rhythmic parameters

$$M_{I_1} : \quad Y_{t,c_i} = \begin{cases} \alpha_{c_i} + \beta_{1,c_i} \cos(\omega t) + \beta_{2,c_i} \sin(\omega t) + \epsilon & i = 1 \\ \alpha_{c_i} + \epsilon & i = 2 \end{cases}$$

$$M_{I_2} : \quad Y_{t,c_i} = \begin{cases} \alpha_{c_i} + \epsilon & i = 1 \\ \alpha_{c_i} + \beta_{1,c_i} \cos(\omega t) + \beta_{2,c_i} \sin(\omega t) + \epsilon & i = 2 \end{cases}$$

$$M_{I_3} : \quad Y_{t,c_i} = \alpha_{c_i} + \beta_1 \cos(\omega t) + \beta_2 \sin(\omega t) + \epsilon. \quad (2.3)$$

The shared model M_{I_3} is notable because it is distinct from the full model M_f . M_{I_3} forces both tissues to have the same amplitude and phase, whereas M_f may have tissue-dependent rhythms. From the point of view of gene regulation, M_{I_3} (shared model, same rhythm) has a simpler explanation than M_f (full model, different rhythms), and we therefore explicitly distinguish these two cases.

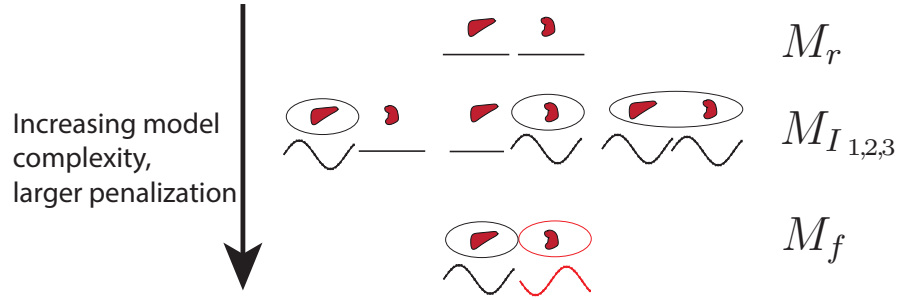


Figure 2.1 – **Illustration of number of models for the case of two tissues.** For two tissues, there are five rhythmic models. Importantly, we allow the tissues to share rhythms, which reduces the number of parameters in the model if the rhythm is synchronized across tissues.

2.2.2 Selecting models by the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and Bayes Factors

After having enumerated the possible rhythmic models, the final step is to fit each model and identify which is the best model, based on the fit of the data as well as the complexity of the model. Typically, we seek to find a model M_γ that minimizes a penalized negative log-likelihood score

$$S_\gamma = -2\ln(\hat{L}_\gamma) + p_\gamma F \quad (2.4)$$

where p_γ is the number of parameters for model γ . $F = 2$ for AIC and $F = \ln(n)$ for BIC. $\ln\hat{L}_\gamma$ is the maximized value of the log loglikelihood. That is, $\hat{L} = p(Y|\hat{\theta}, M_\gamma)$ where $\hat{\theta}$ are the parameter values that maximize the likelihood function.

Although AIC and BIC provide straightforward approaches to selecting the best models, they both have their drawbacks. AIC can be shown to be an inconsistent estimator (i.e. probability of selecting true model does not go to 1 as sample size goes to infinity) (Yang, 2005). BIC can be derived as a Laplace approximation (obtained by finding the mode of the posterior distribution and then fitting a Gaussian centered at the mode) to Bayes factors, which may not always be appropriate, especially for small number of samples (Berger et al., 2003; George and Foster, 2000).

An alternative approach is to evaluate the posterior probabilities of the models, which can be expressed through the Bayes factor K between a candidate model γ and a base model b . In linear regression models, this Bayes factor K form has a convenient closed-form solution in

Chapter 2. Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs

the g-prior method (Equation 2.11).

$$p(M_Y|Y) = \frac{p(M_Y)p(Y|M_Y)}{\sum_{Y'} p(M_{Y'})p(Y|M_{Y'})} = \frac{p(M_Y)K(M_Y:M_b)}{\sum_{Y'} p(M_{Y'})K(M_{Y'}:M_b)} \quad (2.5)$$

where the base model M_b contains only the intercepts, α . A key component from Equation 2.5 is the marginal likelihood of the data given the model:

$$p(Y|M_Y) = \int_{\Theta_Y} p(Y|\vec{\theta}_Y, M_Y) p(\vec{\theta}_Y|M_Y) d\vec{\theta}_Y, \quad (2.6)$$

which requires integrating over the parameters of the model and setting a prior on each of the parameters. For parameters that are common to all models, we may set an improper prior, such as Jeffrey's prior, which is defined up to an arbitrary multiplicative constant. For all other parameters, it is advised to avoid improper priors because the arbitrary constant does not factor out in the posterior calculation.

In linear regression models, the model parameters $\hat{\theta}$ can be written as intercept and regression coefficients, α and $\vec{\beta}$, respectively. The data Y with n data points can be modeled in a linear regression model with Gaussian noise $\epsilon \sim N(0, \sigma^2)$:

$$Y = \vec{1}_n \alpha + X_Y \beta_Y + \epsilon \quad (2.7)$$

A computationally efficient prior for linear regression models is the "g-prior" (Zellner, 1986), defined as

$$p(\alpha, \sigma^2|M_Y) = \frac{1}{\sigma^2}, \quad \beta_Y|\sigma \sim N(0, g\sigma^2(\mathbf{X}^T\mathbf{X}^{-1})) . \quad (2.8)$$

The scalar g controls the spread of the prior in the parameters. Larger g tends to favor parsimonious models (i.e. few large parameters), while smaller g tends to favor saturated models. Note that this prior is related to more standard Bayesian linear regression with Gaussian distribution for coefficients and inverse gamma distribution for the variance:

$$p(\vec{\beta}, \sigma^2) \sim N(\vec{\beta}_0, \sigma^2 V_0) IG(\sigma^2|a_0, b_0) \quad (2.9)$$

and setting $a_0 = b_0 = 0$, corresponding to an uninformative prior for σ^2 and to set $w_0 = 0$ and $V_0 = g(\mathbf{X}^T \mathbf{X})^{-1}$. This reduces the problem to a single hyperparameter, g , and has straightforward closed-form solutions for calculating the Bayes factor, which is helpful for model selection (Liang et al., 2008).

There are many strategies for specifying g as extensively discussed by Liang et al. (2008). We will select a specific value for g based on how we expect the data should behave based on assumptions from circadian biology.

The marginal likelihood is then given in closed form:

$$p(\vec{Y} | M_\gamma, g) = \frac{\Gamma((n-1)/2)}{\sqrt{\pi}^{n-1} \sqrt{n}} \|\vec{Y} - \vec{\bar{Y}}\|^{-(n-1)} \frac{(1+g)^{\frac{n-1-p_\gamma}{2}}}{(1+g(1-R_\gamma^2))^{\frac{n-1}{2}}}. \quad (2.10)$$

The Bayes factor $K(M_\gamma : M_b)$ of model M_γ with respect to a base model M_b can be expressed as the ratio of their marginal likelihoods:

$$K(M_\gamma : M_b) | g = \frac{(1+g)^{\frac{n-1-p_\gamma}{2}}}{(1+g(1-R_\gamma^2))^{\frac{n-1}{2}}}, \quad (2.11)$$

where $R_\gamma^2 = 1 - \frac{\sum_i (y_i - \hat{y})^2}{\sum_i (y_i - \bar{y})^2}$ is the coefficient of determination for model γ , a goodness of fit measurement that can easily be calculated. n is the number of data points.

In this chapter, we apply Equation 2.10 directly by selecting g , which controls the penalty on model complexity. To guide our choice of g , we make an assumption pertinent to circadian biology: slight differences in the amplitude and phase of core clock genes between tissues should still be considered to yield shared rhythms. To achieve this, we choose g with sufficient tolerance to amplitude and phase differences (which may depend on the experiment and technology) such that most of the condition-wide oscillations are incorporated into the shared model (rather than the full model). Plotting the 24h spectral power of genes in the shared model as a function of g allows us to identify the beginnings of a plateau. We interpret this point to be the g value whereby most core clock genes are included in the shared model (Figure 2.2). In our liver and kidney RNA-seq dataset, we found that $g = 1000$ provided biologically interpretable outputs.

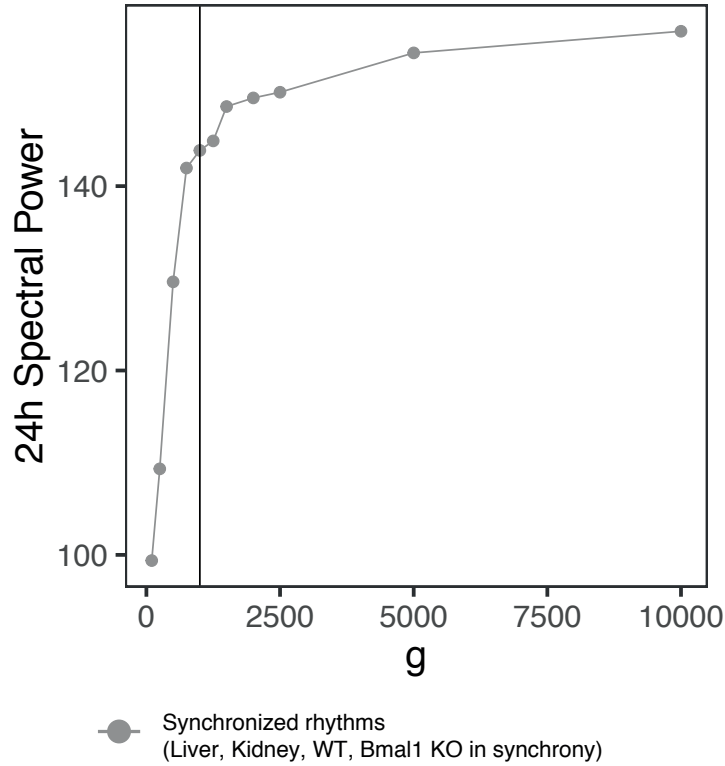


Figure 2.2 – **Spectral variance as a function of g .** For each g , we calculated the 24-hour variance by taking the sum of the 24-hour variance across genes assigned to model that shares a rhythm across all conditions. Note that near $g = 1000$, the variance begins to plateau, suggesting the additional genes that are included $g > 1000$ may likely be noise rather than robust amplitudes.

2.3 Extended methods: complex-valued singular value decomposition

Here, we seek to project temporal gene expression across tissues onto lower projections in order to visualize and summarize high-dimensional data across tissues and time. Circadian datasets across tissues often contain correlation structures that can be represented in lower dimensions. For example, many core circadian clock and clock outputs genes oscillate synchronously in nearly all tissues. This projection is useful as a genome-wide exploration technique (analogous to principal component analysis) or as a downstream analysis of clusters of genes identified using model selection. We found that complex-valued singular value decomposition (cv-SVD) to have useful properties that were biologically interpretable. Chronobiology datasets that sample multiple tissues or conditions over the 24-hour day can be represented by a matrix of complex values where the values represent amplitude (we define it here as min-to-max magnitude rather than mean-to-peak to simplify biological interpretation) and

phase (time at peak expression) for each gene and each condition. This complex number can be calculated by the 24-hour Fourier component corresponding to an angular frequency of $\omega = \frac{2\pi}{24}$ (we checked that the 24-h represents the largest temporal component in Yeung et al. by breaking down the temporal variation into all Fourier components, which showed that the largest temporal component comes from 24-hour rhythms, followed by 12-hours). This approximation works well for genes that can be modeled by cosine and sines such as in models $M_f, M_{I_1}, M_{I_2}, M_{I_3}$ in Equation 2.1 and 2.3.

Since we expect the temporal variance to be well approximated by the 24-hour Fourier space $(\sum_t y_t(t) - \bar{y}(t))^2 \approx |Y_{\omega_{24}}|^2$, we project

$$Y_{g,c,\omega_{24}} = \sum_t Y_{g,c,t}(t) e^{i\omega_{24}t}, \quad (2.12)$$

where $\omega_{24} = \frac{2\pi}{24}$. This projection works well on genes that have temporal signals that can be modeled well by a sine wave of period 24 hours. This approximates the dataset to a complex-valued matrix of G rows and C columns. Each complex-valued element $Y_{g,c}$ corresponds to the amplitude and phase of gene g in condition c . Because many rhythms may be shared across conditions, we can decompose $Y_{g,c,\omega_{24}}$ by standard singular value decomposition (SVD) into sample space (eigentissues) and tissue space (eigensamples), which are also complex values, and the first few singular values should provide a low-dimensional representation of the data. We can then visualize and summarize the eigentissues and eigensamples in the 24-hour Fourier space by noting that $|Y_{g,c,\omega_{24}}|$ corresponds to the amplitude and $\text{Arg}(Y_{g,c,\omega_{24}})$ corresponds to the phase of gene g and tissue c . The amplitude and phase can be naturally visualized in polar coordinates; the radial distance representing amplitude and the phase angle representing time at peak expression.

The SVD technique factorizes $Y_{g,c,\omega_{24}} = U\Sigma V^*$, where V^* is the complex conjugate of V . U and V are unique eigensamples and eigengenes, and are defined up to a unit-phase factor $e^{i\phi}$. To simplify interpretation, we set the phase of the tissue space such that the tissue with the largest amplitude has phase 0 and amplitude 1. Correspondingly, the gene space would be interpreted as the amplitude and phase of the largest amplitude tissue. To get the amplitude and phase of genes in another tissue, one would simply add and multiply the phase and amplitude, respectively, of the loadings for that tissue with the gene space.

2.4 Extended methods: Gene ontology (GO) analysis around the clock

From our model selection process, we identify a set of genes (a module) that have rhythms in a specific combination of tissues (e.g., genes that oscillate in liver but not in kidney). This module have gene expression rhythms that peak at different times of day. One useful downstream analysis is to ask whether certain Gene Ontology (GO) terms are enriched at different times

Chapter 2. Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs

of day. To extend GO term analysis to the 24-hour day, I used foreground genes as genes in the module that are within a time window $[t - 3, t + 3]$ and expressed genes in the dataset as background genes. For each $t \in (1, 2, 3, \dots, 24)$ I calculated a p-value for enrichment using classic Fisher's exact test, as implemented by TopGO package in R. This outputs p-value enrichment over time, which can then be visualized in a polar plot.

Example code can be found on GitHub: <https://github.com/naef-lab/CyclicGO>.

2.5 Contributions

I did all the computational methods and analyses in this project. Jérôme Mermet did all the 4C-seq experiments. RNA-seq experiments were performed by Julien Marquis, Aline Charpagne, and Céline Jouffe.

Research

Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs

Jake Yeung,^{1,5} Jérôme Mermet,^{1,5} Céline Jouffe,² Julien Marquis,³ Aline Charpagne,³ Frédéric Gachon,^{2,4} and Felix Naef¹

¹Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, CH-1015, Switzerland; ²Department of Diabetes and Circadian Rhythms, Nestlé Institute of Health Sciences, CH-1015 Lausanne, Switzerland; ³Functional Genomics, Nestlé Institute of Health Sciences, CH-1015 Lausanne, Switzerland; ⁴Faculty of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Temporal control of physiology requires the interplay between gene networks involved in daily timekeeping and tissue function across different organs. How the circadian clock interweaves with tissue-specific transcriptional programs is poorly understood. Here, we dissected temporal and tissue-specific regulation at multiple gene regulatory layers by examining mouse tissues with an intact or disrupted clock over time. Integrated analysis uncovered two distinct regulatory modes underlying tissue-specific rhythms: tissue-specific oscillations in transcription factor (TF) activity, which were linked to feeding-fasting cycles in liver and sodium homeostasis in kidney; and colocalized binding of clock and tissue-specific transcription factors at distal enhancers. Chromosome conformation capture (4C-seq) in liver and kidney identified liver-specific chromatin loops that recruited clock-bound enhancers to promoters to regulate liver-specific transcriptional rhythms. Furthermore, this looping was remarkably promoter-specific on the scale of less than 10 kilobases (kb). Enhancers can contact a rhythmic promoter while looping out nearby nonrhythmic alternative promoters, confining rhythmic enhancer activity to specific promoters. These findings suggest that chromatin folding enables the clock to regulate rhythmic transcription of specific promoters to output temporal transcriptional programs tailored to different tissues.

[Supplemental material is available for this article.]

A mammalian internal timing system, known as the circadian clock, orchestrates temporal physiology in organs to anticipate daily environmental cycles (Dibner and Schibler 2015). Individual cells within organs contain a molecular oscillator that, together with rhythmic systemic signals such as hormones, temperature, and feeding behavior, collectively drive diurnal oscillations in gene expression and physiology (Lamia et al. 2008; Reinke et al. 2008; Cho et al. 2012; Vollmers et al. 2012). Remarkably, the circadian clock impinges on many gene regulatory layers, from transcriptional and post-transcriptional processes, translation efficiency, to translational and post-translational processes (Mermet et al. 2016).

Transcriptome analysis of different cell types and tissues has highlighted the breadth of tissue-specific transcriptional regulation (Merkin et al. 2012; Yue et al. 2014). However, physiological processes are dynamic at the timescale of hours and often under circadian control, such as hormone secretion, drug and xenobiotic metabolism, and glucose homeostasis (Takahashi et al. 2008). Adding the temporal dimension to tissue-specific gene regulation is needed for an integrated understanding of physiology.

Chronobiology studies have shown that tissues utilize the circadian clock to drive tissue-specific rhythmic gene expression (Storch et al. 2002; Korenčič et al. 2014; Zhang et al. 2014), presum-

ably to schedule physiological functions to optimal times of day. Indeed, genetic ablation of the circadian clock in different tissues can lead to divergent pathologies, such as diabetes in pancreas-specific *Bmal1* knockout (KO) and fasting hypoglycemia in liver-specific *Bmal1* KO, suggesting that the clock interweaves with tissue-specific transcriptional programs (Bass and Lazar 2016), but how diurnal and tissue-dependent regulatory landscapes interact to generate tissue-specific rhythms is poorly understood.

Results

Contributions of tissue, daily time, and circadian clock to global variance in mRNA expression

To estimate the respective contributions of tissues, daily time, and circadian clock to global variance in gene expression, we analyzed available temporal transcriptomes across 11 tissues in WT mice (Zhang et al. 2014) and generated temporal RNA-seq data of liver and kidney from *Bmal1* KO mice and WT littermates (Supplemental Tables S1, S2; Methods). The Zhang et al. data set was obtained under dark-dark (DD), ad libitum feeding, sampled every 2 h. The liver and kidney *Bmal1* KO and WT data sets were obtained under light-dark (LD), night-restricted feeding conditions, sampled every 4 h.

To avoid mixing different experimental designs (e.g., temporal resolution and number of repeats) (Deckard et al. 2013; Li et al.

⁵These authors contributed equally to this work.

Corresponding author: felix.naef@epfl.ch

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.222430.117>. Freely available online through the *Genome Research* Open Access option.

© 2018 Yeung et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Tissue-specific circadian transcription regulation

2015), we analyzed these data sets separately. We performed principal component analysis (PCA) on the entire set of conditions (11 tissues \times 24 time points) to obtain an unbiased overview into the contributions of tissue- and time-specific variance. Most of the variance concerned differences in expression between tissues (Fig. 1A; Supplemental Fig. S1A–D). Temporal variance, in particular, 24-h periodicity, was present among a group of principal components carrying lower amounts of variance (Fig. 1A; Supplemental Fig.

S1E–G). Focusing on genome-wide temporal variation within each tissue, we found that 24-h rhythms constituted the largest contribution of temporal variance, followed by 12-h rhythms, which were close to background levels for many tissues (Fig. 1B; Hughes et al. 2009). We thus focused the rest of our analysis on 24-h rhythms.

We analyzed the peak-to-trough amplitudes (hereafter also referred to as fold change) of 24-h rhythmic transcripts. Metabolic tissues, notably liver, brown fat, and skeletal muscle exhibited more (on the order of 100 transcripts) intermediate- to high-amplitude (between two- and 10-fold) transcript rhythms. Brain tissues showed virtually no rhythmic transcripts above fourfold (Fig. 1C). In liver and kidney of *Bmal1* KO mice, the number of rhythmic mRNAs was reduced by threefold compared to WT littermates. This effect increased for larger amplitudes. Few transcripts in tissues of *Bmal1* KO oscillated by more than 10-fold (Fig. 1D). Thus, a functional circadian clock is required for high-amplitude transcript rhythms across diverse tissues, while systemic signals regulate lower amplitude rhythms that persist in clock-deficient liver (Hughes et al. 2012; Atger et al. 2015; Sobel et al. 2017) and kidney (Nikolaeva et al. 2012).

Combinatorics of rhythmic transcript expression across tissues and genotypes

We reasoned that identifying sets of genes with shared rhythms across subsets of tissues would allow finding underlying regulatory mechanisms. We developed a model selection (MS) algorithm extending harmonic regression (Fisher 1929) to classify genes into modules sharing rhythmic mRNA profiles across subsets of tissues (Fig. 2A; Methods). Phase (time of peak) and amplitude (\log_2 fold change) relationships between genes and tissues are summarized using complex-valued singular value decomposition (SVD) (Fig. 2B; Methods). We applied MS to the 11 tissues, which identified gene modules involving rhythmic mRNA accumulation in nearly all tissues (tissue-wide) (Fig. 2C), in single tissues (tissue-specific), or in several tissues (tissue-restricted) (examples shown in Fig. 2D; Supplemental Fig. S2A; Supplemental Table S3).

The tissue-wide module contained a set of both clock- and system-driven rhythmic mRNAs, as determined by comparing *Bmal1* KO data in liver and kidney (Fig. 2C, left). Moreover, these transcripts oscillated in synchrony across all tissues and peaked at fixed times of day,

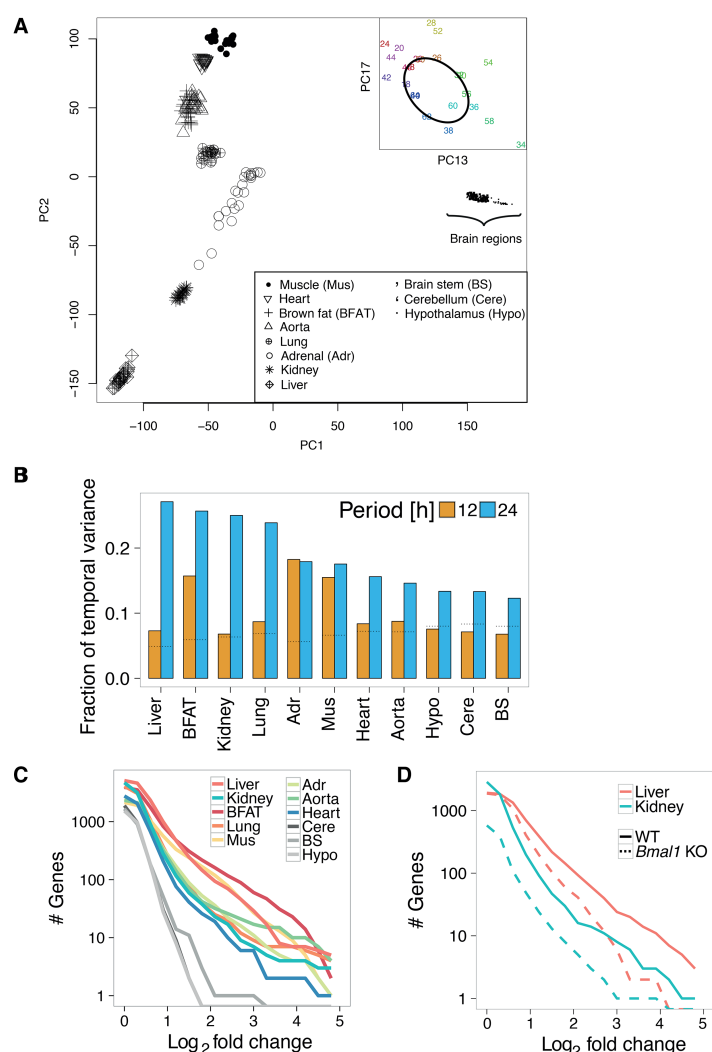


Figure 1. Contribution of tissue, daily time, and circadian clock to global variance in mRNA expression. (A) Principal component analysis (PCA) across 11 WT tissues sampled over 2 d. PC1 and PC2 show clustering of samples by tissues; each point represents a tissue sample (see key) at a specific time point (not labeled). (Inset) Loadings for PC13 and PC17 for the liver samples labeled with circadian time (CT), showing temporal variation along an elliptic path. Labels indicate CT time; samples that are 24 h apart are in the same color. (B) Fractions of temporal variance in each tissue explained by 24- and 12-h periods, obtained by applying spectral analysis genome-wide for each tissue. Dotted horizontal lines represent the expected background level, assuming white noise. (C,D) Cumulative number of rhythmic genes ($P < 0.01$, harmonic regression) with \log_2 fold change larger than the value on the x-axis. (C) Analysis on 11 WT tissues. (D) Analysis on four conditions: *Bmal1* KO mice and WT littermates in liver and kidney.

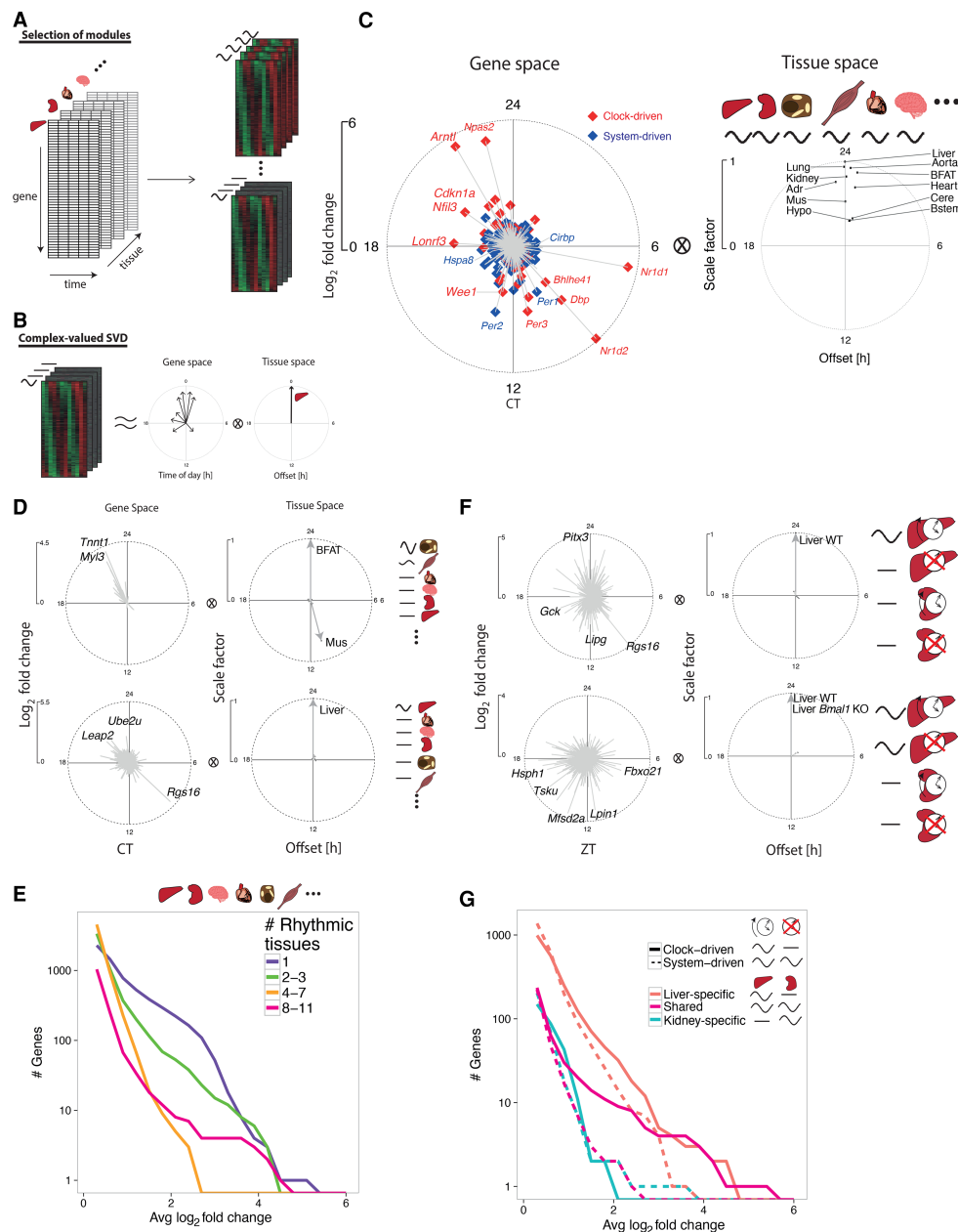


Figure 2. Combinatorics of rhythmic transcript expression across tissues and genotypes. (A) Schema for the model selection (MS) algorithm to identify rhythmic gene expression modules across tissues. Temporal transcriptomes of different tissues represented as a three-dimensional array (left). Gene modules are probabilistically assigned among different combinations of 24-h rhythms across tissues (e.g., tissue-specific or tissue-wide rhythms schematically shown on right). (B) Gene modules are summarized by the first component of complex-valued singular value decomposition (SVD) to highlight phase (peak time shown as the clockwise angle) and amplitude (log₂ fold change shown as the radial distance) relationships between genes (gene space) and between tissues (tissue space). SVD representation is scaled such that the genes show log₂ fold changes, while tissue vectors are scaled such that the highest amplitude tissue has length of 1 and a phase offset of 0 h. (C–E) MS applied to 11 WT tissues. (C) SVD representation of tissue-wide mRNA rhythms from the 11 tissues. Genes are labeled as system-driven (blue) or clock-driven (red) according to the comparison of the corresponding temporal profiles in *Bmal1* KO and WT littermates. (D) Examples of anti-phasic rhythms (brown fat and muscle, *n* = 20, first SVD component explains 81% of variance), and tissue-specific rhythms (liver, *n* = 846, first SVD component explains 59% of variance). Representative genes with large amplitudes are labeled. (E) Number of transcripts showing rhythms (*P*-value < 0.01, harmonic regression) in different numbers of tissues, in function of increasing peak to trough amplitudes on the x-axis. x-axis: average log₂ fold change calculated from the identified rhythmic tissues. (F,G) MS applied to *Bmal1* KO and WT littermates in liver and kidney. (F) SVD representation of clock- (top, *n* = 991, 83% of variance) and system-driven (bottom, *n* = 1395, 84% of variance) liver-specific rhythms. (G) Number of transcripts showing clock- (solid) or system-driven (dotted) rhythms (*P*-value < 0.01, harmonic regression) in liver (red), kidney (blue), or both (magenta).

although their amplitudes varied between tissues, with brain regions showing the smallest amplitudes (Fig. 2C, right). The clock drove synchronized oscillations at high amplitudes, notably clock genes (e.g., *Arntl*, *Npas2*, *Nr1d1,2*; note that *Arntl* and *Nr1d1,2* are also named *Bmal1* and *Rev-erba,b*, respectively), clock output genes (e.g., *Dbp*, *Nfil3*), and cell cycle regulators (*Cdkn1a* and *Wee1*) (Matsuo et al. 2003; Gréchez-Cassiau et al. 2008). Interestingly, clock genes *Per1,2* continued to oscillate in *Bmal1* KO in multiple tissues, extending previous studies in liver (Kornmann et al. 2007). Other clock-independent oscillations included mRNAs of heat- and cold-induced genes, such as *Hspa8* and *Cirbp* (Morf et al. 2012; Gotic et al. 2016), that peaked 12 h apart near CT18 and CT6 (CT: circadian time; CT0 corresponds to subjective dawn and start of the resting phase; CT12 corresponds to subjective dusk and start of the activity phase), concomitantly with highs and lows in body temperature rhythms (Refinetti and Menaker 1992).

Tissue-restricted modules contained rhythmic transcripts that peaked in synchrony, such as in liver and kidney, or with fixed offsets, such as the nearly 12-h shifted rhythms in brown fat and skeletal muscle (Supplemental Fig. S3A). Overall, transcripts with large amplitudes ($FC > 8$) oscillated in either a few tissues (three or less) or tissue-wide (eight or more) (Fig. 2E).

To distinguish clock- and system-driven mRNA rhythms, we applied the MS algorithm to the liver and kidney transcriptomes in WT and *Bmal1* KO mice (Fig. 2F; Supplemental Fig. S3B; Supplemental Table S4). This separation identified clock- and system-driven modules that oscillated in liver but were flat in kidney (Fig. 2F), as exemplified by mRNAs of *Lipg* and *Lpin1* (Supplemental Fig. S2B). Indeed, both transcripts oscillated in WT liver with robust amplitudes, peaking near ZT11, but were flat in kidney (ZT: Zeitgeber time; ZT0 corresponds to onset of lights-on; ZT12 corresponds to onset of lights-off). However, in *Bmal1* KO, *Lpin1* continued to oscillate, while *Lipg* was flat.

Summarizing, we found that shared clock-driven mRNA rhythms, which contained core clock and clock-controlled genes, oscillated with significantly larger amplitudes than system-driven genes (Fig. 2G, magenta solid versus dotted). Similarly, clock-driven liver-specific mRNA rhythms also oscillated at higher amplitudes compared with system-driven mRNA rhythms (Fig. 2G, red solid versus dotted). On the other hand, kidney-specific clock- and system-driven transcripts oscillated with comparable amplitudes (Fig. 2G, blue solid versus dotted) and were less numerous overall, which could reflect the distinct cell types constituting the kidney (Lee et al. 2015). The uncovered diversity of clock- and system-driven mRNA rhythms involving distinct combinations of tissues hints at complex transcriptional or post-transcriptional regulation.

Oscillatory TF activity in one tissue but not others can drive tissue-specific mRNA rhythms

We focused on WT and *Bmal1* KO liver and kidney to identify rhythmic TF activities underlying clock- and system-driven tissue-specific mRNA rhythms. We first analyzed liver-rhythmic genes driven by systemic signals ($n = 1395$, MS) (Fig. 3A), which were associated with feeding and fasting rhythms (Gene Ontology analysis around the clock) (Methods). Indeed, ribosome biogenesis was up-regulated most strongly during the first 6 h of the feeding phase (from ZT12 to ZT18) (Jouffe et al. 2013; Chauvin et al. 2014), while insulin signaling was down-regulated during first 6 h of the fasting phase (from ZT0 to ZT6)

(Ravnskjaer et al. 2013), consistent with daily responses to nutrient fluctuations in liver (Sinturel et al. 2017).

To infer rhythmic TF activities that may underlie these mRNA rhythms, we applied a penalized regression model (MARA) (Balwiercz et al. 2014) that integrates TF binding site predictions near promoters with mRNA accumulation. TF analysis of this module notably identified TFs related to insulin biosynthesis and gluconeogenesis, such as MAFB (Matsuoka et al. 2003) and EGR1 (Matsuoka et al. 2003; Shen et al. 2015), whose activities peaked at ZT11 and ZT3, respectively (Fig. 3B; Supplemental Fig. S4A). Integrating temporal activities of candidate TFs with RNA-seq and our previously described temporal nuclear protein data set (Wang et al. 2017), we found that rhythmic activity of MAFB and EGR1 was supported by rhythmic mRNA abundance, followed by rhythmic nuclear protein abundance (Fig. 3B; Supplemental Fig. S4B), likely reflecting the delayed protein abundance after mRNA accumulation (Mermet et al. 2016).

Next, we analyzed clock-driven transcripts oscillating specifically in the kidney ($n = 156$, MS) (Fig. 3C), among which sodium ion and organic anion transporters peaked near ZT12 and ZT0, respectively. The up-regulation of sodium ion transporters in kidney during the behaviorally active phase may underlie clock-dependent increase of sodium excretion (Nikolaeva et al. 2012). Similarly, the up-regulation of organic anion transporters during the resting phase may explain increased transport activity for precursors of gluconeogenesis, such as pyruvate and lactate, during fasting (Stumvoll et al. 1998; Ekberg et al. 1999). TF analysis predicted TFEB2 to regulate mRNAs that peaked during the resting phase (Fig. 3D; Supplemental Fig. S4C). The predicted TFEB2 activity was anti-phasic with *Tfcp2* mRNA abundance, suggesting a repressive activity, consistent with the ability of TFEB2 to recruit histone deacetylase HDAC1 (Kim et al. 2016).

Finally, liver-specific clock-driven rhythmic transcripts ($n = 991$, MS) were comprised of genes associated with glucose metabolism (enriched at ZT18), such as *Gck* and *Ppp1r3b* (Kelsall et al. 2009; Oosterveer and Schoonjans 2014), as well as lipid, cholesterol, and bile acid metabolism genes (enriched at ZT2), such as *Elovl3*, *Insig2*, *Hsd3b7*, and *Cyp8b1* (Fig. 3E; Shea et al. 2007; Le Martelot et al. 2009; Guillou et al. 2010; Sayin et al. 2013). Predicted activity of ELF oscillated and peaked near ZT3 in WT liver but was flat in *Bmal1* KO (Fig. 3F; Supplemental Fig. S4D; Fang et al. 2014). Interestingly, mRNA abundance of *Elf1*, as well as its nuclear protein abundance, also oscillated in WT, supporting *Elf1* as a potential regulator of oscillating transcripts peaking near ZT6.

Colocalized binding of clock- and liver-specific TFs drives liver-specific mRNA rhythms

To further dissect liver-specific clock-driven rhythms, we reasoned that accessible chromatin regions specific to the liver could harbor regulatory sites for clock TFs, which could then regulate mRNA rhythms liver-specifically. Comparing DNase I hypersensitive sites (DHSs) in liver and kidney (DNase-seq data from ENCODE) (Yue et al. 2014), we found that liver-specific clock-driven genes were enriched with liver-specific DHSs (within 40 kb of promoters), compared to system-driven as well as nonrhythmic genes (Fig. 4A). Using TF binding site predictions underlying these liver-specific DHSs, we applied MARA to predict rhythmic TF activities that explain gene expression of this module (Supplemental Fig. S5A). In WT liver, the predicted activity of RORE oscillated with robust amplitudes and peaked near ZT21. RORE activity became high and flat in *Bmal1* KO liver, consistent with loss of REV-ERB expression and

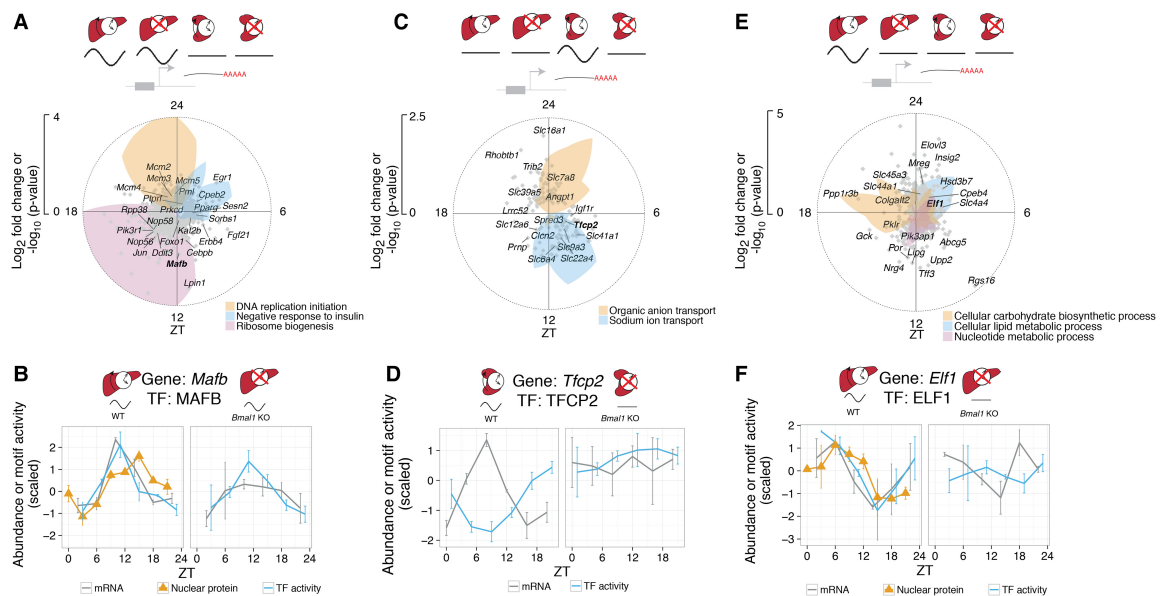


Figure 3. Oscillatory TF activity in one tissue but not others can drive tissue-specific rhythms. (A) Module describing system-driven liver-specific rhythms ($n = 1395$, first SVD component explains 84% of variance). Radial coordinate of the colored polygons represents enrichment of the indicated GO terms at each time point, obtained by comparing the genes falling in a sliding window of ± 3 h to the background set of all 1395 genes assigned to the module (P -value computed from Fisher's exact test). (B) MAFB is a candidate TF for the module in A. Predicted MAFB activity (blue), nuclear protein abundance (orange triangles), and mRNA accumulation (gray) oscillate in WT and *Bmal1* KO, with peak mRNA preceding peak nuclear protein and TF activity. Error bars in nuclear protein, mRNA, and TF activity show SEM ($n = 2$). (C) Clock-driven kidney-specific module ($n = 156$, first SVD component explains 80% of variance). Colored polygons as in A. (D) TF2P2 is a candidate TF for the module in C. The temporal profile of predicted TF2P2 activity (blue) is anti-phasic with *Tfcp2* mRNA accumulation (gray) in WT, and both are flat in *Bmal1* KO. Error bars in mRNA and TF activity show SEM ($n = 2$). (E) Clock-driven liver-specific module ($n = 991$, first SVD explains 83% of variance). (F) ELF is a candidate TF for the module in E. The temporal profile of predicted ELF activity (blue) in WT matches that of nuclear protein abundance in liver (orange triangles), and both are delayed compared to *Elf1* mRNA accumulation (gray). In *Bmal1* KO, ELF activity and *Elf1* mRNA are nonrhythmic. Error bars in nuclear protein, mRNA, and TF activity show SEM ($n = 2$).

consequently de-repression of REV-ERB target genes (Fig. 4B, top; Bugge et al. 2012). Activity of E-box in WT liver peaked at ZT7, consistent with BMAL1:CLOCK activity (Rey et al. 2011), albeit with weaker amplitudes compared to RORE activity, likely reflecting fewer E-box target genes compared to RORE in this module. In *Bmal1* KO mice, E-box activity was low and flat in liver, as expected.

We hypothesized that cooperativity of liver-specific and clock TFs at liver-specific DHSs can regulate liver-specific mRNA rhythms. Pairwise analysis of TF binding sites at liver-specific DHSs found enrichment of co-occurrence between RORE and liver-specific TF motifs, FOXA2, ONECUT, and CUX2 (Fig. 4C). Enrichment of both CUX2 and ONECUT (also named HNF6) is consistent with ONECUT1 binding to both ONECUT and CUX2 motifs (Conforto et al. 2015). mRNAs of genes with co-occurrence of RORE and liver-specific TF motifs peaked near ZT1, consistent with peak RORE activity (near ZT21) preceding peak mRNA abundance of REV-ERB targets (Supplemental Fig. S5B). Analysis of ChIP-exo data sets targeting FOXA2, ONECUT1, and REV-ERBa in liver (Wang et al. 2014; Zhang et al. 2015; Iwafuchi-Doi et al. 2016) confirmed colocalized TF binding at liver-specific DHSs distal from clock-driven liver mRNAs such as *Insig2* and *Slc44a4* (Fig. 4D).

Liver-specific chromatin loops regulate liver-specific mRNA rhythms

To test whether distally located liver-specific DHSs can contact promoters of clock-driven liver-rhythmic genes, we selected the

promoters of *Mreg*, *Pik3ap1*, and *Slc44a1* as baits for 4C-seq experiments in liver and kidney harvested at the time of peak mRNA accumulation for the selected genes (Methods; Fig. 5A; Supplemental Figs. S6A, S7A). Upstream of *Mreg*, the 4C-seq signal, which measures frequency of promoter-enhancer contacts (van de Werken et al. 2012), decayed rapidly to background level in both liver and kidney (Fig. 5B, top). Downstream from *Mreg*, however, the 4C-seq signal showed a tissue-dependent pattern, decaying slowly in the liver but more rapidly in the kidney. This difference in decay suggests increased frequency of promoter-enhancer contacts in the liver compared to the kidney. Indeed, differential analysis identified liver-specific chromatin contacts 40 kb downstream from the promoter (Fig. 5B, bottom). Overlaying the contact data with DNase-seq, we found that liver-specific chromatin contacts downstream from *Mreg* connected liver-specific DHSs with the *Mreg* promoter (Fig. 5C). Furthermore, ChIP-exo showed colocalization of REV-ERBa and FOXA2 binding at liver-specific DHSs contacting the promoters (Fig. 5C). In contrast, accessible regions upstream of the *Mreg* promoter did not show liver-specific chromatin contacts. The 4C-seq data thus suggest that liver-specific chromatin loops can recruit clock-bound distal elements to promoters to regulate liver-specific transcriptional rhythms. Other liver-specific rhythmic transcripts, *Pik3ap1* and *Slc44a1*, also displayed liver-specific chromatin loops between promoter and liver-specific open chromatin regions (Supplemental Figs. S6, S7). In sum, tissue-specific chromatin looping can drive tissue-specific mRNA rhythms.

Tissue-specific circadian transcription regulation

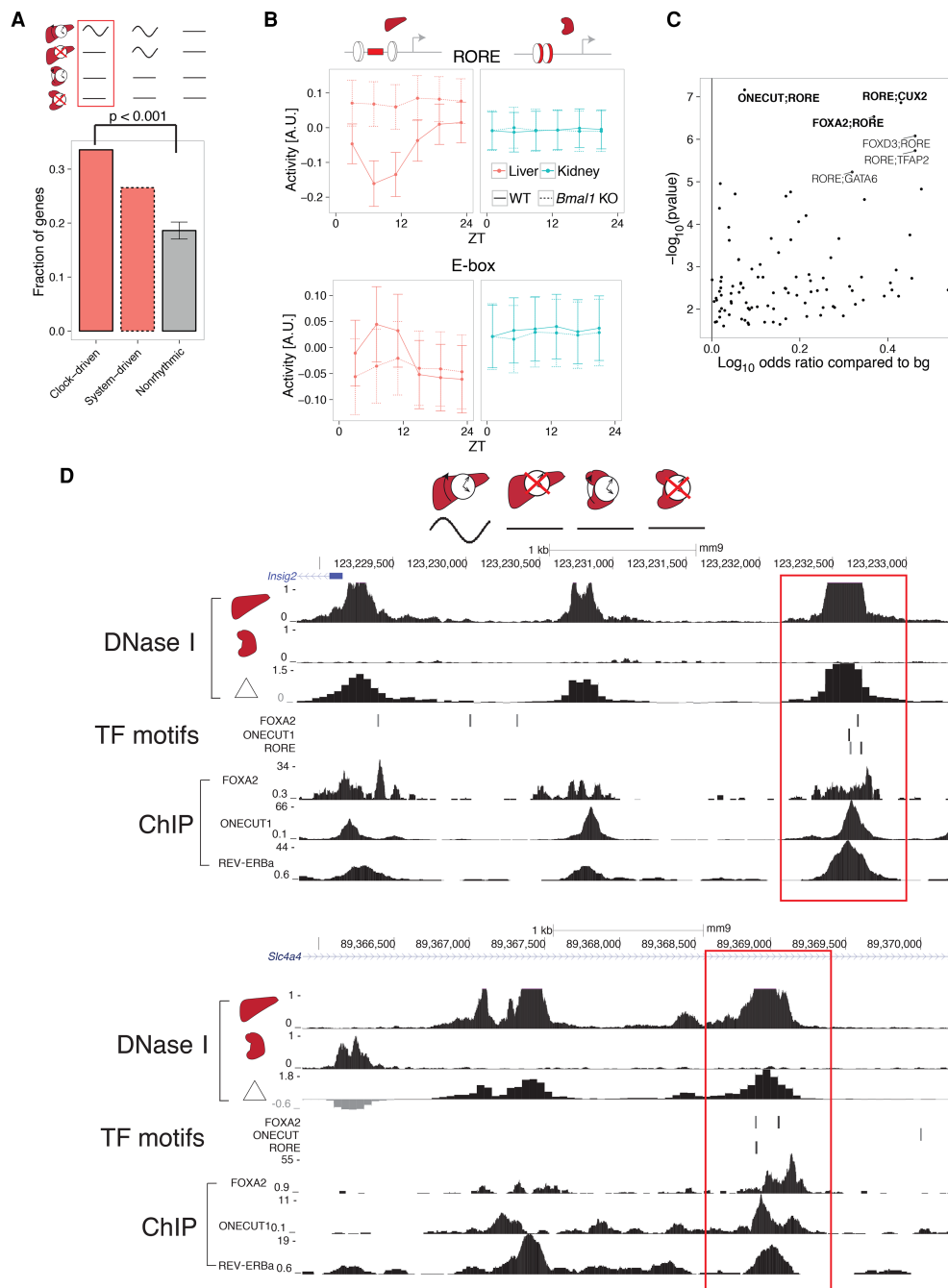


Figure 4. Colocalized binding of clock- and liver-specific TFs underlies liver-specific mRNA rhythms. (A) The fraction of genes containing liver-specific DNase I hypersensitive sites (DHSs) in the clock-driven liver-specific module is higher compared with both nonrhythmic and system-driven liver-specific modules. Error bars and *P*-values calculated from 10,000 bootstrap iterations. (B) Predicted temporal activities of RORE (top) and E-box (bottom) TF motifs located within liver-specific DHSs. Error bars show standard deviation of the estimated activities. (C) Co-occurrence of RORE with all other TFs in the SwissRegulon database (Pachkov et al. 2007) (189 TF motifs). Positive \log_{10} odds ratios (ORs) represent pairs of motifs enriched in the clock-driven liver-specific module compared to the flat module. *P*-values for the motif pairs were calculated from χ^2 tests applied to three-way contingency tables (Mýšicková et al. 2012). Selected pairs are in bold. (D) DNase I hypersensitivity in liver, kidney, and the corresponding differential signal (in \log_2 fold change) near two representative genes (top: *Insig2*; bottom: *Slc4a4*). RORE, ONECUT1, and FOXA2 TF binding motifs (posterior probability > 0.5, MotEvo) co-occur at liver-specific DHSs (red boxes). Predicted TF binding sites correspond to experimentally observed TF binding in publicly available ChIP-exo data sets for REV-ERBa, ONECUT1, and FOXA2 (bottom).

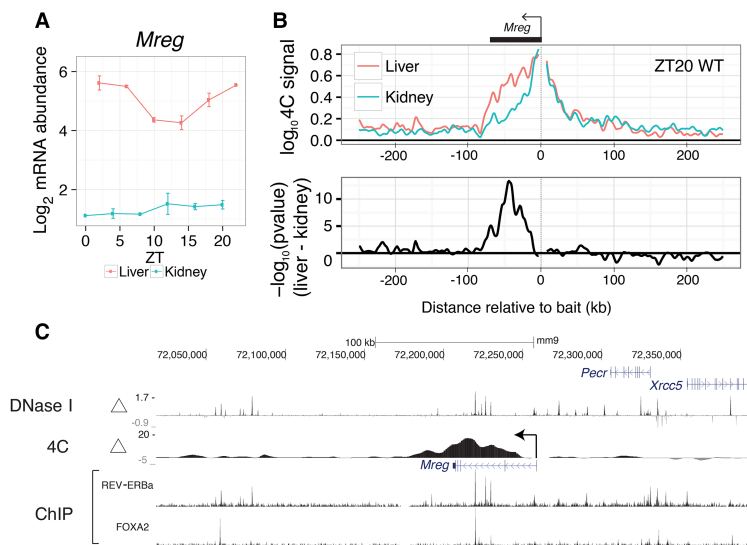


Figure 5. Liver-specific chromatin loops regulate liver-specific mRNA rhythms. (A) Temporal mRNA profile for *Mreg*, a clock-driven liver-rhythmic gene. Error bars are SEM ($n = 2$). (B) 4C-seq profiles (summary from two replicates, each pooling two different mice) using the *Mreg* promoter as a bait in liver and kidney at ZT20. Data are shown in a window of ± 250 kb from the bait (top). Profiles of differential contacts between liver and kidney (bottom) represented as signed log P -values (regularized t -test, positive values denote liver-enriched 4C contacts). (C) Tracks of differential 4C contacts (signed log P -values), log₂ fold change of DNase I hypersensitivity between liver and kidney, and ChIP-seq of REV-ERBa and FOXA2. Regions of significant differential 4C contacts correspond to liver-specific DNase I hypersensitive regions and REV-ERBa binding sites.

Precise promoter-enhancer contacts underlie liver-specific mRNA rhythms

To test whether distinct chromatin loops would form at alternative nearby gene promoters with distinct temporal mRNA profiles, we searched for candidate genes where one promoter was rhythmically transcribed while the alternative one was nonrhythmic (Supplemental Fig. S8). *Slc45a3* has two alternative transcripts using different promoters 8 kb apart. The shorter pre-mRNA oscillated in the liver (rhythmic promoter, *Slc45a3*-short), while the longer did not (flat promoter, *Slc45a3*-long). In kidney, neither *Slc45a3*-short nor *Slc45a3*-long showed robust transcript rhythms (Supplemental Fig. S9). Targeting the *Slc45a3*-short promoter with 4C-seq in liver and kidney showed liver-specific chromatin loops at three distal regions (two upstream, one downstream) (Fig. 6A). These same regions did not form liver-specific chromatin loops with the *Slc45a3*-long promoter (Fig. 6B), suggesting that promoters 8 kb apart can contact distinct enhancers. Overlaying 4C-seq with DNase-seq showed that these chromatin loops link liver-specific DHSs specifically to the *Slc45a3*-short promoter (Fig. 6C). These liver-specific DHSs were bound by liver-specific TFs, FOXA2 and ONECUT1, and clock TF, REV-ERBa, as shown in ChIP-seq. The 4C experiments suggest that enhancers can contact a rhythmic promoter while looping out nearby nonrhythmic alternative promoters, confining rhythmic enhancer activity to specific promoters (Fig. 6D). Furthermore, rhythmically active enhancers can contact promoters in a tissue-specific manner. Thus, chromatin folding not only regulates tissue-specific rhythms but also differentiates between closely spaced promoters to control rhythmic transcription with spatial precision.

Discussion

The mammalian genome encodes transcriptional programs that allow the molecular clock to robustly oscillate across diverse tissue transcriptomes while maintaining flexibility to regulate distinct clock outputs in different combinations of tissues. Here, we identified two regulatory modes underlying tissue-specific transcript rhythms: Regulatory sequences can recruit individual TFs bearing rhythmic activity; coordinated binding of clock- and tissue-specific TFs can generate tissue-specific rhythms. Moreover, we found that clock- and tissue-specific TFs bound at distal enhancers can be recruited to promoters through precise chromatin loops.

Several of our predictions of transcription regulators and regulated genes (e.g., *Egr1*, *Por*, *Upp2*) corroborated with previous analyses of independent data sets (Yan et al. 2008; Bozek et al. 2009; Bhargava et al. 2015). Further analysis incorporating outputs of enhancer activity, such as eRNAs (Fang et al. 2014), across multiple tissues may uncover additional rhythmically active regulators.

Colocalized binding of clock- and tissue-specific TFs at enhancers provides a putative mechanism for the clock to regulate

clock output genes in a tissue-specific manner. In mouse liver, clock TFs can colocalize with liver-specific TFs, such as FOXA2 and ONECUT1, consistent with multiple TFs associating with liver-specific DHSs (Iwafuchi-Doi et al. 2016). Our findings are currently based on sequence-specific DNA binding of TFs, comparison of tissues, and ChIP-seq data sets. Further mechanistic basis for the functional significance of colocalization could be gained, for example, by using inducible knockout models for tissue-specific regulators. Moreover, the observed colocalization does not exclude other cooperative modes, such as tethering of REV-ERBa to ONECUT1 through protein-protein interactions (Zhang et al. 2015).

Our 4C analysis showed that chromatin looping might mediate interaction between clock- and tissue-specific transcriptional programs by recruiting clock-bound distal elements to promoters in a tissue-specific manner. Such loops can surgically discriminate between nearby promoters as close as 8 kb apart, suggesting a way to separate temporal regulation of neighboring promoters. A previous 4C study on a core clock gene enhancer proposed that cohesion-mediated promoter-enhancer looping can compartmentalize rhythmic gene expression within genomic regions spanning 150 kb (Xu et al. 2016). Here, chromatin interactions that differed between tissues were localized to a small genomic region (<10 kb) near promoters (<100 kb). Future studies integrating temporal data across tissues with large-scale promoter-enhancer networks may reveal regulatory sequences that encode promoter-enhancer compatibility and elucidate whether this compatibility is tissue-specific (Li and Noll 1994; Merli et al. 1996; Zabidi et al. 2014; Nguyen et al. 2016).

While our work focused on transcriptional mechanisms, studying other mechanisms such as post-transcriptional,

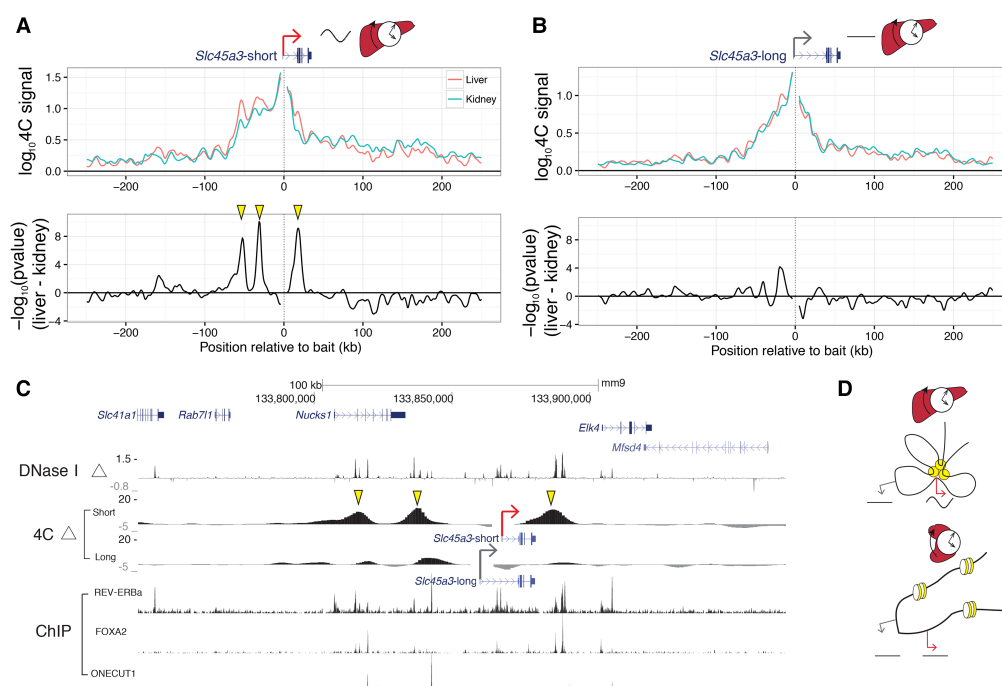


Figure 6. Precise promoter-enhancer contacts underlie liver-specific mRNA rhythms. (A,B) 4C-seq profiles for the (A) *Slc45a3*-short and (B) *Slc45a3*-long isoforms within ± 250 kb around baits targeting the two TSSs (top). Signed log *P*-values for differential contacts between liver and kidney (bottom) as in Figure 5B. TSSs for *Slc45a3*-short and *Slc45a3*-long are 8 kb apart. (C) Differential 4C contacts (signed log *P*-values), log₂ fold change of DNase I hypersensitivity between liver and kidney, and ChIP-exo signal of REV-ERBa, FOXA2, and ONECUT1. Regions of significant differential contacts in *Slc45a3*-short correspond to liver-specific DNase I hypersensitive regions. Yellow arrowheads in A and C show liver-specific distal contacts recruited to the *Slc45a3*-short TSS. These contacts are absent for *Slc45a3*-long TSS (B). (D) Schematic model illustrating enhancer-promoter interactions in liver and kidney that may generate liver-specific rhythms. Yellow circles illustrate liver-active enhancers contacting the rhythmic promoter (red arrow) but not the alternative nonrhythmic promoter (gray arrow). In kidney, the enhancer is not accessible, and both promoters are nonrhythmic.

translational, and post-translational processes using PRO-seq, Ribo-seq, and proteomics data may provide additional insights. Expanding our 24-h analysis to 12-h or other harmonics would broaden the view of tissue-specific temporal gene expression but may require experimental designs of higher temporal resolution (Hughes et al. 2009; Krishnaiah et al. 2017). In sum, integrating the temporal axis into tissue-specific gene regulation offers an integrated understanding of how tissue physiology resonates with daily cycles in the environment.

Methods

Animal experiments

Eight- to 14-wk-old C57Bl/6 mice have been purchased from Charles River Laboratory. *Bmal1* KO mice have been previously described (Jouffe et al. 2013). Without further indications, mice are kept under a 12-h light/12-h dark regimen and ad libitum feeding. All animal care and handling was performed according to the Canton de Vaud (Fred Gachon, authorization no. VD 2720) laws for animal protection.

RNA-seq experiments

To complement the mouse liver WT and *Bmal1* KO RNA-seq data (GSE73554) (Atger et al. 2015), transcriptomes of kidneys

from *Bmal1* KO and WT littermates (12-h light/12-h regimen; night-restricted feeding) were measured from poly(A)-selected mRNA using single-end reads of length 100. mRNA levels were quantified using kallisto version 0.42.4 (mm10) (Bray et al. 2015).

Global temporal variance

For each tissue, we estimated the contribution of temporal variance for each gene, broken down by its Fourier components. We calculated the background level assuming temporally unstructured data (white noise), whose magnitude (strength of the white noise) was estimated from the mean of squared magnitudes of Fourier coefficients that were not submultiples of 24 h (i.e., the mean of 48-, 16-, 9.6-, 6.9-, 5.3-, 4.4-h components).

Model selection

We fitted harmonic regression models that integrated temporal gene expression across different combinations of rhythms in different conditions (Atger et al. 2015). We used a g-prior for the rhythmic parameters β as a penalization scheme (Liang et al. 2008). We set $g = 1000$, which we found to maximize temporal variations captured in the shared rhythms model while minimizing temporal variations captured in the flat model.

Complex singular value decomposition representation of gene and tissue module

We first transformed the time domain to the frequency domain corresponding to 24-h rhythms. The resulting matrix was decomposed using SVD; the first left-singular and first right-singular values were visualized in separate polar plots (Supplemental Methods).

Functional analysis by GO terms

We used Fisher's exact test to assess statistical significance of gene enrichment for each GO term. Foreground genes were genes with phases within a 6-h window. Background genes were genes assigned to a model. For each GO term, we slid the 6-h window with a step size of 1 h and calculated the *P*-value enrichment. GO terms were chosen by visualizing significant GO terms in the tree and choosing GO terms that were comparably deep in the tree.

Chromatin conformation experiments and analysis

C57Bl/6 mice were sacrificed at ZT08 and ZT20 to extract liver and kidneys. Liver and kidney nuclei were prepared as previously described (Ripperger and Schibler 2006) with some minor changes. 4C-seq assays were performed as in Gheldof et al. (2012). Raw read counts for each sample were normalized by library size by the sum of the read counts on the *cis*-chromosome (excluding 10 fragments around the bait). We used a weighted linear model to fit the \log_{10} signal around each fragment *f*. A Gaussian window of standard deviation = 2.5 kb centered on *f* was used to incorporate signal from neighboring fragments (Supplemental Methods). Differential contacts were estimated using *t*-statistics.

Data access

Raw and processed RNA-seq and 4C-seq data generated from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE100457.

Competing interest statement

C.J., J. Marquis, A.C., and F.G. are employees of Nestlé Institute of Health Sciences SA.

Acknowledgments

We thank Eric Paquet for critical reading and Saeed Omid for help and discussions in bioinformatics. This work was supported by Swiss National Science Foundation Grant 31003A-153340, European Research Council Grant ERC-2010-StG-260667, and the Ecole Polytechnique de Lausanne. J.Y. benefits from the Natural Sciences and Engineering Research Council of Canada Postgraduate Studies Doctoral scholarship.

Author contributions: Conceptualization, J.Y., J. Mermet, and F.N.; formal analysis, J.Y. and F.N.; investigation, J. Mermet, C.J., J. Marquis, and A.C.; writing, review, and editing, J.Y., J. Mermet, F.G., and F.N.; supervision, F.N. and F.G.; funding acquisition, F.N. and F.G.

References

Atger F, Gobet C, Marquis J, Martin E, Wang J, Weger B, Lefebvre G, Descombes P, Naef F, Gachon F. 2015. Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc Natl Acad Sci* **112**: E6579–E6588.

Balwierz PJ, Pachkov M, Arnold P, Gruber AJ, Zavolan M, van Nimwegen E. 2014. ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. *Genome Res* **24**: 869–884.

Bass J, Lazar MA. 2016. Circadian time signatures of fitness and disease. *Science* **354**: 994–999.

Bhargava A, Herzel H, Ananthasubramaniam B. 2015. Mining for novel candidate clock genes in the circadian regulatory network. *BMC Syst Biol* **9**: 78.

Bozek K, Relógio A, Kielbasa SM, Heine M, Dame C, Kramer A, Herzel H. 2009. Regulation of clock-controlled genes in mammals. *PLoS One* **4**: e4882.

Bray N, Pimentel H, Melsted P, Pachter L. 2015. Near-optimal RNA-Seq quantification. arXiv:1505.02710 [q-bio.QM].

Bugge A, Feng D, Everett LJ, Briggs ER, Mullican SE, Wang F, Jager J, Lazar MA. 2012. Rev-erb α and Rev-erb β coordinately protect the circadian clock and normal metabolic function. *Genes Dev* **26**: 657–667.

Chauvin C, Koka V, Nouschi A, Mieulet V, Hoareau-Aveilla C, Dreazen A, Cagnard N, Carpentier W, Kiss T, Meyuhos O, et al. 2014. Ribosomal protein S6 kinase activity controls the ribosome biogenesis transcriptional program. *Oncogene* **33**: 474–483.

Cho H, Zhao X, Hatori M, Yu RT, Barish GD, Lam MT, Chong LW, DiTacchio L, Atkins AR, Glass CK, et al. 2012. Regulation of circadian behaviour and metabolism by REV-ERB- α and REV-ERB- β . *Nature* **485**: 123–127.

Conforto TL, Steinhardt GF, Waxman DJ. 2015. Cross talk between GH-regulated transcription factors HNF6 and CUX2 in adult mouse liver. *Mol Endocrinol* **29**: 1286–1302.

Deckard A, Anafi RC, Hogenesch JB, Haase SB, Harer J. 2013. Design and analysis of large-scale biological rhythm studies: a comparison of algorithms for detecting periodic signals in biological data. *Bioinformatics* **29**: 3174–3180.

Dibner C, Schibler U. 2015. Circadian timing of metabolism in animal models and humans. *J Intern Med* **277**: 513–527.

Ekberg K, Landau BR, Wajngot A, Chandramouli V, Efendic S, Brunengraber H, Wahren J. 1999. Contributions by kidney and liver to glucose production in the postabsorptive state and after 60 h of fasting. *Diabetes* **48**: 292–298.

Fang B, Everett LJ, Jager J, Briggs E, Armour SM, Feng D, Roy A, Gerhart-Hines Z, Sun Z, Lazar MA. 2014. Circadian enhancers coordinate multiple phases of rhythmic gene transcription in vivo. *Cell* **159**: 1140–1152.

Fisher RA. 1929. Tests of significance in harmonic analysis. *Proc R Soc Lond Ser A* **125**: 54–59.

Gheldof N, Leleu M, Noordermeer D, Rougemont J, Reymond A. 2012. Detecting long-range chromatin interactions using the chromosome conformation capture sequencing (4C-seq) method. *Methods Mol Biol* **786**: 211–225.

Gotic I, Omid S, Fleury-Olela F, Molina N, Naef F, Schibler U. 2016. Temperature regulates splicing efficiency of the cold-inducible RNA-binding protein gene *Cirbp*. *Genes Dev* **30**: 2005–2017.

Gréchez-Cassiau A, Rayet B, Guillaumond F, Teboul M, Delaunay F. 2008. The circadian clock component BMAL1 is a critical regulator of p21WAF1/CIP1 expression and hepatocyte proliferation. *J Biol Chem* **283**: 4535–4542.

Guillou H, Zdravcov D, Martin PG, Jacobsson A. 2010. The key roles of elongases and desaturases in mammalian fatty acid metabolism: insights from transgenic mice. *Prog Lipid Res* **49**: 186–199.

Hughes ME, DiTacchio L, Hayes KR, Vollmers C, Pulivarthy S, Baggs JE, Panda S, Hogenesch JB. 2009. Harmonics of circadian gene transcription in mammals. *PLoS Genet* **5**: e1000442.

Hughes ME, Hong HK, Chong JL, Indacochea AA, Lee SS, Han M, Takahashi JS, Hogenesch JB. 2012. Brain-specific rescue of clock reveals system-driven transcriptional rhythms in peripheral tissue. *PLoS Genet* **8**: e1002835.

Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts JA, Mahony S, Pugh BF, Lee D, Kaestner KH, Zaret KS. 2016. The pioneer transcription factor FoxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation. *Mol Cell* **62**: 79–91.

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

Kelsall IR, Rosenzweig D, Cohen PTW. 2009. Disruption of the allosteric phosphorylation *a* regulation of the hepatic glycogen-targeted protein phosphatase 1 improves glucose tolerance in vivo. *Cell Signal* **21**: 1123–1134.

Kim JS, Chae JH, Cheon YP, Kim CG. 2016. Reciprocal localization of transcription factors YY1 and CP2c in spermatogonial stem cells and their putative roles during spermatogenesis. *Acta Histochem* **118**: 685–692.

Korenčič A, Košir R, Bordenyugov G, Lehmann R, Rozman D, Herzel H. 2014. Timing of circadian genes in mammalian tissues. *Sci Rep* **4**: 1349–1354.

Korrmann B, Schaad O, Bujard H, Takahashi JS, Schibler U. 2007. System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. *PLoS Biol* **5**: e34.

- Krishnaiah SY, Wu G, Altman BJ, Growe J, Rhoades SD, Coldren F, Venkataraman A, Olererin-George AO, Francey LJ, Mukherjee S, et al. 2017. Clock regulation of metabolites reveals coupling between transcription and metabolism. *Cell Metab* **25**: 961–974.e4.
- Lamia KA, Storch KF, Weitz CJ. 2008. Physiological significance of a peripheral tissue circadian clock. *Proc Natl Acad Sci* **105**: 15172–15177.
- Le Martelot G, Claudel T, Gatfield D, Schaad O, Kornmann B, Lo Sasso G, Moschetta A, Schibler U. 2009. REV-ERB α participates in circadian SREBP signaling and bile acid homeostasis. *PLoS Biol* **7**: e1000181.
- Lee JW, Chou CL, Knepper MA. 2015. Deep sequencing in microdissected renal tubules identifies nephron segment-specific transcriptomes. *J Am Soc Nephrol* **26**: 2669–2677.
- Li X, Noll M. 1994. Compatibility between enhancers and promoters determines the transcriptional specificity of gooseberry and gooseberry neuro in the *Drosophila* embryo. *EMBO J* **13**: 400–406.
- Li J, Grant GR, Hogenesch JB, Hughes ME. 2015. Considerations for RNA-seq analysis of circadian rhythms. *Methods Enzymol* **551**: 349–367.
- Liang F, Paulo R, Molina G, Clyde MA, Berger JO. 2008. Mixtures of *g* priors for Bayesian variable selection. *J Am Stat Assoc* **103**: 410–423.
- Matsuo T, Yamaguchi S, Mitsui S, Emi A, Shimoda F, Okamura H. 2003. Control mechanism of the circadian clock for timing of cell division in vivo. *Science* **302**: 255–259.
- Matsuoka TA, Zhao L, Artner I, Jarrett HW, Friedman D, Means A, Stein R. 2003. Members of the large Maf transcription family regulate insulin gene transcription in islet β cells. *Mol Cell Biol* **23**: 6049–6062.
- Merkin J, Russell C, Chen P, Burge CB. 2012. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* **338**: 1593–1599.
- Merli C, Bergstrom DE, Cygan JA, Blackman RK. 1996. Promoter specificity mediates the independent regulation of neighboring genes. *Genes Dev* **10**: 1260–1270.
- Mermel J, Yeung J, Naef F. 2016. Systems chronobiology: global analysis of gene regulation in a 24-hour periodic world. *Cold Spring Harb Perspect Biol* **9**: a028720.
- Morf J, Rey G, Schneider K, Stratmann M, Fujita J, Naef F, Schibler U. 2012. Cold-inducible RNA-binding protein modulates circadian gene expression posttranscriptionally. *Science* **338**: 379–383.
- Myšičková A, Vingron M. 2012. Detection of interacting transcription factors in human tissues using predicted DNA binding affinity. *BMC Genomics* **13**(Suppl 1): S2.
- Nguyen TA, Jones RD, Snaveley AR, Pfenning AR, Kirchner R, Hemberg M, Gray JM. 2016. High-throughput functional comparison of promoter and enhancer activities. *Genome Res* **26**: 1023–1033.
- Nikolaeva S, Pradervand S, Centeno G, Zavadova V, Tokonami N, Maillard M, Bonny O, Firsov D. 2012. The circadian clock modulates renal sodium handling. *J Am Soc Nephrol* **23**: 1019–1026.
- Oosterveer MH, Schoonjans K. 2014. Hepatic glucose sensing and integrative pathways in the liver. *Cell Mol Life Sci* **71**: 1453–1467.
- Pachkov M, Erb I, Molina N, van Nimwegen E. 2007. SwissRegulon: a database of genome-wide annotations of regulatory sites. *Nucleic Acids Res* **35**(Database issue): D127–D131.
- Ravnskjaer K, Hogan MF, Lackey D, Tora L, Dent SY, Olefsky J, Montminy M. 2013. Glucagon regulates gluconeogenesis through KAT2B- and WDR5-mediated epigenetic effects. *J Clin Invest* **123**: 4318–4328.
- Refinetti R, Menaker M. 1992. The circadian rhythm of body temperature. *Physiol Behav* **51**: 613–637.
- Reinke H, Saini C, Fleury-Olela F, Dibner C, Benjamin IJ, Schibler U. 2008. Differential display of DNA-binding proteins reveals heat-shock factor 1 as a circadian transcription factor. *Genes Dev* **22**: 331–45.
- Rey G, Cesbron F, Rougemont J, Reinke H, Brunner M, Naef F. 2011. Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. *PLoS Biol* **9**: e1000595.
- Ripperger JA, Schibler U. 2006. Rhythmic CLOCK-BMAL1 binding to multiple E-box motifs drives circadian *Dbp* transcription and chromatin transitions. *Nat Genet* **38**: 369–374.
- Sayin SI, Wahlström A, Felin J, Jäntti S, Marschall HU, Bamberg K, Angelin B, Hyötyläinen T, Orešič M, Bäckhed F. 2013. Gut microbiota regulates bile acid metabolism by reducing the levels of tauro- β -muricholic acid, a naturally occurring FXR antagonist. *Cell Metab* **17**: 225–235.
- Shea HC, Head DD, Setchell KDR, Russell DW. 2007. Analysis of HSD3B7 knockout mice reveals that a 3 α -hydroxyl stereochemistry is required for bile acid function. *Proc Natl Acad Sci* **104**: 11526–11533.
- Shen N, Jiang S, Lu JM, Yu X, Lai SS, Zhang JZ, Zhang JL, Tao WW, Wang XX, Xu N, et al. 2015. The constitutive activation of Egr-1/C/EBP α mediates the development of type 2 diabetes mellitus by enhancing hepatic gluconeogenesis. *Am J Pathol* **185**: 513–523.
- Sinturel F, Gerber A, Mauvoisin D, Wang J, Gatfield D, Stubblefield JJ, Green CB, Gachon F, Schibler U. 2017. Diurnal oscillations in liver mass and cell size accompany ribosome assembly cycles. *Cell* **169**: 651–663.e14.
- Sobel JA, Krier I, Andersin T, Raghav S, Canella D, Gilardi F, Kalantzi AS, Rey G, Weger B, Gachon F, et al. 2017. Transcriptional regulatory logic of the diurnal cycle in the mouse liver. *PLoS Biol* **15**: e2001069.
- Storch KF, Lipan O, Leykin I, Viswanathan N, Davis FC, Wong WH, Weitz CJ. 2002. Extensive and divergent circadian gene expression in liver and heart. *Nature* **417**: 78–83.
- Stumvoll M, Meyer C, Perriello G, Kreider M, Welle S, Gerich J. 1998. Human kidney and liver gluconeogenesis: evidence for organ substrate selectivity. *Am J Physiol* **274**: E817–E826.
- Takahashi JS, Hong HK, Ko CH, McDearmon EL. 2008. The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nat Rev Genet* **9**: 764–775.
- van de Werken HJ, Landan G, Holwerda SJ, Hoichman M, Klous P, Chachik R, Splinter E, Valdes-Quezada C, Oz Y, Bouwman BA, et al. 2012. Robust 4C-seq data analysis to screen for regulatory DNA interactions. *Nat Methods* **9**: 969–972.
- Vollmers C, Schmitz RJ, Nathanson J, Yeo G, Ecker JR, Panda S. 2012. Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell Metab* **16**: 833–845.
- Wang L, Chen J, Wang C, Uusküla-Reimand L, Chen K, Medina-Rivera A, Young EJ, Zimmermann MT, Yan H, Sun Z, et al. 2014. MACE: model based analysis of ChIP-exo. *Nucleic Acids Res* **42**: e156.
- Wang J, Mauvoisin D, Martin E, Atger F, Galindo AN, Dayon L, Sizzano F, Palini A, Kussmann M, Waridel P, et al. 2017. Nuclear proteomics uncovers diurnal regulatory landscapes in mouse liver. *Cell Metab* **25**: 102–117.
- Xu Y, Guo W, Li P, Zhang Y, Zhao M, Fan Z, Zhao Z, Yan J. 2016. Long-range chromosome interactions mediated by cohesin shape circadian gene expression. *PLoS Genet* **12**: e1005992.
- Yan J, Wang H, Liu Y, Shao C. 2008. Analysis of gene regulatory networks in the mammalian circadian rhythm. *PLoS Comput Biol* **4**: 647–676.
- Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD, et al. 2014. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**: 355–364.
- Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O, Stark A. 2014. Enhancer–core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* **518**: 556–559.
- Zhang R, Lahens NF, Ballance HI, Hughes ME, Hogenesch JB. 2014. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci* **111**: 16219–16224.
- Zhang Y, Fang B, Emmett MJ, Damle M, Sun Z, Feng D, Armour SM, Remsberg JR, Jager J, Soccio RE, et al. 2015. Discrete functions of nuclear receptor Rev-erb α couple metabolism to the clock. *Science* **348**: 1488–1492.

Received June 23, 2017; accepted in revised form December 11, 2017.



Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs

Jake Yeung, Jérôme Mermet, Céline Jouffe, et al.

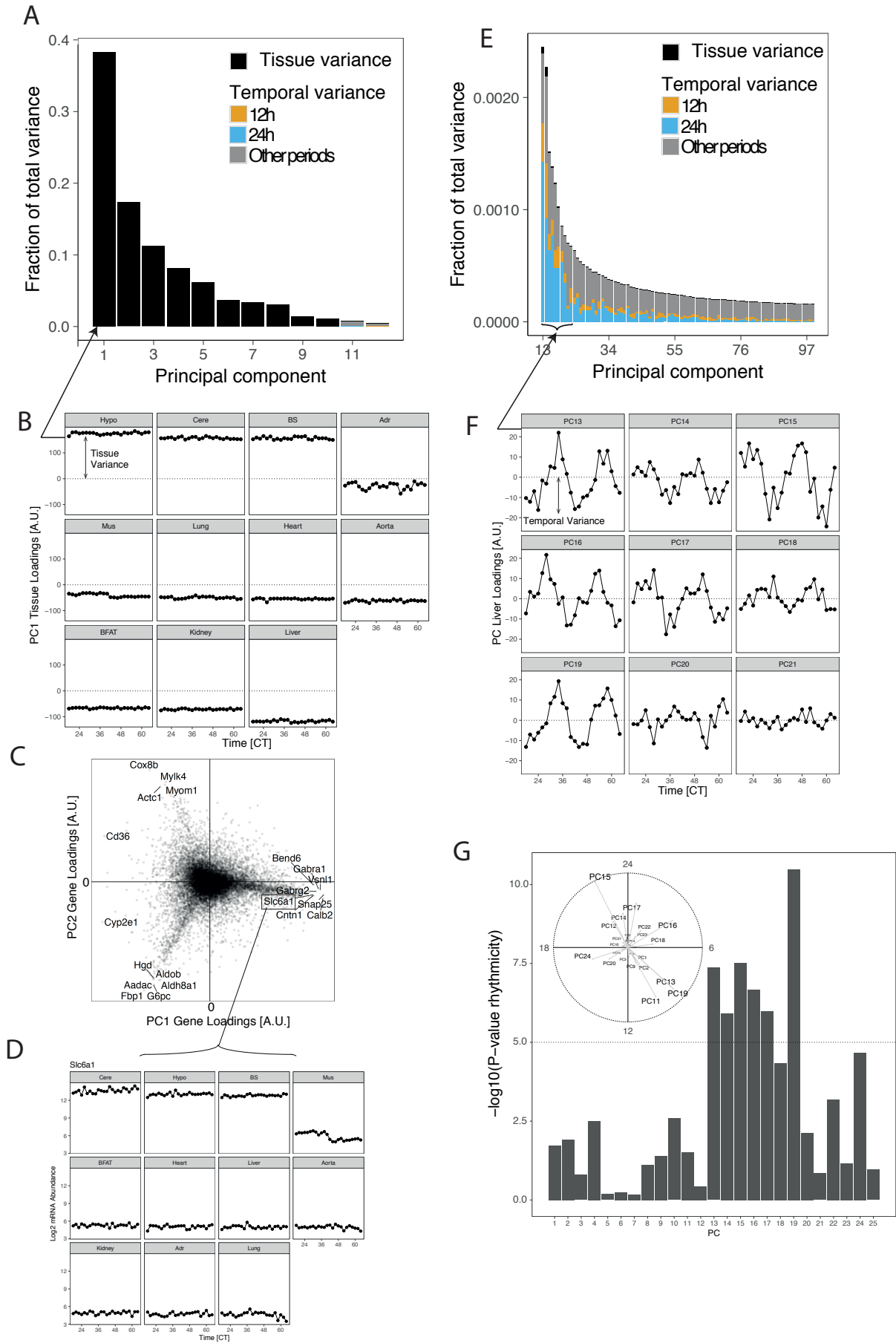
Genome Res. 2018 28: 182-191 originally published online December 18, 2017

Access the most recent version at doi:[10.1101/gr.222430.117](https://doi.org/10.1101/gr.222430.117)

Supplemental Material	http://genome.cshlp.org/content/suppl/2018/01/10/gr.222430.117.DC1
References	This article cites 67 articles, 20 of which can be accessed free at: http://genome.cshlp.org/content/28/2/182.full.html#ref-list-1
Open Access	Freely available online through the <i>Genome Research</i> Open Access option.
Creative Commons License	This article, published in <i>Genome Research</i> , is available under a Creative Commons License (Attribution 4.0 International), as described at http://creativecommons.org/licenses/by/4.0/ .
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here .

To subscribe to *Genome Research* go to:
<http://genome.cshlp.org/subscriptions>

Supplemental Figure S1



Supplemental Figure S1 - Contribution of tissue and temporal variance across 11 tissues

(A) Fraction of total variance explained by principal components 1 to 12. Colors represent the contribution of different temporal periods (Fourier coefficients) in each principal component. Components 1 to 12 show predominantly tissue differences.

(B) Sample loadings of 11 tissues over time in PC1. Contribution of tissue variance to the principal component is calculated by the sum of the squared difference between tissue mean and the global mean. Variance in PC1 consists of mostly tissue variance.

(C) Scatterplot of gene loadings from PC1 and PC2. Genes with large PC1 loadings are mainly brain-specific genes.

(D) mRNA abundance across tissues and time of a brain-specific GABA transporter, *Slc6a1*.

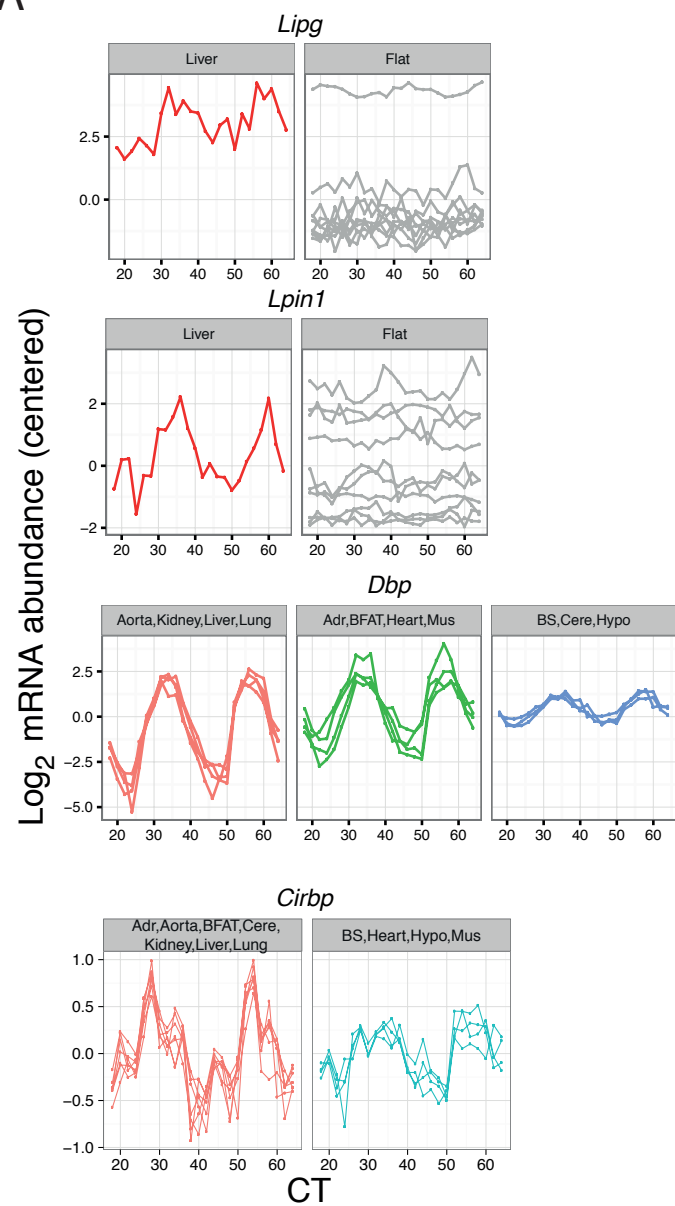
(E) *idem* as A, for principal components 13 to 100.

(F) Liver loadings over time for PCs 13 to 21. Contribution of temporal variance to the principal component is calculated by the sum of the squared difference between each time point and the tissue mean.

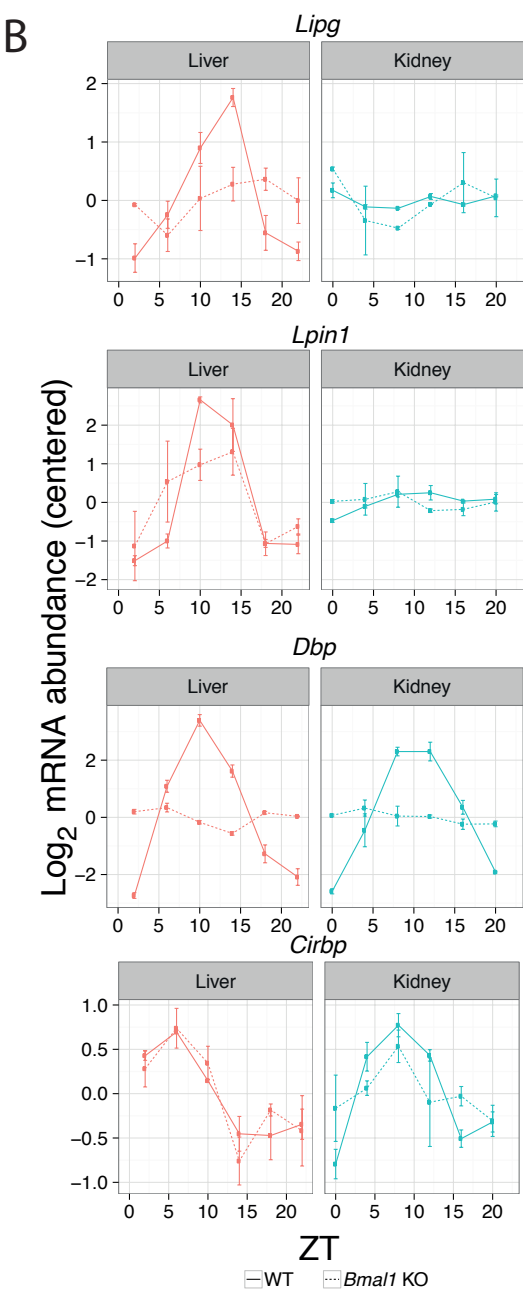
(G) Analysis of 24h rhythmicity of liver loadings over time (PC1 to PC25, harmonic regression). Inset: Amplitude (radial coordinate) and phase (clockwise angle) of liver loadings in PC1 and PC25.

Supplemental Figure S2

A



B



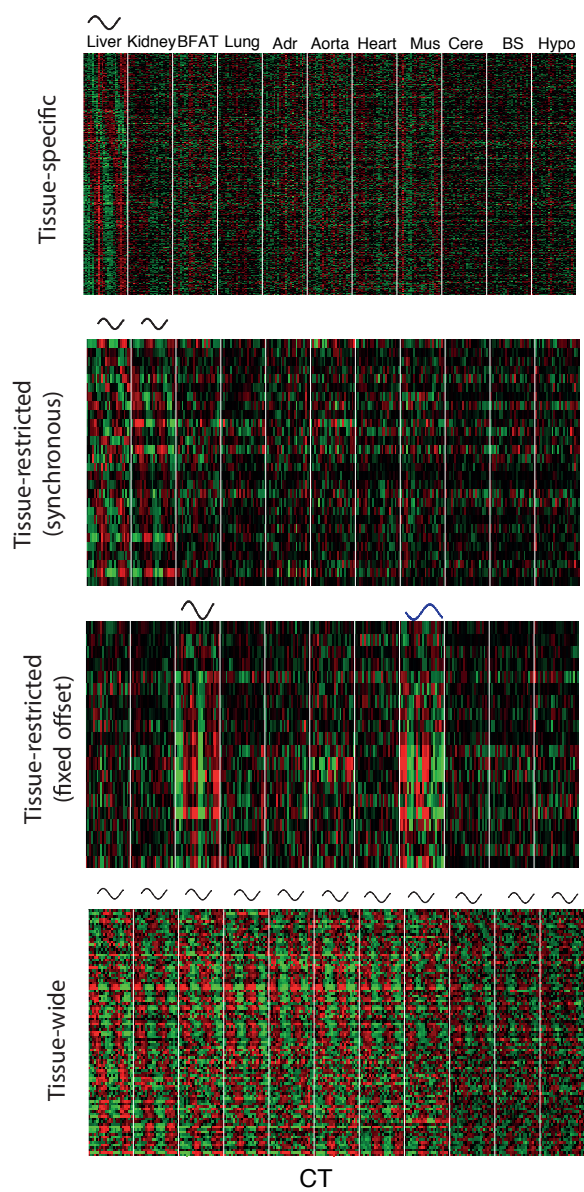
Supplemental Figure S2 – Temporal profiles of mRNA accumulation across tissues and in clock-disrupted conditions

(A) Examples of genes whose expression is rhythmic in different combinations of tissues. Flat indicates nonrhythmic tissues. Rhythms can be further classified as clock- or system-driven by analysis in WT and *Bmal1* KO data, shown in (B).

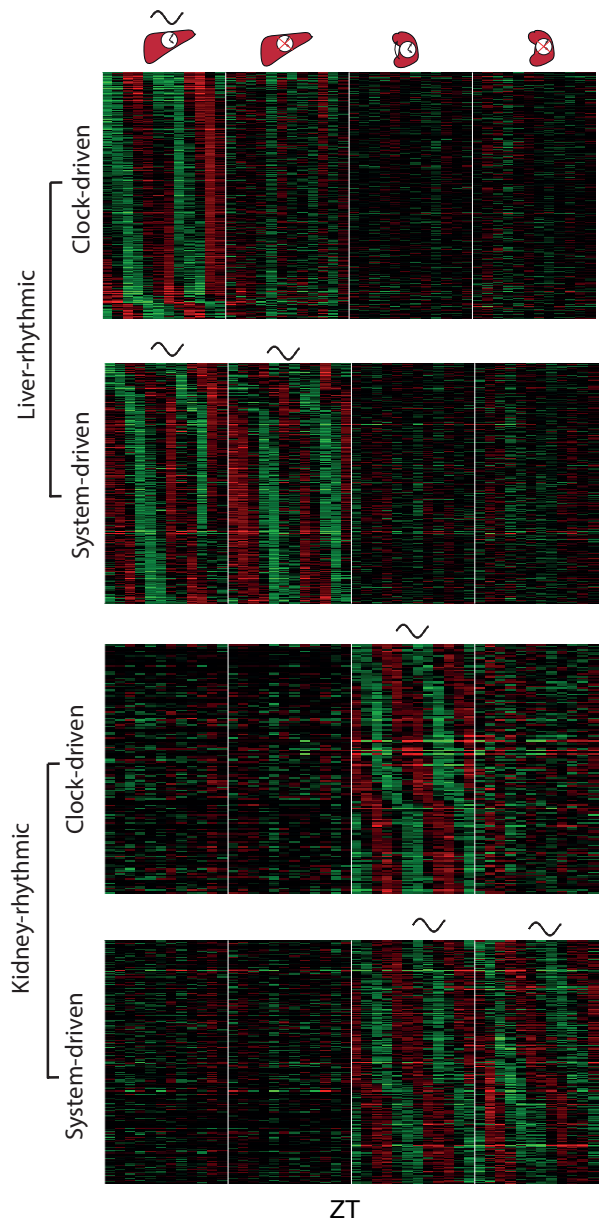
(B) Examples of tissue-specific (*Lipg*, *Lpin1*) and tissue-wide (*Dbp*, *Cirbp*) genes whose oscillations can be clock-dependent (*Lipg*, *Dbp*) or –independent (*Lpin1*, *Cirbp*) manner.

Supplemental Figure S3

A



B

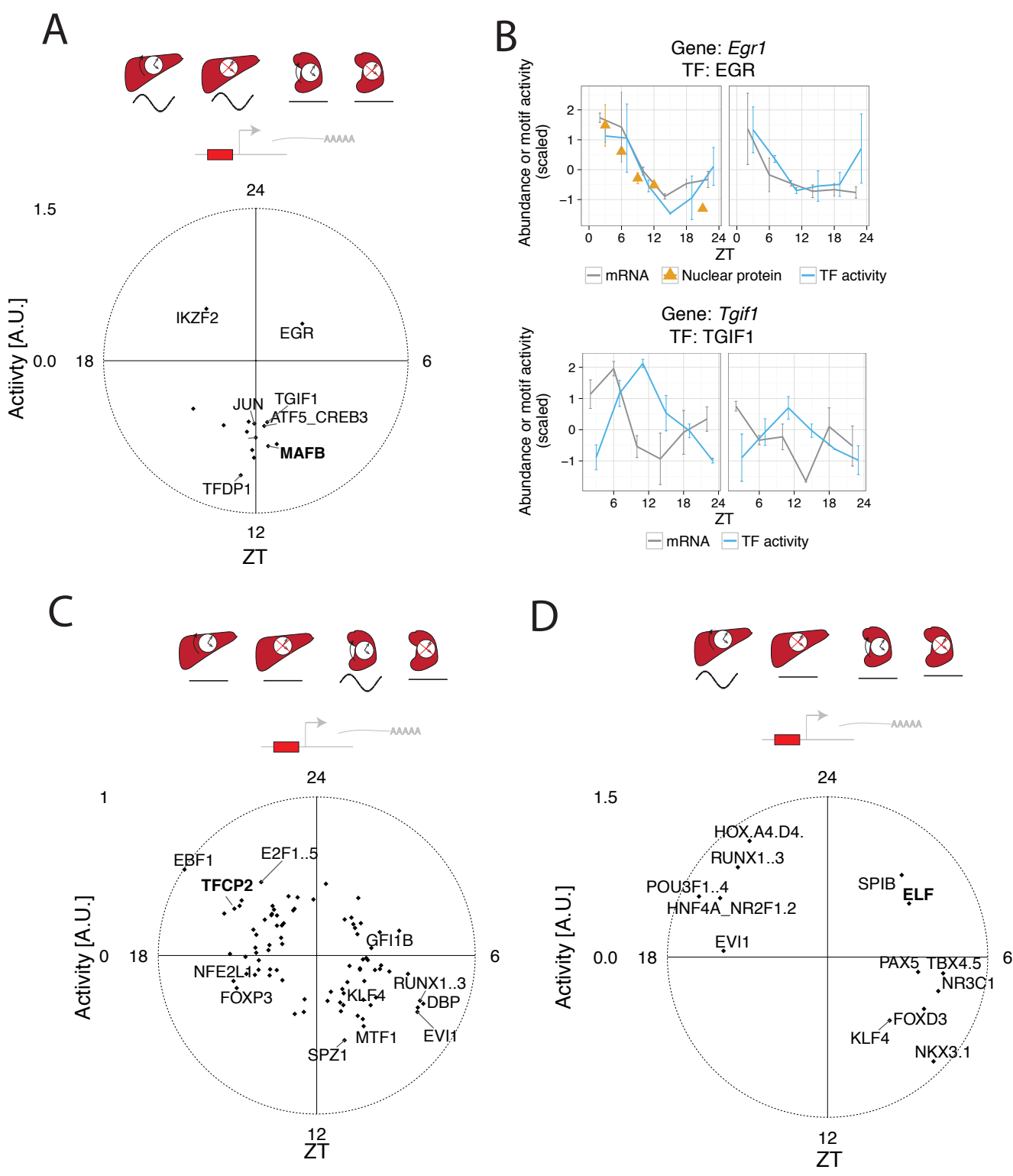


Supplemental Figure S3 – Modules of rhythmic gene expression across different subsets of tissues

(A) Heatmaps for the modules corresponding to genes rhythmic in different subsets of tissues. Each rectangle, demarcated by white vertical lines, represents normalized mRNA accumulation over 48 hours (red denotes low expression, green denotes high) for each gene in module (y-axis) in a single tissue. Schematic of the combination of tissues in which genes are rhythmic are shown above each heatmap. For clarity, genes in tissue-wide module (bottom) are filtered for amplitudes with average log₂ fold change greater than 0.8.

(B) Heatmaps representing modules of liver- and kidney-rhythmic genes that are driven by the local clock or systemic cues.

Supplemental Figure S4



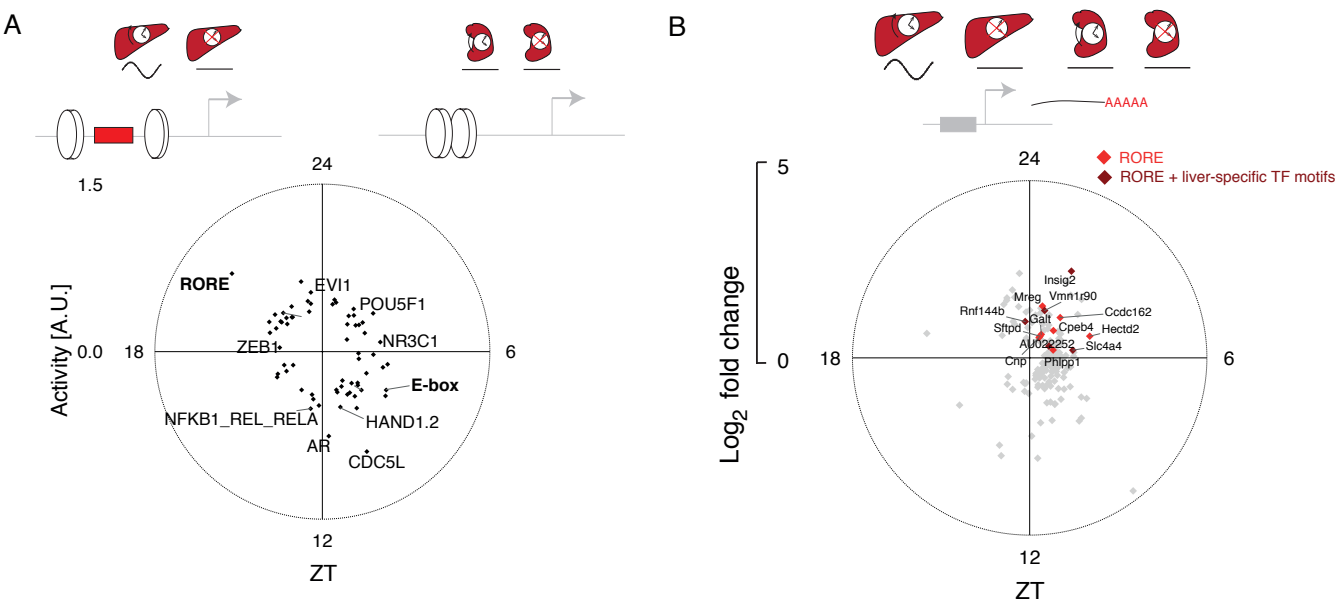
Supplemental Figure S4 - TF regulators associated with tissue-specific rhythmic gene expression

(A) Predicted activities of TF motifs associated with system-driven liver-rhythmic module (left).

(B) Predicted TF activity (solid line), nuclear protein abundance (if available, triangles), and mRNA accumulation (dotted) oscillate in both WT and *Bmal1* KO. Error bars in nuclear protein, mRNA, and TF activity are SEM (n=2).

(C,D) Predicted activities of TF motifs associated with clock-driven kidney-rhythmic (C), and clock-driven liver-rhythmic (D) module represented by the first component of complex-valued SVD. Motifs with z-score > 1.25 are shown. Candidate TFs shown in Figure 4B,D,F are displayed in bold.

Supplemental Figure S5

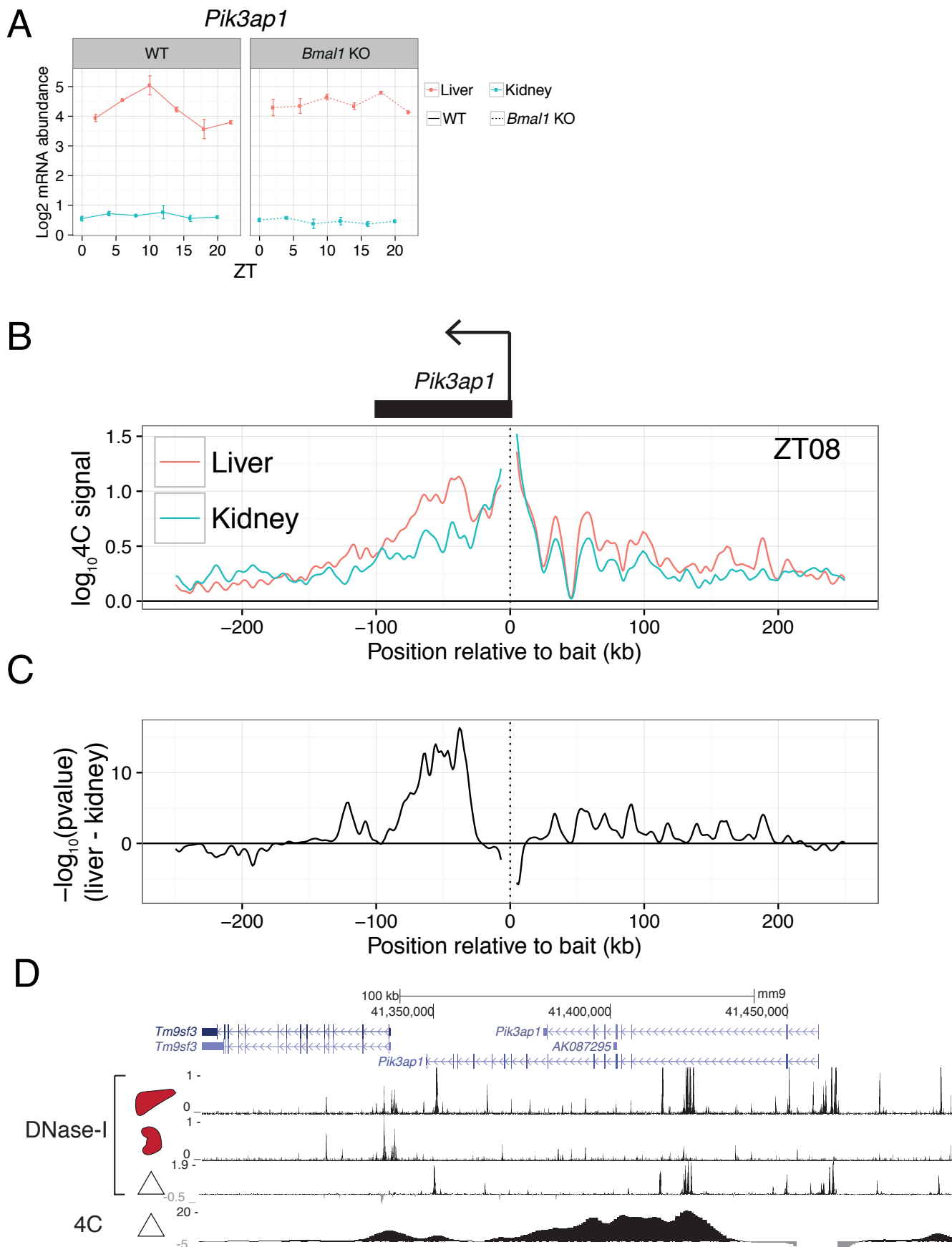


Supplemental Figure S5 – Liver-specific accessible regions harboring clock TF binding sites underlie clock-driven liver-specific rhythms

(A) Activities of TF motifs associated with clock-driven liver-rhythmic module, predicted using TF binding site occurrences at liver-specific DHSs.

(B) Genes containing RORE (red) or co-occurrence of RORE and liver-specific TF (ONECUT1, CUX2, or FOXA2, dark red) in clock-driven liver-rhythmic module. mRNA abundances of genes with liver-specific DHSs harboring RORE motifs peak approximately 3 hours after peak RORE activity.

Supplemental Figure S6



Supplemental Figure S6 - Chromatin loops between liver-specific enhancers and promoter of *Pik3ap1*

(A) Temporal mRNA abundance profiles of *Pik3ap1* in liver and kidney of mice with (left) and without (right) a functioning clock. Nuclei were extracted from WT liver and kidney from 4 mice at ZT8 to perform 4C-Seq.

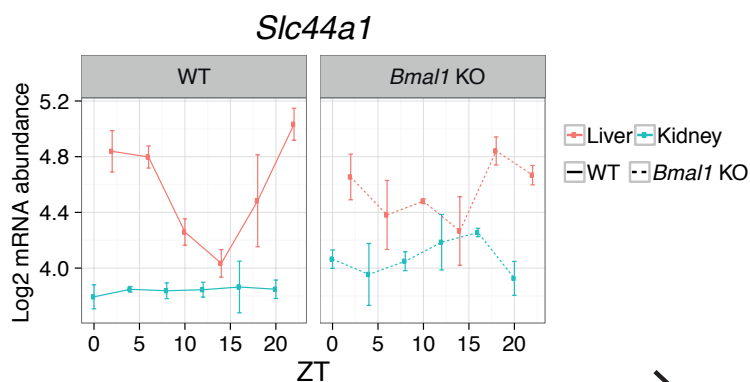
(B) 4C-Seq profiles (summary from 2 samples, each pooled from 2 mice) using the *Pik3ap1* promoter as a bait in liver and kidney within a window of 500 kb.

(C) Profiles of differential contacts for liver versus kidney, shown as signed log p-values (regularized t-test, positive values show liver-enriched 4C contacts).

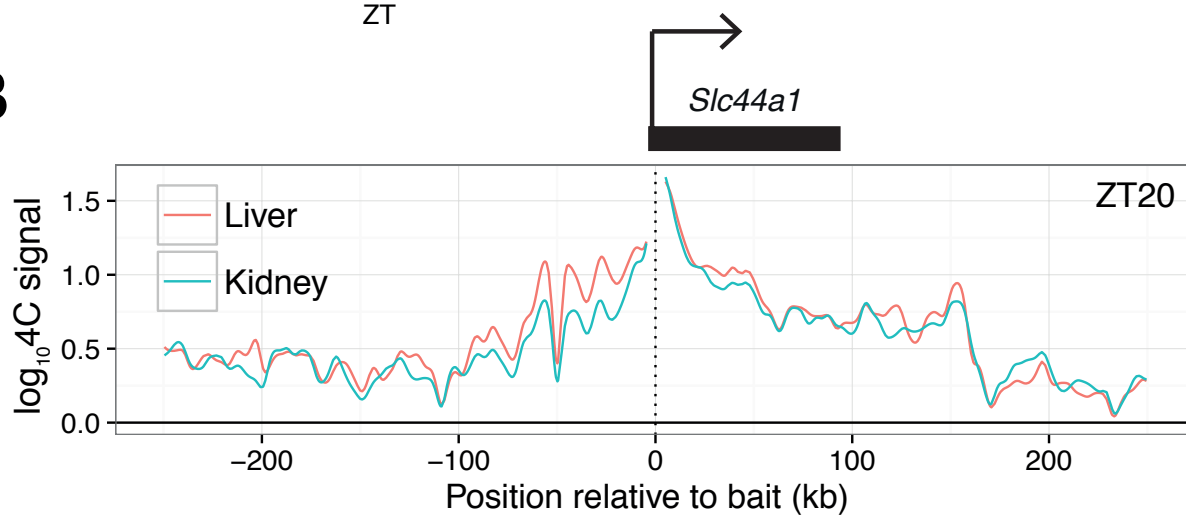
(D) Tracks of differential contacts (signed log p-values), DNase-I hypersensitivity in liver, kidney, and their log₂ fold change. Regions of significant differential contacts correspond to liver-specific DHS regions.

Supplemental Figure S7

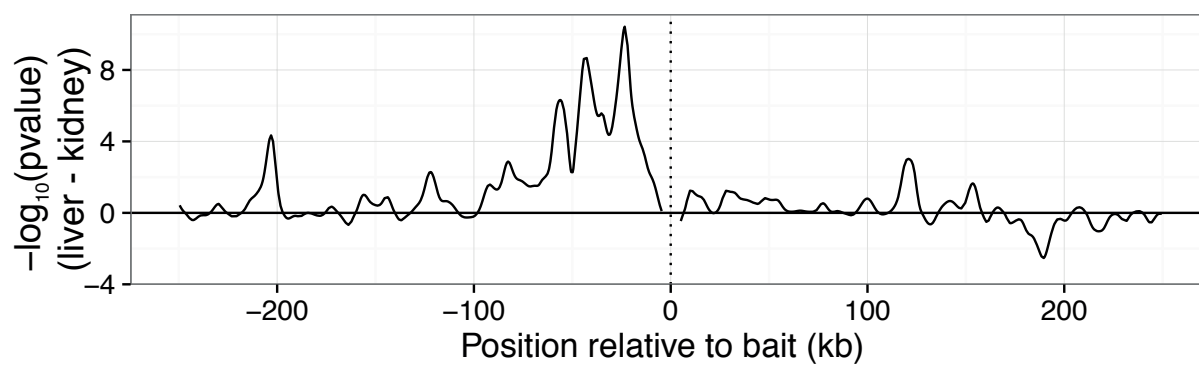
A



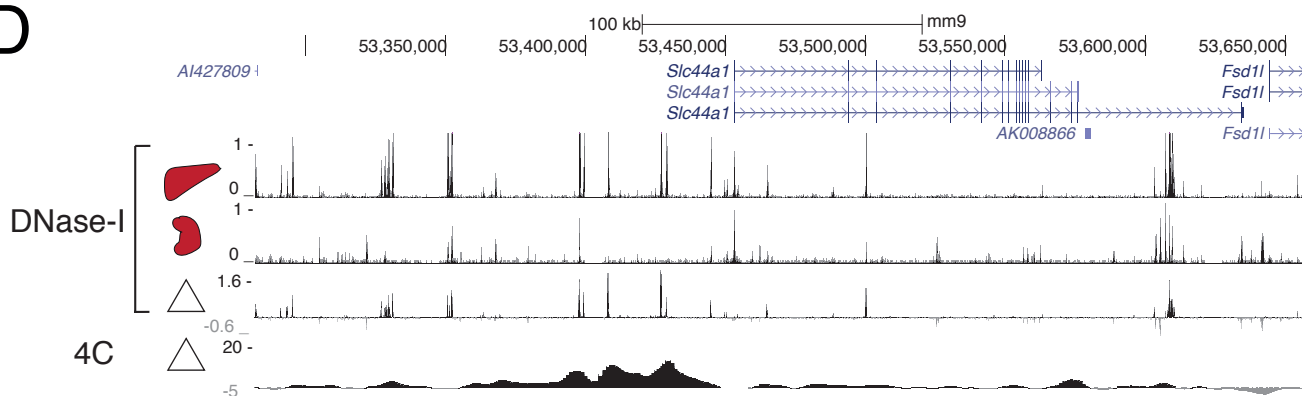
B



C



D



Supplemental Figure S7 - Chromatin loops between liver-specific enhancers and promoters of *Slc44a1* transcript

(A) Temporal mRNA abundance profiles of *Slc44a1* in liver and kidney of mice with (left) and without (right) a functioning clock. Nuclei were extracted from WT liver and kidney from 4 mice at ZT20 to perform 4C-Seq.

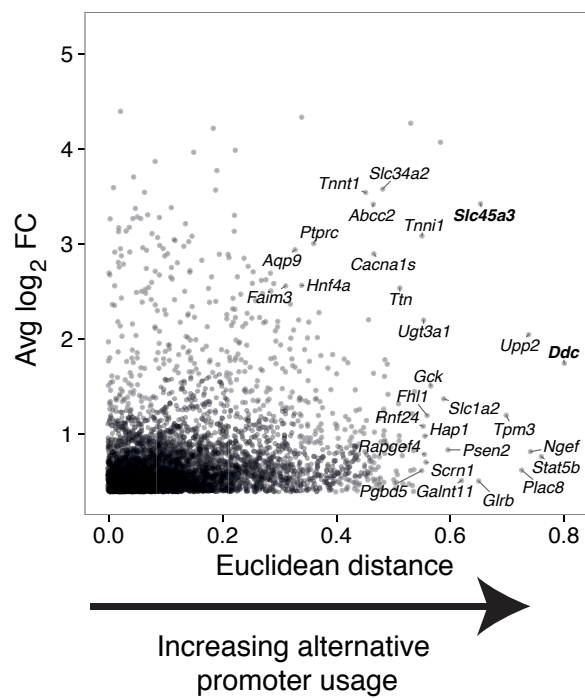
(B) 4C-Seq profiles (summary from 2 samples, each pooled from 2 mice) using the *Slc44a1* promoter as a bait in liver and kidney within a window of 500- kb.

(C) Profiles of differential contacts between liver versus kidney, shown as signed log p-values (regularized t-test, positive values show liver-enriched 4C contacts).

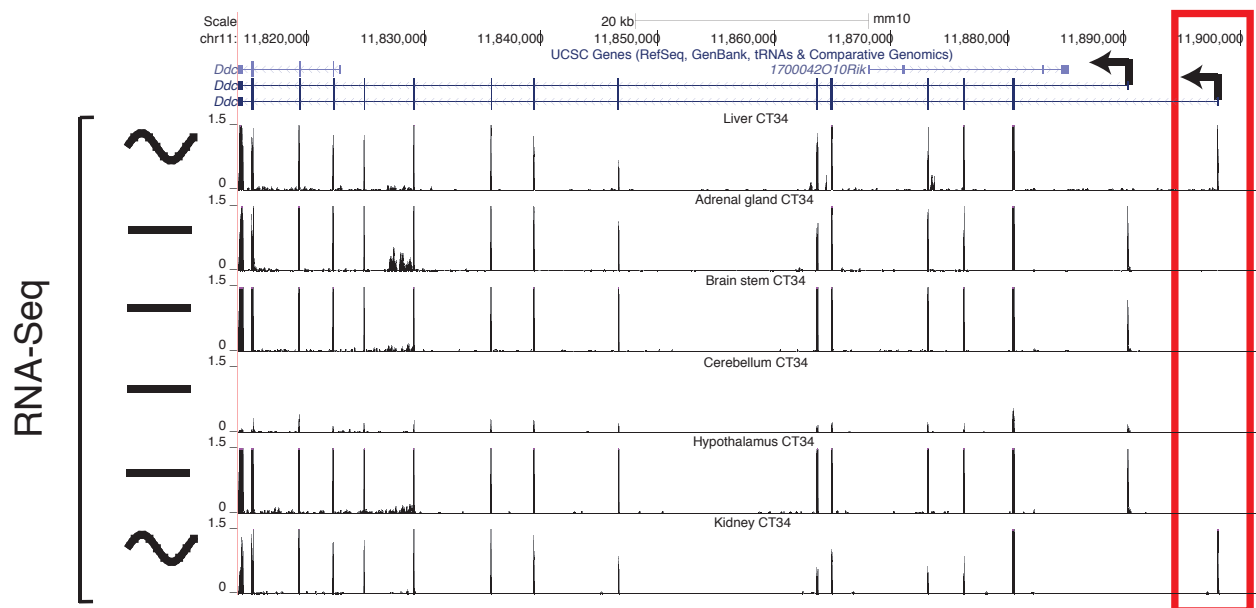
(D) Tracks of differential contacts (signed log p-values), DNase-I hypersensitivity in liver, kidney, and their log2 fold change. Regions of significant differential contacts correspond to liver-specific DHS regions.

Supplemental Figure S8

A



B

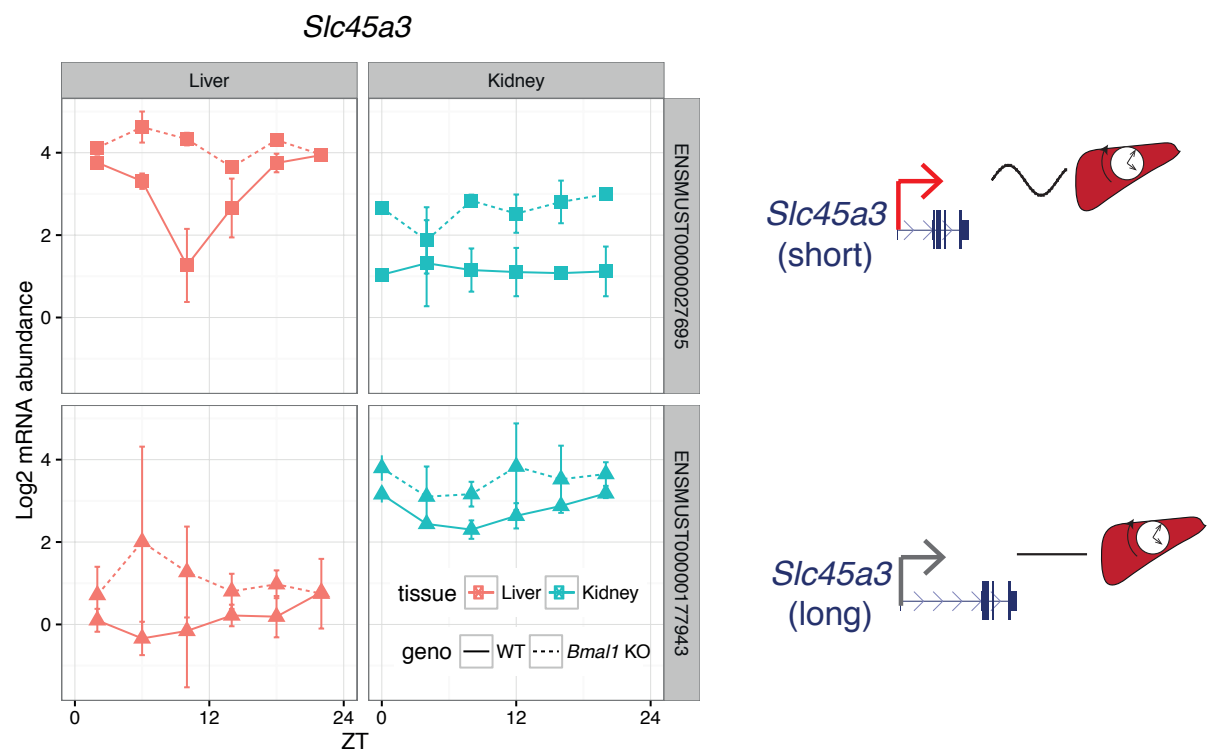


Supplemental Figure S8 - Correlations of alternative transcript start site (TSS) usage and tissue-specific rhythms in mRNA accumulation

(A) Scatterplot of alternative promoter usage versus log₂ fold-change (FC) of rhythmic tissues. Alternative TSS usage defined by calculating the Euclidean distance of transcript expression in tissues with rhythmic transcript versus tissues with nonrhythmic transcripts.

(B) Example of alternative TSS usage in rhythmic gene expression. RNA-Seq of *Ddc* at CT34 shows that rhythmic tissues (liver and kidney) use an upstream promoter whereas nonrhythmic tissues use a downstream promoter.

Supplemental Figure S9



Supplemental Figure S9 — Liver and kidney use different TSSs in *Slc45a3*, a clock-driven liver transcript

Temporal abundance profiles of two transcripts of *Slc45a3* in liver and kidney of mice with and without a functioning clock. *Slc45a3*-short isoform is rhythmic specifically in the liver; *Slc45a3*-long isoform is not robustly rhythmic in liver or kidney.

2.6 Conclusion and perspectives

This paper investigates transcriptional regulatory modes that allow diurnal gene expression to be tissue-specific. Our study reveals that the circadian clock is embedded into tissue-specific regulatory networks. Tissue-specific chromatin interactions can establish this regulation, allowing circadian clock transcription factors to regulate a gene in a tissue-specific manner. In fact, genes that oscillate with 24-hour rhythm in all tissues are a small minority compared tissue-specific or tissue-restricted oscillations.

One mode by which the clock can embed into tissue-specific regulatory networks is through tissue-specific chromatin interactions. Our data suggest that circadian clock transcription factors (TFs) can bind to tissue-specific enhancers, which allow clock TFs to regulate gene expression oscillations in a tissue-specific manner through promoter-enhancer looping. Further questions and extensions for understanding tissue-specific circadian gene expression include:

1. Incorporate mean expression to comparisons of oscillatory gene expression across tissues.
2. Extend method to 12-hour rhythms.
3. Investigate post-transcriptional regulation underlying tissue-specific gene expression

3 Clock-dependent chromatin topology modulates circadian transcription and behavior

3.1 Introduction

This chapter looks into the circadian dynamics of promoter-enhancer looping and the function of rhythmically active enhancers in gene transcription as well as circadian locomotor activity. The work has been published in *Genes & Development*, 2018 under the Creative Commons license and is reproduced here. The main text as well as supplemental figures are attached. Due to the size of supplemental tables, they are not attached but can be found on the online version of the *Genes & Development* article. No changes were made to the main text or supplemental figures, which were downloaded from doi:10.1101/gad.312397.118.

3.2 Contributions

J  rome Mermet did the 4C-seq and RNA-seq experiments, with help from C  line Jouffe, Damien Nicolas, and Daniel Mauvoisin. I performed the analysis of the 4C-seq, RNA-seq, and ChIP-seq data, with help from Kyle Gustafson and Felix Naef in the initial stages of the project. Cl  mence performed and analyzed the smRNA-FISH experiments, with some help from me in the analysis. Yann Emmenegger from the Paul Franken lab performed the locomotor activity experiments.

Clock-dependent chromatin topology modulates circadian transcription and behavior

Jérôme Mermet,^{1,4} Jake Yeung,^{1,4} Clémence Hurni,¹ Daniel Mauvoisin,¹ Kyle Gustafson,¹ Céline Jouffe,² Damien Nicolas,¹ Yann Emmenegger,³ Cédric Gobet,^{1,2} Paul Franken,³ Frédéric Gachon,^{1,2} and Félix Naef¹

¹School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland; ²Nestle Institute of Health Sciences, CH-1015 Lausanne, Switzerland; ³Center for Integrative Genomics, University of Lausanne, CH-1015 Lausanne, Switzerland

The circadian clock in animals orchestrates widespread oscillatory gene expression programs, which underlie 24-h rhythms in behavior and physiology. Several studies have shown the possible roles of transcription factors and chromatin marks in controlling cyclic gene expression. However, how daily active enhancers modulate rhythmic gene transcription in mammalian tissues is not known. Using circular chromosome conformation capture (4C) combined with sequencing (4C-seq), we discovered oscillatory promoter–enhancer interactions along the 24-h cycle in the mouse liver and kidney. Rhythms in chromatin interactions were abolished in arrhythmic *Bmal1* knockout mice. Deleting a contacted intronic enhancer element in the *Cryptochrome 1* (*Cry1*) gene was sufficient to compromise the rhythmic chromatin contacts in tissues. Moreover, the deletion reduced the daily dynamics of *Cry1* transcriptional burst frequency and, remarkably, shortened the circadian period of locomotor activity rhythms. Our results establish oscillating and clock-controlled promoter–enhancer looping as a regulatory layer underlying circadian transcription and behavior.

[**Keywords:** circadian rhythms; chromatin topology; promoter–enhancer loops; DNA regulatory elements; transcriptional bursting]

Supplemental material is available for this article.

Received January 26, 2018; revised version accepted March 2, 2018.

The circadian clock, encoded in a core genetic network, governs rhythms in behavior and physiology (Schibler et al. 2015), such as nocturnal activity in mice and oscillations in carbohydrate and lipid metabolism in the liver (Bass and Lazar 2016). This clock also orchestrates the daily rhythmic synthesis of thousands of transcripts by impinging on multiple gene regulatory layers (Zhang et al. 2014). These rhythmic transcripts often coincide with rhythms in chromatin modifications, DNA accessibility, enhancer activity, and transcription factor (TF) binding at promoter-proximal and promoter-distal regions (Mermet et al. 2017; Takahashi 2017), suggesting that chromatin interactions play a role in regulating circadian gene expression.

Chromatin architecture in the nucleus is organized over multiple scales (Dekker et al. 2013). At the fine scale, this organization involves the interactions between gene pro-

motors and enhancer DNA elements through promoter–enhancer looping (Fulco et al. 2016). The remodeling of such DNA contacts and the accompanying dynamics of transcriptional responses have been investigated in the context of signal-dependent gene induction, cell differentiation, and developmental transitions (Palstra et al. 2003; Ghavi-Helm et al. 2014; Kuznetsova et al. 2015). However, little is known about the dynamics of DNA looping along the recurring daily 24-h cycle and the consequences on clock-dependent gene expression in animals.

Cell culture models investigating genes of interest have suggested that nuclear compartmentalization modulates cyclic gene expression (Zhao et al. 2015) and that oscillatory contacts between gene promoters and genomic regions on *trans* chromosomes accompany rhythmic mRNA expression (Aguilar-Arnal et al. 2013). Recently, we described tissue-specific chromatin interactions selectively associated with rhythmically expressed clock output transcripts (Yeung et al. 2018), but, in general, the circadian dynamics of DNA interactions, including their

⁴These authors contributed equally to this work.

Corresponding author: felix.naef@epfl.ch

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.312397.118>. Freely available online through the *Genes & Development* Open Access option.

© 2018 Mermet et al. This article, published in *Genes & Development*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

Mermet et al.

regulation of core clock function and control of circadian gene expression, remain an open question. Indeed, rhythmic transcription could be regulated over an established static promoter–enhancer network (Chavi-Helm et al. 2014; Xu et al. 2016), or, conversely, the clock could drive dynamic promoter–enhancer looping for high-amplitude daily oscillations in transcription.

Here we monitored promoter–enhancer contacts of a core clock and metabolic clock output gene across time and genotypes in mouse tissues and discovered that contact frequencies oscillated along the 24-h cycle. In arrhythmic *Bmal1* knockout animals, these oscillations were abolished. Deletion in mice of an enhancer that was rhythmically recruited to the *Cryptochrome 1* (*Cry1*) promoter led to a short period phenotype in locomotor activity. Moreover, this deletion compromised rhythmic chromatin topology in the liver and led to reduced peak *Cry1* mRNA expression levels. Finally, single-molecule RNA fluorescent in situ hybridization (smRNA-FISH) showed that the abolished rhythmic chromatin contact reduced the daily dynamics of *Cry1* transcriptional burst frequency.

Results

Rhythmic local chromatin interactions in mouse livers

We focused on two genes representing key temporally regulated hepatic functions: a gene essential for the core circadian oscillator, *Cry1* (Griffin et al. 1999; van der Horst et al. 1999), and a liver-specific clock-controlled gene, *Glycogen Synthase 2* (*Gys2*) (Doi et al. 2010), which

encodes the rate-limiting enzyme in hepatic glycogen synthesis (Irimia et al. 2010). These transcripts are rhythmically expressed in the liver at opposite times of day, *Cry1* peaking during the night at Zeitgeber time 20 (ZT20) and *Gys2* peaking during the day at ZT08 (with ZT0 corresponding to lights on and ZT12 corresponding to lights off) (Supplemental Fig. S1A). Using circular chromosome conformation capture (4C) combined with sequencing (4C-seq) (Gheldof et al. 2012), we estimated the interaction frequencies of DNA bait fragments placed near the transcription start sites (TSSs) of *Cry1* and *Gys2* versus the entire genome in livers of wild-type mice collected at ZT08 and ZT20 ($n = 4$ per time point). 4C-seq signals around the *Cry1* and *Gys2* TSSs decayed to background levels following a power law (Supplemental Fig. S1B,C; Supplemental Table S1; Sanborn et al. 2015) and did not exceed background on *trans* chromosomes (Supplemental Fig. S1D,E; Supplemental Table S1). The high proportion of chromatin interactions within the first 2 Mb surrounding the baits on the *cis* chromosome (*Cry1* TSS: 41% of total *cis* contacts at ZT08 and 46% at ZT20; *Gys2* TSS: 54% at ZT08 and 57% at ZT20) indicated that *Cry1* and *Gys2* regulatory contacts were contained within this signal-rich region (Sanyal et al. 2012). To compare 4C-seq profiles across conditions, we normalized the data and applied locally weighted multilinear regression (LWLR), which uses a Gaussian window ($\sigma = 2500$ kb) centered on each fragment for local smoothing (Materials and Methods). For *Cry1*, the 4C-seq profiles after LWLR were similar between ZT08 and ZT20 except in a region downstream from the *Cry1* promoter, where the contact frequency was increased at ZT20 (Fig. 1A). While the

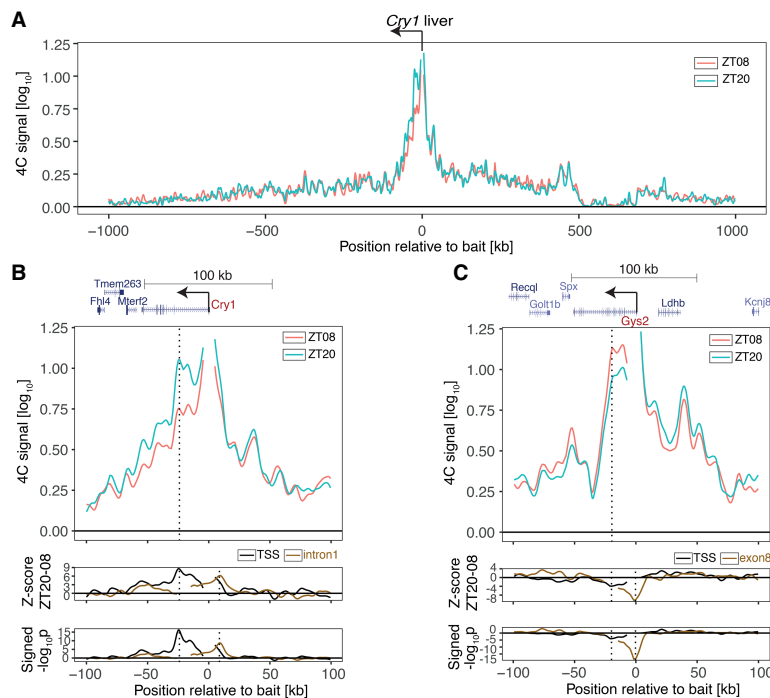


Figure 1. Rhythmic chromatin interactions in mouse livers. (A) 4C-seq data (LWLR summarizes $n = 4$ animals per group) in a 2-Mb genomic region surrounding *Cry1* at ZT08 and ZT20. (B) 4C-seq signals in a 200-kb genomic region surrounding *Cry1* at ZT08 and ZT20. (Bottom tracks) 4C-score and signed $-\log_{10}(p)$ show rhythmic contacts between the promoter region and the intronic region. (Black) *Cry1* TSS bait ($P < 10^{-16}$ at peak); (brown) *Cry1* intron1 bait ($P < 10^{-8}$ at peak). (C) Same as B, targeting the *Gys2* promoter. (Bottom tracks) Same as B for *Gys2* TSS bait ($P < 10^{-4}$ at peak). (Brown) *Gys2* exon8 bait ($P < 10^{-18}$ at peak). Vertical dotted lines show the positions locally of maximal differential chromatin interactions.

differential signal covered the entire *Cry1* locus, the largest difference was localized—peaking 26 kb downstream from the TSS in the first *Cry1* intron—and highly significant ($P < 5.5 \times 10^{-17}$ at the peak) (Fig. 1B, bottom tracks, vertical dotted line at the left). A secondary peak was observed near the 3' end of the *Cry1* transcript.

To further validate the time-dependent contacts, we placed a bait at the +26-kb intronic site (reciprocal 4C-seq). The reciprocal 4C-seq confirmed the increased contact frequency with the *Cry1* promoter region at ZT20 compared with ZT08 (Fig. 1B bottom tracks, brown solid line; Supplemental Fig. S2A). In fact, the reciprocal differential signal peaked 7 kb upstream of the *Cry1* TSS, a site that was also differentially contacted by the *Cry1* TSS bait ($P < 1.9 \times 10^{-9}$) (Fig. 1B, bottom tracks, vertical dotted line at the right; Supplemental Fig. S2A; Supplemental Table S1). Thus, these 4C-seq data in the liver suggested dynamic contacts between the *Cry1* promoter and the +26-kb intronic site as well as the -7-kb upstream site. Since *Cry1* mRNA accumulated rhythmically in the kidney (Supplemental Fig. S3A), we also performed 4C-seq in kidneys. Consistent with the liver data, these sites were also recruited to the *Cry1* promoter more frequently at ZT20 than at ZT08 (Supplemental Fig. S3B,C).

Opposite to *Cry1*, the *Gys2* promoter contacted an intragenic region more frequently at ZT08 versus ZT20 (Fig. 1C), with a peak 21 kb downstream from the TSS in exon 8 ($P < 8.7 \times 10^{-5}$ at peak) (Fig. 1C, bottom tracks, black solid line, vertical dotted line at the left), consistent with its anti-phasic rhythmic mRNA accumulation (Supplemental Fig. S1A). This significant differential signal was validated by reciprocal 4C-seq using the exon 8 as bait ($P < 2.3 \times 10^{-19}$ at peak) (Fig. 1C, bottom tracks, brown solid line, vertical dotted line at the right; Supplemental Fig. S2B). In the kidney, where *Gys2* mRNA accumulation was constant and low, this differential signal was absent (Supplemental Fig. S3D–F). Thus, both gene promoters formed DNA loops with neighboring intragenic regions in *cis* that coincided with the timing of the respective peaks in *Cry1* and *Gys2* mRNA expression.

The dynamics of chromatin topology depend on BMAL1

To test whether these dynamic contacts depended on a functional circadian clock, we performed 4C-seq in the livers of clock-deficient animals (*Bmal1* knockout) in which *Cry1* and *Gys2* lost rhythmic expression and were constantly expressed at high and low levels, respectively (Supplemental Fig. S4A,B). In *Bmal1* knockout, the *Cry1* +26-kb intronic and -7-kb upstream regions contacted the promoter at comparable frequencies at ZT20 and ZT08, suggesting static chromatin loops (Fig. 2A,B). For *Gys2*, the profile between the exon 8 region and the promoter was also static (Fig. 2C,D). Comparing wild-type and *Bmal1* knockout at both time points revealed that for *Cry1*, the loop was locked in a closed conformation (Supplemental Fig. S4C, constitutively high frequencies), and for *Gys2*, it was locked in an open conformation (Supplemental Fig. S4D, constitutively low frequencies). Thus, the closed and open states of DNA loops concurred

with high and low transcription, respectively (Supplemental Fig. S4, cf, A,C and B,D). We note that these 4C profiles suggested a BMAL1-independent interaction upstream of *Gys2* (Fig. 2C, lower panels), but this effect was less robust compared with the BMAL1-dependent intragenic looping. As a negative control, we targeted the *Hoxd4* locus, which is a transcriptionally silent region in the adult liver. As expected, chromatin contact profiles at the *Hoxd4* locus remained static over time in both wild-type and *Bmal1* knockout livers (Fig. 2E; Supplemental Fig. S4E). These data thus showed that rhythmic loops in *Cry1* and *Gys2* depended on the clock TF BMAL1.

Rhythmic DNA loops connect gene promoters with daily active enhancers

To characterize the interacting genomic regions, we integrated temporal data on DNase-I hypersensitivity sites (DHSs) with ChIP-seq (chromatin immunoprecipitation [ChIP] combined with high-throughput sequencing) data for RNA polymerase II (Pol II), the activity-related chromatin mark H3K27ac (Sobel et al. 2017), and rhythmically active TFs (Rey et al. 2011; Zhang et al. 2015). This allowed us to assess whether the rhythms in DNA contacts coincided with rhythms in activity-related chromatin marks. For *Cry1*, RNA Pol II and H3K27ac signals peaked near ZT20 (Fig. 3), while, for *Gys2*, they peaked near ZT08 (Fig. 4). However, while RNA Pol II signals extended throughout the gene bodies, H3K27ac signals were spatially confined around the largest differential contact precisely at sites marked with DHSs. Furthermore, both the 26-kb downstream intronic site and the 7-kb upstream site of the *Cry1* TSS contained a RORE-responsive element (RRE) and were bound by the circadian TFs REV-ERBa and ROR γ (Fig. 3; Supplemental Table S4; Zhang et al. 2015). In mouse fibroblasts, the intronic RRE is required for proper timing of *Cry1* expression (Ukai-Tadenuma et al. 2011). The interacting *Gys2* exon 8 site was bound by the clock regulator BMAL1 at ZT06 (Rey et al. 2011) and by REV-ERBa at ZT10 (Fig. 4). This indicated that DNA contacts connected local rhythmically active enhancer elements with the promoters of *Cry1* and *Gys2*.

*Deleting the *Cry1* intronic enhancer in mice shortens the circadian locomotor period*

To study the function of the rhythmic chromatin interactions, we generated a mouse strain (*Cry1 Δ e*) with a 300-base-pair (bp) deletion covering the *Cry1* intronic enhancer (Supplemental Fig. S5A,B). We measured spontaneous locomotor activity in constant darkness and observed that *Cry1 Δ e* animals had an endogenous circadian period that was significantly shorter ($P < 1.1 \times 10^{-5}$, *t*-test) by 15 min compared with wild-type littermates (Fig. 5A; Supplemental Fig. S5C,D). Such period shortening is in the range of classic short period core clock mutants such as *Per1* (Cermakian et al. 2001) and *Clock* (Debruyne et al. 2006). As *Cry1* loss of function shortens the circadian period by 1.2-h (van der Horst et al. 1999), our noncoding DNA deletion suggests a *Cry1* hypomorph.

Mermet et al.

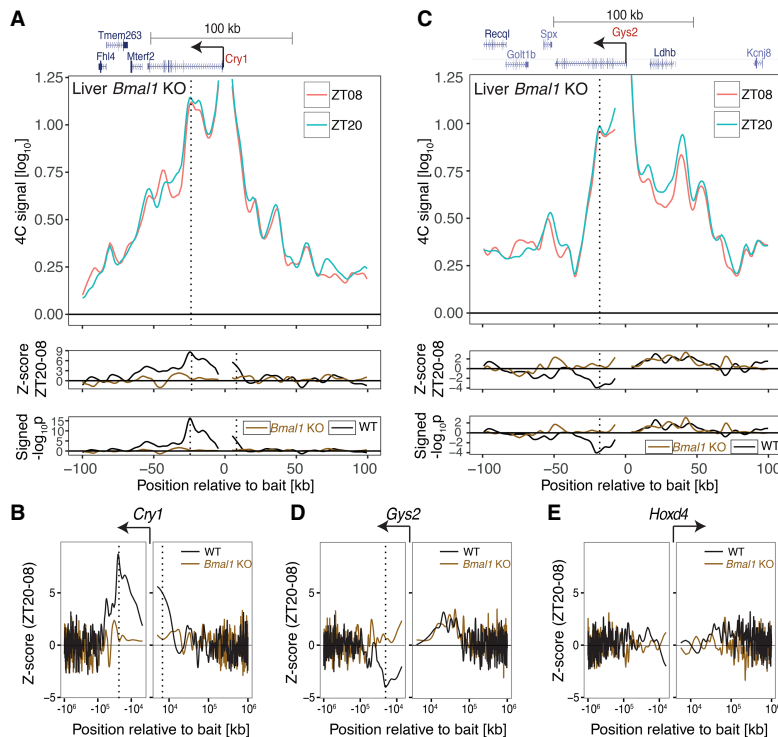


Figure 2. The dynamics of chromatin topology depend on BMAL1. (A, top) 4C-seq signal targeting *Cry1* from the livers of *Bmal1* knockout mice at ZT20 versus ZT08 shows loss of rhythms in chromatin interactions. (Bottom) Z-score and signed $-\log_{10}(p)$ of differential 4C-seq signal (ZT20–ZT08) in wild-type versus *Bmal1* knockout. Vertical lines show BMAL1-dependent rhythmic contacts. (B) Z-score in a 2-Mb genomic region surrounding *Cry1* in wild-type versus *Bmal1* knockout. (C) Same as in A but for *Gys2* bait. (D,E) Same as in B but for *Gys2* (D) and *Hoxd4* (E) baits. B and D show that the BMAL1-dependent rhythmic contacts are localized within 100 kb of the bait.

Expression of *Cry1*, clock, and clock output genes is perturbed in *Cry1Δe*

To investigate the link between the deletion, promoter-enhancer looping, and *Cry1* expression in livers and kidneys, we first generated temporal RNA sequencing (RNA-seq) data in *Cry1Δe* and wild-type littermates under an entraining light–dark cycle. The transcriptomes in *Cry1Δe* and wild-type littermates were comparable overall in both tissues (Supplemental Fig. S6A). While *Cry1* mRNA levels remained rhythmic in both genotypes, likely driven by further regulatory sites (e.g., the TSS and –7-kb sites), the peak expression at ZT20 was significantly reduced by 27% in the livers (15% in the kidneys) of *Cry1Δe* animals compared with wild type (Supplemental Fig. S6B). Quantifying the intronic reads as a proxy for transcription showed that *Cry1* transcription was also phase-advanced in *Cry1Δe* animals (Supplemental Fig. S6C). Moreover, CRY1 protein abundance in the liver was lower in *Cry1Δe* compared with wild type, consistent with a reduction in mRNA levels (Supplemental Fig. S6D,E).

As is known in chronobiology, entraining a short period circadian oscillator by an external light–dark cycle leads to a phase advance of internal timing markers (Aschoff and Pohl 1978). This prediction was confirmed in the transcriptome data. Indeed, core clock and clock-controlled genes (Supplemental Table S5) were phase-advanced by, on average, 30 min in the livers of *Cry1Δe* animals compared with wild type ($P < 0.01$ binomial test) (Supplemen-

tal Fig. S6F), with *Cry1* showing the largest phase advance ($P = 0.011$ for livers; $P = 0.047$ for kidneys, bootstrap test) (Supplemental Fig. S6B).

The *Cry1Δe* mutation disrupts rhythmic chromatin topology

Next, we explored the dynamics of chromatin topology along the 24-h cycle in liver sampled every 4 h in wild type and *Cry1Δe* ($n = 3$ per time point). First, we confirmed oscillatory chromatin interactions in *Gys2* in wild type. Indeed, the *Gys2* promoter rhythmically recruited the +21-kb enhancer, peaking near ZT08 in both the TSS bait and exon 8 bait ($P < 10^{-6}$ at the peak harmonic regression) (Supplemental Fig. S7A–C). As negative control, the *Hoxd4* bait measured around the clock did not show oscillatory contacts (Supplemental Fig. S7D). For *Cry1* wild type, the frequency of contacts between the promoter and the +26-kb enhancer significantly oscillated, peaking near ZT20 ($P < 10^{-8}$ at the peak) (Fig. 5B–D; Supplemental Fig. S8A,B). In contrast, in *Cry1Δe*, the contact frequencies in this region were lower at all time points compared with wild type, and the oscillation was compromised (Fig. 5B–D; Supplemental Fig. S8A,B). Finally, we also estimated chromatin contacts for a bait placed at the –7-kb upstream enhancer (Fig. 1B, bottom tracks, vertical dotted line at the right), showing oscillation in contact frequency peaking around ZT20 with the +26-kb intronic enhancer in wild type but

Rhythmic DNA loops tune transcription and behavior

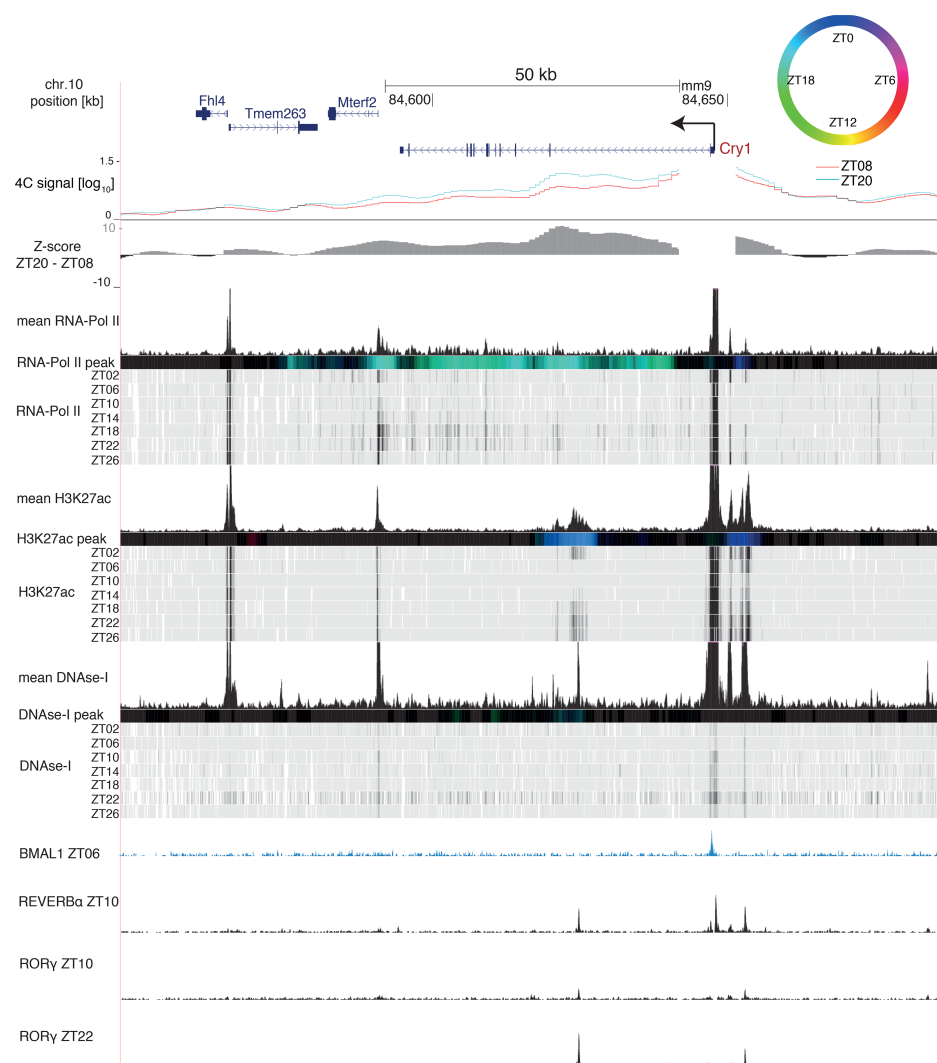


Figure 3. The rhythmic *Cry1* loop connects the promoter with a H3K27ac-marked enhancer. The *Cry1* genomic region containing 4C-seq signals from *Cry1* TSS at ZT08 (red) and ZT20 (blue) and Z-score (ZT20–ZT08) in wild-type livers. RNA Pol II loadings (ChIP-seq), H3K27ac mark (ChIP-seq), and DNase-I signal are from Sobel et al. (2017). Temporally averaged signals and temporal signals of each mark are plotted. Colored bars represent peak times according to the color legend at the top right; black signifies no rhythm (Materials and Methods). BMAL1 ChIP-seq signal is from Rey et al. (2011), and REV-ERBa and RORγ ChIP-seq signals are from Zhang et al. (2015).

nonrhythmic and overall lower contact frequency in *Cry1Δe* (Supplemental Fig. S8C,D). Decreased contact frequency in *Cry1Δe* mice indicates that the RRE-containing 300-bp fragment drives the promoter–enhancer loop.

Overall, these data demonstrate robust rhythmic chromatin topology for *Cry1* and *Gys2*, where the frequency of enhancer–promoter contacts is modulated with time of day. Furthermore, deleting a localized noncoding DNA enhancer element (300 bp) in the *Cry1* gene could disrupt such rhythms.

The *Cry1* intronic enhancer modulates transcriptional burst frequency

To analyze whether the *Cry1* intronic enhancer modulates transcription, we estimated transcriptional parameters by smRNA-FISH against *Cry1* pre-mRNA in the livers of wild-type and *Cry1Δe* animals at ZT08 and ZT20 (Fig. 6A). Mammalian promoters are irregularly transcribed (transcriptional bursting), as characterized by the burst size and burst frequency (Suter et al. 2011; Bahar Halpern et al. 2015). Taking into account the ploidy of

Mermet et al.

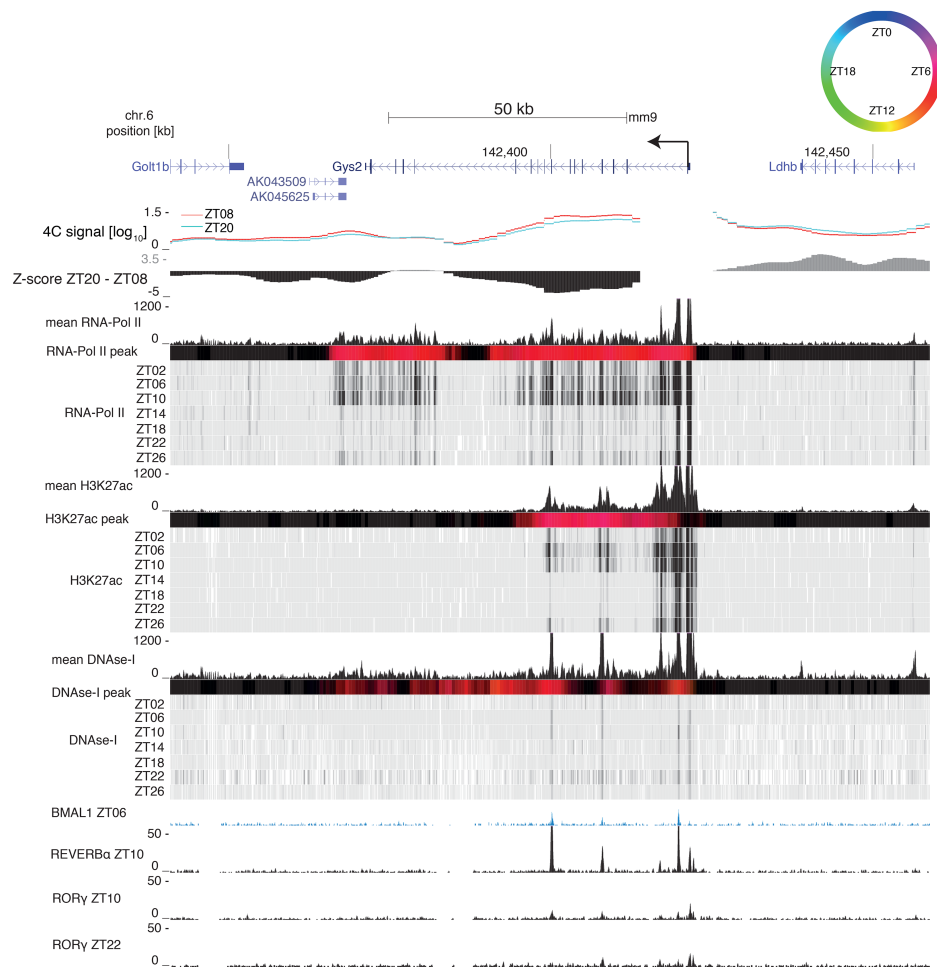


Figure 4. The rhythmic *Gys2* loop connects the promoter with a H3K27ac-marked enhancer. The *Gys2* genomic region containing 4C-seq signals from the *Gys2* TSS at ZT08 (red) and ZT20 (blue) and Z-score (ZT20–ZT08) in wild-type livers. RNA Pol II loadings (ChIP-seq), H3K27ac mark (ChIP-seq), and DNase-I signal are from Sobel et al. (2017). Temporally averaged signals and temporal signals of each mark are plotted. Colored bars represent peak times according to the color legend at the top right; black signifies no rhythm (Materials and Methods). The BMAL1 ChIP-seq signal is from Rey et al. (2011), and the REV-ERB α and ROR γ ChIP-seq signals are from Zhang et al. (2015).

liver nuclei (Supplemental Fig. S9A–D), smRNA-FISH showed that *Cry1* burst fraction (fraction of active transcription sites in each nucleus, which is proportional to the burst frequency per allele) was 2.2-fold higher at ZT20 compared with ZT08 in wild type (Fig. 6B). Importantly, the burst fraction was reduced by 28% in *Cry1 Δ e* animals at ZT20 (Fig. 6B). In contrast, the burst intensity (proportional to the burst size) was similar in all conditions (Fig. 6C). Thus, the lowered *Cry1* mRNA levels in *Cry1 Δ e* at ZT20 can be quantitatively explained by the reduced burst fraction. In sum, dynamic enhancer loops modulate transcriptional bursting in mammalian tissues (Bartman et al. 2016; Fukaya et al. 2016); in particular, rhythmic DNA loops involving clock enhancers control burst frequency while maintaining burst size.

Discussion

In animals, developmental transitions occurring on the time scales of days have been shown to involve remodeled DNA contacts and promoter–enhancer loop formation (Noordermeer et al. 2014). While such dynamics are typically irreversible, we here discovered that chromatin topology in mouse tissues can be locally (100 kb, in *cis*) plastic, exhibiting temporal dynamics that are regulated by daily time and the circadian oscillator and thus recur within a 24-h period. While previous work in cell culture reported dynamic chromatin contacts on larger genomic scales, notably between the *Dbp* gene and DNA regions on *trans* chromosomes (Aguilar-Arnal et al. 2013), the genes analyzed here did not show rhythmic chromatin

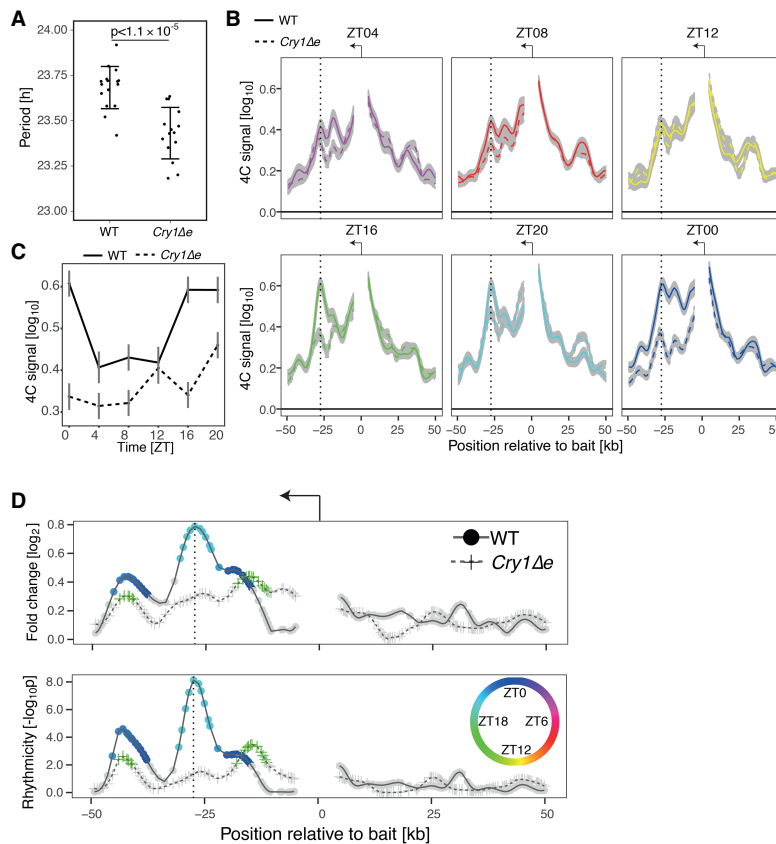


Figure 5. Deleting the *Cry1* intronic enhancer in mice shortens the period of the clock and disrupts oscillations in *Cry1* promoter-enhancer contact frequencies. (A) The circadian period of spontaneous locomotor activity is significantly different between *Cry1Δe* and wild-type littermates. The mean period and standard deviation were calculated from 16 wild-type and 15 *Cry1Δe* littermates. $P = 1.1 \times 10^{-5}$, *t*-test. (B) 4C-seq signal for *Cry1* TSS bait over time in livers (LWMM summarizes $n = 3$ animals per group; gray shade shows \pm standard error) in wild-type versus *Cry1Δe* littermates. Vertical lines show the +26-kb intronic enhancer. (C) 4C-seq signal over time adjacent to the intronic enhancer. (D) \log_2 fold change and $-\log_{10}(p)$ from rhythmicity analysis of 4C-seq signal over time. $P < 10^{-8}$ at peak, LWMM, χ^2 test. Fragments with $P < 0.01$ are colored by time of peak contact frequency (color legend is shown at the right).

interactions on such scales. We then showed genetically that these rhythmic DNA contacts depend on the clock protein BMAL1 and, in the case of *Cry1*, a 300-bp intronic RRE-containing enhancer sequence.

How is BMAL1 involved in the formation of these dynamic loops? In the case of *Gys2* in *Bmal1* knockout mice, the loop is constitutively open, and *Gys2* mRNA expression is constitutively low. Combined with the binding of BMAL1 at the looping site, these data strongly argue for a direct involvement of BMAL1. For *Cry1*, the activator ROR γ and the repressor REV-ERB α bind to the *Cry1* intronic enhancer at the expected peak (ZT20) and trough (ZT08) activities, as is typical of functional RREs. We note that while the expression of the RRE-binding repressors *Rev-Erba*/ β is low in *Bmal1* knockout, the corresponding activator *Ror* γ is constitutively high (Atger et al. 2015). Therefore, the constitutively closed *Cry1* loop in the *Bmal1* knockout most likely reflects an indirect effect via perturbed REV-ERB and ROR activities. This is further corroborated by the constitutively open state of the *Cry1* promoter-enhancer loop in *Cry1Δe* mice, showing chromatin interactions that are constantly below wild-type trough levels, indicating that loop-promoting factors (for example, RORs) act within the 300-bp element. Therefore, our data suggest a canonical mechanism of enhancer-promoter looping by which sequence-specific TFs

help recruit transcription complexes, which facilitate the function of Pol II at core promoters (Levine and Tjian 2003).

To investigate the effects of the dynamic looping on transcriptional parameters, we complemented bulk 4C-seq and RNA-seq experiments with single-molecule transcript analysis in situ, which revealed that the abolished rhythmic chromatin contact in *Cry1Δe* mice reduced *Cry1* transcriptional burst frequency. These results in mammalian tissues contribute to our current understanding of how enhancer loops modulate transcriptional bursting (Bartman et al. 2016; Fukaya et al. 2016). In particular, we showed that rhythmically active clock enhancers can increase burst frequency while not changing burst size.

The ablation of the *Cry1* noncoding regulatory element even led to a short period phenotype in locomotor activity. While noncoding genetic variation in humans has been associated recently with circadian clock-related and sleep phenotypes (Allebrandt et al. 2010; Hu et al. 2016), no demonstration of such variation on circadian transcription or behavior has yet been provided. Indeed, previously characterized mutations impacting mammalian circadian behavior have concerned protein-coding regions (Vitaletta et al. 1994; Toh et al. 2001). Here, we provided evidence that noncoding regulatory elements within the

Mermet et al.

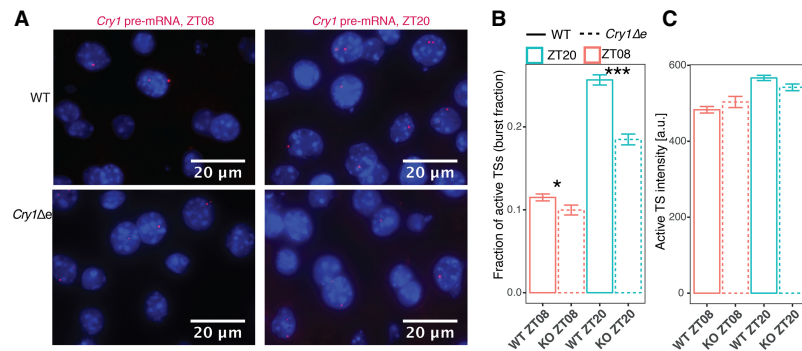


Figure 6. The oscillatory *Cry1* promoter–enhancer loop modulates *Cry1* transcriptional bursting. (A) smRNA-FISH against *Cry1* pre-mRNA in the livers of wild-type (top) and *Cry1* Δe (bottom) animals at ZT08 (left) and ZT20 (right). Burst fractions (B) and burst intensities (C) measured from images of smRNA-FISH performed against *Cry1* pre-mRNA in *Cry1* Δe (dashed) and wild-type (solid) livers at ZT08 (red) and ZT20 (blue). Burst fraction is the number of active transcription sites in each nucleus divided by the ploidy. (B,C) Shown are the means and standard errors over nuclei collected and pooled from two animals in each of the four conditions (individual animals are analyzed in Supplemental Fig. S9C,D). $n = 2191$ wild-type ZT08 nuclei; $n = 983$ *Cry1* Δe ZT08 nuclei; $n = 2150$ wild-type ZT20 nuclei; $n = 1473$ *Cry1* Δe ZT20 nuclei. In B, (*) $P < 0.05$; (***) $P < 0.001$, t -test. In C, differences between genotypes are not significant.

core circadian regulatory network can drive dynamic promoter–enhancer looping, modulate temporal transcription, and regulate circadian locomotor behavior.

Materials and methods

Animal and ethics statement

All animal care and handling were performed according to Canton de Vaud laws for animal protection (authorization VD2801 [Frédéric Gachon] and VD3109 [Félix Naef]). All experiments were performed on males between 8 and 10 wk old. *Bmal1* knock-out animals were described previously in Jouffe et al. (2013).

Mouse genome editing by direct knockout using CRISPR–Cas9

Px-330 plasmids targeting upstream of and downstream from the *Cry1* intron1 regulatory region were injected into pronuclei and then transplanted into B6D2F1 pseudopregnant mice at the Ecole Polytechnique Fédérale de Lausanne (EPFL) Transgenic Core Facility (<http://tcf.epfl.ch>). Pups from the first generation (F0) were then screened for the deletion using the PCR primers indicated in Supplemental Table S2. F0 animals of interest were backcrossed on C57/BL6J wild-type mice, and F1 animals were screened for transmission of the mutation. Heterozygous animals were crossed together to obtain all genotypes of interest. The Ethical Committee of the State of Vaud Veterinary Office, Switzerland, approved all experiments.

Nucleus purification and fixation

Immediately after sacrifice, 5 mL of 1 \times PBS was perfused through the spleen to flush blood from the liver. Livers and kidneys from individual animals were homogenized and fixed in 4 mL of 1 \times PBS, including 1.5% formaldehyde, for 10 min at room temperature. The cross-linking reaction was stopped by adding 25 mL of ice-cold stop reaction buffer (2.2 M sucrose, 150 mM glycine, 10 mM HEPES at pH 7.6, 15 mM KCl, 2 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 0.5 mM DTT, 0.5 mM PMSF) to the homogenates and was kept for 5 min on ice. Homogenates were then

loaded on top of 10 mL of cushion buffer (2.05 M sucrose, 10% glycerol, 125 mM glycine, 10 mM HEPES at pH 7.6, 15 mM KCl, 2 mM EDTA, 0.15 mM spermine, 0.5 mM spermidine, 0.5 mM DTT, 0.5 mM PMSF) and centrifuged at 10^5g for 45 min at 4°C. Nuclei were washed twice in 1 \times PBS and immediately frozen.

4C-seq

4C template preparation 4C templates were prepared as in Gheldof et al. (2012). Nuclei were resuspended in 1 mL of a buffer containing 10 mM Tris-HCl (pH 8.0), 10 mM NaCl, 0.2% NP-40, and 1 \times protease inhibitor cocktail (Complete Mini EDTA-free protease inhibitor cocktail; Sigma-Aldrich); kept for 15 min on ice; and washed twice with 1 \times DpnII buffer (New England Biolabs). Thirty million nuclei were resuspended in 1 \times DpnII buffer (New England Biolabs) containing 0.1% SDS and incubated for 10 min at 65°C. Triton X-100 was added to 1% final concentration. Chromatin was digested overnight with 400 U of DpnII (New England Biolabs) at 37°C with shaking. After heat inactivation, digestion efficiency was evaluated by both DNA visualization on agarose gels and quantitative PCR using primer pairs covering multiple restriction sites. Chromatin was then ligated with 3000 U of T4 DNA ligase (New England Biolabs) in an 8-mL final volume for 4 h at 16°C plus 1 h at room temperature. The cross-linking reaction was reverted by the addition of 50 μ L of 10 mg/mL proteinase K and incubation overnight at 65°C. DNA was purified by multiple phenol/chloroform extractions, resuspended in TE buffer (pH 8.0) containing RNase A, and incubated for 30 min at 37°C. Ligation efficiency was evaluated by loading DNA on an agarose gel. Libraries were digested with 1 U of NlaIII per microgram of template (New England Biolabs) overnight at 37°C, and digestion was controlled by visualization on an agarose gel. After heat inactivation, digested products were ligated with 2000 U of T4 DNA ligase (New England Biolabs) for 4 h at 16°C in a 14-mL final volume. Circularized products were purified and resuspended in TE buffer (pH 8.0). 4C templates were prepared in four biological replicates in wild-type mouse livers and kidneys and three biological replicates in the livers of *Bmal1* knockout and *Cry1* Δe and wild-type littermates (Supplemental Table S1).

Inverse PCR and sequencing in wild-type and *Bmal1* knockout mouse livers and kidneys Six-hundred nanograms of 4C template was used for PCR amplification using Sigma-Aldrich long-template PCR system with bait-specific inverse primers conjugated to Illumina sequencing adaptors (primer sequences are in Supplemental Table S3) in a final volume of 50 μ L in the following PCR program: 2 min at 94°C followed by 30 cycles of 15 sec at 94°C, 1 min at 55°C, and 3 min at 68°C and a final extension of 7 min at 68°C. PCR were performed in parallel reactions with 6 \times 100 ng of template for each sample. PCR products were purified with the AMPure XP beads system (Beckman Coulter), and amplification profiles were analyzed by fragment analyzer and then sequenced on Illumina HiSeq 2000 machines using single-end 100-bp read length.

Inverse PCR and sequencing in the livers of *Cry1 Δ e* and wild-type littermates Six-hundred nanograms of 4C template was used for PCR amplification using Sigma-Aldrich long-template PCR system with two-step PCR system from Illumina. Bait-specific inverse primers conjugated to Illumina sequencing adaptors (primer sequences are in the Supplemental Table S3) were used in a first PCR reaction in a final volume of 50 μ L with the following program: 2 min at 94°C followed by 20 cycles of 15 sec at 94°C, 1 min at 55°C, and 3 min at 68°C and a final extension of 7 min at 68°C. PCRs were performed in parallel reactions with 6 \times 100 ng of template for each sample. PCR products were purified with the AMPure XP beads system (Beckman Coulter). Purified products were pooled and used as the template of a second PCR reaction with Nextera XT index kit version2 primers (FC-131-2004) in a final volume of 50 μ L with the following program: 2 min at 94°C followed by 10 cycles of 15 sec at 94°C, 1 min at 55°C, and 3 min at 68°C and a final extension of 7 min at 68°C. PCR products were purified with the AMPure XP beads system (Beckman Coulter) and then sequenced on NextSeq 500 machines using single-end 75-bp read length.

4C-seq analysis

Preprocessing computational methods Demultiplexed Fastq files were mapped to the mouse genome (mm9) using Bowtie2 with default HTSstation parameters (<http://htsstation.epfl.ch>). Since each restriction fragment contained two mapping sites (two ends of the fragment), the fragment score was computed as the average of the number of reads per mapping site.

Quality control of 4C-seq data Samples with $\geq 75\%$ of restriction fragments without any counts in a window of ± 1 Mb upstream of and downstream from each bait were not analyzed (Supplemental Table S1). The first five fragments upstream of and downstream from the bait (10 total) were not considered in the analysis because they mostly contained partially digested and self-ligated products.

Normalization and LWMR We follow a method developed recently in Yeung et al. (2018) with minor modifications. Briefly, raw read counts for each sample were library size-rescaled by the normalized sum of the read counts on the *cis* chromosome (excluding 10 restriction fragments around the bait). To control the variability of low signals, in subsequent analyses, the fragment counts c in each sample were log transformed using the variable

$$Y = \log_{10}\left(\frac{c}{p} + 1\right),$$

with $P = 500$. A weighted linear model was then fit locally using a Gaussian window ($\sigma_G = 2500$ bp) centered on the fragment of

interest. For each position, nearby 4C-seq signals (Y) were modeled with fragment effects a_i and condition effects b_j (which can be time, tissue, or genotype). In LWMR, these parameters were estimated by minimizing the weighted sum S of squared residuals across replicates r : $S = \arg\min_{a,b} \sum_{i,j,r} W_{i,j} (Y_{i,j,r} - a_i - b_j)^2$, with weights $W_{i,j}$ defined as $W_{i,j} = w_{g,i} \times w_{s,j}$, where $w_{g,i}$ is the Gaussian smoothing kernel at position i , and $w_{s,j}$ is a condition weight based on the number of samples with nonzero counts on fragment i . Specifically, we used $w_s = 0.5, 1.5, 2.5, 3.5$, or 4.5 for fragments with zero, one, two, three, or four replicates showing nonzero counts, which down-weights positions with high dropout rates. To estimate the statistical significance for differential contacts (for example, ZT20 vs. ZT08), we propagated the estimated uncertainty (standard errors for locally weighted regression) in the corresponding b values to calculate Z-scores and used regularized t statistics with $n-p$ degrees of freedom (DOF; n is the number of data points within window, and p is the number of parameters). For the analysis of 24-h rhythmicity in contacts (weighted harmonic regression), we proceeded analogously by propagating the uncertainty in the b s for the six time points to that in the squared 24-h Fourier coefficient and used the χ^2 test with two DOF (owing to the real and imaginary parts). For each set of samples, we computed the regularized residual variance as

$$\hat{\sigma}^2 = \hat{\sigma}^2 + \sigma_{\min}^2 \exp\left(-\frac{\bar{b}}{b_s}\right),$$

with $\hat{\sigma}^2$ as the estimator of the squared residuals, \bar{b} as the estimated signal across samples, and $b_s = \log_{10}(2)$. σ_{\min}^2 prevents artificially small variance from positions of high dropout rates and is estimated from the distribution of $\hat{\sigma}^2$ across all fragments. σ_{\min} ranges from 0.06 to 0.16 (same units as Y), depending on the bait (Supplemental Table S1).

H3K27ac and RNA Pol II ChIP-seq and DNase-I-seq analysis

Bam files from GSE60578 (Sobel et al. 2017) were analyzed in genomic regions ± 1 Mb from the 4C-seq baits. There, read counts were binned in 500-bp intervals and normalized by the library size. The amplitude and phase of the \log_2 read counts of each of the three signals were calculated for each bin after applying a running average of seven bins (three bins upstream, three bins downstream, and one bin in the center) to smooth the signal. Obtained rhythmic amplitudes and phases were compared with differential 4C-seq signals. The rhythmic signal in each bin [phase, amplitude, and $-\log_{10}(p)$] was mapped to a color using the hue, saturation, and value (HSV) color scheme. Hue h was defined by the phase of the oscillation, with blue as ZT0. The saturation s was set to 1. The value v was set to a color if both amplitude X_a and $-\log_{10}(p)$ X_p were beyond thresholds $k_a = 1, k_p = 4.5$; otherwise, the color was set to black. To obtain smooth transitions, v was calculated using a Hill function with Hill coefficient $n = 5$ and

$$v = \min_{i \in (a,p)} \left(\frac{-\log(x_i)^5}{k_i^5 - \log(x_i)^5} \right).$$

For TF-binding site predictions (Supplemental Table S4), we used weight matrices of TFs defined by SwissRegulon (Pachkov et al. 2007; <http://swissregulon.unibas.ch/cgi/sr/downloads>).

RNA-seq in the livers and kidneys of *Cry1 Δ e* and wild-type littermates

Parts of the livers and kidneys from the animals used for temporal 4C-seq experiments were frozen in liquid nitrogen immediately after sacrifice. Organs were homogenized in 4 M guanidine thiocyanate, 25 mM sodium citrate, 1% β -mercaptoethanol, and 0.2

Mermet et al.

M sodium acetate. Nucleic acids were extracted with phenol: chloroform:isoamylalcohol, and RNA was precipitated with 4 M LiCl. RNA concentration and purity were measured using nanodrop, and the quality was controlled by fragment analyzer. Poly-A-selected RNA was sequenced on NextSeq 500 machines using single-end 75-bp read length. mRNA levels were quantified using kallisto version 0.42.4 (mm10) (Bray et al. 2016).

RNA-seq in the livers and kidneys of Bmal1 knockout and wild-type mice

To complement the mouse liver wild-type and *Bmal1* knockout RNA-seq data (GSE73554), transcriptomes of kidneys from wild-type animals were measured following the same protocol as in Atger et al. (2015). mRNA levels were quantified using the same method as in Atger et al. (2015).

Circadian period estimation in Cry1Δe animals and wild-type littermates

Estimation of the circadian period was performed as in Diessler et al. (2017). Briefly, 8- to 10-wk-old males were single-caged and kept under 12 h/12 h light/dark cycle for 14 d and switched to constant darkness for 21 d. During the 5 wk of the experiment, the locomotor activity was recorded with passive infrared sensors. Data were sampled with 5-min resolution and analyzed using the χ^2 periodogram function in the ClockLab software (ActiMetrics). Food and water were available ad libitum during the entire experiment.

Western blotting

Liver cytoplasmic extracts were prepared as described previously (Jouffe et al. 2013). Protein extract concentrations were quantified using a BCA protein assay kit (Thermo Fisher Scientific), and 20 μ g of liver protein extract was resolved by SDS-PAGE using standard procedures. Densitometry analyses of the blots were performed using the ImageJ software. Naphtol blue and black staining of the membranes was used as a loading control and served as a reference for normalization of the quantified values. CRY1 antibody (1/500) was from Abcam (ab104736).

smRNA-FISH on mouse liver sections

Parts of the livers from the same animals used in the 4C-seq and RNA-seq were collected, immediately embedded in O.C.T. compound (Tissue-Tek, Sakura-Finetek USA), and snap-frozen. The RNA-FISH was done on 8- μ m cryosections using a RNAscope probe for *Cry1* pre-mRNA (*Cry1_intron1*, catalog no. 500231) according to the manufacturer's instructions for the RNAscope fluorescent multiplex assay (Advanced Cell Diagnostics). Nuclei were counterstained with DAPI, and sections were mounted with ProLong Gold anti-fade mountant (Molecular Probes).

Microscope image acquisition, quantification, and ploidy assignment

The sections were imaged using a Leica DM5500 wide-field microscope equipped with a CCD camera (DFC 3000) for fluorescence (Leica Microsystem) and a motorized stage. Z-stacks were acquired (0.2 μ m between each Z position, 40 images per frame) with an oil immersion 63 \times objective. The images were quantified using ImageJ. To detect the fluorescent RNA-FISH spots, a Laplacian filter was applied on a maximal projection, and local maxima were computed. Transcription site fluorescent intensities (burst size) were quantified on the sum projection of the nine best-focused stacks per image. Total transcription site signals were com-

puted using a mask of 3 \times 3 pixels. Nuclei were detected using filters, thresholding, and watershed transformation. Ploidy (2N, 4N, or 8N) was assigned to the nuclei based on their diameter (Bahar Halpern et al. 2015). A four-component Gaussian mixture model was fitted to the diameter distribution (package "mixtools" in R). Nuclei with a probability of >0.7 to belong to one of the three inferred populations with the smallest means were assigned to 2N, 4N, and 8N, respectively. The Gaussian distribution with the largest variance captured outliers in nucleus diameters (>15–18 μ m) and were discarded. Burst fraction was calculated as the number of active transcription sites in each nucleus divided by its estimated ploidy, and these fractions were then averaged over the entire populations of nuclei (Fig. 5B,C). For Supplemental Figure S10C, we modeled the number of active transcription sites with genotype-dependent slopes and compared it with a reduced model without a genotype effect (lme4 function in R, likelihood ratio test). For Supplemental Figure S10D, we modeled the mean intensity of intronic dots with genotype-dependent intercepts and compared it with a reduced model with a single intercept.

Data availability

Raw and processed sequencing data generated from this study (4C-seq and RNA-seq) have been submitted to Gene Expression Omnibus under accession number GSE101423.

Acknowledgments

We thank Isabelle Barde at the École Polytechnique Fédérale de Lausanne (EPFL) Transgenic Core Facility for help generating the *Cry1Δe* mouse strain, Jessica Dessimoz at the EPFL Histology Core Facility for smRNA-FISH experiments, Jacques Rougemont and Marion Leuleu for help with analysis of 4C-seq, Keith Harshman and Corinne Peter from the Genomic Technologies Facility at University of Lausanne, and Bastien Mangeat at the EPFL Gene Expression Core Facility for 4C-seq and RNA-seq. This work was supported by Swiss National Science Foundation grant 31003A-153340 (to F.N.), European Research Council grant ERC-2010-StG-260667 (to F.N.), and École Polytechnique Fédérale de Lausanne. J.Y. benefits from the Natural Sciences and Engineering Research Council of Canada Post-graduate Studies Doctoral scholarship. Computations and analyses were performed at Vital-IT (<http://www.vital-it.ch>).

Author contributions: J.M., J.Y., and F.N. conceived the study; J.Y., K.G., C.G., and F.N. performed the formal analyses; J.M., C.H., D.M., C.J., D.N., and Y.E. performed the investigations; J.M., J.Y., and F.N. wrote the manuscript; J.M., J.Y., P.F., F.G., and F.N. reviewed and edited the manuscript; P.F., F.G., and F.N. supervised the study; and P.F., F.G., and F.N. acquired the funding.

Note added in proof

While this manuscript was in review, similar 24-h rhythmic promoter-enhancer chromatin interactions at the *Cry1* locus were reported (Kim et al. 2018).

References

Aguilar-Arnal L, Hakim O, Patel VR, Baldi P, Hager GL, Sassone-Corsi P. 2013. Cycles in spatial and temporal chromosomal

- organization driven by the circadian clock. *Nat Struct Mol Biol* **20**: 1206–1213.
- Allebrandt KV, Teder-Laving M, Akyol M, Pichler I, Muller-Myhsok B, Pramstaller P, Merrow M, Meitinger T, Metspalu A, Roenneberg T. 2010. CLOCK gene variants associate with sleep duration in two independent populations. *Biol Psychiatry* **67**: 1040–1047.
- Aschoff J, Pohl H. 1978. Phase relations between a circadian-rhythm and its Zeitgeber within range of entrainment. *Naturwissenschaften* **65**: 80–84.
- Atger F, Gobet C, Marquis J, Martin E, Wang J, Weger B, Lefebvre G, Descombes P, Naef F, Gachon F. 2015. Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc Natl Acad Sci* **112**: E6579–E6588.
- Bahar Halpern K, Tanami S, Landen S, Chapal M, Szlak L, Hutzler A, Nizhberg A, Itzkovitz S. 2015. Bursty gene expression in the intact mammalian liver. *Mol Cell* **58**: 147–156.
- Bartman CR, Hsu SC, Hsiung CCS, Raj A, Blobel GA. 2016. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol Cell* **62**: 237–247.
- Bass J, Lazar MA. 2016. Circadian time signatures of fitness and disease. *Science* **354**: 994–999.
- Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**: 525–527.
- Cermakian N, Monaco L, Pando MP, Dierich A, Sassone-Corsi P. 2001. Altered behavioral rhythms and clock gene expression in mice with a targeted mutation in the *Period1* gene. *EMBO J* **20**: 3967–3974.
- Debruyne JP, Noton E, Lambert CM, Maywood ES, Weaver DR, Reppert SM. 2006. A clock shock: mouse CLOCK is not required for circadian oscillator function. *Neuron* **50**: 465–477.
- Dekker J, Marti-Renom MA, Mirny LA. 2013. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* **14**: 390–403.
- Diessler S, Kostic C, Arsenijevic Y, Kawasaki A, Franken P. 2017. Rail frees mice from the repression of active wake behaviors by light. *eLife* **6**: e23292.
- Doi R, Oishi K, Ishida N. 2010. CLOCK regulates circadian rhythms of hepatic glycogen synthesis through transcriptional activation of *Gys2*. *J Biol Chem* **285**: 22114–22121.
- Fukaya T, Lim B, Levine M. 2016. Enhancer control of transcriptional bursting. *Cell* **166**: 358–368.
- Fulco CP, Munschauer M, Anyoha R, Munson G, Grossman SR, Perez EM, Kane M, Cleary B, Lander ES, Engreitz JM. 2016. Systematic mapping of functional enhancer–promoter connections with CRISPR interference. *Science* **354**: 769–773.
- Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EE. 2014. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **512**: 96–100.
- Gheldof N, Leleu M, Noordermeer D, Rougemont J, Reymond A. 2012. Detecting long-range chromatin interactions using the chromosome conformation capture sequencing (4C-seq) method. *Methods Mol Biol* **786**: 211–225.
- Griffin EA Jr, Staknis D, Weitz CJ. 1999. Light-independent role of CRY1 and CRY2 in the mammalian circadian clock. *Science* **286**: 768–771.
- Hu Y, Shmygelska A, Tran D, Eriksson N, Tung JY, Hinds DA. 2016. GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nat Commun* **7**: 10448.
- Irimia JM, Meyer CM, Peper CL, Zhai L, Bock CB, Previs SF, McGuinness OP, DePaoli-Roach A, Roach PJ. 2010. Impaired glucose tolerance and predisposition to the fasted state in liver glycogen synthase knock-out mice. *J Biol Chem* **285**: 12851–12861.
- Jouffe C, Cretenet G, Symul L, Martin E, Atger F, Naef F, Gachon F. 2013. The circadian clock coordinates ribosome biogenesis. *PLoS Biol* **11**: e1001455.
- Kim YH, Marhon SA, Zhang Y, Steger DJ, Won KJ, Lazar MA. 2018. Rev-erba dynamically modulates chromatin looping to control circadian gene transcription. *Science* doi: 10.1126/science.aao6891.
- Kuznetsova T, Wang SY, Rao NA, Mandoli A, Martens JH, Rother N, Aartse A, Groh L, Janssen-Megens EM, Li G, et al. 2015. Glucocorticoid receptor and nuclear factor κ B affect three-dimensional chromatin organization. *Genome Biol* **16**: 264.
- Levine M, Tjian R. 2003. Transcription regulation and animal diversity. *Nature* **424**: 147–151.
- Mermet J, Yeung J, Naef F. 2017. Systems chronobiology: global analysis of gene regulation in a 24-hour periodic world. *Cold Spring Harb Perspect Biol* **9**: a028720.
- Noordermeer D, Leleu M, Schorderet P, Joye E, Chabaud F, Duboule D. 2014. Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. *eLife* **3**: e02557.
- Pachkov M, Erb I, Molina N, van Nimwegen E. 2007. SwissRegulon: a database of genome-wide annotations of regulatory sites. *Nucleic Acids Res* **35**: D127–D131.
- Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, de Laat W. 2003. The β -globin nuclear compartment in development and erythroid differentiation. *Nat Genet* **35**: 190–194.
- Rey G, Cesbron F, Rougemont J, Reinke H, Brunner M, Naef F. 2011. Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. *PLoS Biol* **9**: e1000595.
- Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, Jewett AJ, Bochkov ID, Chinnappan D, Cutkosky A, Li J, et al. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci* **112**: E6456–E6465.
- Sanyal A, Lajoie BR, Jain G, Dekker J. 2012. The long-range interaction landscape of gene promoters. *Nature* **489**: 109–113.
- Schibler U, Gotic I, Saini C, Gos P, Curie T, Emmenegger Y, Sirturel F, Gosselin P, Gerber A, Fleury-Olela F, et al. 2015. Clock-talk: interactions between central and peripheral circadian oscillators in mammals. *Cold Spring Harb Symp Quant Biol* **80**: 223–232.
- Sobel JA, Krier I, Andersin T, Raghav S, Canella D, Gilardi F, Kalantzi AS, Rey G, Weger B, Gachon F, et al. 2017. Transcriptional regulatory logic of the diurnal cycle in the mouse liver. *PLoS Biol* **15**: e2001069.
- Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**: 472–474.
- Takahashi JS. 2017. Transcriptional architecture of the mammalian circadian clock. *Nat Rev Genet* **18**: 164–179.
- Toh KL, Jones CR, He Y, Eide EJ, Hinze WA, Virshup DM, Ptacek LJ, Fu YH. 2001. An hPer2 phosphorylation site mutation in familial advanced sleep phase syndrome. *Science* **291**: 1040–1043.
- Ukai-Tadenuma M, Yamada RG, Xu H, Ripperger JA, Liu AC, Ueda HR. 2011. Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell* **144**: 268–281.
- van der Horst GT, Muijtjens M, Kobayashi K, Takano R, Kanno S, Takao M, de Wit J, Verkerk A, Eker AP, van Leenen D, et al.

Mermet et al.

1999. Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature* **398**: 627–630.
- Vitaterna MH, King DP, Chang AM, Kornhauser JM, Lowrey PL, McDonald JD, Dove WF, Pinto LH, Turek FW, Takahashi JS. 1994. Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science* **264**: 719–725.
- Xu Y, Guo W, Li P, Zhang Y, Zhao M, Fan Z, Zhao Z, Yan J. 2016. Long-range chromosome interactions mediated by cohesin shape circadian gene expression. *PLoS Genet* **12**: e1005992.
- Yeung J, Mermet J, Jouffe C, Marquis J, Charpagne A, Gachon F, Naef F. 2018. Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs. *Genome Res* **28**: 182–191.
- Zhang R, Lahens NF, Ballance HI, Hughes ME, Hogenesch JB. 2014. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci* **111**: 16219–16224.
- Zhang Y, Fang B, Emmett MJ, Damle M, Sun Z, Feng D, Armour SM, Remsberg JR, Jager J, Soccio RE, et al. 2015. Discrete functions of nuclear receptor Rev-erba couple metabolism to the clock. *Science* **348**: 1488–1492.
- Zhao H, Sifakis EG, Sumida N, Millan-Arino L, Scholz BA, Svensson JP, Chen X, Ronnegren AL, Mallet de Lima CD, Varnosfaderani FS, et al. 2015. PARP1- and CTCF-mediated interactions between active and repressed chromatin at the lamina promote oscillating transcription. *Mol Cell* **59**: 984–997.



Clock-dependent chromatin topology modulates circadian transcription and behavior

Jérôme Mermet, Jake Yeung, Clémence Hurni, et al.

Genes Dev. 2018, **32**: originally published online March 23, 2018
Access the most recent version at doi:[10.1101/gad.312397.118](https://doi.org/10.1101/gad.312397.118)

Supplemental Material <http://genesdev.cshlp.org/content/suppl/2018/03/23/gad.312397.118.DC1>

Related Content **A day in the life of chromatin: how enhancer promoter loops shape daily behavior**
Benjamin J. Weidemann, Kathryn Moynihan Ramsey and Joseph Bass
[Genes Dev. March , 2018 32: 321-323](#)

References This article cites 44 articles, 17 of which can be accessed free at:
<http://genesdev.cshlp.org/content/32/5-6/347.full.html#ref-list-1>
Articles cited in:
<http://genesdev.cshlp.org/content/32/5-6/347.full.html#related-urls>

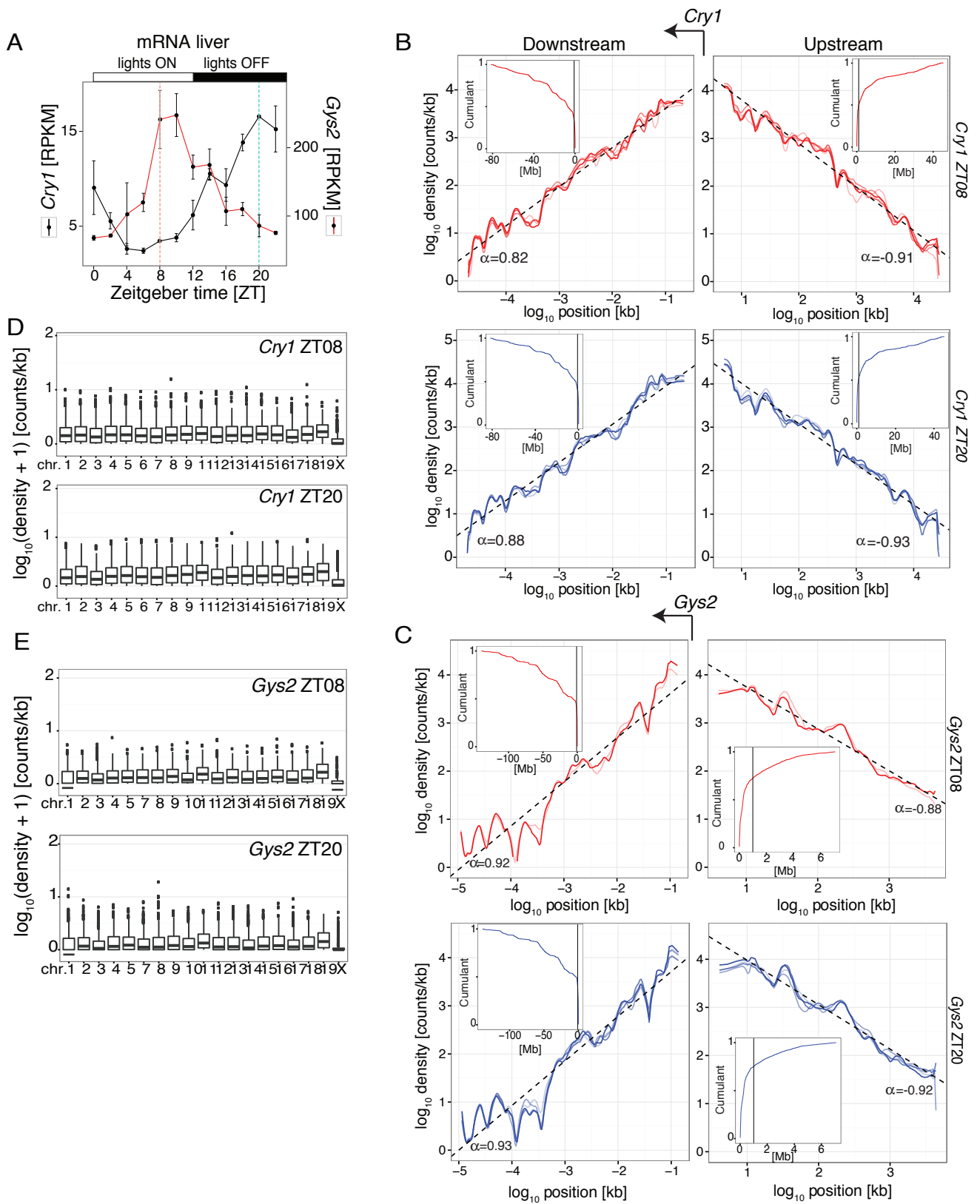
Creative Commons License This article, published in *Genes & Development*, is available under a Creative Commons License (Attribution 4.0 International), as described at
<http://creativecommons.org/licenses/by/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Boost NGS microRNA profiling.
Read about 3 methods tested

EXIQON
Now a QIAGEN company

The QIAGEN logo consists of a grid of blue dots of varying sizes, with the word 'QIAGEN' in red capital letters below it.

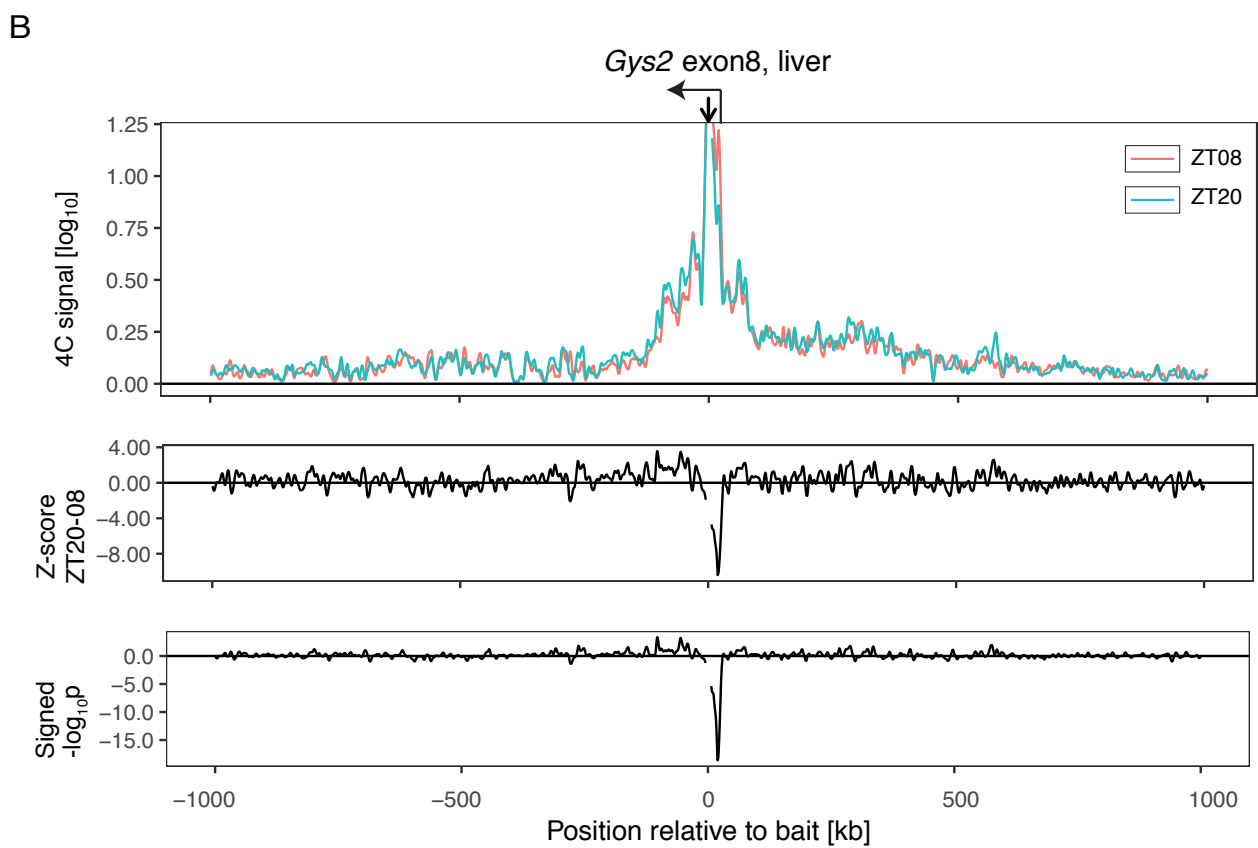
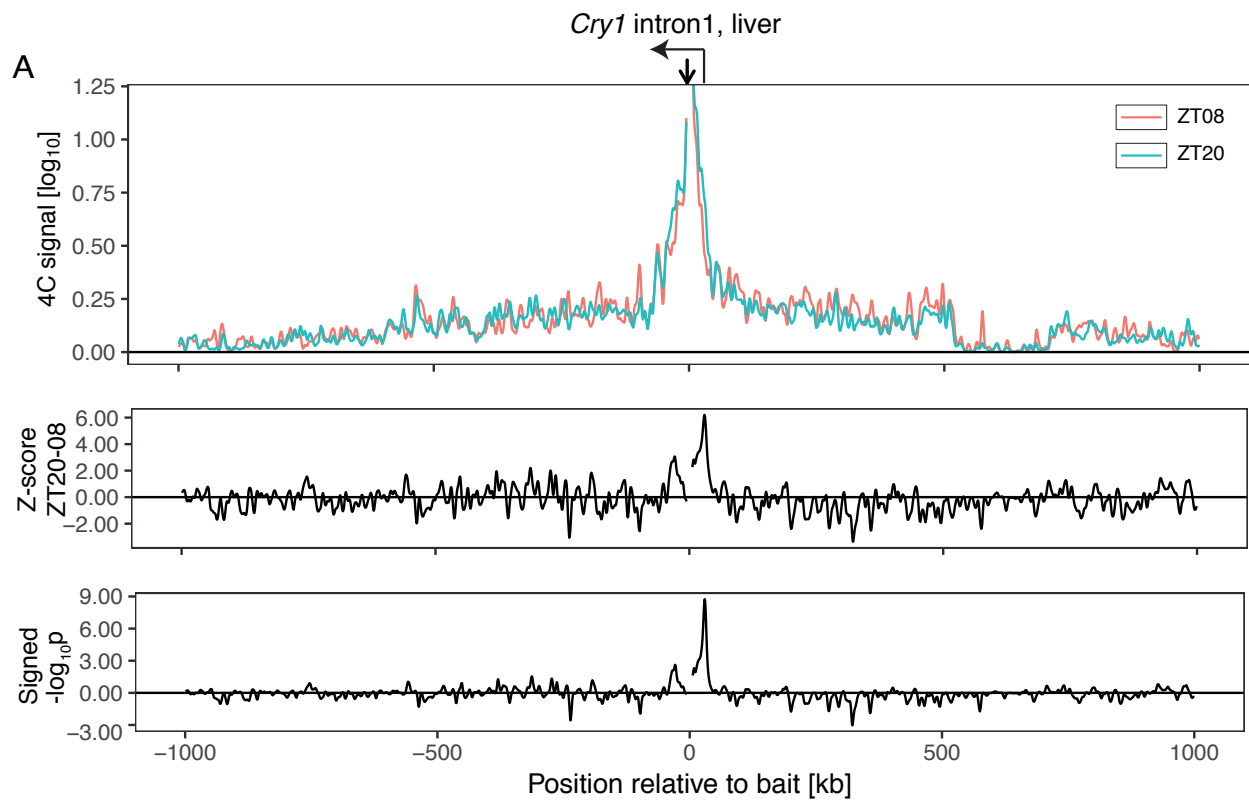


1 **Supplemental Figure 1: 4C-seq signal on *cis* and *trans* on baits targeting**
2 ***Cry1* and *Gys2*.** (A) Temporal *Cry1* and *Gys2* mRNA accumulations in mouse liver
3 (Atger et al., 2015) (reads per kilobase per million, RPKM); error bars: standard
4 deviation (SD) of two animals; red and blue dashed lines: 4C-seq time points.
5 (B,C) 4C-seq counts density vs. genomic position from *Cry1* TSS (B) and *Gys2* TSS
6 (C) for each biological replicate at ZT08 (red) and ZT20 (blue) in the WT mouse
7 liver. Dashed lines: power-law fit. α are decay exponents. Insets: cumulative
8 counts on the *cis* chromosome from the bait position to chromosome end. Black
9 vertical line shows 1 Mb from bait. (D, E) *Cry1* TSS (D) and *Gys2* TSS (E) 4C-seq
10 counts density on *trans* chromosomes.

11

12

Mermet Supplemental Figure 2



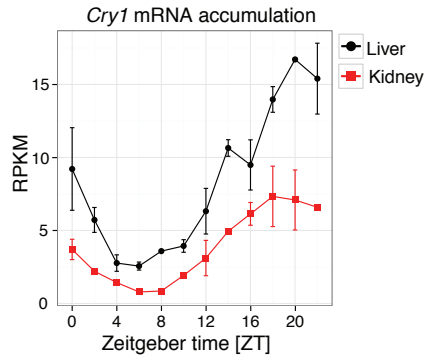
13 **Supplemental Figure 2: Validation of rhythmic chromatin interactions at**
 14 ***Cry1* and *Gys2* by targeting their respective enhancers. (A,B)** 4C-seq signals
 15 from the *Cry1* intron1 bait (A) and *Gys2* exon8 bait (B) in mouse liver at ZT08
 16 (red) and ZT20 (blue) and the corresponding Z-score (ZT20-ZT08) and signed –
 17 $\log_{10}(p)$ in a genomic window of 2 Mb. Largest differential signal occurs at the
 18 promoter region of the respective genes. Vertical arrows show location of bait
 19 relative to TSS of the respective gene.

20

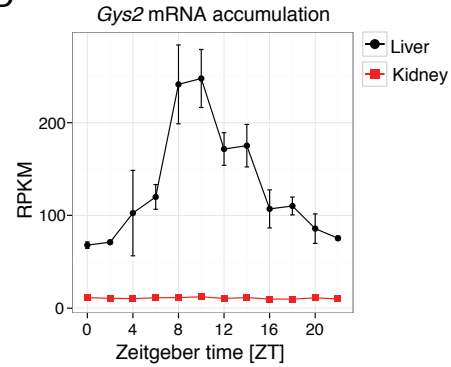
21

Mermet Supplemental Figure 3

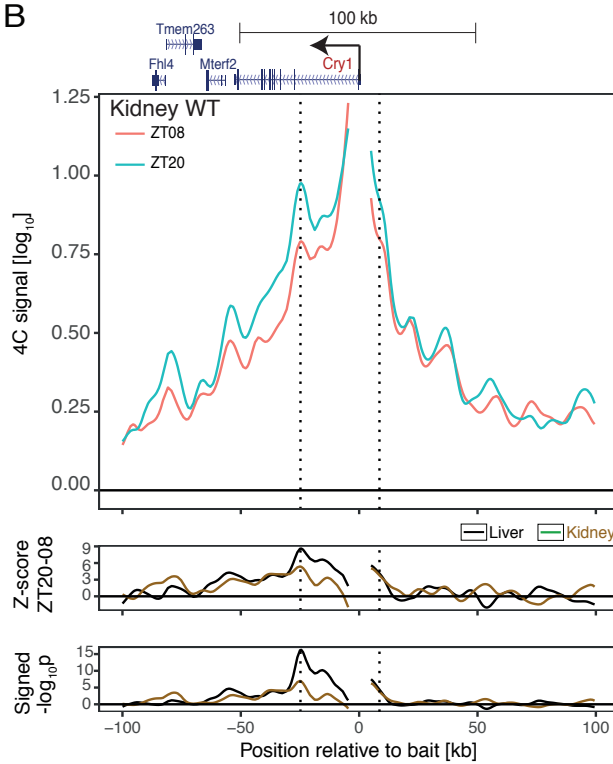
A



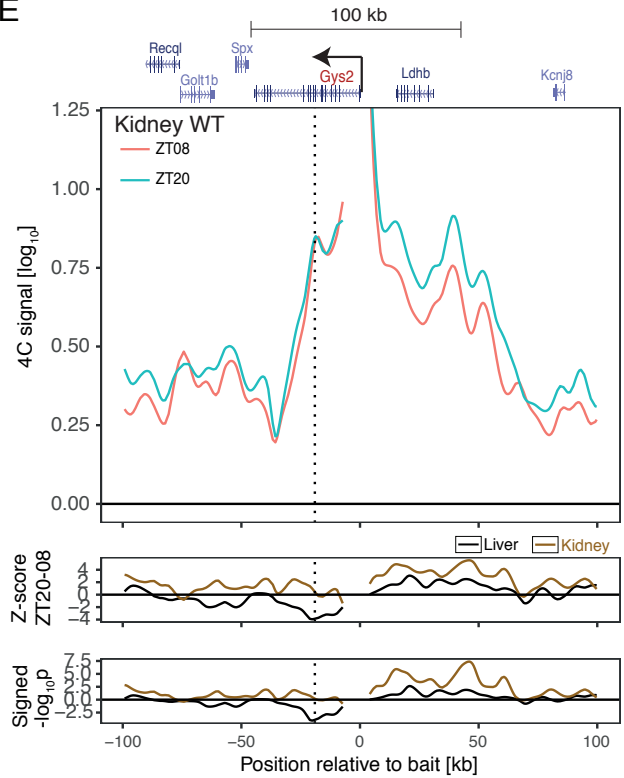
D



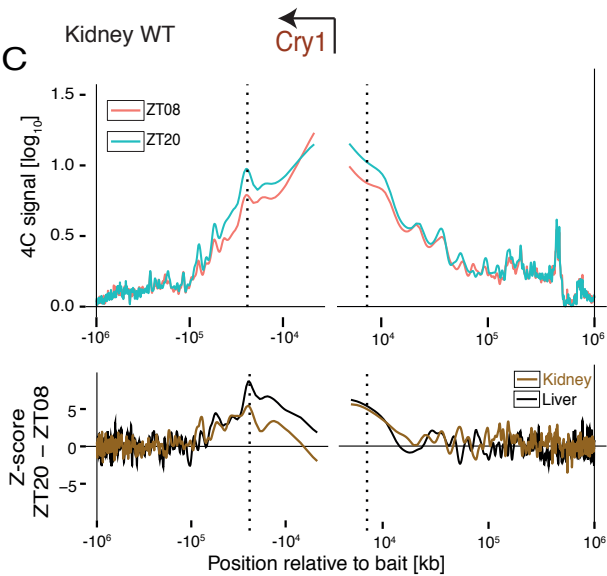
B



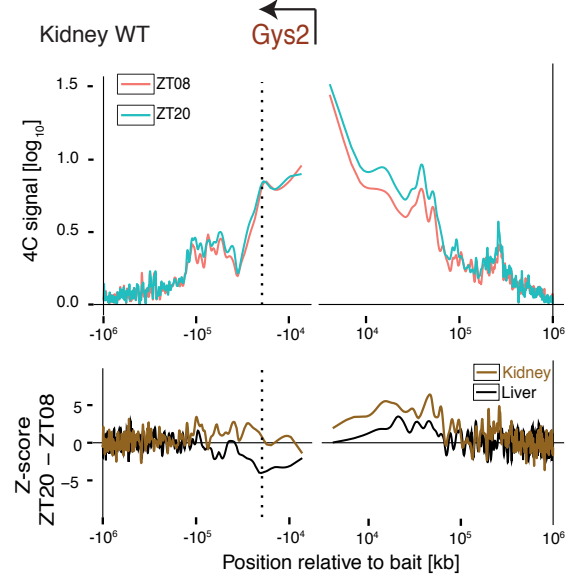
E



C



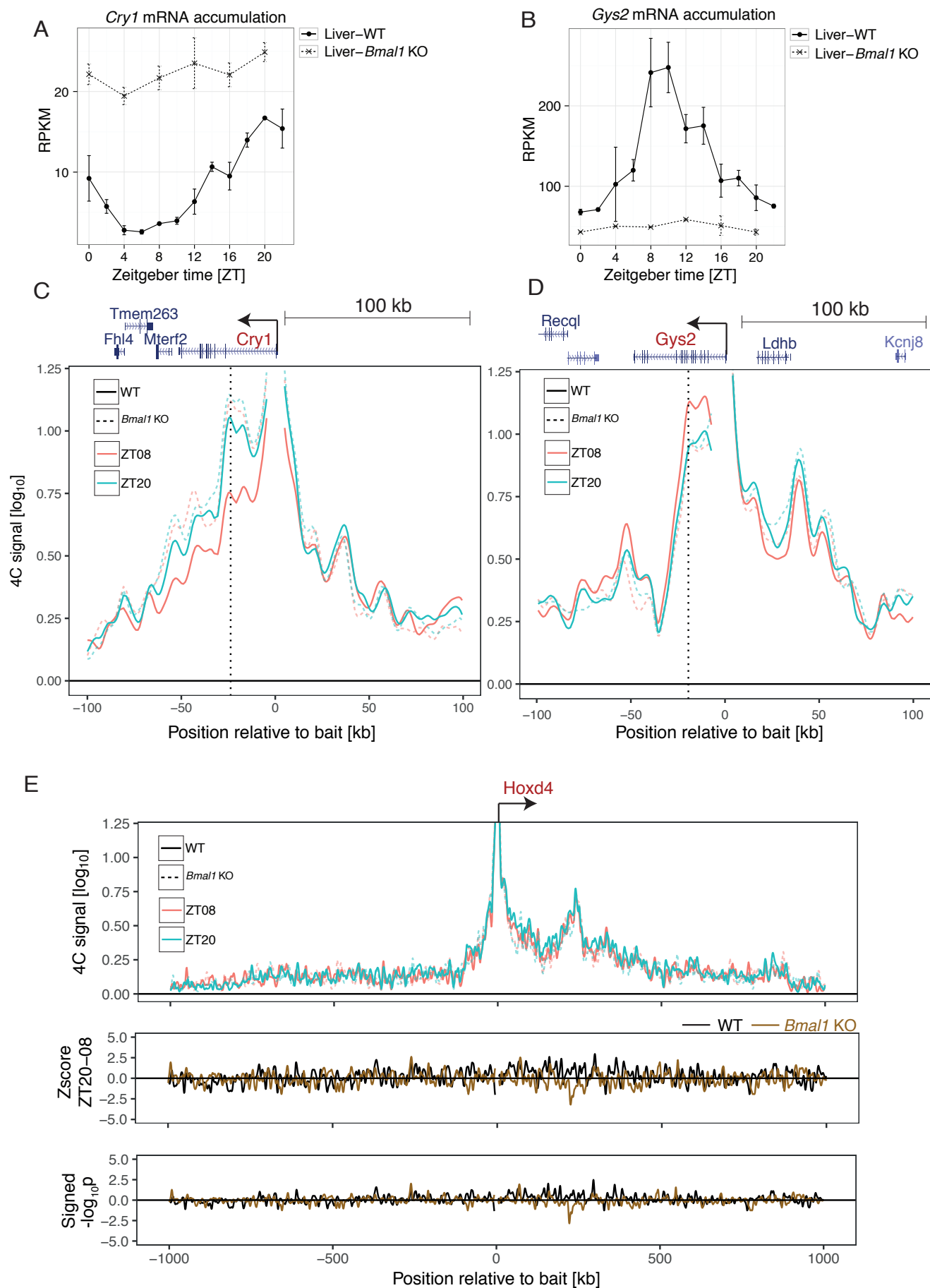
F



Supplemental Figure 3: 4C-seq in kidney show rhythmic chromatin interactions at *Cry1* TSS but not at *Gys2* TSS.

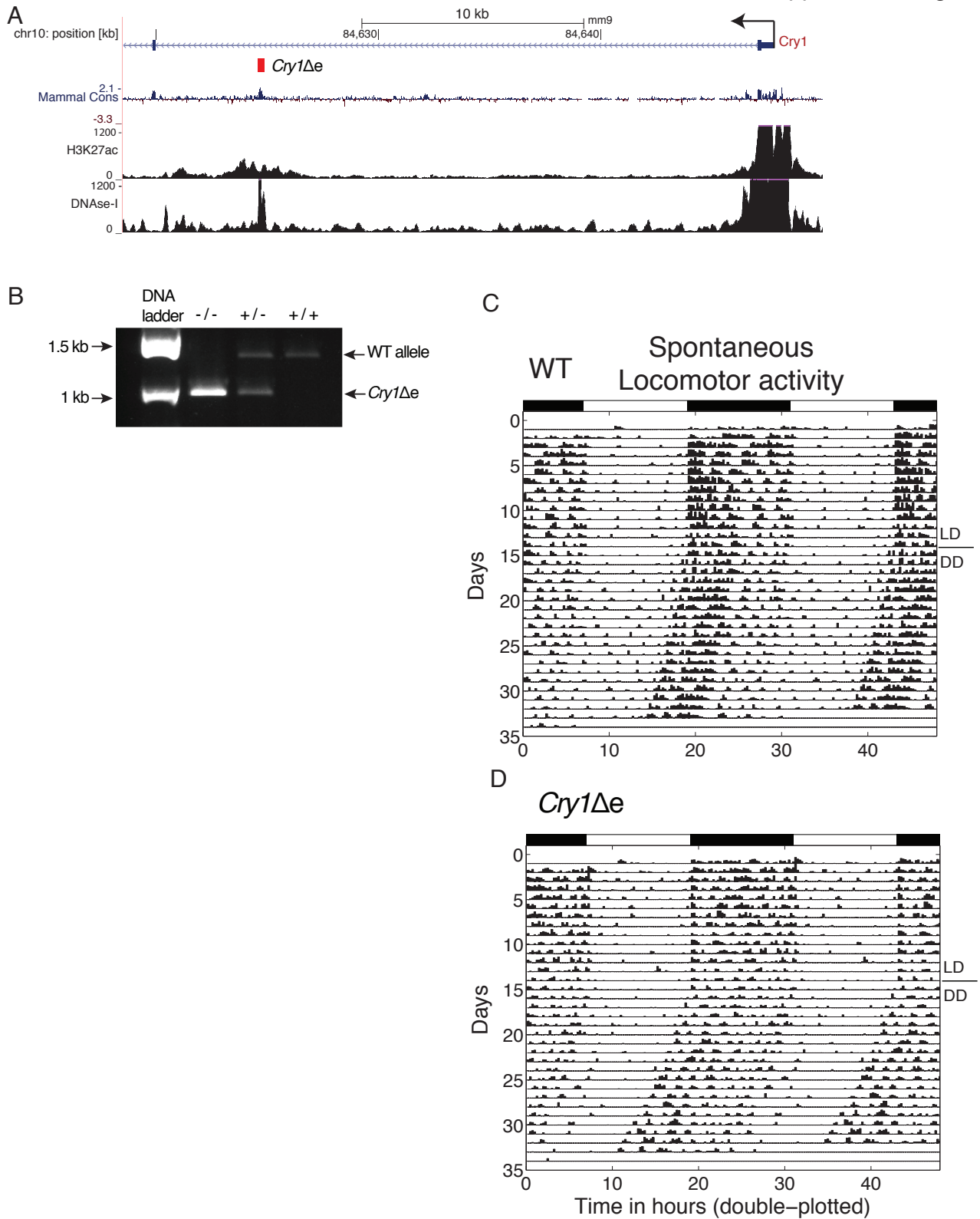
(A) *Cry1* mRNA profile in WT mouse liver (black) (Atger et al., 2015) and kidney (red); error bars: SD of two animals. (B) 4C-seq signal from *Cry1* TSS bait in WT kidney at ZT08 (red) and ZT20 (blue) in a genomic window of 200 kb. Vertical lines for *Cry1* show rhythmic contacts in kidney and liver. Z-score (ZT20 vs 08) and signed $-\log_{10}(p)$ for rhythmic contacts in liver (black) and kidney (red). (C) Same as (B) using a 2 MB window. (D) same as (A) for *Gys2* mRNA. (E) same as (B) for *Gys2* TSS. Vertical lines for *Gys2* show contacts that are static in kidney but rhythmic in liver. (F) same as (C) for *Gys2* TSS.

Mermet Supplemental Figure 4



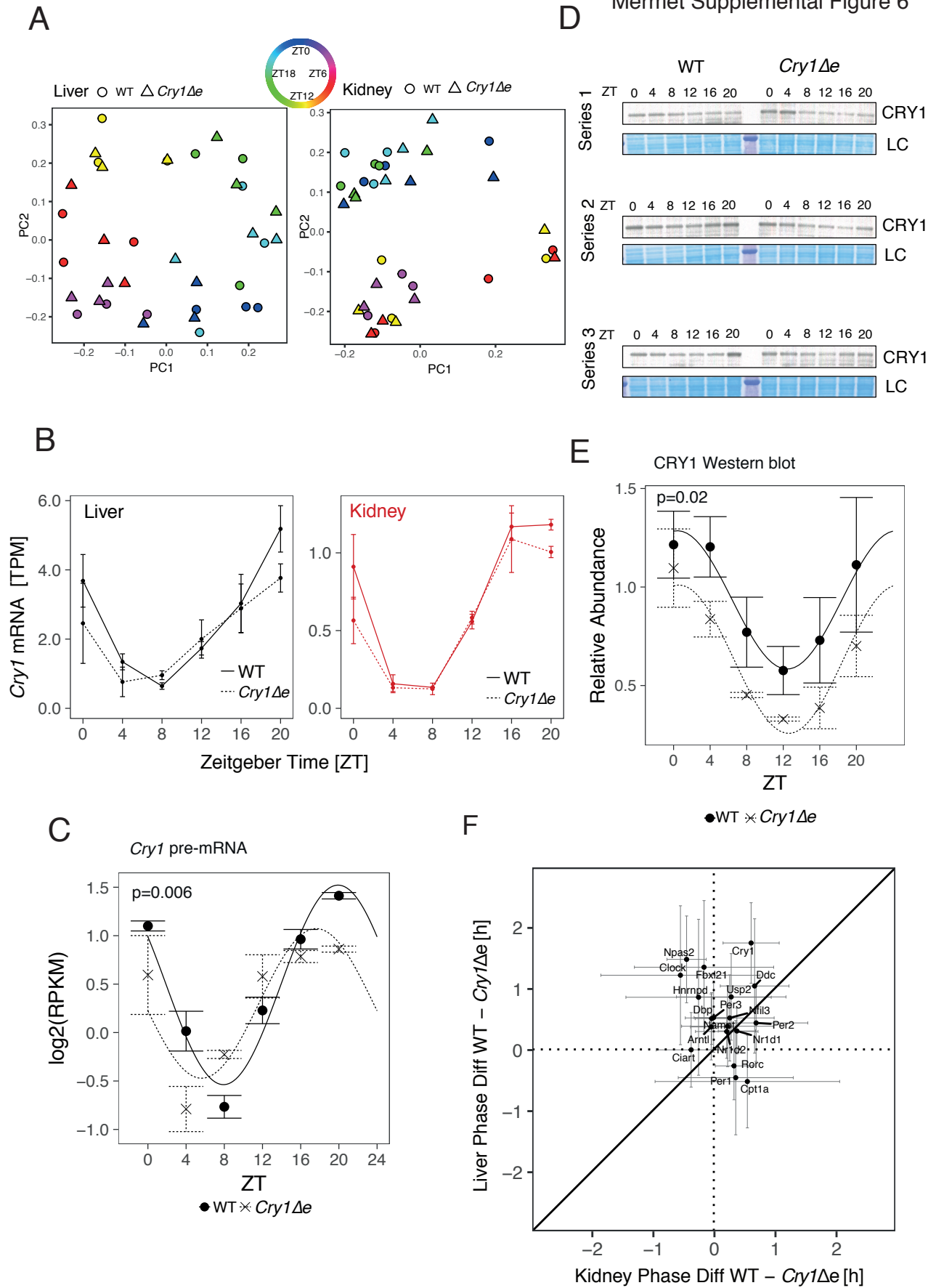
34 **Supplemental Figure 4: Dynamics of chromatin topology depends on**
 35 **BMAL1. (A,B)** *Cry1* (A) and *Gys2* (B) mRNA profile in WT and *Bmal1* KO mouse
 36 liver along the diurnal cycle from (Atger et al., 2015); error bars: SD of two
 37 animals. **(C,D)** 4C-seq signals from *Cry1* TSS (C) and *Gys2* TSS (D) bait in WT and
 38 *Bmal1* KO livers at ZT08 and ZT20. Vertical lines show BMAL1-dependent
 39 rhythmic contacts. **(E)** 4C-seq signals from negative control region *Hoxd4* in WT
 40 and *Bmal1* KO livers show static chromatin topology. Bottom: Z-score (ZT20 vs
 41 08) and signed $-\log_{10}(p)$ for WT and *Bmal1* KO for *Hoxd4*.
 42
 43

Mermet Supplemental Figure 5



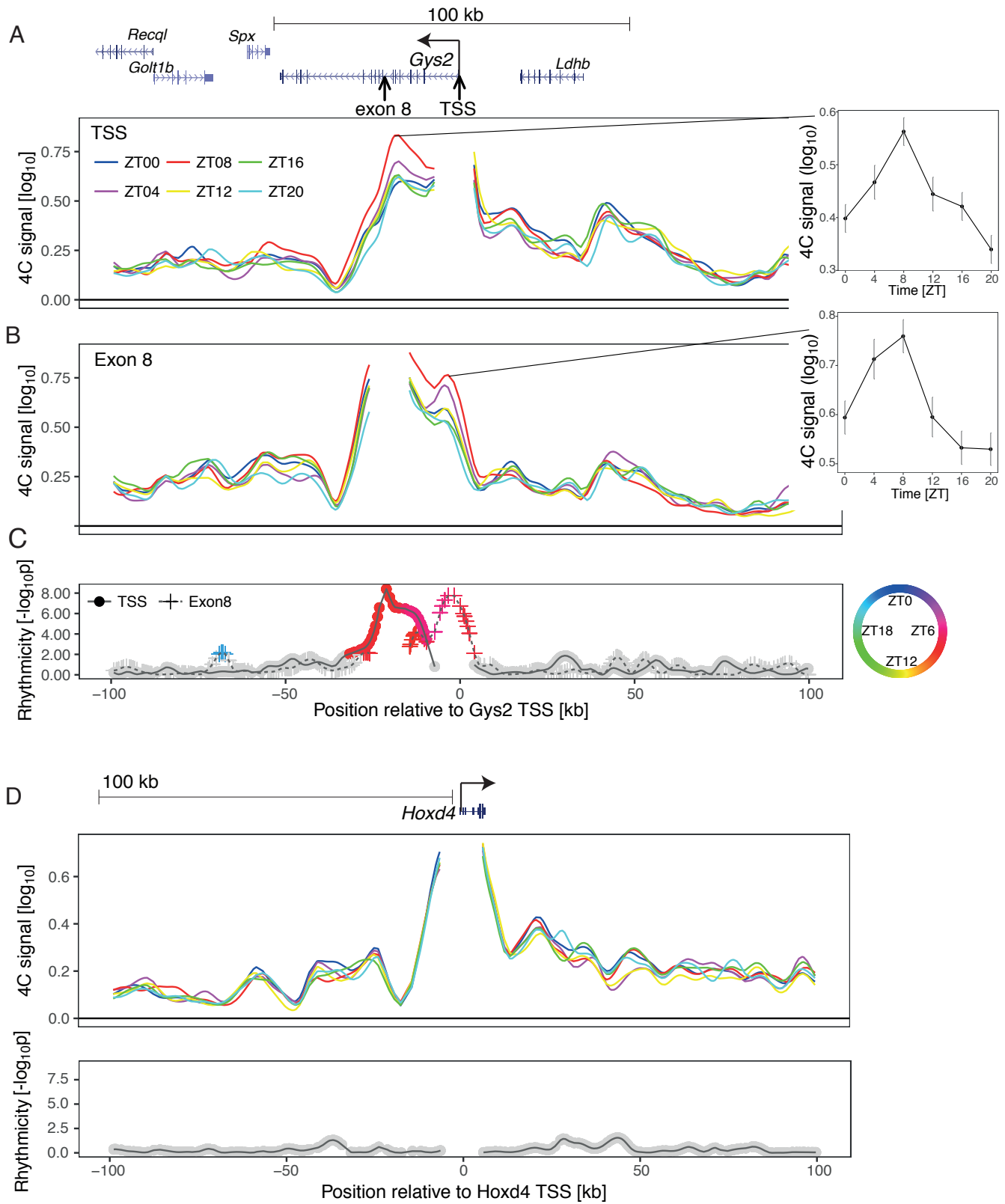
44 **Supplemental Figure 5: Deleting *Cry1* intronic enhancer *in vivo* shortens**
 45 **the free-running period and phase advances clock and clock-controlled**
 46 **genes. (A)** Genome browser view showing the CRISPR-Cas9 mediated deletion of
 47 the *Cry1* intronic enhancer in mouse (*Cry1Δe*). The evolutionary conserved
 48 300bp deletion (red square) covers the DHS containing a RRE described in
 49 mouse fibroblasts (Ukai-Tadenuma et al., 2011) within the H3K27ac marked
 50 intronic region. Note that this sequence is as conserved as *Cry1* exonic
 51 sequences. **(B)** Agarose gel showing the deletion. **(C-D)** Actograms showing the
 52 spontaneous locomotor activity recorded for 14 days in 12 hours/12 hours
 53 light/dark (LD) cycles followed by 21 days in constant darkness (DD) in a
 54 representative WT (C) and a *Cry1Δe* animal (D).
 55
 56

Mermet Supplemental Figure 6

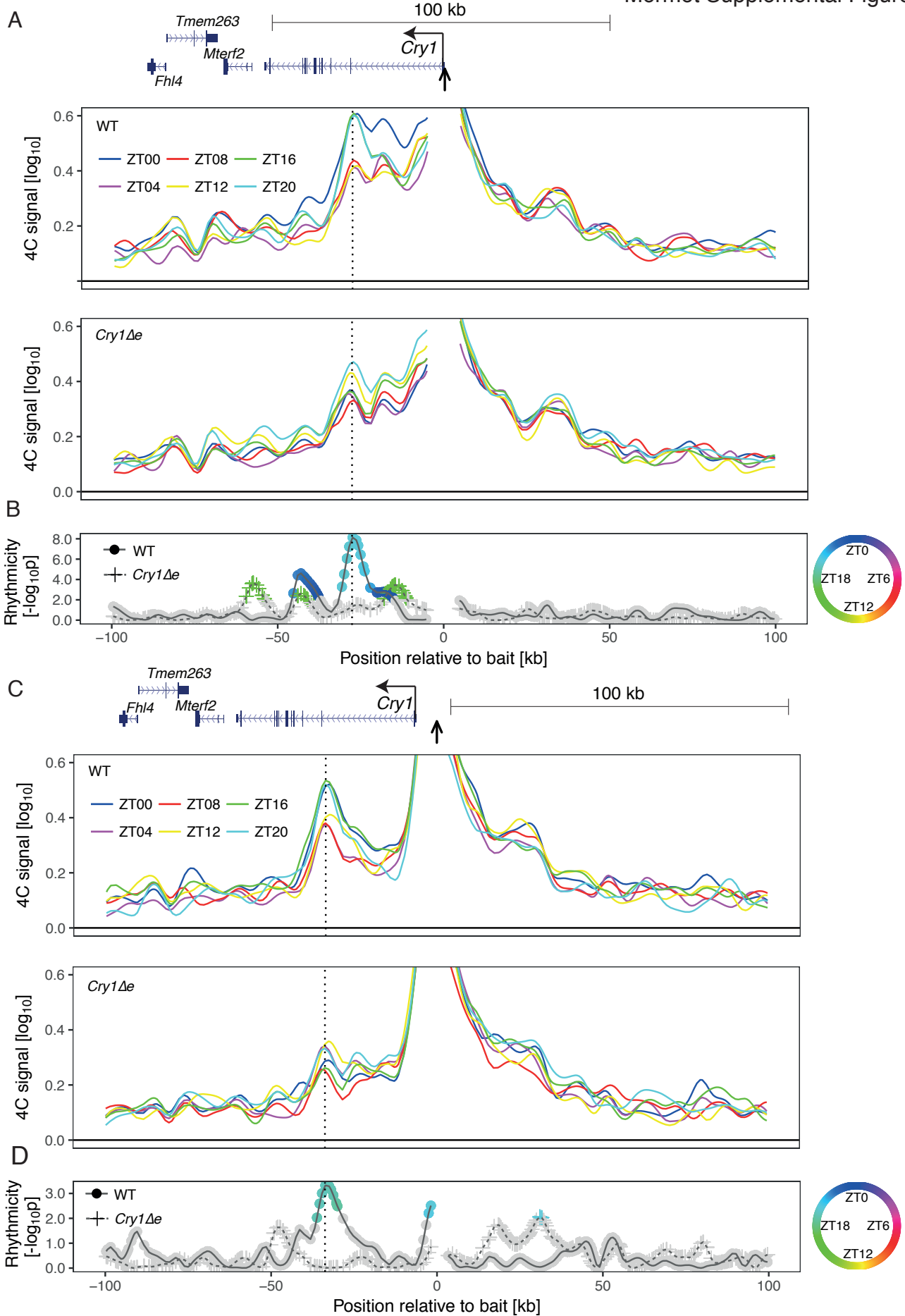


Supplemental Figure 6: Deleting *Cry1* intronic enhancer *in vivo* shortens phase advances the clock and clock-controlled genes.

(A) PCA analysis of temporal RNA-seq data in the liver (left panel) and the kidney (right panel) of *Cry1Δe* and WT littermates. Right: color code for ZT time. (B) *Cry1* mRNA profile (in Transcripts Per Million, TPM) in liver (black) and kidney (red) of *Cry1Δe* (dashed line) and WT (solid line) littermates. Error bars: SD of 3 animals. At ZT20, *Cry1* mRNA levels are significantly different between WT versus *Cry1Δe* for liver and kidney (respectively $p=0.045$ and $p=0.0037$, t-test). *Cry1* mRNA is phase advanced in *Cry1Δe* versus WT ($p=0.011$ and $p=0.0047$ for liver and kidney, respectively, bootstrap test). (C) *Cry1* intronic reads around the clock for WT and *Cry1Δe* mice. The oscillations in WT versus *Cry1Δe* are significantly different ($p=0.0063$, F-test test) (D,E) Temporal expression of CRY1 protein in the liver cytoplasm of WT and *Cry1Δe* littermates. Western blot (D) and quantification of CRY1 relative abundance normalized by the Naptho blue black coloration of the membranes (E). CRY1 shows a decreased protein abundance in *Cry1Δe* ($p=0.02$, F-test). (F) Differential phase between *Cry1Δe* versus WT in liver and kidney. Genes selected for core-clock and clock-controlled transcripts ($p<0.01$, harmonic regression in both liver and kidney) based on Gene Ontology (GO) annotation indicated in table S5.

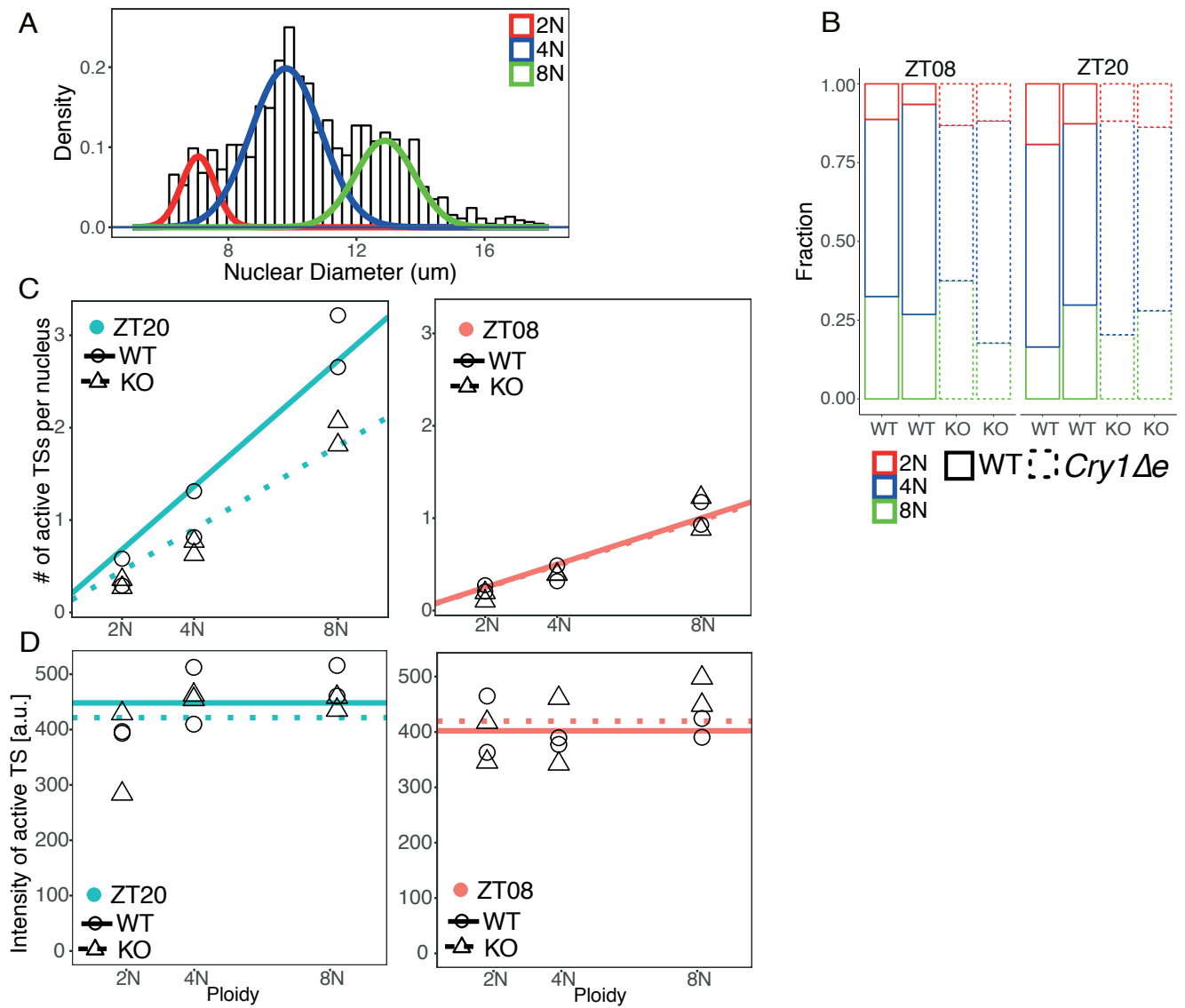


80 **Supplemental Figure 7: 4C-seq around the clock at *Gys2* TSS and at *Gys2***
 81 **exon8 demonstrates rhythmic chromatin topology.**
 82 **(A-B)** 4C-seq signals (LWMR summarizing n=3 animals per condition) from *Gys2*
 83 TSS (A) and exon8 (B) baits across time. Right: 4C signal over time at oscillatory
 84 chromatin contacts. **(C)** Statistical significance of rhythmic amplitudes of
 85 rhythmic contacts ($p < 10^{-7}$ for TSS and exon8, chi-squared test). Fragments with
 86 $p < 0.01$ are colored by their time of peak contact frequency according to color
 87 legend (right). **(D)** 4C-seq around the clock on negative control region, *Hoxd4*, in
 88 WT animals shows static chromatin topology.
 89
 90



91 **Supplemental Figure 8: Deleting the *Cry1* intronic enhancer disrupts**
 92 **rhythmic chromatin topology. (A)** 4C-seq signals (LWMR summarizing n=3
 93 animals per condition) from *Cry1* TSS bait across time in WT (top panel) and
 94 *Cry1*Δ*e* livers (middle panel). **(B)** Statistical significance of amplitudes of
 95 rhythmic contacts (chi-squared test, bottom panel) for *Cry1* TSS and *Cry1*Δ*e*
 96 baits. Fragments with p<0.01 are colored by their time of peak contact frequency
 97 according to color legend (right). **(C)** *Idem* as (A) targeting the RRE bait -7 kb
 98 upstream *Cry1* TSS (*Cry1* upstream) **(D)** *Idem* as (B) for *Cry1* upstream. Vertical
 99 lines show rhythmic promoter-enhancer interactions that are disrupted in
 100 *Cry1*Δ*e* liver. Vertical arrows show location of the bait relative to the *Cry1* TSS.
 101
 102

Mermet Supplemental Figure 9



103 **Supplemental Figure 9: smRNA-FISH against *Cry1* pre-mRNA in WT versus**
104 ***Cry1Δe* livers. (A)** Size distribution of nuclei for a representative animal.
105 Colored curves show fitted Gaussian mixture model corresponding to
106 populations with 2N, 4N, 8N ploidy. **(B)** Fraction of nuclei for each animal (n=2
107 per condition) assigned to different ploidy. **(C)** Number of active transcription
108 sites (TSs) averaged per animal increases with ploidy. At ZT20, *Cry1Δe* animals
109 show reduced number of TSs compared to WT: lines show mixed effect linear
110 model with genotype-dependent slopes ($p(H_0:\text{equal slopes})=0.00014$, F-test); at
111 ZT08, the slopes are not different ($p=0.84$). **(D)** Active TS intensity averaged per
112 animal shows comparable intensity across ploidy and conditions: lines show
113 mixed effect model with genotype-dependent intercepts, intercept comparisons
114 at both ZT08 ($p(H_0:\text{equal intercept})=0.53$, F-test) and ZT20 ($p=0.41$) are not
115 significant.
116

117 **Supplemental Table 1: Distribution of 4C-seq counts on the genome.** Sheet
 118 1: Distribution of 4C-seq raw counts for each bait and each biological replicate in
 119 the genome. Exponent of the power-law fit is indicated. Sheet 2: σ_{min} used to
 120 regularize the residual variance for each bait.

121 **Supplemental Table 2: Sequences of CRISPR-Cas9 RNA guides.** Sequences of
 122 CRISPR-Cas9 RNA guides targeting upstream and downstream the *Cry1* intronic
 123 enhancer element. The sequence of the PCR primers used to screen the deletion
 124 is indicated.

125

126 **Supplemental Table 3: 4C-seq primers.** Sequence of the PCR primers used for
 127 the inverse PCR step during 4C-seq libraries preparation.

128

129 **Supplemental Table 4: Motif counts.** List of motif counts for TF motifs from
 130 SwissRegulon. TFBS site count probabilities calculated using Motevo. Genomic
 131 coordinates from mm9 annotations.

132

133 **Supplemental Table 5: GO term annotation.** List of GO term annotation used
 134 to analyze temporal RNA-seq data in the liver and kidney of *Cry1* Δ e and WT
 135 littermates.

136

3.3 Conclusion and perspectives

This study establishes dynamic chromatin interactions as a novel regulatory layer underlying circadian gene expression and behavior. By performing 4C-seq around the 24-hour day, we find that the proper timing of *Cry1* gene expression depends on a distal enhancer located in the first intron. We establish the function of this distal enhancer by deleting the *Cry1* enhancer. Looking at single-cell transcription with smRNA FISH, we link this deletion with changes in transcriptional bursting, specifically the burst frequency but not burst size. Finally, we find a behavioral phenotype from mice with the deleted enhancer; the period of the circadian locomotor activity was decreased by 15 minutes. This 15 minutes is expected because deleting the entire *Cry1* gene shortens locomotor activity by approximately 1.2 hours (Ko and Takahashi, 2006).

1. How are the chromatin conformation data relate to what is happening at the single cells? Does transcription occur immediately after a promoter-enhancer contact?
2. How do we incorporate publicly available Hi-C datasets onto our study to predict more loci with circadian dynamics?

4 Dynamic gene expression and regulation from circadian and sleep-wake processes in the mouse cortex

4.1 Introduction

This collaborative project with the Franken lab at the University of Lausanne investigates the gene expression dynamics in response to acute sleep deprivation. The computational analysis of RNA-seq profiles also use model selection, as in Yeung et al. 2018 but the models go beyond simply periodic functions. We use the EEG data to predict sleep-wake processes. We use a model selection framework to distinguish between sleep-wake and cosine dynamics as well as interactions between the two. Surprisingly, we find that core clock genes are sensitive to sleep deprivation; their amplitudes become damped for at least 48 hours after sleep deprivation treatment. Overall, our results highlight surprising dynamics underlying how the circadian clock embeds into gene regulatory networks regulating sleep/wakefulness.

This chapter is a manuscript in preparation and should be kept confidential.

4.2 Contributions

Charlotte Hor conceived the project, performed the RNA-seq, ATAC-seq, and EEG experiments, with help from Yann Emmenegger and Jeffrey Hubbard. I developed the theoretical and computational framework for analyzing the temporal dataset. I analyzed the RNA-seq data, with help from Charlotte Hor and Maxime Jan. Maxime Jan analyzed the ATAC-seq data over time to identify genomic loci with dynamic ATAC-seq signal. I analyzed the ATAC-seq data to identify transcription factor binding motifs linked to gene expression dynamics.

Dynamic gene expression and regulatory outputs of circadian and sleep homeostat processes in the mouse cortex

Charlotte N. Hor^{*1}, Jake Yeung^{*2,3}, Yann Emmenegger¹, Maxime Jan^{1,3}, Jeffrey Hubbard¹, Felix Naef^{§2,3}, Paul Franken^{§1}

¹Centre for Integrative Genomics, University of Lausanne, Lausanne, Switzerland

²Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

³Swiss Institute of Bioinformatics, Lausanne, Switzerland

* These authors contributed equally to this work.

§ These authors contributed equally to this work.

Abstract

The timing and quality of sleep are thought to be regulated by the interaction of two processes. First, a sleep homeostat accumulates pressure for sleep during wake, and is relieved during sleep; second, the circadian timekeeping mechanism sets preferential times for sleep intervals. Although studies have implicated sleep-wake driven and circadian clock genes in the regulation of sleep, a systematic survey of how sleep and time-of-day govern gene expression dynamics and its underlying transcriptional regulation has not been established. Here, we sleep-deprived mice for 6 hours and tracked sleep-wake history, gene expression, and chromatin accessibility over more than 48 hours. Integrating the electroencephalography (EEG) data with RNA-seq and ATAC-seq, we developed a model selection framework to systematically infer gene expression dynamics that are driven by sleep-wake, time-of-day, or interactions of the two processes. From our model selection, we found that sleep-wake history, as measured by EEG, explained the largest fraction of variance, followed by rhythmic gene expression that was unaffected by sleep deprivation (SD). Remarkably, we found that the majority of circadian clock genes oscillated with damped amplitudes following SD, suggesting that circadian gene expression can adapt to perturbations. This damping was sustained for more than 48 hours after exposure. Sleep-wake driven dynamics in gene expression were accompanied by dynamics in chromatin accessibility sites. Open chromatin regions of immediate early genes were enriched for serum response factor (SRF) binding motifs. Furthermore, the predicted dynamics in SRF motif activity explained the upregulation of expression for many immediate early genes during SD. Our findings indicate that sleep-wake history and time-of-day interact to output diverse transcriptional dynamics in response to sleep deprivation.

Introduction

Sleep regulation is classically viewed as the interaction of sleep homeostatic process and circadian process (Paul Franken 2013). The sleep homeostat increases pressure for sleep during wake and releases pressure during sleep. This sleep need can be monitored by electroencephalogram (EEG)-derived variable delta power (Mongrain et al. 2010). Although the timing at which sleep preferentially occurs is modulated by the circadian process, this time-of-day modulation does not affect the need for sleep. Arrhythmic mice with lesioned suprachiasmatic nuclei (SCN) still maintain the quantitative relationship between sleep-wake distribution and EEG delta power during NREM sleep (Trachsel et al. 1992).

The circadian clock is encoded through negative feedback loops that generate robust oscillations even in constant conditions (Takahashi 2017). This ubiquitous oscillator is involved in a variety of physiological processes such as metabolism, blood pressure, wound healing, and locomotor activity (Lamia et al. 2011; Zuber et al. 2009; Hoyle et al. 2017; Mermet et al. 2018). Surprisingly, although lesioning the SCN of mice did not affect homeostatic sleep need, *Cry1/Cry2* double knockout mice, which are also arrhythmic, showed increased time in NREM sleep, suggesting the molecular clock also regulates sleep (Wisor et al. 2002). Thus, both the circadian and sleep-homeostatic processes both contribute to sleep timing and sleep-wake distribution.

Changes in gene expression can be dependent or independent of what time of day the sleep perturbation took place (Maret et al. 2007; T Curie et al. 2013), which is consistent with the two-process model of sleep homeostasis (Paul Franken 2013). Under this model, the homeostatic process (Process S) tracks the time spent awake and thus the sleep need, while the circadian process (also referred to as Process C) modulates the timing of sleep with regards to time of day. Therefore, under normal conditions where environmental changes occur periodically and predictably, the expression of genes displaying a nycthemeral (*i.e.*, rhythm with a period of 24-hours) rhythm can be regulated by either of these processes or both in combination.

Sleep deprivation (SD) experiments allow to uncouple/disentangle the sleep homeostat from circadian processes (P. Franken and Dijk 2009). In fact, SD beginning at the dark-to-light transition will cause mice to stay awake during a time when they normally sleep. Thus, sleep-wake driven genes will respond acutely to SD whereas dynamics that are robust to SD will be comparable to control mice. Studies comparing gene expression levels during SD with controls have identified many differentially expressed genes (Maret et al. 2007; Mongrain et al. 2010; Paul Franken et al. 2007; T Curie et al. 2013; Diessler et al. 2018). However, assessing the many possible dynamics requires comparing beyond time-matched control conditions. Sampling during and after SD and analyzing the entire time course can reveal rich dynamic processes. Thus, identifying the regulatory contributions from the sleep homeostat and circadian clock require considering gene expression dynamics along a sufficiently long time-course.

Here, we measured gene expression and chromatin accessibility in adult C57BL6/J mice before, during, and over 48 hours following one 6-hour session of total sleep deprivation (SD), as well as 7 days after the intervention. This time course can then be compared with a baseline control day. We also report a long-term electroencephalographic (EEG) recording of vigilance states and correlates of sleep need over the same time interval.

Our analysis integrates EEG, transcriptome, and chromatin accessibility data over 78 hours sampled at least every 6 hours. Analyzing the entire time series using both unsupervised and supervised methods, we find that the largest group of gene expression dynamics come from a sleep-wake driven process, which could be modeled from the EEG data. We also find unexpected diversity of gene expression dynamics. Notably, the (peak to trough) amplitudes of most core clock gene expressions, such as *Arntl*, *Nr1d1*, and *Nr1d2*, were prolongedly damped after SD, suggesting that SD can have long-term effects on gene expression dynamics in the cortex beyond the initial treatment. Our time course data also reveal interactions between the sleep homeostat and time-of-day processes. Combining both the sleep homeostat and time-of-day allows genes to respond with larger fold changes relative to baseline compared to genes that are only sleep-wake driven. These gene expression dynamics are accompanied by dynamics in chromatin accessibility. Finally, we find sleep-wake driven activity of serum response factor (SRF) underlies the gene expression dynamics of many

immediate early genes previously implicated in sleep response, such as *Arc*, *Egr2*, *Fos*, and *Junb*.

Materials and Methods

Animals

C57BL/6J male mice were purchased from Charles River France (Lyon, France) and allowed to habituate to our sleep study facility for 2-4 weeks prior to habituation to the experimental setting, and the experimentation. Animals were kept in accordance to the Swiss Animal Protection Act, and all experimental procedures were approved by the local veterinary authorities (Authorization nb. VD3037).

Surgery and EEG recording

The EEG cohort consisted of 6 male C57BL/6J mice 10-12 weeks at the time of SD. Surgical implantation of electrodes, EEG recording and data collection were performed according to our standard procedure (Mang and Franken 2012). EEG was recorded from 2 days prior to SD (which were averaged to constitute a 24-hour baseline) until 7 days after SD. Electrophysiological signals were captured at 2000Hz, transformed from analog to digital, and downsampled and stored at 200Hz (EMBLA A10 and Somnologica-3; Medcare Flaga; Thornton). Sleep and wake states were annotated according to established criteria based on the properties of the EEG and EMG signals (Mang and Franken 2012). To determine spectral composition, EEG signals (0 to 90 Hz) underwent a discrete Fourier transformation, using a window of 4-seconds (Hamming function), to determine power spectral density. Delta power (1-4Hz) was extracted for NREM epochs and averaged across the experiment. To counteract differences in absolute EEG power between individuals, power spectral density for each 0.25 Hz bin was expressed as a percentage of the 2 baseline days from ZT8-12, when NREM delta power is lowest. SD and recovery time points were compared to baseline by means of a t-test.

Sleep deprivation and tissue collection

Mice for tissue collection were divided into two experimental cohorts, sleep deprived (SD) and non-sleep deprived (controls, or Ctr). After a one-week habituation to the experimental setting, at the age of 11-12 weeks, the SD mice were sleep-deprived by gentle handling for 6 hours starting at light onset (Zeitgeber time ZT0-ZT6) as

described in (Mang and Franken 2012), and allowed to recover according to the tissue collection schedule. Mice were anesthetised with isoflurane prior to decapitation. Cortex was rapidly dissected and flash frozen in liquid nitrogen. NSD mice were sacrificed at ZT0, ZT3, ZT6, ZT12, ZT18 of the first day of experimentation (samples T0-T18), serving as a baseline day (BL). SD mice were sacrificed at the same time of day on Recovery Day 1 (R1, samples T24-T42), at ZT0, ZT6, ZT12, ZT18 on Recovery Day 2 (R2, samples T48-66), as well as ZT0 and ZT6 on the 3rd day (samples T72-78), see Figure 1A. Finally, two groups of mice were allowed to recover for 7 days after SD, before being sacrificed at ZT0 and ZT6 (samples T192-198). We refer to each time point in hours from the start of the baseline day (T0), with sleep deprivation occurring from T24 to T30. In the analysis, samples at T0 and T24 were assigned evenly across two different batch runs of ZT0 NSD. We collected 3-4 replicates per time point and condition, and 8 replicates of ZT0 NSD.

Tissue processing and sequencing library preparation

Frozen cortex of each individual was ground in liquid nitrogen and stored at -80°C until further use. Total RNA was extracted using the miRNeasy kit (Qiagen; Hilden, Germany) following the manufacturer's instructions.

RNA-seq libraries were prepared using 1000 ng of total RNA and the Illumina TruSeq Stranded mRNA reagents (Illumina; San Diego, CA, USA) on a Sciclone liquid handling robot (PerkinElmer; Waltham, MA, USA) using a PerkinElmer-developed automated script. Libraries were sequenced on the Illumina HiSeq 2500 sequencer, producing >36 million (median 55 million) single-end 100 bp reads.

ATAC-seq was performed with minor modifications from (Jason D Buenrostro et al. 2015). 100'000 nuclei were treated with 2.5 µl Tagment DNA enzyme (Nextera DNA Sample Preparation Kit, Illumina) in transposition buffer (10mM Tris Base, 5mM MgCl₂, 10% DMSO, pH 7.6, adapted from (Wang et al. 2013)) at 37°C for 30 minutes, followed by cleanup on a Qiagen Minelute column. Fragments >1kb in size were removed using AmpureXP beads (Beckman Coulter Life Sciences; Indianapolis, IN, USA) using 0.6X and 1X volumes. DNA fragments were subjected to 11 cycles of PCR amplification with Nextera index primers (Illumina) and NEBNext High Fidelity 2X PCR Master Mix (New England Biolabs; Ipswich, MA, USA). PCR reactions were cleaned up with one volume AmpureXP beads, quantified by Qubit (ThermoFisher Scientific; Waltham, MA, USA) and quality controlled by Fragment Analyzer

(Advanced Analytical Technologies; Ankeny, IA, USA). Libraries were sequenced on the Illumina HiSeq 2500 sequencer, producing >25 million (median 41 million) 50 bp paired-end reads per sample after removal of duplicate and mitochondrial sequences. Due to sequencing failure, the ATAC-seq data of two out of three replicates of T66 were excluded from the analysis.

Sequencing data analysis

Transcript abundance was quantified by *kallisto* version 0.43.0 (Bray et al. 2016) using the GRCm38 reference transcriptome (mm10) and the parameters `--single -l 100 -s 20 -b 100`. The abundances were further processed using *sleuth* version 0.29.0 (Pimentel et al. 2017), starting with merging transcript abundance into gene counts. We applied a detection cutoff of 5.5 on the mean gene counts across samples in the time series, yielding a set of 13'842 expressed genes which were used for further analysis. Batch effects were corrected by ComBat (R package *sva*_version 3.25.4 (Leek et al. 2012)).

ATAC-seq reads were aligned to the mouse genome (mm10) using bowtie2 (Langmead and Salzberg 2012) in paired-end mode, with the parameters recommended for open chromatin (`--very-sensitive --maxins 2000 --no-mixed --no-discordant`). Duplicate sequences were removed using *samtools* rmdup (Li et al. 2009).

Differential gene expression

To identify genes displaying a statistically significant effect over time, we used a likelihood ratio test implemented by *sleuth* version 0.29.0 (Pimentel et al. 2017), comparing a full model of time plus batch effects with a null model (no time effect) plus batch effects. We used FDR-adjusted p-value of 0.001 to identify 3461 statistically significant genes, which were used in the clustering and modeling analysis.

mRNA profile clusters

To uncover temporal patterns of mRNA abundance, we performed k-means clustering on statistically significant temporal gene expression (3461 genes). For a range of number of clusters, k , we calculated the within cluster variation as the sum of the Euclidean distance between data points and their assigned cluster centroids and

empirically chose $k=10$ as a balance between variance explained and generalizability of each cluster. Mean $-\log_{10}(p\text{-value})$ and effect sizes across genes in each cluster are represented as shaded rectangles above each plotted cluster, p -values are calculated from likelihood ratio test.

mRNA time course analysis

We used a model selection approach to classify temporal log mRNA abundance $m(t)$ of expressed genes into the scenarios described in Results. The models can be expressed as stated below. For models 3, 5 and 6, sleep-wake history was used to model the synthesis rate of mRNA as the S process (P Franken, Chollet, and Tafti 2001) using EEG data from $n=15$ C57BL/6J mice (this study and (Diessler et al. 2018)).

1: Flat model with constant μ and noise ϵ (flat)

$$m(t) = \mu + \epsilon$$

2: Sinusoidal, oscillatory model with 24 h periodic rhythmic parameters a and b

$$m(t) = \mu + a \cos(\omega t) + b \sin(\omega t) + \epsilon \quad (\text{Cosine})$$

$$\text{Angular frequency } \omega = 2\pi/24 \text{ h}^{-1}$$

3: Sleep-wake driven model (S)

$$\frac{dm(t)}{dt} = S(t, U, L, \tau_w, \tau_s) - \gamma_{mrna}(m(t)), \text{ where } S \text{ is defined recursively:}$$

$$S(t, U, L, \tau_w, \tau_s) = \begin{cases} U - (U - S(t-1)) \exp(-\Delta t/\tau_w) & \text{if awake} \\ L + (S(t-1) - L) \exp(-\Delta t/\tau_s) & \text{if sleep} \end{cases} \quad \text{with}$$

$$S(t_0) = S_0 : \text{initial value at time } t=0$$

Δt : mean period of continuous wake or sleep, defined by EEG data from 15 mice

U : asymptotic value for long periods of wake

L : asymptotic value for long periods of sleep

γ_{mrna} : inferred degradation rate of mRNA. Slow degradation rates damp fluctuations from EEG data, while fast degradation rates follow the fluctuations.

We solved the differential equation for $m(t)$ using the Euler method with a time step of 0.1 h.

We will call the solution of this differential equation $D(t, \vec{\theta}_{sleep})$ where $\vec{\theta}_{sleep}$ are the sleep parameters, U, L, τ_w, τ_s .

4: Sinusoidal model with change in amplitude (A).

$m(t) = \mu + C(t)(a \cos(\omega t) + b \sin(\omega t)) + \epsilon$, $C(t) = \begin{cases} 1 & \text{for } t \leq 33h \\ c & \text{for } t > 33h \end{cases}$, where $t=33h$ corresponds to 3 h after the end of sleep deprivation. Thus, in this model the amplitude is changed by a factor c after $t = 33 h$.

5: Sleep-wake and oscillatory model (S+A)

$m(t) = D(t, \vec{\theta}_{sleep}) + a \cos(\omega t) + b \sin(\omega t) + \epsilon$, where $D(t, \vec{\theta}_{sleep})$ is the solution to the differential equation in the sleep model.

6: Combined with change in amplitude model (S+A)

$m(t) = D(t, \vec{\theta}_{sleep}) + C(t)(a \cos(\omega t) + b \sin(\omega t)) + \epsilon$, where D and C are defined as above.

7: Model with one parameter for each sampled time point (this is the most complex model, termed generic).

$m(t) = \beta(t) + \epsilon$; $t \in \{0, 3, 6, 12, 18, 24, 27, 30, 36, 42, 48, 54, 60, 66, 72, 78\}$

For models that are nonlinear (models 3-6) with respect to the parameters, we fitted the model with optim in R using the L-BFGS-B method. Linear models (models 1, 2, and 7) were solved using the lm() function in R. The gene expression was fit in the log scale.

For each gene, we estimated the posterior probability of each model by first calculating the Bayesian Information Criterion (BIC) scores:

$$B_i = -2 * L_i + k_i \log(n)$$

where L is the log likelihood. A better fit will improve (decrease) the BIC, while a more complex model will penalize (increase) the BIC. Intuitively, an optimal model will fit the data while not using an excessive number of parameters. We assume the model

errors are independent and identically distributed following a Gaussian distribution with variance estimated from the fits:

$$\widehat{\sigma^2} = \frac{1}{n} \sum_i (m_i - \widehat{m}_i)^2$$

Exponentiating the BIC scores yields Schwarz weights w_i :

$$P(M_i|D) \approx w_i = \frac{\exp(-B_i/2)}{\sum_j \exp(-B_j/2)}$$

and we then assigned each gene to the model i corresponding to the largest w_i .

w_i assigns a probability to each model, and this probability measurement takes into account the number of parameters k in the model through the BIC score (i.e. complex models with large k are penalized by having a larger B , which would have smaller w).

For comparison with previously published sleep-related gene lists, we opted to remove close calls between models by selecting the genes with a weight w_i in any model of at least 0.7, which yielded a set of 9620 confidently assigned genes.

ATAC-seq peak detection:

ATAC-seq data files were processed before peak calling as follows. Alignment files were converted into bed files and tags were extracted using bedtools version 2.26.0. Each tag position was shifted from +4 base pairs on the positive strand and -5 base pairs on the negative strand to center tags on transposase binding events as suggested by (J D Buenrostro et al. 2013). The peak calling was performed on pooled tags for replicates using Macs2 version 2.1.1 (Zhang et al. 2008) [`--nomodel --shift -75 --extsize 150`], and peaks were filtered using a 0.05 FDR cutoff. Peak boundaries were merged between time points and conditions in order to build a common peak mapping reference among all samples, encompassing a total of 215'045 peaks. Finally, peak coverage was quantified using HTSeq version 0.6.1 for each sample using the

common mapping reference. A region was considered positive if it passed a $q=0.05$ FDR threshold vs. random noise. As a quality control, we probed whether genes within accessible regions were enriched in cortex/brain tissue. To this end, we used the Bgee database and topAnat (Komljenovic et al. 2016) to look for significant enrichment, and found that the top 20 enriched tissues were all nervous system structures (FDR p-value $< 10e-8$). The proximity in the PCA of the two technical replicates at T24 attests the reproducibility of ATAC-seq over different batches of sequencing.

ATAC-seq differential accessibility analysis:

To identify peaks with differential accessibility, we first normalized count data using a TMM normalization, applied a 10 read count threshold, and used a likelihood ratio test implemented in edgeR. We compared chromatin accessibility of sleep deprivation samples (T24-198) with the corresponding ZT during baseline (T0-18, see Figure 1A). Thus, for differential accessibility at ZT3, we compare T27 with T3, at ZT6, T30 and T6, etc. P-values were adjusted using the Benjamini & Hochberg (FDR) method (Benjamini and Hochberg 1995). K-means clustering was performed using $k=10$ clusters.

Genomic distribution of ATAC-seq peaks:

The annotation of the detected ATAC-seq peaks was performed using PAVIS with the *Ensembl_GRCm38/mm10 all genes* reference annotation (Huang et al. 2013).

Peak-to-gene expression association:

To associate gene expression dynamics with chromatin accessibility dynamics, we used a pearson correlation coefficient across the samples and limited the possible association test within previously defined topological interaction domains (TADs), which were computed from cortex tissue by (Dixon et al. 2012). TAD boundaries position were originally detected using the mm9 reference genome, so we converted them to mm10 using CrossMap 0.2.6 (Zhao et al. 2014). For association statistics, we used a strategy similar to that implemented within FastQTL (Ongen et al. 2016). Each pair consisting of a peak and a gene within the same TAD were associated using the pearson correlation coefficient. For each gene, only the top correlation to a peak was kept. To control for multiple associations within a TAD and adjust nominal p-values, we used 1000 permutations per gene and modelled the null distribution fitting a beta distribution. The

parameters were estimated using a maximum likelihood approach (R/MASS::fitdistr). Finally, a genome-wide p-value adjustment was computed using a qvalue procedure (R/qvalue). Of the 10894 genes mapping within a TAD, 3386 were associated to an ATAC-seq peak within the same TAD.

Predictions of transcription factor binding site (TFBS) activity in promoters

To predict the activity of TFBSs, we used position weight matrices of 179 mouse transcription factors (TFs) defined by SwissRegulon on mm9 (<http://swissregulon.unibas.ch/fcgi/sr/downloads>). For each of the 179 position weight matrices, we scanned genomic regions (500 bp windows) 15 kb from promoters using MotEvo (Arnold et al. 2012). We filtered regions containing ATAC-seq counts greater than 0.1 RPM (reads per million mapped reads). The sitecount matrix of each motif was scaled across genes so that ranges in sitecounts were comparable across motifs. We inferred TF activity using the TF binding site predictions and the temporal mRNA abundance. To infer the TF activity, we applied a penalized regression model as previously described (Balwierz et al. 2014; Yeung et al. 2017) using an L_2 norm penalty for regularization (ridge regression). Prior to the regression, we mean-centered the input matrix of temporal mRNA abundances, standardized the columns of the sitecount matrix (each motif across genes), and excluded genes that were assigned to the flat model.

Motif search

We used MEME-CHIP of the MEME suite (www.meme-suite.org) (Bailey et al. 2009) to search for motifs within differentially active regions.

Results

Study design

We subjected mice to 6 hours of total sleep deprivation (SD), starting at light onset of the 12-hour light-dark cycle and collected cerebral cortex during SD and over the following 48 hours. We used Zeitgeber time (ZT) to indicate time since light onset. ZT0 indicates light on while ZT12 indicates lights on. Mice are nocturnal and tend to be active during dark (ZT12 to ZT0/ZT24) and sleep during light (ZT0 to ZT12). A second group of mice was allowed to sleep *ad libitum*, undisturbed in another room, and

sacrificed at the same times of day as SD mice, serving as non-sleep deprived controls or baseline day (Figure 1A). We profiled gene expression by RNA sequencing of polyadenylated transcripts (RNAseq) and accessible portions of the genome by ATAC-seq (J D Buenrostro et al. 2013). To characterize the behavioural response and recovery after SD, we recorded EEG data in a separate group of 6 mice over 9 days, including 2 days of baseline prior to SD, the day of SD and 6 days of recovery.

Behavioral response and recovery after SD

To assess the long-term dynamics of the sleep phenotype, we quantified non-rapid eye movement sleep (NREMS) in 6-hour intervals, as well as hourly EEG delta power, a well-known correlate of sleep need (P Franken, Chollet, and Tafti 2001). We observed the typical distribution of sleep over 24 hours in baseline, with mice spending most of the time of the light period asleep, while being predominantly awake during the dark period (Figure 1B, bottom). Delta power in NREMS (Figure 1B, top) followed the amount of sleep pressure, high after spontaneous waking in the dark phase and low during the light phase. We also observed the well-known effects of sleep deprivation. NREMS rebounded during the 12 hours following SD and an increase in delta power immediately after the end of SD. We found that values stopped differing from baseline already during the second half of the dark phase after SD for NREMS, and as of Recovery day 2 for delta power (Figure 1B). REM sleep was affected in the same manner as NREMS (not shown).

Principal component analysis reveals sleep-wake state as main driver of transcriptome dynamics

We asked whether the fast recovery of the phenotype in the EEG data could be observed at the gene expression level, and whether novel dynamic patterns beyond the EEG dynamics could be observed. We therefore analyzed the temporal dynamics of transcriptomes.

We first examined the detected fraction of our transcriptome dataset (13'842 genes) using principal component analysis (PCA, Figure 1C). We observed that samples formed three groups along the first principal component (PC1) axis. The right-most group clustered time points during the light phase where mice generally spend more time asleep, while the middle group represents time points taken during the dark phase of the LD cycle where mice are predominantly awake. Surprisingly, ZT3 and ZT6

during SD (T27 and T30) separate far from their time-matched baseline (T3 and T6) towards the awake group, suggesting that PC1 dynamics follow sleep-wake history rather than Zeitgeber time.

To illustrate the sleep-wake-driven dynamics underlying PC1, we overlaid PC1 with the average EEG signal over time (Figure 1D). This integration showed that sleep-wake history explains PC1. During periods of wake, PC1 increases, while during periods of sleep, PC1 decreases. Importantly, the six-hour sleep deprivation period, PC1 reaches its maximum. PC1 thus reflects the amount of sleep prior to sample collection, and illustrates the impact of sleep-wake distribution on mRNA expression.

No significant differences in gene expression 7 days after exposure

We next examined the temporal dynamics of gene expression. We noted that expression levels on Day 7 (T192 and T198) were not significantly different from baseline at T0 and T6 according to our differential gene expression analysis (FDR adjusted p val > 0.05 , Supplemental Figure 1A). We therefore focused our analysis and modeling to the baseline day 0 and recovery days 1 and 2 (T0 to T78).

Clustering analysis reveals dynamics that are driven by sleep-wake history and also robust to sleep-wake history

To uncover general temporal patterns in our data, we performed an exploratory analysis using k-means clustering. We clustered temporal expression of 3461 genes identified as displaying statistically significant temporal variation from T0 to T78 (FDR-adjusted p -value < 0.001 , Methods).

We observed distinct dynamics over 78 hours (Figure 2). Genes in clusters 1-6 displayed an immediate response to SD, many of which showed statistically significant differences at T27 and T30 (p -values across genes in each cluster are summarized for in Figure 2). In cluster 7, the response is longer, many genes showing differences up to T36. Cluster 8 showed a delayed response; the largest differences occur 18 hours after the end of SD.

Clusters 9 and 10 showed prominent 24-hour rhythm and, on average, the genes are not perturbed by SD (mean p values across genes > 0.24). However, comparing individual

genes at days 1 and 2 against baseline day 0 found damped rhythms after SD such as *Arntl*, *Fabp7*, *Obecn*, and *Lfng*. (Supplemental Figure 2).

Concerning the recovery properties, clusters 1-4 showed a fast recovery, at T36 the expression on average has returned to baseline. By contrast, genes in clusters 5 and 6 reverted more slowly on average, returning to baseline at T42. Genes in cluster 7 also displayed slow recovery, returning back to baseline levels 12 hours after peak response. Finally, in cluster 8 we observed a pattern compatible with a "rebound", in the form of an exacerbated increase at T48 following the initial downregulation at T36. By T54, cluster 8 has returned to baseline.

Generally, we observed that SD can have three effects on dynamics relative to baseline. First, SD can up or downregulate expression of genes that are normally down or upregulated during ZT0-6 in baseline, respectively. This effect explains a significant variance (20%) of the transcriptome, as shown in our PCA analysis (cluster 1-6). Second, this down or upregulation can extend to ZT12 (cluster 7 and 8), suggesting slower dynamics or downstream effects from the first effect. Third, SD can dampen 24 h periodic oscillations in mRNA abundance.

Overview of identified models

Explicitly modeling the temporal dynamics of mRNA profiles can offer advantages over unsupervised methods. From our downstream analysis of clusters, we devised 6 models to explain the log gene expression dynamics of 13842 detected genes (Supplemental Figure 3): (1) constant flat model; (2) sleep-wake history modeled from the EEG data (S); (3) cosine dynamics with 24-hour rhythm (C); (4) cosine with amplitude change after SD (A); (5) sleep-wake + cosine (S+C); (6) sleep-wake + cosine with amplitude change (S+A). The parameters of the model can give useful insights to the underlying dynamics, competing models can be systematically compared, and explicit hypotheses can be tested. For example, using a sleep-wake model will unify dynamics that appear in separate clusters (e.g., Cluster 1 and 4 both may have genes that are driven by sleep-wake). The C and A models can separate dynamics that appear in the same cluster (e.g., Cluster 9 contains both SD-robust and SD-sensitive dynamics). To select competing models, we used the Bayesian Information Criterion (BIC) to balance model fit and model complexity (Methods). Of note, we also included a generic

model where the gene expression is modeled as the mean expression across replicates at each time point to assess the possibility of more complex dynamics not explained by any of the models. We found that the BIC weight w was always lower for the generic model than the other 6 models, meaning no genes were selected in the generic model, and therefore not included in subsequent analyses.

For each gene, we fit temporal gene expression to 6 models (examples shown in Figure 3A). We summarized the genes assigned to each model genome-wide (Figure 3B). We found that, out of all temporal models (*i.e.*, excluding flat model, Figure 3C), sleep-wake driven model had the most genes assigned to it (Model S Figure 3D). This large number of genes is consistent with our PCA and clustering analysis, where sleep-wake driven was also observed in the first component in PCA (Figure 1C) and in clusters 1-6 (Figure 2) in clustering analysis.

The cosine model contained the second highest number of genes (Model C, Figure 3E), which is consistent with the clustering analysis (cluster 9 Figure 2). We found *Hif3a*, hypoxia-inducible factor 3, as a prominent example of robust oscillations despite SD. This robust model suggests a significant fraction of dynamic genes is unaffected by SD.

Interestingly, the cosine with amplitude change model had the third most genes, suggesting SD can affect dynamics beyond the 3-hour SD (Model A Figure 3B). For *Nr1d1* (Figure 3A), the best model based on BIC (BIC weight w) was cosine with amplitude change, suggesting SD damped its amplitude (Figure 3A bold line). SD also damped the amplitude of *Arntl*, another core clock gene (Takahashi 2017). Surprisingly, we found that SD damped the amplitudes of many clock genes, suggesting that the circadian clock is adaptive to sleep-wake history.

Finally, our approach identified more complex models involving both a sleep-wake plus cosine effect (Model S+C Figure 3G) or a sleep-wake plus cosine with amplitude change (Model S+A Figure 3H). These novel dynamics suggest a significant number of genes have expression dynamics that integrate sleep-wake history with time-of-day.

Model parameters predict novel transcriptional dynamics

We found the overwhelming majority of known sleep-wake driven genes to be correctly assigned by our model selection method. 58 out of 61 genes previously described as sleep-wake driven (Mongrain et al. 2010) were fully or partially explained by the EEG data (Model S, S+C, S+A, Supplemental Figure 4A). Similarly, we looked at a list of 92 genes previously described as affected by SD (Maret et al. 2007), and found that 77/92 of the genes were also inferred to be affected by SD in our model select (Supplemental Figure 4B). These corroborations from independent datasets suggest that our model selection recapitulated known SD-affected genes.

Analyzing sleep-wake driven genes, we found that SD can upregulate or downregulate genes (Figure 4A) in the sleep-wake model. The largest fold changes ($\log_2 \text{FC} > 2.5$, or nearly six fold) were exclusively upregulated genes. The sleep model also predicted an mRNA half-life which buffers fluctuations from the EEG data and calculates the mRNA abundance (Figure 4A) (Methods). We found that large fold changes corresponded to short half-lives, consistent with fast dynamics (REF here? E.g. Wang 2018 or/and Zeisel MSB?). Finally, the fitted time constants describing the S process corresponding to wake tended to be longer than the one for sleep, as expected from mice spending more time awake than sleep (Figure 4B). Overall, the inferred parameters from sleep-wake driven genes were consistent with fast dynamics and relatively large fold changes, most of them upregulated during wake.

In contrast to the sleep-wake model, the cosine dynamics are robust to SD over 78 hours. Analysis of the fold change and peak times from the model found that large amplitude oscillations tended to occur near the dark-light or light-dark transition (Figure 4C). Notable genes such as *Sgk1*, glucocorticoid regulated kinase, and *Cldn5*, principal tight junction protein in blood-brain barrier, showed 24-hour oscillations in gene expression.

We found that SD can alter amplitudes in rhythms, suggesting effects that last beyond the initial 3 hour SD. Notable examples with damped amplitudes include circadian clock genes *Nr1d1*, *Nr1d2*, *Arntl*, and *Per3* (Figure 4D, Figure 4E). This model contained also genes with increased amplitudes after sleep deprivation such as *ErbB3*, *Evalb*, *Zfp473*, and *Akr1cl*. In sum, SD can have long-term effects lasting at least 48 hours after sleep deprivation.

Finally, we found interactions between the cosine and sleep model (additive in the log scale). These combined models (S+C and S+A) in general allow genes to be upregulated or downregulated with significant fold changes compared to sleep model (Figure 4F left, Supplemental Figure 5A,B). This dynamic process can be coupled with decreased fold changes after SD (S+A model). Examples here include *Per2*, a gene shown to have complex interactions between the two processes (Thomas Curie et al. 2013), as well as clock output genes *Dbp*, *Nfil3*, and *Bhlhe41* (Supplemental Figure 5C-F). The large proportion of core clock and clock output genes found to show interactions between sleep-wake history and time-of-day highlight how these two fundamental processes interact to generate complex temporal dynamics. On day 2, aggregate analysis for models S, S+C, and S+A, showed that the amplitudes are comparable between day 0 (Figure 4E right).

Genome-wide analysis of ATAC-seq shows sleep-wake driven dynamics in chromatin accessibility

We next asked which regulatory elements are involved in the response to SD, a compartment that has hitherto not been explored in this context. We identified a total of 215'045 ATAC-positive regions (read counts > 10 in at least one time point). Principal component analysis of accessibility levels in these regions separated samples by sleep-wake history (Figure 5A), consistent with the RNA-seq analysis (Figure 1C).

Accessibility peaks from all time points and conditions were mainly located in intronic or intergenic regions (Supplemental Figure 6A). When considering only peaks that were differentially active at specific time points (Supplemental Figure 6B-D), the proportion of intergenic regions was increased, suggesting that SD influences accessibility of distal rather than genic or proximal elements. Genes associated with DAS were enriched among models involving sleep-wake dynamics (Supplemental Figure 6E-J) compared to all peaks at all ZT (p-values < 2e-10, chi-square test) except ZT12 (p-value = 0.48).

To probe the general dynamics of chromatin accessibility, we performed a K-means clustering analysis (Figure 5B) and found sleep-wake driven dynamics. The strongest differential signal relative to baseline occurred during the 6h SD (T27 vs T3 and T30 vs T6). We examined the differential accessibility at each time point after SD compared

to its corresponding baseline day time. 1542 peaks were differentially accessible in SD at ZT3 (T27, after 3h SD) and 1906 at ZT6 (T30, end of 6h SD) and 678 at ZT12 (T36, after 6h of recovery, (Figure 5C). Chromatin accessibility signal tended to increase during sleep deprivation (ZT3 and ZT6). Overall, dynamics in chromatin accessibility was most pronounced during sleep deprivation, and returned to baseline by 12 hours after SD.

Chromatin accessibility correlate with gene expression dynamics

We correlated ATAC-seq peaks to genes by calculating the Pearson correlation across samples (Methods). We restricted the possible peaks associated with each gene within topologically associated domains (TADs) defined from Hi-C data generated from the mouse cortex (Dixon et al. 2012). This analysis identified chromatin interactions that correlated with gene expression dynamics (Figure 6A-F). For example, we found *Hif3a*, a gene relatively robust to SD, with a correlated ATAC-seq peak 2 kb from the promoter (Figure 6C). The promoter of *Ciart*, a gene that has been reported to be directly regulated by the clock (Anafi et al. 2014), had peak with correlated dynamics, which appeared to have damped amplitudes on the first day after SD and a partial amplitude recovery in the second day. We found that ATAC-seq dynamics in mouse cortex can show about a fold change of about 1.5, comparable with changes in the RNA-seq data. In sum, gene expression dynamics can be accompanied by dynamics in chromatin accessibility.

Activity of Serum Response Factor underlies immediate early gene expression

To infer which transcription factors drive the gene expression dynamics, we performed a motif activity analysis using the MARA (Balwierz et al. 2014). Briefly, we used the SwissRegulon database (179 mouse TF motifs) to search for TF motifs within ATAC-positive 500 bp bins within a distance of 15 kb of promoter of genes. We included genes that were assigned to any dynamic model (i.e., S, C, A, S+C, S+A). Using the sitecount matrix as covariates and the log gene expression as observed data, MARA uses a ridge regression framework to infer the temporal activity of each motif that could explain the gene expression data.

We found the SRF motif to be the most statistically significant candidate underlying the temporal gene expression dataset (Figure 7A). The inferred temporal activity of

SRF increased during sleep deprivation and rapidly returned to baseline, suggesting that its activity could be predicted by the sleep-wake history of the mouse (Figure 7B). Indeed, our model selection method on the SRF motif activity confirmed a sleep-wake-driven response. Furthermore, analysis of TF motifs near (<5 kb) promoters of sleep-wake-driven genes, such as *Egr2*, *Fos*, *Egr1*, and *Arc*, predicted SRF binding sites (Figure 7C). These SRF binding sites corresponded to open chromatin regions in the cortex. SRF ChIP-seq data in fibroblasts (C Esnault et al. 2014) showed binding of SRF to the SRF motifs present in those genes (Figure 7C, Supplemental Figure 7A). The ATAC-seq signal overlapping the SRF motif and ChIP-seq data is flat over time (Figure 7D, Supplemental Figure 7B), consistent with the regulation of many SRF target genes requiring cofactors (Cyril Esnault et al. 2017). SRF has previously been reported to regulate wake-dependent immediate early gene expression (Ramanan et al. 2005), consistent with SRF activity increasing during SD. In sum, our analysis suggests that SRF may underlie the immediate early response of many sleep-wake-driven transcripts.

Discussion

We characterized the dynamics of transcriptome and regulatory elements over time before, during and after a 6 hours of sleep deprivation. By integrating EEG data into the analysis, we developed a model selection approach to systematically identify modules of gene expression with distinct dynamics. This framework separated genes that could be explained by the EEG data (S), cosine dynamics (C), cosine with amplitude change (A), as well as interactions between sleep-wake and time-of-day (S+C and S+A). We attributed the largest gene expression variance to S dynamics, followed by C, A, S+C, and S+A.

The EEG data explained sleep-wake gene expression dynamics. These dynamics often show large fold changes in baseline. During SD, gene expression responds divergently relative to baseline and rapidly recover to baseline within 12 hours after SD. These dynamics reflect the sleep-wake behavior as determined by EEG. Remarkably, the first principal component could be explained by the EEG data, suggesting that sleep-wake behavior explains the largest variance in our dataset. We inferred SRF as a potential transcriptional regulator underlying dynamics of immediate early genes, which are upregulated during SD.

We also found that many SRF target genes had accessibility that was open but static, consistent with the role of cofactors in SRF target gene regulation (Cyril Esnault et al. 2017). Overall, we found that the largest fold changes (> 2.5) due to SD tended to come from SRF target genes (e.g., *Arc*, *Egr2*, *Fos*, *Egr1*, *Junb*, *Nr4a1*, *Homer1*), suggesting that SRF may play a role in sleep regulation. Our integrated analysis also found widespread chromatin accessibility dynamics that coupled with transcriptome, suggesting that sleep-wake history can alter transcription factor binding and mediate gene expression dynamics.

In a significant fraction of genes (734 genes), SD had a prolonged effect (>48 hours after SD) on mRNA levels. Remarkably, clock genes These long-term effects may eventually recover, but our simple cosine model with amplitude change after SD adequately explained expression of 734 genes, suggesting that these genes did not fully recover back to the baseline within 48 hours. Among the 734 genes are many clock genes, such as *Nr1d1*, *Nr1d2*, *Cry2*, and *Arntl*. Our results demonstrate that SD can disrupt oscillations in circadian clock expression by damping amplitudes, and this damping can be sustained long after exposure.

Finally, the S+C and S+A models combined EEG data with time-of-day highlight the interaction between sleep-wake history, sleep deprivation, and time-of-day. S+C allowed baseline fold changes to be modest and increase substantially during sleep deprivation.

The S+A model incorporated an amplitude change, often damping, following sleep deprivation. Many clock output genes such as *Dbp*, *Tef*, *Nfil3*, *Bhlhe41* showed damped oscillations after SD. This model also included complex interactions between the two processes, exemplified by *Per2* (Thomas Curie et al. 2013). Overall, we found the dynamics of most clock and clock output genes to be affected by sleep deprivation, many of them through damping of amplitudes. These damped amplitudes persist long after SD. Further studies will determine exactly how this disruption effect on other parts of physiology, such as metabolism, or alter response in subsequent environmental exposures.

Acknowledgments

We thank Shanaz Diessler, Marieke Hoekstra, Konstantinos Kompotis, Dessislava Petrova and the Lausanne Genomics Technologies Facility for technical assistance. J.Y. benefits from the Natural Sciences and Engineering Research Council of Canada Postgraduate Studies Doctoral scholarship. Work in the F.N. laboratory was supported by Swiss National Science Foundation Grant [310030_173079](#) and the École Polytechnique Fédérale de Lausanne.

Bibliography

- Anafi, Ron C., Yool Lee, Trey K. Sato, Anand Venkataraman, Chidambaram Ramanathan, Ibrahim H. Kavakli, Michael E. Hughes, Julie E. Baggs, Jacqueline Growe, Andrew C. Liu, Junhyong Kim, John B. Hogenesch, JS Takahashi, H-K Hong, CH Ko, EL McDearmon, J Bass, S Panda, JB Hogenesch, et al. 2014. "Machine Learning Helps Identify CHRONO as a Circadian Clock Component." Edited by Ueli Schibler. *PLoS Biology* 12 (4): e1001840. <https://doi.org/10.1371/journal.pbio.1001840>.
- Arnold, Phil, Ionas Erb, Mikhail Pachkov, Nacho Molina, and Erik van Nimwegen. 2012. "MotEvo: Integrated Bayesian Probabilistic Methods for Inferring Regulatory Sites and Motifs on Multiple Alignments of DNA Sequences." *Bioinformatics (Oxford, England)* 28 (4): 487–94. <https://doi.org/10.1093/bioinformatics/btr695>.
- Bailey, T L, M Boden, F A Buske, M Frith, C E Grant, L Clementi, J Ren, W W Li, and W S Noble. 2009. "MEME SUITE: Tools for Motif Discovery and Searching." *Nucleic Acids Res* 37 (Web Server issue): W202-8. <https://doi.org/10.1093/nar/gkp335>.
- Balwierz, Piotr J, Mikhail Pachkov, Phil Arnold, Andreas J Gruber, Mihaela Zavolan, and Erik van Nimwegen. 2014. "ISMARA: Automated Modeling of Genomic Signals as a Democracy of Regulatory Motifs." *Genome Research* 24 (5): 869–84. <https://doi.org/10.1101/gr.169508.113>.
- Benjamini, Y, and Y Hochberg. 1995. "Controlling the False Discovery Rate - a

- Practical and Powerful Approach to Multiple Testing.” *J Roy Stat Soc B Met* 57 (1): 289–300.
- Bray, Nicolas L, Harold Pimentel, Páll Melsted, and Lior Pachter. 2016. “Near-Optimal Probabilistic RNA-Seq Quantification.” *Nature Biotechnology* 34 (5): 525–27. <https://doi.org/10.1038/nbt.3519>.
- Buenrostro, J D, P G Giresi, L C Zaba, H Y Chang, and W J Greenleaf. 2013. “Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position.” *Nat Methods* 10 (12): 1213–18. <https://doi.org/10.1038/nmeth.2688>.
- Buenrostro, Jason D, Beijing Wu, Howard Y Chang, and William J Greenleaf. 2015. “ATAC-Seq: A Method for Assaying Chromatin Accessibility Genome-Wide.” In *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc. <https://doi.org/10.1002/0471142727.mb2129s109>.
- Curie, T, V Mongrain, S Dorsaz, G M Mang, Y Emmenegger, and P Franken. 2013. “Homeostatic and Circadian Contribution to EEG and Molecular State Variables of Sleep Regulation.” *Sleep* 36 (3): 311–23. <https://doi.org/10.5665/sleep.2440>.
- Curie, Thomas, Valérie Mongrain, Stéphane Dorsaz, Géraldine M Mang, Yann Emmenegger, and Paul Franken. 2013. “Homeostatic and Circadian Contribution to EEG and Molecular State Variables of Sleep Regulation.” *Sleep* 36 (3): 311–23. <https://doi.org/10.5665/sleep.2440>.
- Diessler, S, M Jan, Y Emmenegger, N Guex, B Middleton, D J Skene, M Ibberson, F Burdet, L Gotz, M Pagni, M Sankar, R Liechti, C N Hor, I Xenarios, and P Franken. 2018. “A Systems Genetics Resource and Analysis of Sleep Regulation in the Mouse.” *PLoS Biol* 16 (8): e2005750. <https://doi.org/10.1371/journal.pbio.2005750>.
- Dixon, Jesse R, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S Liu, and Bing Ren. 2012. “Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions.” *Nature* 485 (7398): 376–80. <https://doi.org/10.1038/nature11082>.
- Esnault, C, A Stewart, F Gualdrini, P East, S Horswell, N Matthews, and R Treisman. 2014. “Rho-Actin Signaling to the MRTF Coactivators Dominates

- the Immediate Transcriptional Response to Serum in Fibroblasts." *Genes Dev* 28 (9): 943–58. <https://doi.org/10.1101/gad.239327.114>.
- Esnault, Cyril, Francesco Gualdrini, Stuart Horswell, Gavin Kelly, Aengus Stewart, Phil East, Nik Matthews, and Richard Treisman. 2017. "ERK-Induced Activation of TCF Family of SRF Cofactors Initiates a Chromatin Modification Cascade Associated with Transcription." *Molecular Cell* 65 (6): 1081–1095.e5. <https://doi.org/10.1016/J.MOLCEL.2017.02.005>.
- Franken, P., and D.-J. Dijk. 2009. "Circadian Clock Genes and Sleep Homeostasis." *European Journal of Neuroscience* 29 (9): 1820–29. <https://doi.org/10.1111/j.1460-9568.2009.06723.x>.
- Franken, P, D Chollet, and M Tafti. 2001. "The Homeostatic Regulation of Sleep Need Is under Genetic Control." *J Neurosci* 21 (8): 2610–21. <http://www.ncbi.nlm.nih.gov/pubmed/11306614>.
- Franken, Paul. 2013. "A Role for Clock Genes in Sleep Homeostasis." *Current Opinion in Neurobiology* 23 (5): 864–72. <https://doi.org/10.1016/j.conb.2013.05.002>.
- Franken, Paul, Ryan Thomason, H Craig Heller, and Bruce F O'Hara. 2007. "A Non-Circadian Role for Clock-Genes in Sleep Homeostasis:A Strain Comparison." *BMC Neuroscience* 8 (1): 87. <https://doi.org/10.1186/1471-2202-8-87>.
- Hoyle, Nathaniel P, Estere Seinkmane, Marrit Putker, Kevin A Feeney, Toke P Krogager, Johanna E Chesham, Liam K Bray, Justyn M Thomas, Ken Dunn, John Blaikley, and John S O'Neill. 2017. "Circadian Actin Dynamics Drive Rhythmic Fibroblast Mobilization during Wound Healing." *Science Translational Medicine* 9 (415): eaal2774. <https://doi.org/10.1126/scitranslmed.aal2774>.
- Huang, W, R Loganantharaj, B Schroeder, D Fargo, and L Li. 2013. "PAVIS: A Tool for Peak Annotation and Visualization." *Bioinformatics* 29 (23): 3097–99. <https://doi.org/10.1093/bioinformatics/btt520>.
- Komljenovic, A, J Roux, M Robinson-Rechavi, and F B Bastian. 2016. "BgeeDB, an R Package for Retrieval of Curated Expression Datasets and for Gene List Expression Localization Enrichment Tests." *F1000Research* 5: 2748. <https://doi.org/10.12688/f1000research.9973.1>.

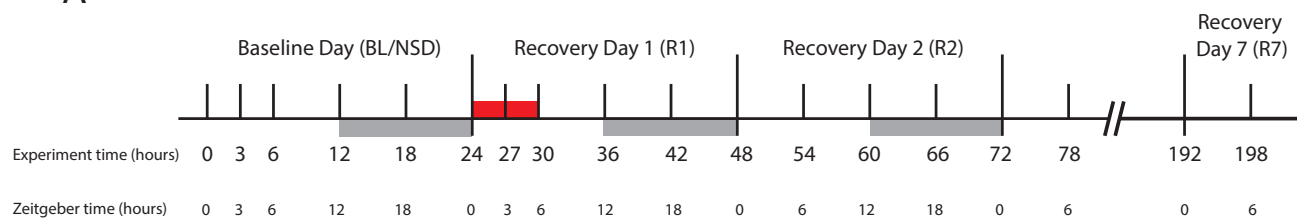
- Lamia, Katja A., Stephanie J. Papp, Ruth T. Yu, Grant D. Barish, N. Henriette Uhlenhaut, Johan W. Jonker, Michael Downes, and Ronald M. Evans. 2011. "Cryptochromes Mediate Rhythmic Repression of the Glucocorticoid Receptor." *Nature* 480 (7378): 552–56. <https://doi.org/10.1038/nature10700>.
- Langmead, B, and S L Salzberg. 2012. "Fast Gapped-Read Alignment with Bowtie 2." *Nat Methods* 9 (4): 357–59. <https://doi.org/10.1038/nmeth.1923>.
- Leek, Jeffrey T., W. Evan Johnson, Hilary S. Parker, Andrew E. Jaffe, and John D. Storey. 2012. "The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments." *Bioinformatics* 28 (6): 882–83. <https://doi.org/10.1093/bioinformatics/bts034>.
- Li, H, B Handsaker, A Wysoker, T Fennell, J Ruan, N Homer, G Marth, G Abecasis, R Durbin, and Subgroup Genome Project Data Processing. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79. <https://doi.org/10.1093/bioinformatics/btp352>.
- Mang, G M, and P Franken. 2012. "Sleep and EEG Phenotyping in Mice." *Current Protocols in Mouse Biology* 2: 55–74.
- Maret, Stéphanie, Stéphane Dorsaz, Laure Gurcel, Sylvain Pradervand, Brice Petit, Corinne Pfister, Otto Hagenbuchle, Bruce F O'Hara, Paul Franken, and Mehdi Tafti. 2007. "Homer1a Is a Core Brain Molecular Correlate of Sleep Loss." *Proceedings of the National Academy of Sciences of the United States of America* 104 (50): 20090–95. <https://doi.org/10.1073/pnas.0710131104>.
- Mermet, Jérôme, Jake Yeung, Clémence Hurni, Daniel Mauvoisin, Kyle Gustafson, Céline Jouffe, Damien Nicolas, Yann Emmenegger, Cédric Gobet, Paul Franken, Frédéric Gachon, and Félix Naef. 2018. "Clock-Dependent Chromatin Topology Modulates Circadian Transcription and Behavior." *Genes & Development*, March. <https://doi.org/10.1101/gad.312397.118>.
- Mongrain, V, S A Hernandez, S Pradervand, S Dorsaz, T Curie, G Hagiwara, P Gip, H C Heller, and P Franken. 2010. "Separating the Contribution of Glucocorticoids and Wakefulness to the Molecular and Electrophysiological Correlates of Sleep Homeostasis." *Sleep* 33 (9): 1147–57. <http://www.ncbi.nlm.nih.gov/pubmed/20857860>.
- Ongen, H, A Buil, A A Brown, E T Dermitzakis, and O Delaneau. 2016. "Fast and

- Efficient QTL Mapper for Thousands of Molecular Phenotypes.” *Bioinformatics* 32 (10): 1479–85.
<https://doi.org/10.1093/bioinformatics/btv722>.
- Pimentel, H, N L Bray, S Puente, P Melsted, and L Pachter. 2017. “Differential Analysis of RNA-Seq Incorporating Quantification Uncertainty.” *Nat Methods* 14 (7): 687–90. <https://doi.org/10.1038/nmeth.4324>.
- Ramanan, Narendrakumar, Ying Shen, Sarah Sarsfield, Thomas Lemberger, Günther Schütz, David J Linden, and David D Ginty. 2005. “SRF Mediates Activity-Induced Gene Expression and Synaptic Plasticity but Not Neuronal Viability.” *Nature Neuroscience* 8 (6): 759–67.
<https://doi.org/10.1038/nn1462>.
- Takahashi, Joseph S. 2017. “Transcriptional Architecture of the Mammalian Circadian Clock.” *Nature Reviews Genetics* 18 (3): 164–79.
<https://doi.org/10.1038/nrg.2016.150>.
- Trachsel, Lorenz, Dale M. Edgar, Wesley F. Seidel, and H. Craig Heller. 1992. “Sleep Homeostasis in Suprachiasmatic Nuclei-Lesioned Rats: Effects of Sleep Deprivation and Triazolam Administration.” *Brain Research* 589 (2): 253–61. [https://doi.org/10.1016/0006-8993\(92\)91284-L](https://doi.org/10.1016/0006-8993(92)91284-L).
- Wang, Q, L Gu, A Adey, B Radlwimmer, W Wang, V Hovestadt, M Bahr, S Wolf, J Shendure, R Eils, C Plass, and D Weichenhan. 2013. “Tagmentation-Based Whole-Genome Bisulfite Sequencing.” *Nat Protoc* 8 (10): 2022–32.
<https://doi.org/10.1038/nprot.2013.118>.
- Wisor, Jonathan P, Bruce F O’Hara, Akira Terao, Chris P Selby, Thomas S Kilduff, Aziz Sancar, Dale M Edgar, and Paul Franken. 2002. “A Role for Cryptochromes in Sleep Regulation.” *BMC Neuroscience* 3 (1): 20.
<https://doi.org/10.1186/1471-2202-3-20>.
- Yeung, J, J Mermet, C Jouffe, J Marquis, A Charpagne, F Gachon, and F Naef. 2017. “Transcription Factor Activity Rhythms and Tissue-Specific Chromatin Interactions Explain Circadian Gene Expression across Organs.” *Genome Res.*
<https://doi.org/10.1101/gr.222430.117>.
- Zhang, Yong, Tao Liu, Clifford A Meyer, Jérôme Eeckhoute, David S Johnson, Bradley E Bernstein, Chad Nusbaum, Richard M Myers, Myles Brown, Wei Li, and X Shirley Liu. 2008. “Model-Based Analysis of ChIP-Seq (MACS).”

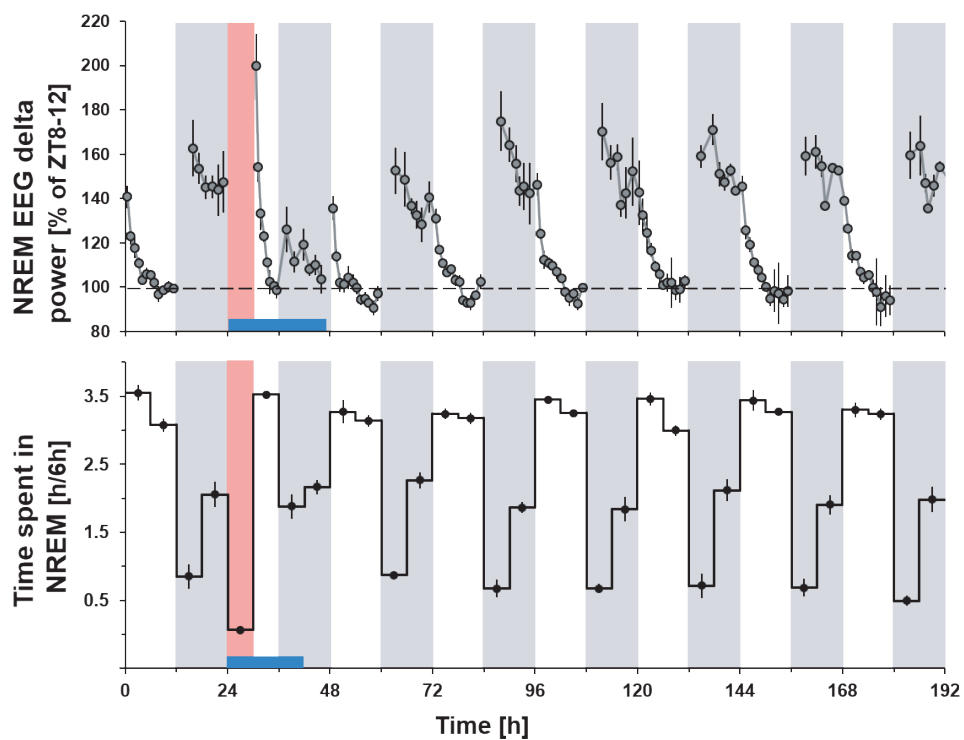
- Genome Biology* 9 (9): R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
- Zhao, H, Z Sun, J Wang, H Huang, J P Kocher, and L Wang. 2014. "CrossMap: A Versatile Tool for Coordinate Conversion between Genome Assemblies." *Bioinformatics* 30 (7): 1006–7. <https://doi.org/10.1093/bioinformatics/btt730>.
- Zuber, Annie Mercier, Gabriel Centeno, Sylvain Pradervand, Svetlana Nikolaeva, Lionel Maquelin, Léonard Cardinaux, Olivier Bonny, and Dmitri Firsov. 2009. "Molecular Clock Is Involved in Predictive Circadian Adjustment of Renal Function." *Proceedings of the National Academy of Sciences of the United States of America* 106 (38): 16523–28. <https://doi.org/10.1073/pnas.0904890106>.

Figure 1

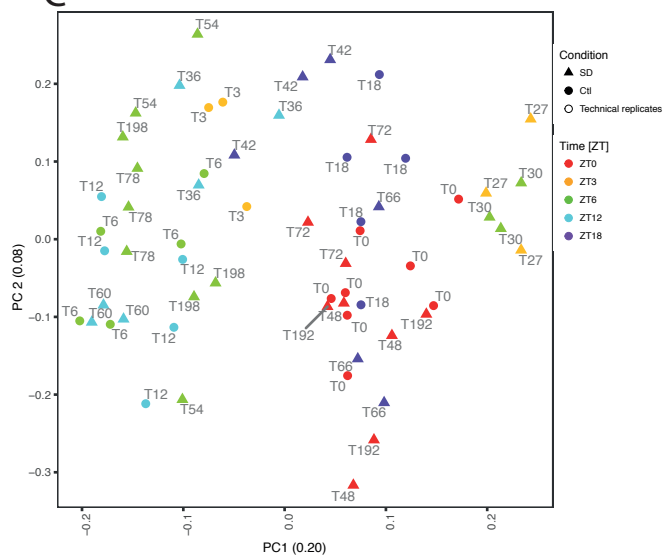
A



B



C



D

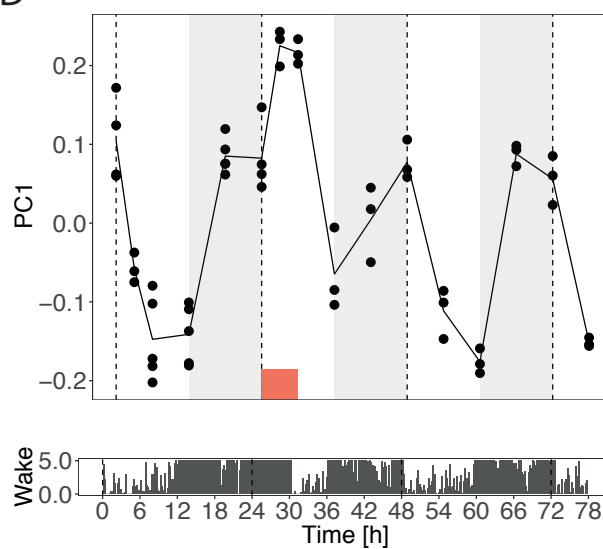


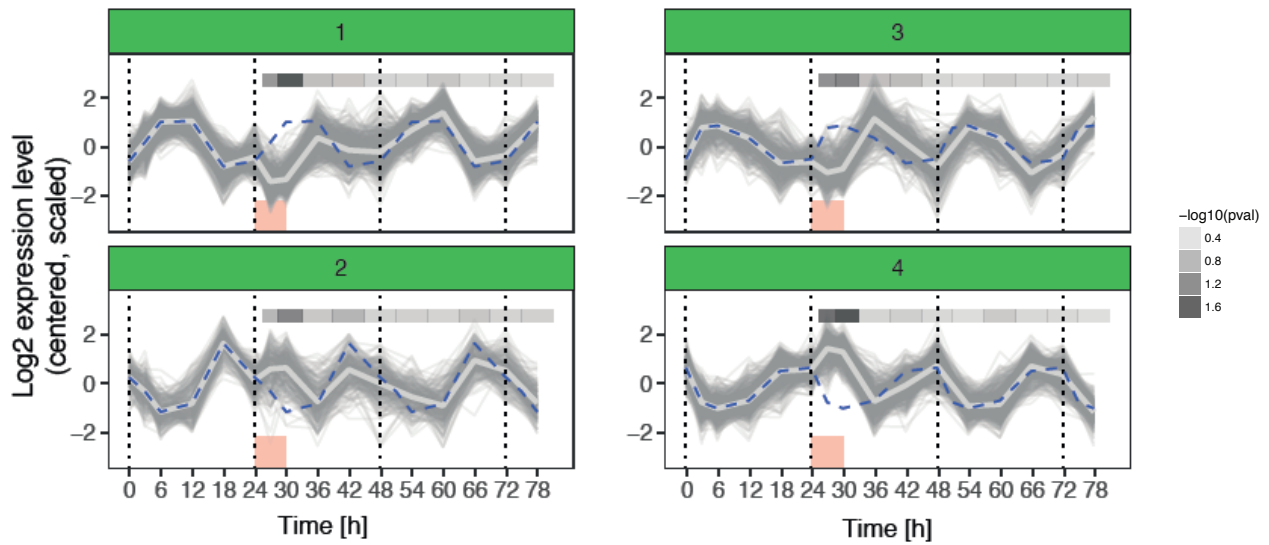
Figure 1

Overview of experimental design, sleep-wake activity, and gene expression data.

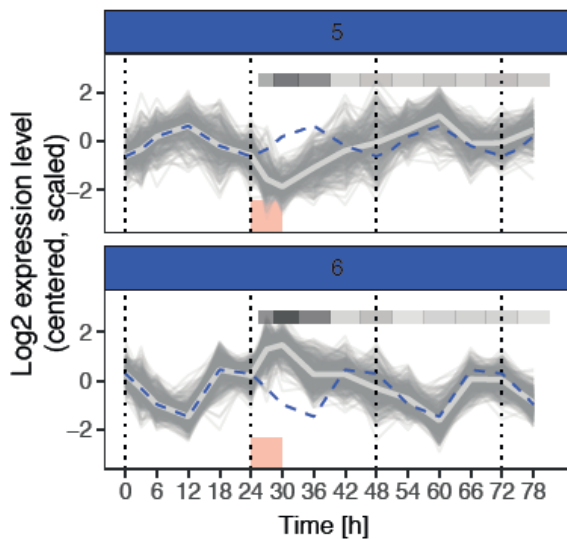
A. Tissue collection schedule with time from beginning of the experiment and corresponding ZT. White and grey bars below the timeline represent the 12h:12h light/dark cycle. Red bar: sleep deprivation. **B.** Long-term effects of sleep deprivation on NREMS delta power and quantity. Following a 6-hour SD (red shade), mice display a significant increase in delta power (1-4Hz; top), which returns to baseline levels within 6 hours. Subsequent recovery days (starting at T48) show no differences from baseline and persists until the end of the recording period, 6 days after SD (T192). Quantification of NREM sleep showed similar patterns across all recording days (bottom). Mean delta power values (\pm SEM) are expressed as the percentage of intraindividual deviations from the period of baseline with the lowest overall power (average across 2 days, ZT8-12). Blue bars represent significant differences from corresponding baseline values (t-test, $p < 0.05$, $n = 6$). **C.** Principal component analysis of expression of 13842 detected genes in RNA-seq data. Parenthesis of axis label denotes fraction of variance explained by the component. Colors denote Zeitgeber time (ZT0 to ZT12: light period, ZT12 to ZT18: dark period). Text labels denote time of experiment, in hours. **D.** First principal component plotted over time (top). Temporal EEG data over time averaged across 15 mice (bottom). EEG y-axis denotes number of minutes awake over last 5 minutes. PC1 increases during wake and decreases during sleep.

Figure 2

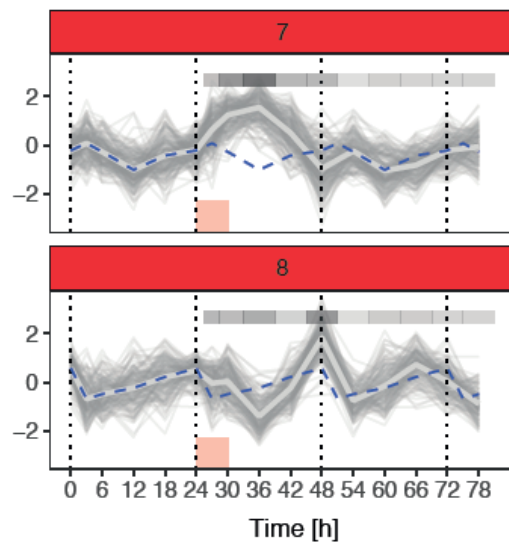
Immediate response, fast recovery



Immediate response, slow recovery



Prolonged/delayed response



No/weak response

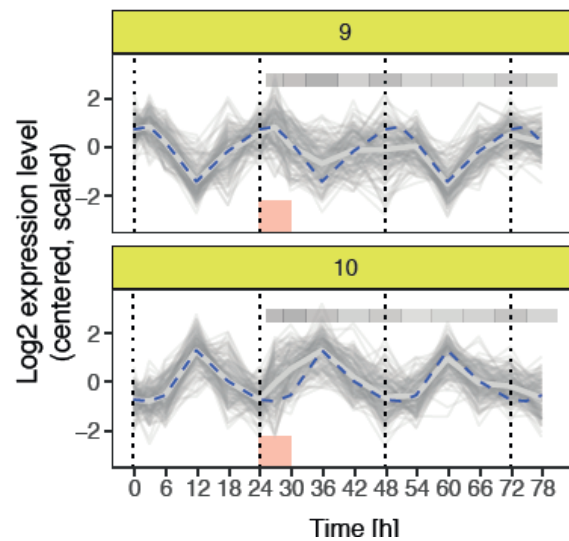


Figure 2

Clustering analysis of gene expression data reveals diverse responses to sleep deprivation.

K-means clusters of 3461 genes (filtered with FDR adjusted p val < 0.001 , likelihood ratio test). Blue dashed line: average of the cluster under baseline, repeated over the three days of the experiment. Light grey thick line: cluster average. Red shaded box: time of sleep deprivation. Grey shaded bar: Mean statistical significance across genes using t-test between sleep-deprived and baseline conditions at the same Zeitgeber time.

Figure 3

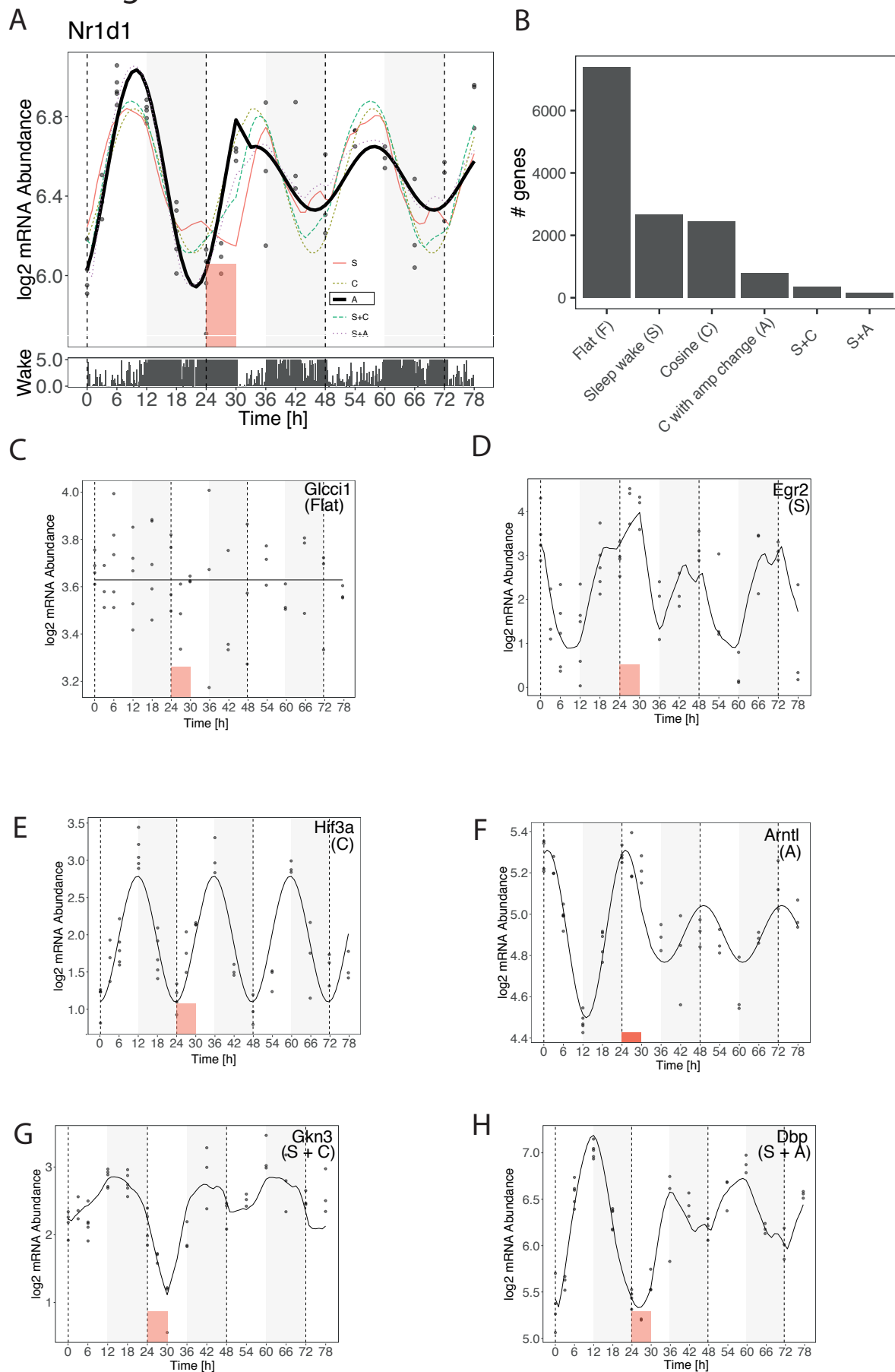


Figure 3

Model selection identifies sleep-wake dynamics and amplitude changes.

A: Example of model fitting on *Nr1d1*. Dots represent RNA level data points, the red box the SD. Data points and all models, listed on bottom right. Best fitting model (cosine with amp change) is highlighted in dark bold.. Lower bar show EEG data for minutes spent awake in 5-minute bins. **B.** Number of genes per model. **C-H.** Examples of 6 genes with their best fit model shown in solid line.

Figure 4

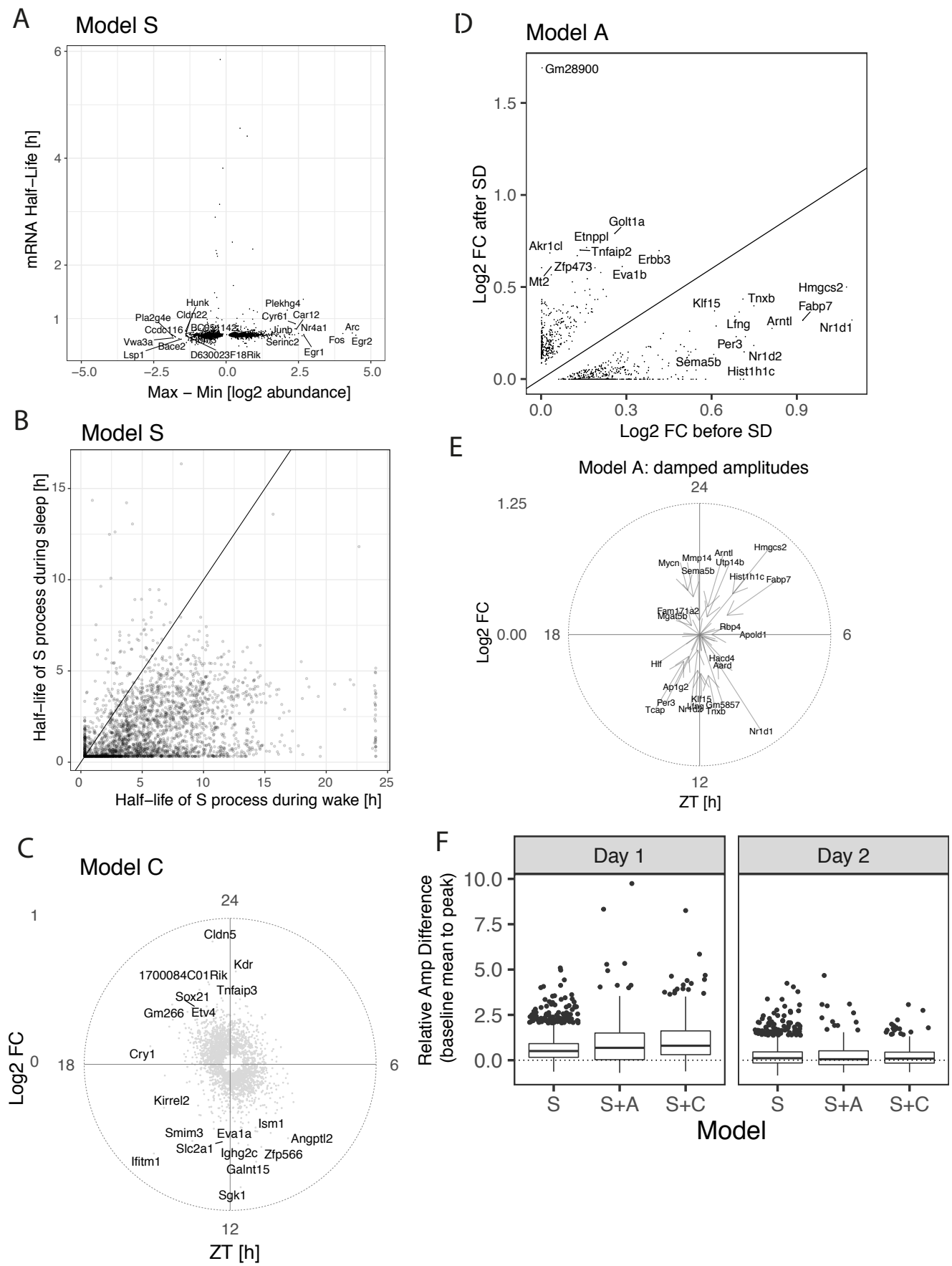


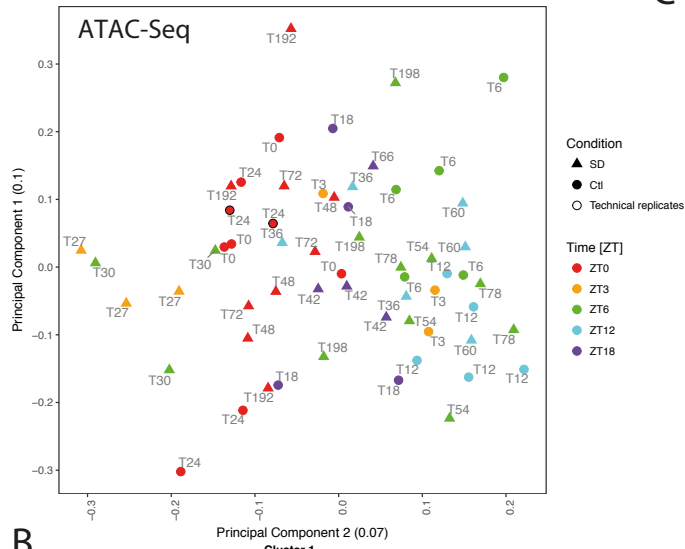
Figure 4

Summary of gene modules by model parameters.

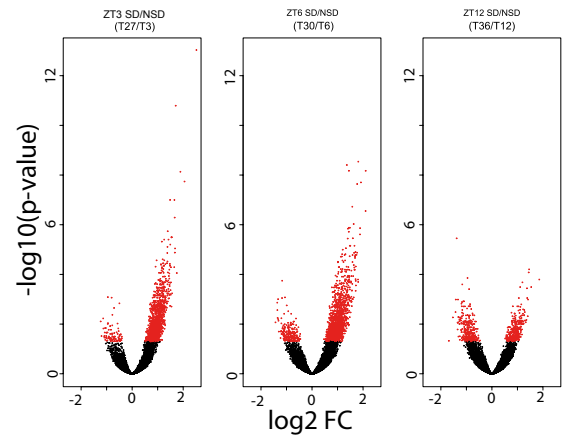
- A.** Summary of log2 abundance (positive values denote increased during sleep deprivation) as a function of predicted mRNA half-life in the sleep-wake driven model (S). **B.** Time constants of S process show a trend where the rate of change during wake is slower than during sleep.
- C.** Polar plot of log2 fold change (radial distance) and time of peak expression (clockwise angle) for genes in cosine model (C).
- D.** Log2 fold change before (x-axis) and after (y-axis) sleep deprivation treatment for genes in cosine with amp change (A).
- E.** Polar plot of log2 fold change before and after of selected genes with damped amplitudes.
- F.** Relative change in amplitude relative to day 0 (baseline) in sleep-wake driven models. Left panel show relative changes for day 1, right panel for day 2.

Figure 5

A



C



B

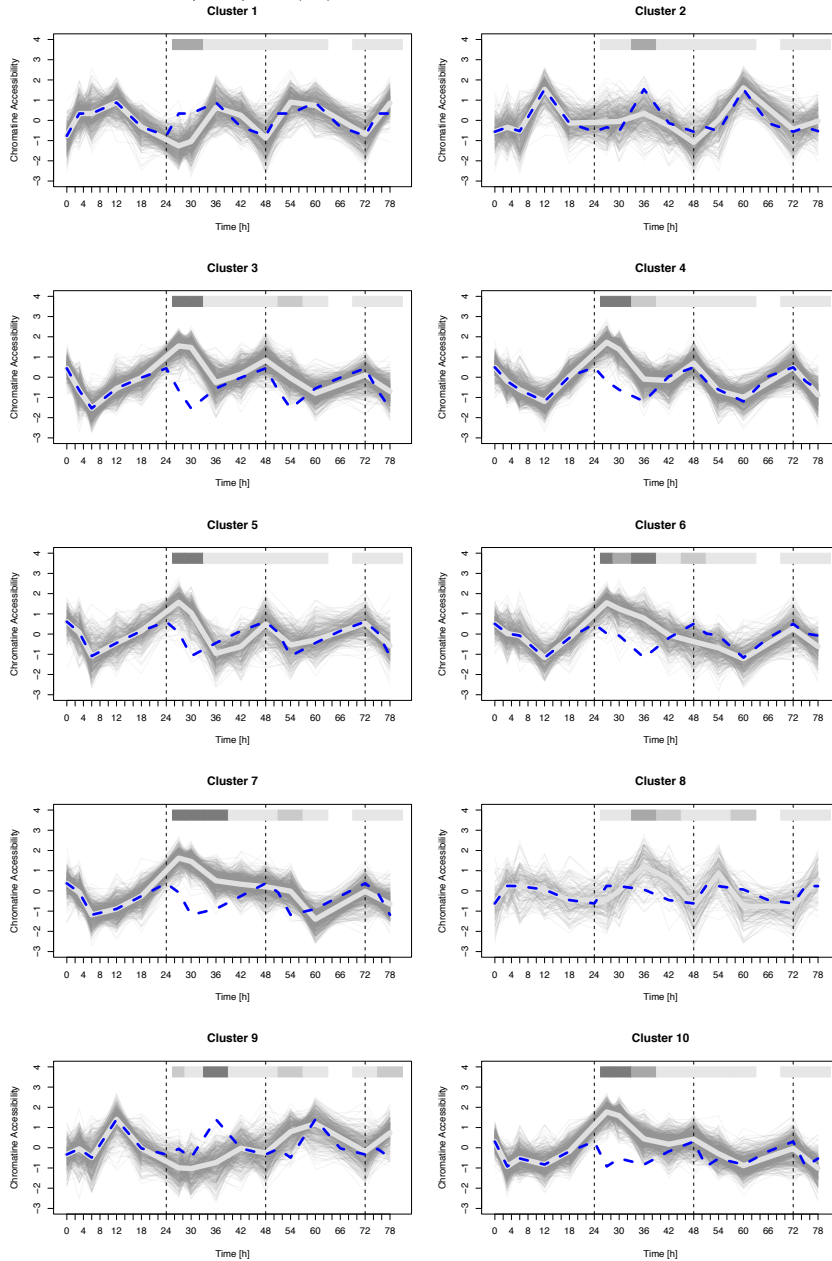


Figure 5

Dynamics in chromatin accessibility and response to sleep deprivation.

A. PCA of heatmap of accessibility levels in SD and NSD samples at ZT6 and ZT12 (215'045 ATAC-seq loci). **B.** K-means clustering analysis of ATAC-seq signal (K=10) showing sleep-wake driven dynamics. **C.** Volcano plots comparing ATAC-seq differences at T3 vs. T27 (3 hours of SD), T6 vs. T30 (6 hours of SD), and T12 vs T36 (6 hours after SD).

Figure 6

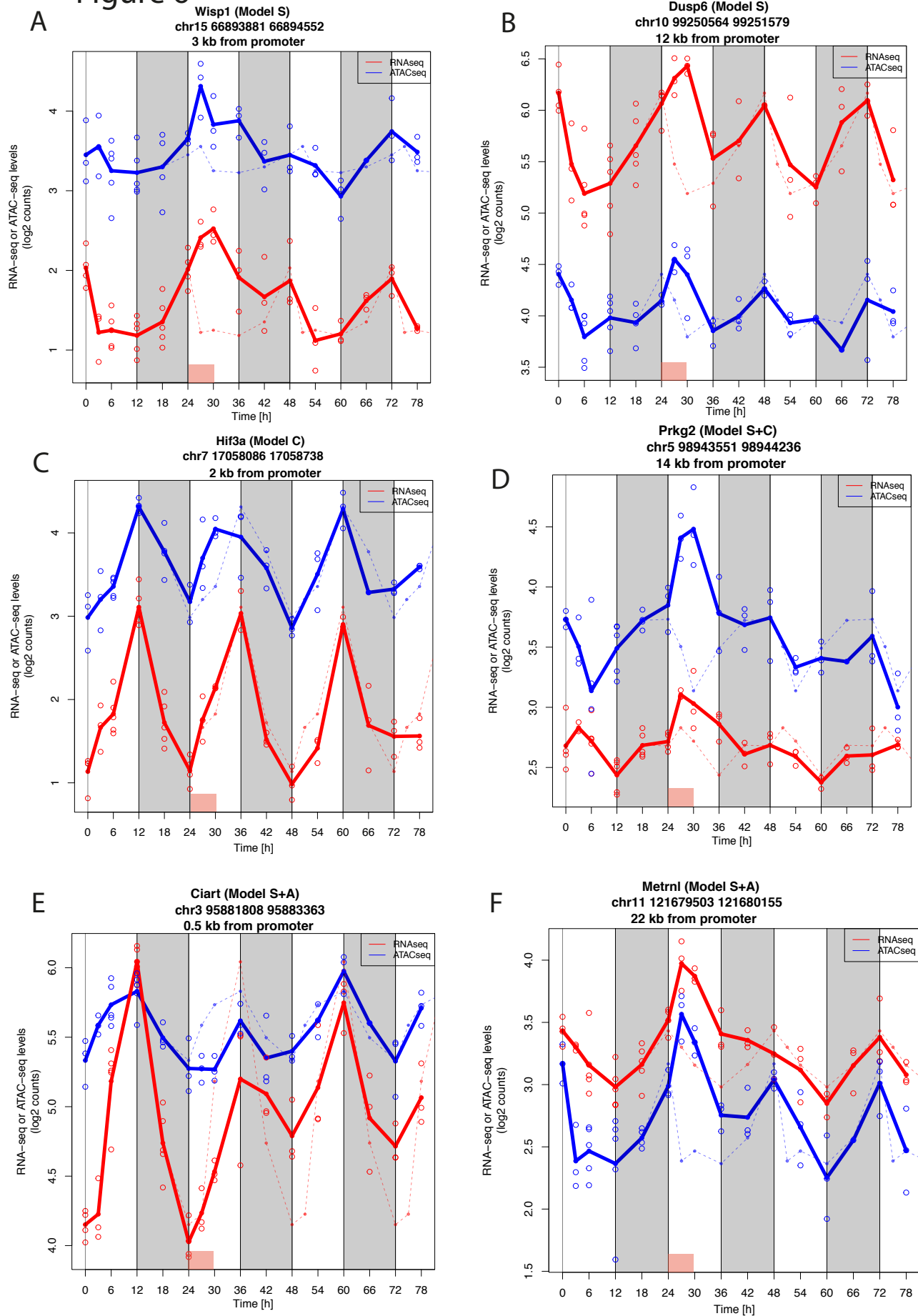


Figure 6

Chromatin accessibility correlate with gene expression dynamics

(A-F) Examples of ATAC-seq peaks with dynamics that correlate with gene expression. Plot titles highlight gene name and its respect model, peak location (mm10), and distance from promoter. The log RNA-seq (red) and ATAC-seq (blue) counts are plotted on the same y-axis. Solid lines represent mean across biological replicates. Dashed lines represent the baseline signal. Open circles are signal from individual mice.

Figure 7

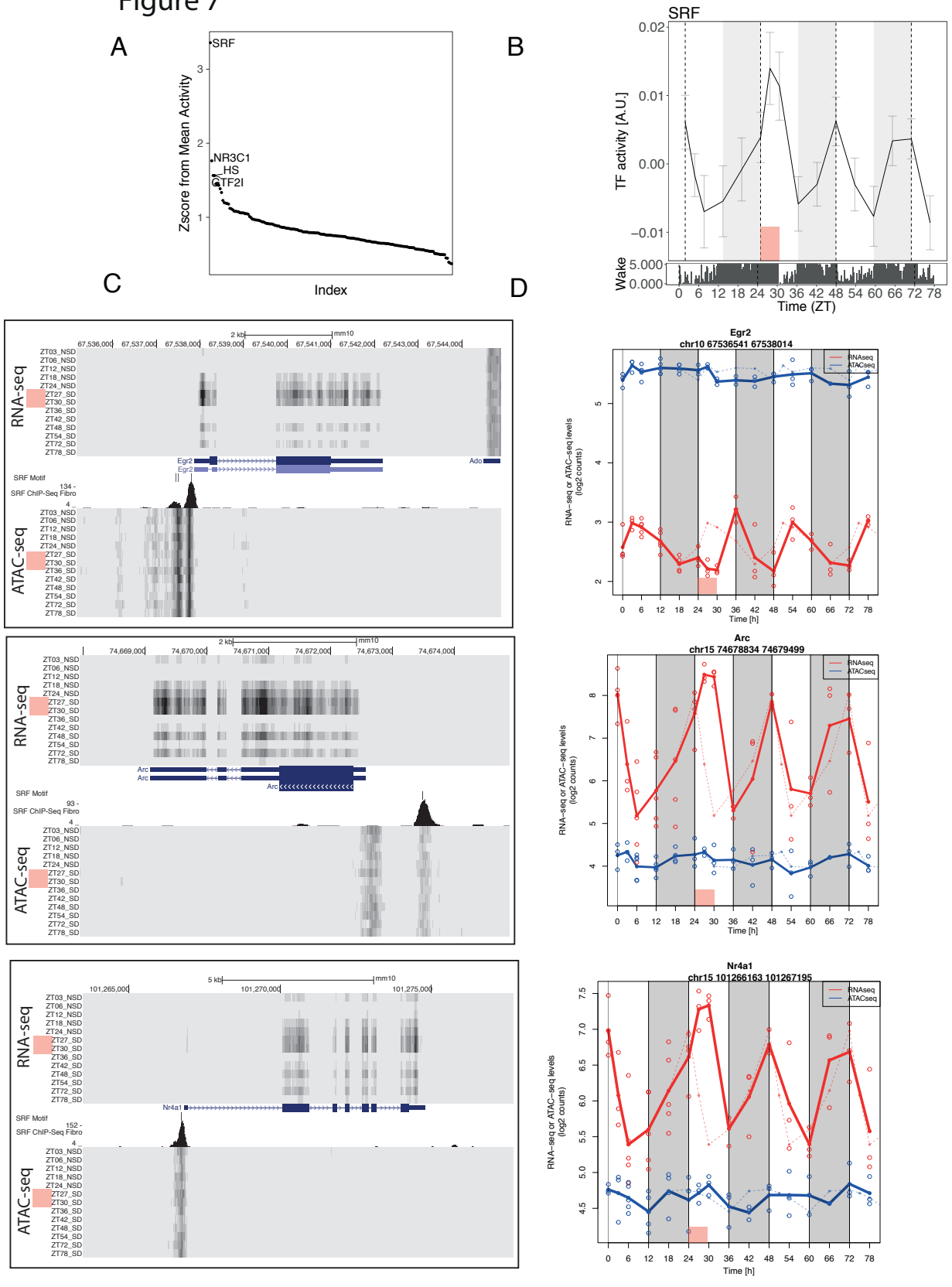


Figure 7

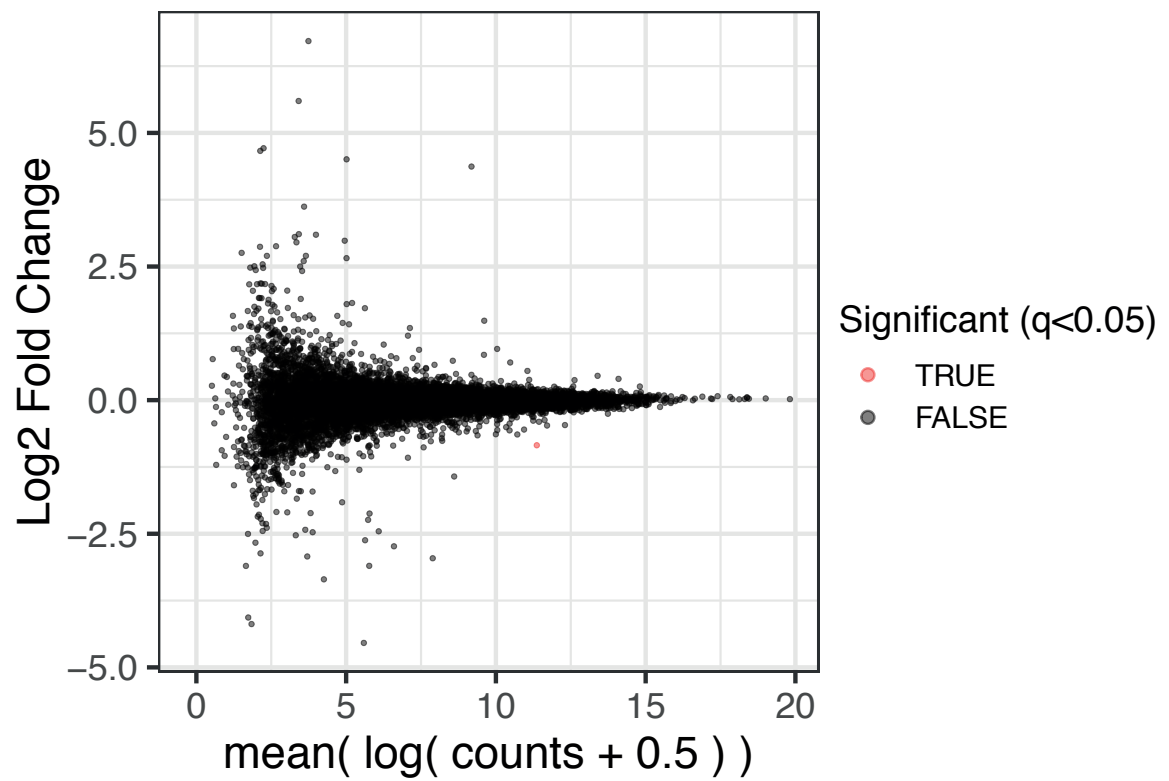
Serum response factor underlies immediate early gene expression dynamics.

A. 179 TF motifs ranked by z-score that explains the temporal dynamics in the RNA-seq dataset. **B.** Inferred temporal activity of SRF. Error bars are standard deviations of the activity estimates. Lower bar denotes time spent awake in 5-minute bins. **C.** Candidate SRF target genes. RNA-seq (top) and ATAC-seq (bottom) signal near SRF target genes. Additional tracks: SRF motif; predicted SRF motif instances. SRF ChIP-seq: ChIP-seq targeting SRF from (Esnault et al. 2014). **D.** ATAC-seq (blue) and RNA-seq (red) levels over time at candidate SRF target genes. Solid lines represent mean across biological replicates. Open circles represent individual mice. Dashed line shows the baseline signal.

Supplemental Figure 1

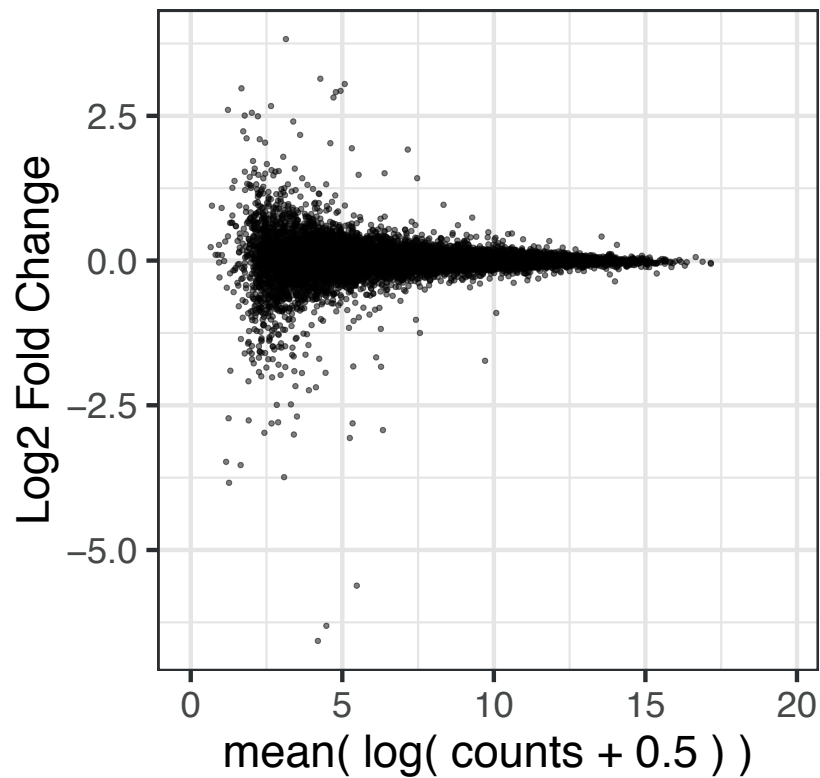
A

T6 vs. T198 (ZT6)



B

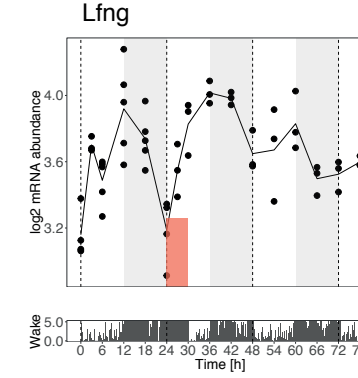
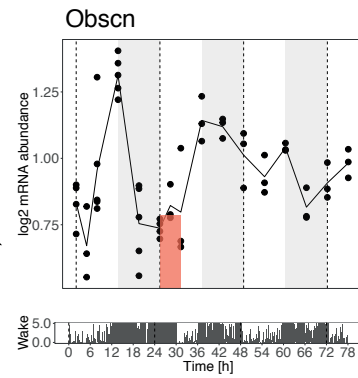
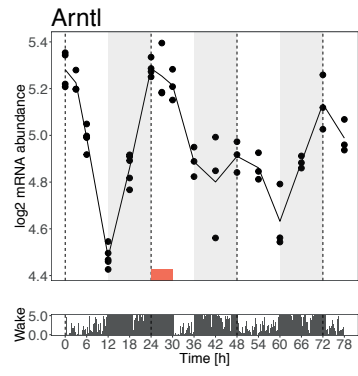
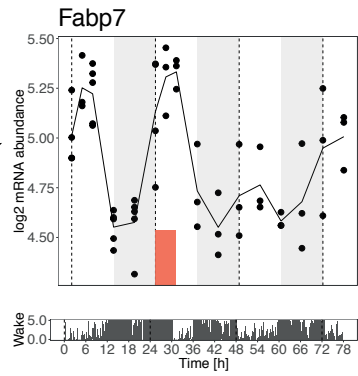
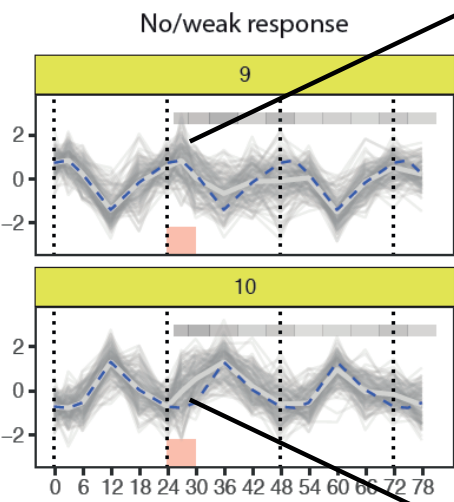
T0 vs. T192 (ZT0)



Supplemental Figure 1

MA plot showing almost no differential expression 7 days after the beginning and end of sleep deprivation.

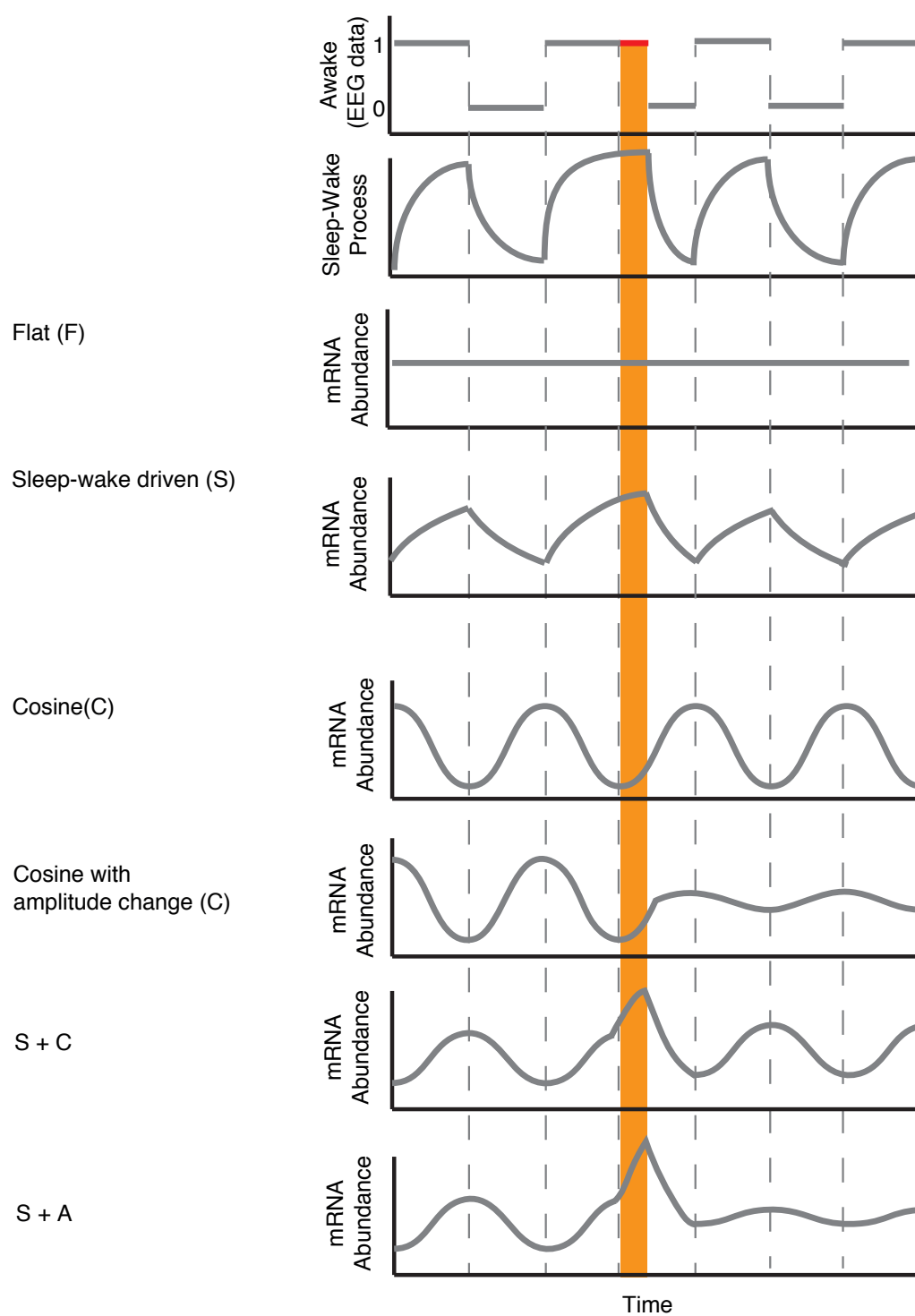
Supplemental Figure 2



Supplemental Figure 2

Examples of individual genes in cluster 9 and 10 with damped amplitudes after sleep deprivation treatment.

Supplemental Figure 3



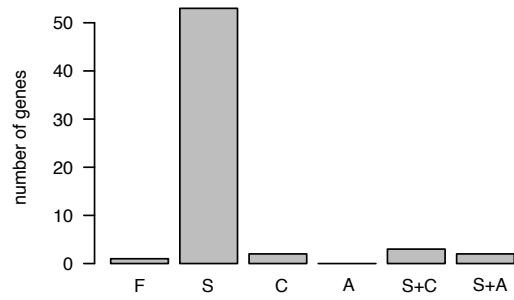
Supplemental Figure 3

Schematic of the 6 models used for model selection.

Supplemental Figure 4

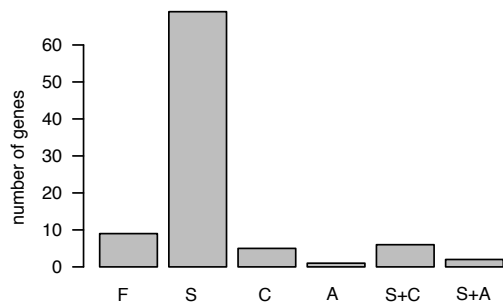
A

Sleep-wake driven genes from Mongrain et al. 2010



B

Genes affected by SD from Maret et al. 2007



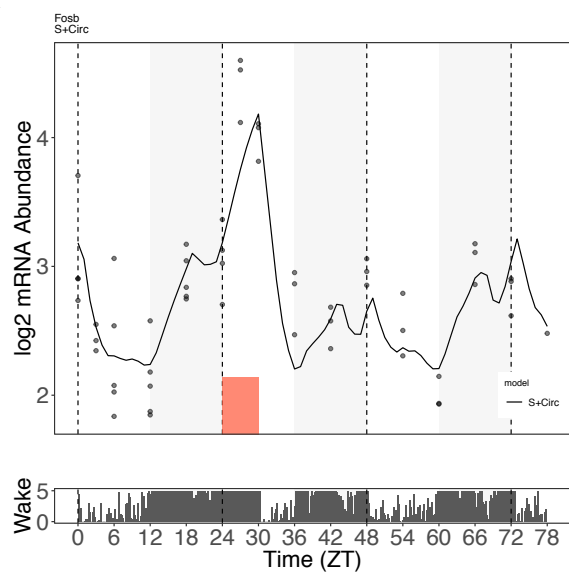
Supplemental Figure 4

A. Distribution of 61 sleep-wake driven genes previously identified from Mongrain et al. 2010 to the 6 models used in this study.

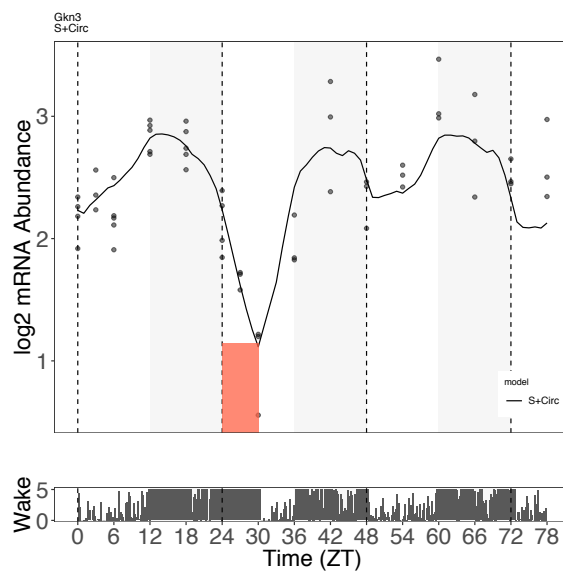
B. Distribution of 92 previously identified to be affected by sleep deprivation from Maret et al. 2007 to the 6 models used in this study.

Supplemental Figure 5

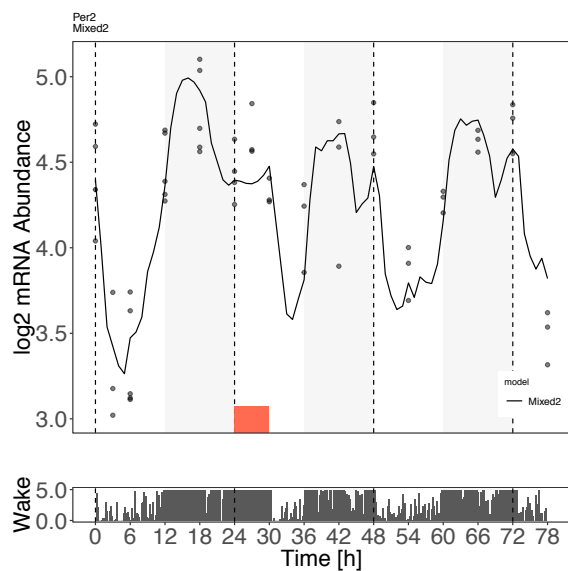
A



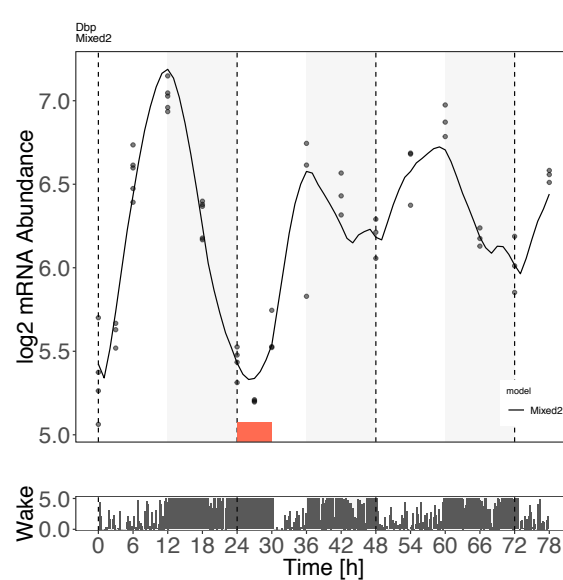
B



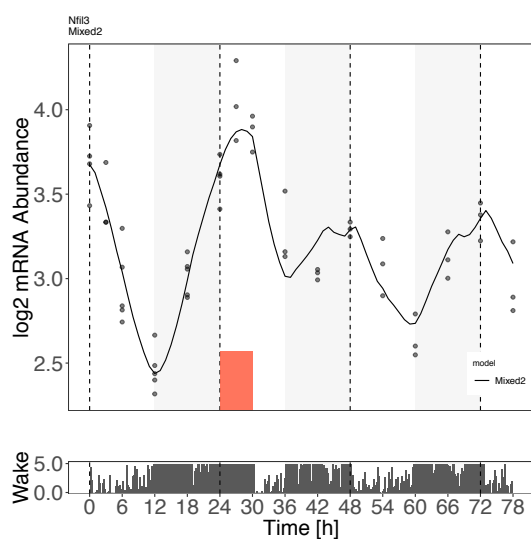
C



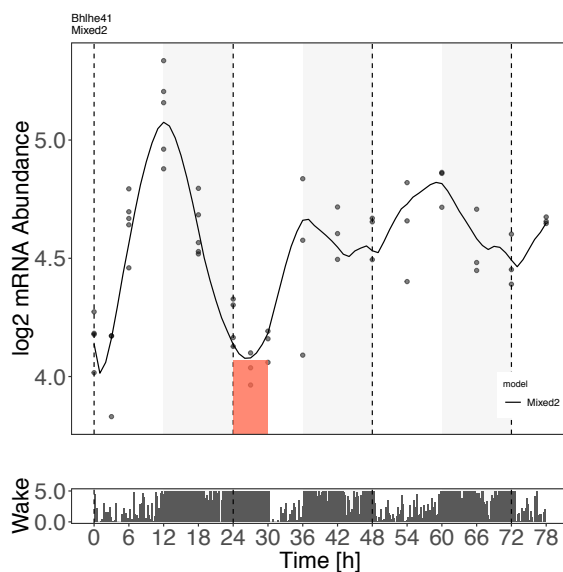
D



E



F

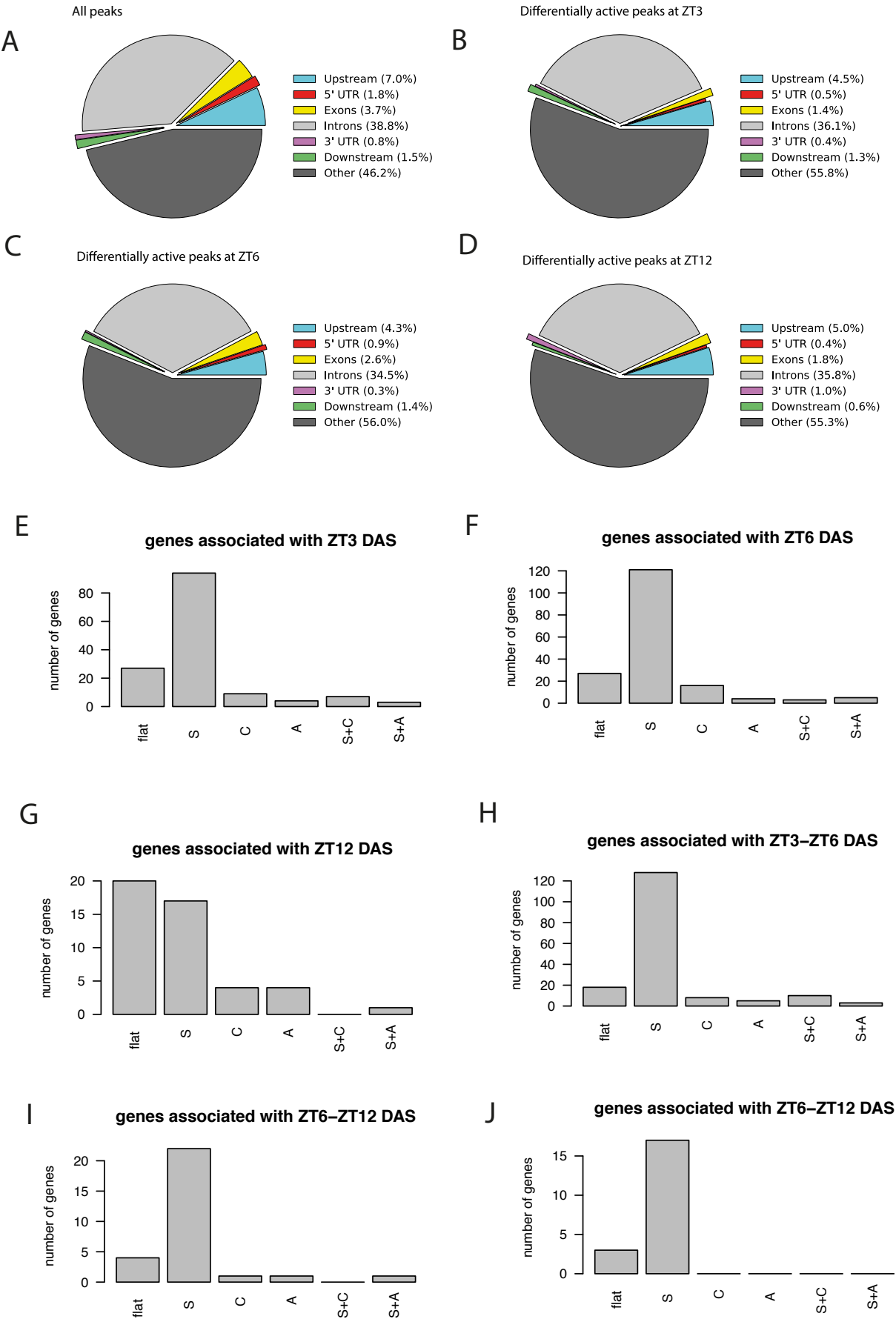


Supplemental Figure 5

A-B. Examples of genes in S+C model showing upregulation (A) and downregulation (B) during sleep deprivation.

C-F. Examples of genes in S+A model. Clock and clock output genes show immediate response to sleep deprivation and damped amplitudes to after sleep deprivation.

Supplemental Figure 6



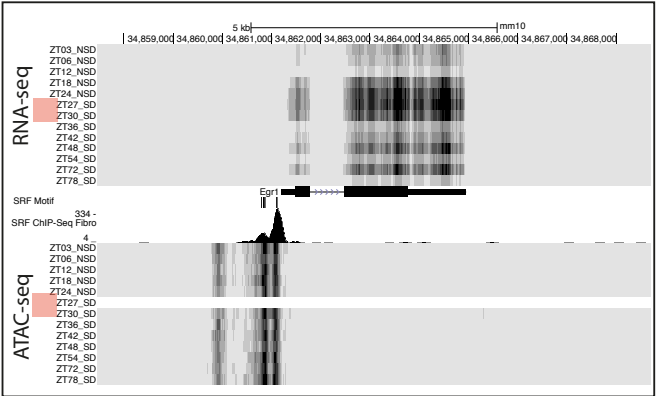
Supplemental Figure 6

A-D. Genomic features of ATAC-seq peaks. A. All peaks. B-D Peaks differentially active at T27 vs T3 (B), at T30 vs T6 (C), and at T36 vs T12 (D).

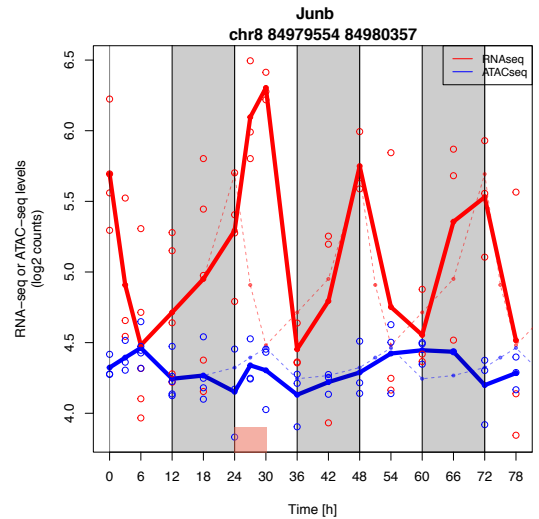
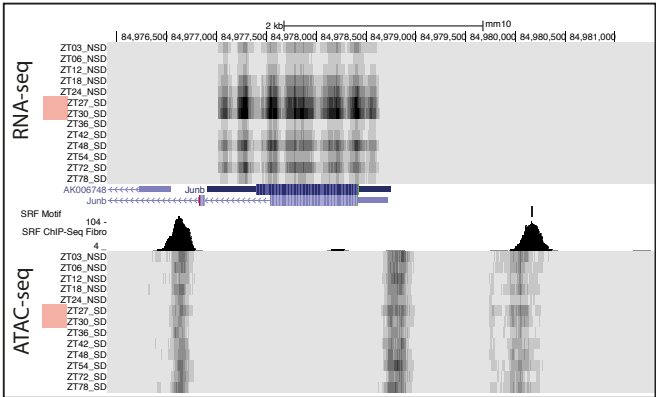
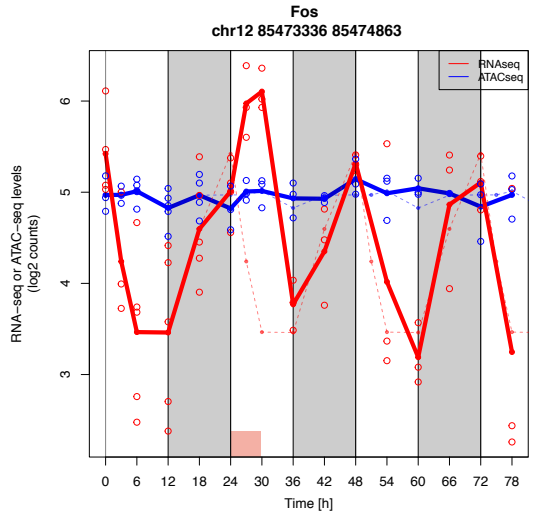
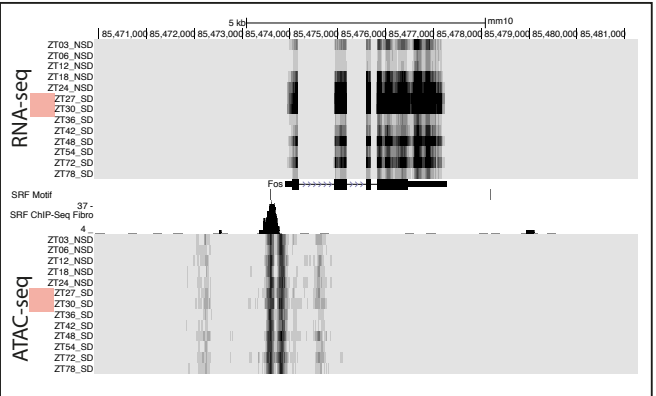
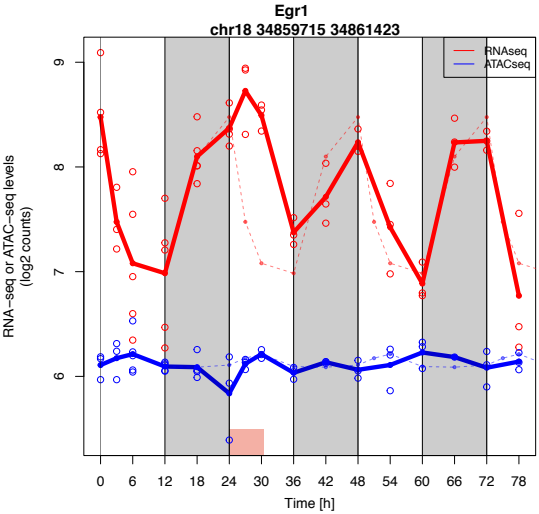
E-J. Distribution of models for genes associated by nearest differentially active sites (DAS). E. ZT3 compares T27 vs. T3. F. ZT6 compares T30 vs. T6. G. ZT12 compares T12 vs. T36.

Supplemental Figure 7

A



B



Supplemental Figure 7

Serum response factor underlies immediate early gene expression dynamics.

A. Candidate SRF target genes. RNA-seq (top) and ATAC-seq (bottom) signal near SRF target genes. Additional tracks: SRF motif: predicted SRF motif instances. SRF ChIP-seq: ChIP-seq targeting SRF from (Esnault et al. 2014). **B.** ATAC-seq (blue) and RNA-seq (red) levels over time at candidate SRF target genes. Solid lines represent mean across biological replicates. Open circles represent individual mice. Dashed line shows the baseline signal.

4.3 Conclusion and perspectives

This study systematically analyzes how dynamics in the transcriptome are regulated by the sleep homeostat (Process S), time-of-day (Process C), as well as interactions between the two. Sleep deprivation can have complex effects on circadian gene expression. Robust oscillations in clock and clock outputs genes become damped after sleep deprivation, and this damping can last more than 48 hours. Thus, our results highlight that observing clock genes during short exposures can have surprising responses. We also introduce a statistical framework to infer the types of dynamics that occur in gene expression. Our work also motivates future questions:

1. How do these non-periodic temporal functions manifest in other systems such as liver and feeding?
2. What are the single-cell dynamics and are there cell-types with specific dynamics within the cortex?
3. What underlying processes make dynamics robust or sensitive to sleep deprivation? And is it possible to perturb or reinforce rhythms using drugs?

5 Conclusions and Perspectives

This thesis addressed three specific questions in chronobiology:

1. How does the circadian clock regulate gene expression in a tissue-specific manner?
2. What is the role of chromatin interactions in regulating circadian gene expression and circadian rhythms?
3. How does diurnal gene expression respond to acute perturbations such as sleep deprivation?

I explored the first question in Chapter 2 by studying transcriptional regulatory modes that enable gene expression to oscillate in one tissue but not others. I developed new methods to analyze high-dimensional data with periodic structure, namely complex-valued singular value decomposition (cvSVD) in combination with model selection. This combination of unsupervised methods (cvSVD) and supervised methods (model selection) allowed me to explore circadian gene expression data at different levels of granularity: from whole transcriptome to gene modules and to individual genes.

Future extensions from this study will deepen our understanding of dynamic gene regulation across tissues. First, we can extend to posttranscriptional regulation to ask how the proportions of transcriptional and posttranscriptional regulation vary across tissues. I participated in the development of methods to systematically compare proportion of genes regulated by transcriptional and posttranscriptional regulation in mouse liver (Wang et al., 2018). Applying these methods to multiple tissues may reveal insights on dynamic gene expression of different regulatory layers across tissues. Second, the question of how gene products from one tissue influence gene expression in others have not been systematically explored. Inferring dynamics of inter-tissue communication requires the analysis of temporal gene expression data across tissues. Tissue-specific activation of ligand-receptor pathways can be predicted by analyzing RNA-seq data of tissues, using computational frameworks such as MARA. These predictions could then be validated by using agonists to activate candidate pathways. Moving

beyond mRNA abundance towards protein abundance and activity (e.g., protein phosphorylation) (Wang et al., 2017) will provide a more direct readout of pathway activation. Finally, the 4C-seq analysis across tissues found that enhancers can contact a rhythmic promoter while looping out nearby nonrhythmic alternative promoters, confining rhythmic enhancer activity to specific promoters. How this specificity arises poses an intriguing open question for future work. Scaling up this analysis to many promoters using techniques such as promoter capture Hi-C (Mifsud et al., 2015) will allow different promoter structures (e.g., CpG islands, CG content, presence of core promoter motifs) to be analyzed to infer links between enhancer-promoter contacts and promoter sequences.

For the second question, discussed in Chapter 3, Jerome and I established chromatin interactions as a fundamental layer that enables proper timing of the circadian clock as well as robustness of circadian locomotor activity. Deleting a regulatory enhancer in *Cry1* showed not only disrupted 24-hour dynamics in chromatin looping, but also shortened period of locomotor activity. The period decreased by 15 minutes, which is expected considering that deleting the entire *Cry1* gene shortens the locomotor activity by approximately 1.2 hours (Ko and Takahashi, 2006). This study incorporated analysis at the level of single cells, whole tissue, and behavior. Nevertheless, how robust oscillations at the tissue level result from noisy gene expression in single cells, which involve promoter-enhancer interactions, remain fascinating open questions. Single-cell technologies that combine single-molecule RNA FISH with DNA labeling are beginning to uncover the temporal relationships between chromatin interactions and gene expression (Chen et al., 2018). Extending these techniques to investigate gene expression dynamics will elucidate how intrinsic and extrinsic noise components influence dynamics.

Third, Chapter 4 looks beyond periodic gene expression by investigating how the plasticity of the circadian clock responds to sleep deprivation over multiple days. This unique study design allowed us to incorporate sleep-wake history of mice to model long-term dynamics in gene expression. Our findings that many core clock and clock output genes have damped amplitudes and perturbed dynamics in response to sleep deprivation suggest that clock genes are highly adaptive to physiological states and acute environmental changes. We have gained quantitative insights in the role of sleep homeostat and the circadian clock on gene expression dynamics, and further investigations that include other perturbations such as feeding and temperature would reveal the different pathways in which the circadian clock responds to environmental signals. These other perturbations would look into other tissues such as liver and link with how different tissues may coordinate physiology to adapt to environmental changes. The tissue heterogeneity in the mouse cortex warrants further investigation into the single-cell dynamics and how different cell-types may have different dynamics within the cortex. Our findings also open up questions to whether robustness or sensitivity of sleep deprivation could be modulated pharmacologically.

We expect advances in experimental techniques as well as in computational methods to uncover general principles of how organisms have embraced environmental periodicity

across all scales of biological. For example, theoretical frameworks of how rhythms in enzyme activities, metabolic flux, and metabolites coordinate to propagate rhythms may uncover novel insight into the regulation of temporal metabolism (Thurley et al., 2017). Assaying enzyme activities in a high-throughput manner, which is yet not available, would enable systematic analysis of the relationships between gene expression, protein abundance, and enzyme activities.

While most molecular chronobiology findings have relied on studying model organisms, there are currently exciting opportunities for human chronobiology. Genome-wide association and candidate gene sequencing studies have reported genetic variations associated with circadian clock-related and sleep phenotypes (Hu et al., 2016; Allebrandt et al., 2010; Shi et al., 2017). Beyond self-reported questionnaires, temporal activity patterns or other behaviors, such as food intake, can be directly measured through smartphone apps (Gill and Panda, 2015; Roenneberg, 2017; Aledavood et al., 2015). Time stamping assays such as the BodyTime assay combines both computational methods and, critically, a robust experimental assay to accurately determine internal circadian time from a blood sample (Wittenbrink et al., 2018).

Analysis of large consortia of gene expression across human tissues (Ardlie and Guigó, 2017) can be used to discover circadian gene expression in human by applying machine learning methods discussed above (Anafi et al., 2017; Ruben et al., 2018). Taking into account natural variation may reveal coding and noncoding variants that affect circadian gene expression. Noncoding variants may be associated with disruption of TF binding (Deplancke et al., 2016), which can reveal links between genome variation and gene regulatory mechanisms underlying behavioral phenotypes. A first indication that non-coding DNA may have phenotypic consequence on the mammalian circadian clock was reported in mouse (Mermet et al., 2018), and it will be interesting to assess how this generalizes to human.

Bibliography

- Victoria A. Acosta-Rodríguez, Marleen H.M. de Groot, Filipa Rijo-Ferreira, Carla B. Green, and Joseph S. Takahashi. Mice under Caloric Restriction Self-Impose a Temporal Restriction of Food Intake as Revealed by an Automated Feeder System. *Cell Metabolism*, 26(1):267–277.e2, jul 2017. ISSN 1550-4131. doi: 10.1016/J.CMET.2017.06.007. URL <https://www.sciencedirect.com/science/article/pii/S1550413117303492>.
- Agorastos Agorastos, Richard L Hauger, Donald A Barkauskas, Tobias Moeller-Bertram, Paul L Clopton, Uzair Haji, James B Lohr, Thomas D Geraciotti, Piyush M Patel, George P Chrousos, and Dewleen G Baker. Circadian rhythmicity, variability and correlation of interleukin-6 levels in plasma and cerebrospinal fluid of healthy men. *Psychoneuroendocrinology*, 44:71–82, jun 2014. ISSN 1873-3360. doi: 10.1016/j.psyneuen.2014.02.020. URL <http://www.ncbi.nlm.nih.gov/pubmed/24767621>.
- Lorena Aguilar-Arnal, Ofir Hakim, Vishal R Patel, Pierre Baldi, Gordon L Hager, and Paolo Sassone-Corsi. Cycles in spatial and temporal chromosomal organization driven by the circadian clock. *Nature Structural & Molecular Biology*, 20(10):1206–1213, oct 2013. ISSN 1545-9993. doi: 10.1038/nsmb.2667. URL <http://www.nature.com/articles/nsmb.2667>.
- Talayeh Aledavood, Sune Lehmann, and Jari Saramäki. Digital daily cycles of individuals. *Frontiers in Physics*, 3:73, oct 2015. ISSN 2296-424X. doi: 10.3389/fphy.2015.00073. URL <http://journal.frontiersin.org/Article/10.3389/fphy.2015.00073/abstract>.
- Karla V. Allebrandt, Maris Teder-Laving, Mahmut Akyol, Irene Pichler, Bertram Müller-Myhsok, Peter Pramstaller, Martha Merrow, Thomas Meitinger, Andreas Metspalu, and Till Roenneberg. CLOCK Gene Variants Associate with Sleep Duration in Two Independent Populations. *Biological Psychiatry*, 67(11):1040–1047, jun 2010. ISSN 0006-3223. doi: 10.1016/J.BIOPSYCH.2009.12.026. URL <https://www.sciencedirect.com/science/article/pii/S0006322309015273?via%3Dihubhttps://www.sciencedirect.com/science/article/pii/S0006322309015273?via%3Dihub#!>
- Ron C Anafi, Lauren J Francey, John B Hogenesch, and Junhyong Kim. CYCLOPS reveals human transcriptional rhythms in health and disease. *Proceedings of the National Academy of Sciences of the United States of America*, 114(20):5312–5317, may 2017. ISSN 1091-6490. doi: 10.1073/pnas.1619320114. URL <http://www.ncbi.nlm.nih.gov/pubmed/28439010http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5441789>.

Bibliography

- Kristin G. Ardlie and Roderic Guigó. Data resources for human functional genomics. *Current Opinion in Systems Biology*, 1:75–79, feb 2017. ISSN 2452-3100. doi: 10.1016/J.COISB.2016.12.019. URL <https://www.sciencedirect.com/science/article/pii/S2452310017300136>.
- Florian Atger, Cédric Gobet, Julien Marquis, Eva Martin, Jingkui Wang, Benjamin Weger, Grégory Lefebvre, Patrick Descombes, Felix Naef, and Frédéric Gachon. Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proceedings of the National Academy of Sciences*, 112(47):E6579–E6588, nov 2015. ISSN 0027-8424. doi: 10.1073/pnas.1515308112. URL <http://www.pnas.org/lookup/doi/10.1073/pnas.1515308112>.
- Keren Bahar Halpern, Sivan Tanami, Shanie Landen, Michal Chapal, Liran Szlak, Anat Hutzler, Anna Nizhberg, and Shalev Itzkovitz. Bursty Gene Expression in the Intact Mammalian Liver. *Molecular Cell*, 58(1):147–156, apr 2015. ISSN 1097-2765. doi: 10.1016/J.MOLCEL.2015.01.027. URL <https://www.sciencedirect.com/science/article/pii/S1097276515000507><http://www.ncbi.nlm.nih.gov/pubmed/25728770><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4500162><http://www.sciencedirect.com/science/article/pii/S1097276515000507>.
- Caroline R. Bartman, Sarah C. Hsu, Chris C.-S. Hsiung, Arjun Raj, and Gerd A. Blobel. Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell*, 62(2):237–247, apr 2016. ISSN 1097-2765. doi: 10.1016/J.MOLCEL.2016.03.007. URL <https://www.sciencedirect.com/science/article/pii/S1097276516001854>.
- Joseph Bass and Mitchell A. Lazar. Circadian time signatures of fitness and disease. *Science*, 354(6315), 2016. URL <http://science.sciencemag.org/content/354/6315/994.full>.
- Joseph Bass and Joseph S Takahashi. Circadian integration of metabolism and energetics. *Science (New York, N.Y.)*, 330(6009):1349–54, dec 2010. ISSN 1095-9203. doi: 10.1126/science.1195027. URL <http://www.sciencemag.org/content/330/6009/1349.short>.
- Robert A. Beagrie, Antonio Scialdone, Markus Schueler, Dorothee C. A. Kraemer, Mita Chotalia, Sheila Q. Xie, Mariano Barbieri, Inês de Santiago, Liron-Mark Lavitas, Miguel R. Branco, James Fraser, Josée Dostie, Laurence Game, Niall Dillon, Paul A. W. Edwards, Mario Nicodemi, and Ana Pombo. Complex multi-enhancer contacts captured by genome architecture mapping. *Nature*, 543(7646):519–524, mar 2017. ISSN 0028-0836. doi: 10.1038/nature21411. URL <http://www.nature.com/doifinder/10.1038/nature21411>.
- Deborah Bell-Pedersen, Vincent M. Cassone, David J. Earnest, Susan S. Golden, Paul E. Hardin, Terry L. Thomas, and Mark J. Zoran. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature reviews. Genetics*, 6(7):544–56, jul 2005. ISSN 1471-0056. doi: 10.1038/nrg1633. URL <http://www.nature.com/doifinder/10.1038/nrg1633><http://dx.doi.org/10.1038/nrg1633>.

James O. Berger, Jayanta K. Ghosh, and Nitai Mukhopadhyay. Approximations and consistency of Bayes factors as model dimension grows. *Journal of Statistical Planning and Inference*, 112(1):241–258, 2003. ISSN 03783758. doi: 10.1016/S0378-3758(02)00336-1.

Joshua R Beytebiere, Alexandra J Trott, Ben Greenwell, Collin A Osborne, H  l  ne Marine Vitet, Jessica Spence, Seung-Hee Yoo, Zheng Chen, Joseph S Takahashi, Noushin Ghaffari, and Jerome S Menet. Tissue-specific BMAL1 cistromes reveal that enhancer-enhancer interactions regulate rhythmic transcription. *bioRxiv*, page 319319, may 2018. doi: 10.1101/319319. URL <https://www.biorxiv.org/content/early/2018/05/10/319319.1>.

Britta AM Bouwman and Wouter de Laat. Getting the genome in shape: the formation of loops, domains and compartments. *Genome Biology*, 16(1):154, dec 2015. ISSN 1474-760X. doi: 10.1186/s13059-015-0730-1. URL <http://genomebiology.com/2015/16/1/154>.

Hongtao Chen, Michal Levo, Lev Barinov, Miki Fujioka, James B. Jaynes, and Thomas Gregor. Dynamic interplay between enhancer-promoter topology and gene activity. *Nature Genetics*, page 1, jul 2018. ISSN 1061-4036. doi: 10.1038/s41588-018-0175-z. URL <http://www.nature.com/articles/s41588-018-0175-z>.

G. Chow. Test of Equality between Sets of Coefficients in Two Linear Regression. *Econometrica*, 28:591–605, 1960.

Maria J Costa, B  rbel Finkenst  dt, V  ronique Roche, Francis L  vi, Peter D Gould, Julia Foreman, Karen Halliday, Anthony Hall, and David A Rand. Inference on periodicity of circadian time series. *Biostatistics (Oxford, England)*, 14(4):792–806, sep 2013. ISSN 1468-4357. doi: 10.1093/biostatistics/kxt020. URL <http://biostatistics.oxfordjournals.org/content/14/4/792.full><http://www.ncbi.nlm.nih.gov/pubmed/23743206><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3988453>.

Gaspard Cretenet, Mika  l Le Clech, and Fr  d  ric Gachon. Circadian Clock-Coordinated 12 Hr Period Rhythmic Activation of the IRE1   Pathway Controls Lipid Metabolism in Mouse Liver. *Cell Metabolism*, 11(1):47–57, jan 2010. ISSN 1550-4131. doi: 10.1016/J.CMET.2009.11.002. URL [https://www.sciencedirect.com/science/article/pii/S1550413109003398?](https://www.sciencedirect.com/science/article/pii/S1550413109003398?via%3Dihub)<https://www.sciencedirect.com/science/article/pii/S1550413109003398>.

C A Czeisler, J F Duffy, T L Shanahan, E N Brown, J F Mitchell, D W Rimmer, J M Ronda, E J Silva, J S Allan, J S Emens, D J Dijk, and R E Kronauer. Stability, precision, and near-24-hour period of the human circadian pacemaker. *Science (New York, N.Y.)*, 284(5423):2177–81, jun 1999. ISSN 0036-8075. doi: 10.1126/SCIENCE.284.5423.2177. URL <http://www.ncbi.nlm.nih.gov/pubmed/10381883>.

F Damiola, N Le Minh, N Preitner, B Kornmann, F Fleury-Olela, and U Schibler. Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes & development*, 14(23):2950–61, dec 2000. ISSN 0890-9369. doi: 10.1101/GAD.183500. URL <http://www.ncbi.nlm.nih.gov/pubmed/11114885><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC317100>.

Bibliography

- Job Dekker, Karsten Rippe, Martijn Dekker, and Nancy Kleckner. Capturing chromosome conformation. *Science (New York, N.Y.)*, 295(5558):1306–11, feb 2002. ISSN 1095-9203. doi: 10.1126/science.1067799. URL <http://www.ncbi.nlm.nih.gov/pubmed/11847345>.
- Bart Deplancke, Daniel Alpern, and Vincent Gardeux. The Genetics of Transcription Factor DNA Binding Variation. *Cell*, 166(3):538–554, jul 2016. ISSN 0092-8674. doi: 10.1016/J.CELL.2016.07.012. URL <https://www.sciencedirect.com/science/article/pii/S0092867416309187>.
- Elzo de Wit, Erica S.M. Vos, Sjoerd J.B. Holwerda, Christian Valdes-Quezada, Marjon J.A.M. Verstegen, Hans Teunissen, Erik Splinter, Patrick J. Wijchers, Peter H.L. Krijger, and Wouter de Laat. CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell*, 60(4): 676–684, nov 2015. ISSN 1097-2765. doi: 10.1016/J.MOLCEL.2015.09.023. URL <https://www.sciencedirect.com/science/article/pii/S1097276515007625>.
- Charna Dibner, Ueli Schibler, and Urs Albrecht. The Mammalian Circadian Timing System: Organization and Coordination of Central and Peripheral Clocks. *Annual Review of Physiology*, 72(1):517–549, mar 2010. ISSN 0066-4278. doi: 10.1146/annurev-physiol-021909-135821. URL <http://www.annualreviews.org/doi/full/10.1146/annurev-physiol-021909-135821>.
- Jesse R. Dixon, Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376–380, apr 2012. ISSN 0028-0836. doi: 10.1038/nature11082. URL <http://www.nature.com/doifinder/10.1038/nature11082>.
- J Christopher Ehlen, Allison J Brager, Julie Baggs, Lennisha Pinckney, Cloe L Gray, Jason P DeBruyne, Karyn A Esser, Joseph S Takahashi, and Ketema N Paul. Bmal1 function in skeletal muscle regulates sleep. *eLife*, 6, 2017. ISSN 2050-084X. doi: 10.7554/eLife.26557. URL <http://www.ncbi.nlm.nih.gov/pubmed/28726633><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5574702>.
- R. A. Fisher. Tests of significance in Harmonic analysis. *Proc. Roy. Soc. London Ser. A*, 125: 54–59, 1929.
- Paul Franken, Ryan Thomason, H Craig Heller, and Bruce F O’Hara. A non-circadian role for clock-genes in sleep homeostasis:a strain comparison. *BMC Neuroscience*, 8(1):87, oct 2007. ISSN 1471-2202. doi: 10.1186/1471-2202-8-87. URL <http://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-8-87>.
- Edward I. George and Dean P. Foster. Calibration and empirical Bayes variable selection. *Biometrika*, 87(4):731–747, dec 2000. ISSN 0006-3444. doi: 10.1093/biomet/87.4.731. URL <http://biomet.oupjournals.org/cgi/doi/10.1093/biomet/87.4.731>.
- Alan Gerber, Cyril Esnault, Gregory Aubert, Richard Treisman, François Pralong, and Ueli Schibler. Blood-borne circadian signal stimulates daily oscillations in actin dynamics and SRF activity. *Cell*, 152(3):492–503, jan 2013. ISSN 1097-4172. doi: 10.1016/j.cell.2012.12.027. URL <http://www.sciencedirect.com/science/article/pii/S0092867412015498><https://www.sciencedirect.com/science/article/pii/S0092867412015498>.

- Johan H Gibcus, Kumiko Samejima, Anton Goloborodko, Itaru Samejima, Natalia Nau-mova, Johannes Nuebler, Masato T Kanemaki, Linfeng Xie, James R Paulson, William C Earnshaw, Leonid A Mirny, and Job Dekker. A pathway for mitotic chromosome formation. *Science (New York, N.Y.)*, 359(6376):eaao6135, jan 2018. ISSN 1095-9203. doi: 10.1126/science.aao6135. URL <http://www.ncbi.nlm.nih.gov/pubmed/29348367><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5924687>.
- Johan H. Gibcus and Job Dekker. The Hierarchy of the 3D Genome. *Molecular Cell*, 49(5): 773–782, mar 2013. ISSN 1097-2765. doi: 10.1016/J.MOLCEL.2013.02.011. URL <https://www.sciencedirect.com/science/article/pii/S1097276513001391>.
- Shubhroz Gill and Satchidananda Panda. A Smartphone App Reveals Erratic Diurnal Eating Patterns in Humans that Can Be Modulated for Health Benefits. *Cell Metabolism*, 22(5): 789–798, nov 2015. ISSN 1550-4131. doi: 10.1016/J.CMET.2015.09.005. URL <https://www.sciencedirect.com/science/article/pii/S1550413115004623>.
- Xiaoping Han, Renying Wang, Yincong Zhou, Lijiang Fei, Huiyu Sun, Shujing Lai, Assieh Saadatpour, Ziming Zhou, Haide Chen, Fang Ye, Daosheng Huang, Yang Xu, Wentao Huang, Mengmeng Jiang, Xinyi Jiang, Jie Mao, Yao Chen, Chenyu Lu, Jin Xie, Qun Fang, Yibin Wang, Rui Yue, Tiefeng Li, He Huang, Stuart H. Orkin, Guo-Cheng Yuan, Ming Chen, and Guoji Guo. Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell*, 172(5):1091–1107.e17, feb 2018. ISSN 0092-8674. doi: 10.1016/J.CELL.2018.02.001. URL <https://www.sciencedirect.com/science/article/pii/S0092867418301168>.
- Michael H. Hastings, Elizabeth S. Maywood, and Marco Brancaccio. Generation of circadian rhythms in the suprachiasmatic nucleus. *Nature Reviews Neuroscience*, page 1, jun 2018. ISSN 1471-003X. doi: 10.1038/s41583-018-0026-z. URL <http://www.nature.com/articles/s41583-018-0026-z>.
- Nathaniel P Hoyle, Estere Seinkmane, Marrit Putker, Kevin A Feeney, Toke P Krogager, Johanna E Chesham, Liam K Bray, Justyn M Thomas, Ken Dunn, John Blaikley, and John S O'Neill. Circadian actin dynamics drive rhythmic fibroblast mobilization during wound healing. *Science translational medicine*, 9(415):eaal2774, nov 2017. ISSN 1946-6242. doi: 10.1126/scitranslmed.aal2774. URL <http://www.ncbi.nlm.nih.gov/pubmed/29118260><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5837001>.
- Youna Hu, Alena Shmygelska, David Tran, Nicholas Eriksson, Joyce Y. Tung, and David A. Hinds. GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. *Nature Communications*, 7:10448, feb 2016. ISSN 2041-1723. doi: 10.1038/ncomms10448. URL <http://www.nature.com/doifinder/10.1038/ncomms10448>.
- M. E. Hughes, J. B. Hogenesch, and K. Kornacker. JTK_CYCLE: An Efficient Nonparametric Algorithm for Detecting Rhythmic Components in Genome-Scale Data Sets. *Journal of Biological Rhythms*, 25(5):372–380, oct 2010. ISSN 0748-7304. doi: 10.1177/0748730410379711. URL <http://jbr.sagepub.com/cgi/doi/10.1177/0748730410379711>.

Bibliography

- Michael E. Hughes, Luciano DiTacchio, Kevin R. Hayes, Christopher Vollmers, S. Pulivarthy, Julie E. Baggs, Satchidananda Panda, John B. Hogenesch, AM Curtis, GA Fitzgerald, MH Hastings, AB Reddy, ES Maywood, EB Klerman, F Levi, U Schibler, F Halberg, G Cornelissen, W Ulmer, M Blank, W Hrushesky, CH Ko, JS Takahashi, M Stratmann, U Schibler, B Kornmann, O Schaad, H Bujard, JS Takahashi, U Schibler, SH Yoo, S Yamazaki, PL Lowrey, K Shimomura, CH Ko, U Schibler, P Sassone-Corsi, KF Storch, C Paz, J Signorovitch, E Raviola, B Pawlyk, KF Storch, O Lipan, I Leykin, N Viswanathan, FC Davis, SL Harmer, JB Hogenesch, M Straume, HS Chang, B Han, S Panda, MP Antoch, BH Miller, AI Su, AB Schook, HR Ueda, A Matsumoto, M Kawamura, M Iino, T Tanimura, Y Lin, M Han, B Shimada, L Wang, TM Ghibler, GE Duffield, JD Best, BH Meurers, A Bittner, JJ Loros, MF Ceriani, JB Hogenesch, M Yanovsky, S Panda, M Straume, RA Akhtar, AB Reddy, ES Maywood, JD Clayton, VM King, MJ McDonald, M Rosbash, A Claridge-Chang, H Wijnen, F Naef, C Boothroyd, N Rajewsky, E Nagoshi, C Saini, C Bauer, T Laroche, F Naef, JE Baggs, TS Price, L DiTacchio, S Panda, GA FitzGerald, M Hughes, L Deharo, SR Pulivarthy, J Gu, K Hayes, M Straume, S Wichert, K Fokianos, K Strimmer, JD Storey, R Tibshirani, JD Storey, W Xiao, JT Leek, RG Tompkins, RW Davis, C Grundschober, F Delaunay, A Puhlfhofer, G Triqueneaux, V Laudet, GJ Menger, GC Allen, N Neuendorff, SS Nahm, TL Thomas, CA Feillet, JA Ripperger, MC Magnone, A Dulloo, U Albrecht, KA Stokkan, S Yamazaki, H Tei, Y Sakaki, M Menaker, U Klingmuller, A Bauer, S Bohl, PJ Nickel, and K Breitkopf. Harmonics of Circadian Gene Transcription in Mammals. *PLoS Genetics*, 5(4):e1000442, apr 2009. ISSN 1553-7404. doi: 10.1371/journal.pgen.1000442. URL <http://dx.plos.org/10.1371/journal.pgen.1000442>.
- Michael E. Hughes, Hee-Kyung Hong, Jason L. Chong, Alejandra A. Indacochea, Samuel S. Lee, Michael Han, Joseph S. Takahashi, and John B. Hogenesch. Brain-Specific Rescue of Clock Reveals System-Driven Transcriptional Rhythms in Peripheral Tissue. *PLoS Genetics*, 8(7):e1002835, jul 2012. ISSN 1553-7404. doi: 10.1371/journal.pgen.1002835. URL <http://dx.plos.org/10.1371/journal.pgen.1002835>.
- Jacob J Hughey, Trevor Hastie, and Atul J Butte. ZeitZeiger: supervised learning for high-dimensional data from an oscillatory system. *Nucleic acids research*, 44(8):e80, may 2016. ISSN 1362-4962. doi: 10.1093/nar/gkw030. URL <http://www.ncbi.nlm.nih.gov/pubmed/26819407><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4856978>.
- Alan L. Hutchison, Mark Maienschein-Cline, Andrew H. Chiang, S. M. Ali Tabei, Herman Gudjonson, Neil Bahroos, Ravi Allada, and Aaron R. Dinner. Improved Statistical Methods Enable Greater Sensitivity in Rhythm Detection for Genome-Wide Data. *PLOS Computational Biology*, 11(3):e1004094, mar 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1004094. URL <http://dx.plos.org/10.1371/journal.pcbi.1004094>.
- Alon Kalo, Itamar Kanter, Amit Shrager, Jonathan Sheinberger, Hadar Tzemach, Noa Kinor, Robert H. Singer, Timothée Lionnet, and Yaron Shav-Tal. Cellular Levels of Signaling Factors Are Sensed by β -actin Alleles to Modulate Transcriptional Pulse Intensity. *Cell Reports*, 11(3):419–432, apr 2015. ISSN 2211-1247. doi: 10.1016/J.CELREP.2015.03.039. URL <http://www.sciencedirect.com/science/article/pii/S2211124715003083>.

- Maren Keller, Jeannine Mazuch, Ute Abraham, Gina D Eom, Erik D Herzog, Hans-Dieter Volk, Achim Kramer, and Bert Maier. A circadian clock in macrophages controls inflammatory immune responses. *Proceedings of the National Academy of Sciences of the United States of America*, 106(50):21407–12, dec 2009. ISSN 1091-6490. doi: 10.1073/pnas.0906361106. URL <http://www.ncbi.nlm.nih.gov/pubmed/19955445><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2795539>.
- Yong Hoon Kim, Sajid A Marhon, Yuxiang Zhang, David J Steger, Kyoung-Jae Won, and Mitchell A Lazar. Rev-erba dynamically modulates chromatin looping to control circadian gene transcription. *Science (New York, N.Y.)*, 359(6381):1274–1277, mar 2018. ISSN 1095-9203. doi: 10.1126/science.aao6891. URL <http://www.ncbi.nlm.nih.gov/pubmed/29439026><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5995144>.
- Caroline H. Ko and Joseph S. Takahashi. Molecular components of the mammalian circadian clock. *Human molecular genetics*, 15 Spec No(suppl_2):R271–7, oct 2006. ISSN 0964-6906. doi: 10.1093/hmg/ddl207. URL <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddl207>http://academic.oup.com/hmg/article/15/suppl_2/R271/624758/Molecular-components-of-the-mammalian-circadianhttp://hmg.oxfordjournals.org/content/15/suppl_2/R271.full.
- Anja Korenčič, Rok Košir, Grigory Bordyugov, Robert Lehmann, Damjana Rozman, Hanspeter Herzel, J. Bass, J. S. Takahashi, H. R. Ueda, S. Panda, M. E. Hughes, L. Fu, C. C. Lee, X. Yang, M. Keller, F. Levi, U. Schibler, J. Yan, H. Wang, Y. Liu, C. Shao, K.-F. Storch, R. Košir, H. Oster, S. Damerow, R. A. Hut, G. Eichele, D. B. Forger, C. S. Peskin, H. P. Mirsky, A. C. Liu, D. K. Welsh, S. A. Kay, F. J. Doyle, A. Relogio, A. Korenčič, L. Bintu, L. V. Sharova, C. C. Friedel, L. Doelken, Z. Ruzsics, U. H. Koszinowski, R. Zimmer, D. M. Suter, C. Lee, J. P. Etchegaray, F. R. Cagampang, A. S. Loudon, S. M. Reppert, N. Preitner, E. E. Hamilton, S. A. Kay, G. Rey, K. Vanselow, R. Košir, A. C. A. Meireles-Filho, A. F. Bardet, J. O. Yanez-Cuna, G. Stampfel, A. Stark, A. S. Hansen, E. K. O'Shea, N. Koike, A. Bugge, D. Feng, H. Cho, S. Gery, H. Reinke, A. Gerber, S. Kojima, E. L. Sher-Chen, C. B. Green, C. Jouffe, P. O. Westermarck, H. Herzel, M. Ukai-Tadenuma, K. Oishi, K. Bozek, D. W. Huang, B. T. Sherman, R. A. Lempicki, M. Ohsugi, X. Luo, Y. Ikeda, K. L. Parker, S. Mora, J. E. Pessin, S. Pikkarainen, H. Tokola, R. Kerkela, H. Ruskoaho, T. Hai, M. G. Hartman, K. Oishi, J. S. Menet, J. Rodriguez, K. C. Abruzzi, M. Rosbash, B. Schwanhauser, M. S. Robles, J. Cox, M. Mann, D. Mauvoisin, Z. Ouyang, Q. Zhou, and W. H. Wong. Timing of circadian genes in mammalian tissues. *Scientific Reports*, 4:1349–1354, jul 2014. ISSN 2045-2322. doi: 10.1038/srep05782. URL <http://www.nature.com/articles/srep05782>.
- Benoît Kornmann, Olivier Schaad, Hermann Bujard, Joseph S Takahashi, and Ueli Schibler. System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. *PLoS biology*, 5(2):e34, feb 2007. ISSN 1545-7885. doi: 10.1371/journal.pbio.0050034. URL <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0050034>.
- Anna Kriebs, Sabine D Jordan, Erin Soto, Emma Henriksson, Colby R Sandate, Megan E Vaughan, Alanna B Chan, Drew Duglan, Stephanie J Papp, Anne-Laure Huber, Megan E

Bibliography

- Afetian, Ruth T Yu, Xuan Zhao, Michael Downes, Ronald M Evans, and Katja A Lamia. Circadian repressors CRY1 and CRY2 broadly interact with nuclear receptors and modulate transcriptional activity. *Proceedings of the National Academy of Sciences of the United States of America*, 114(33):8776–8781, jul 2017. ISSN 1091-6490. doi: 10.1073/pnas.1704955114. URL <http://www.ncbi.nlm.nih.gov/pubmed/28751364><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5565439>.
- Katja A Lamia, Kai-Florian Storch, and Charles J Weitz. Physiological significance of a peripheral tissue circadian clock. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39):15172–7, sep 2008. ISSN 1091-6490. doi: 10.1073/pnas.0806717105. URL <http://www.pnas.org/content/105/39/15172.short><http://www.ncbi.nlm.nih.gov/pubmed/18779586><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2532700>.
- Katja A. Lamia, Stephanie J. Papp, Ruth T. Yu, Grant D. Barish, N. Henriette Uhlenhaut, Johan W. Jonker, Michael Downes, and Ronald M. Evans. Cryptochromes mediate rhythmic repression of the glucocorticoid receptor. *Nature*, 480(7378):552–6, dec 2011. ISSN 1476-4687. doi: 10.1038/nature10700. URL <http://www.nature.com/doifinder/10.1038/nature10700><http://www.ncbi.nlm.nih.gov/pubmed/22170608><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3245818><http://dx.doi.org/10.1038/nature10700>.
- Gwendal Le Martelot, Donatella Canella, Laura Symul, Eugenia Migliavacca, Federica Giarlardi, Robin Liechti, Olivier Martin, Keith Harshman, Mauro Delorenzi, Béatrice Desvergne, Winship Herr, Bart Deplancke, Ueli Schibler, Jacques Rougemont, Nicolas Guex, Nouria Hernandez, Felix Naef, and the CycliX Consortium. Genome-Wide RNA Polymerase II Profiles and RNA Accumulation Reveal Kinetics of Transcription and Associated Epigenetic Changes During Diurnal Cycles. *PLoS Biology*, 10(11):e1001442, nov 2012. ISSN 1545-7885. doi: 10.1371/journal.pbio.1001442. URL <http://dx.plos.org/10.1371/journal.pbio.1001442>.
- Ning Leng, Li-Fang Chu, Chris Barry, Yuan Li, Jeeha Choi, Xiaomao Li, Peng Jiang, Ron M Stewart, James A Thomson, and Christina Kendziorski. Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. *Nature Methods*, 12(10):947–950, aug 2015. ISSN 1548-7091. doi: 10.1038/nmeth.3549. URL <http://www.nature.com/articles/nmeth.3549><http://www.nature.com/doifinder/10.1038/nmeth.3549>.
- Feng Liang, Rui Paulo, German Molina, Merlise A Clyde, and Jim O Berger. Mixtures of g -Priors for Bayesian Variable Selection. *Journal of the American Statistical Association*, 103(481):410–423, mar 2008. ISSN 0162-1459. doi: 10.1198/016214507000001337. URL <http://www.tandfonline.com/doi/abs/10.1198/016214507000001337>.
- Erez Lieberman-Aiden, Nynke L van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, Bryan R Lajoie, Peter J Sabo, Michael O Dorschner, Richard Sandstrom, Bradley Bernstein, M A Bender, Mark Groudine, Andreas Gnirke, John Stamatoyannopoulos, Leonid A Mirny, Eric S Lander, and Job Dekker. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New*

- York, N.Y.), 326(5950):289–93, oct 2009. ISSN 1095-9203. doi: 10.1126/science.1181369. URL <http://www.ncbi.nlm.nih.gov/pubmed/19815776><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2858594>.
- Phillip L Lowrey and Joseph S Takahashi. Mammalian circadian biology: elucidating genome-wide levels of temporal organization. *Annual review of genomics and human genetics*, 5:407–41, jan 2004. ISSN 1527-8204. doi: 10.1146/annurev.genom.5.061903.175925. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3770722&tool=pmcentrez&rendertype=abstract>.
- S Luck, K Thurley, P F Thaben, and P O Westermark. Rhythmic degradation explains and unifies circadian transcriptome and proteome data. *Cell Rep*, 9(2):741–751, 2014. doi: 10.1016/j.celrep.2014.09.021. URL <http://www.ncbi.nlm.nih.gov/pubmed/25373909>.
- Biliana Marcheva, Kathryn Moynihan Ramsey, Ethan D. Buhr, Yumiko Kobayashi, Hong Su, Caroline H. Ko, Ganka Ivanova, Chiaki Omura, Shelley Mo, Martha H. Vitaterna, James P. Lopez, Louis H. Philipson, Christopher A. Bradfield, Seth D. Crosby, Lellean JeBailey, Xiaozhong Wang, Joseph S. Takahashi, and Joseph Bass. Disruption of the clock components CLOCK and BMAL1 leads to hypoinsulinaemia and diabetes. *Nature*, 466(7306):627–631, jul 2010. ISSN 0028-0836. doi: 10.1038/nature09253. URL <http://www.nature.com/articles/nature09253>.
- Daniel Mauvoisin, Florian Atger, Loïc Dayon, Antonio Núñez Galindo, Jingkui Wang, Eva Martin, Laetitia Da Silva, Ivan Montoliu, Sebastiano Collino, Francois-Pierre Martin, Joanna Ratajczak, Carles Cantó, Martin Kussmann, Felix Naef, and Frédéric Gachon. Circadian and Feeding Rhythms Orchestrate the Diurnal Liver Acetylome. *Cell Reports*, 20(7):1729–1743, aug 2017. ISSN 22111247. doi: 10.1016/j.celrep.2017.07.065. URL <http://www.ncbi.nlm.nih.gov/pubmed/28813682><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5568034><http://linkinghub.elsevier.com/retrieve/pii/S2211124717310586>.
- Jérôme Mermet, Jake Yeung, Clémence Hurni, Daniel Mauvoisin, Kyle Gustafson, Céline Jouffe, Damien Nicolas, Yann Emmenegger, Cédric Gobet, Paul Franken, Frédéric Gachon, and Félix Naef. Clock-dependent chromatin topology modulates circadian transcription and behavior. *Genes & development*, mar 2018. ISSN 1549-5477. doi: 10.1101/gad.312397.118. URL <http://www.ncbi.nlm.nih.gov/pubmed/29572261>.
- Borbala Mifsud, Filipe Tavares-Cadete, Alice N Young, Robert Sugar, Stefan Schoenfelder, Lauren Ferreira, Steven W Wingett, Simon Andrews, William Grey, Philip A Ewels, Bram Herman, Scott Happe, Andy Higgs, Emily LeProust, George A Follows, Peter Fraser, Nicholas M Luscombe, and Cameron S Osborne. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nature Genetics* 2015 47:6, 47(6):598, may 2015. ISSN 1546-1718. doi: 10.1038/ng.3286. URL <https://www.nature.com/articles/ng.3286>.
- Jennifer A Mohawk, Carla B Green, and Joseph S Takahashi. Central and peripheral circadian clocks in mammals. *Annual review of neuroscience*, 35:445–62, jan 2012. ISSN 1545-4126.

Bibliography

- doi: 10.1146/annurev-neuro-060909-153128. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3710582&tool=pmcentrez&rendertype=abstract>.
- Emi Nagoshi, Camille Saini, Christoph Bauer, Thierry Laroche, Felix Naef, and Ueli Schibler. Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell*, 119(5):693–705, nov 2004. ISSN 0092-8674. doi: 10.1016/j.cell.2004.11.015. URL <http://www.sciencedirect.com/science/article/pii/S0092867404010542>.
- Damien Nicolas, Benjamin Zoller, David M Suter, and Felix Naef. Modulation of transcriptional burst frequency by histone acetylation. *Proceedings of the National Academy of Sciences of the United States of America*, page 201722330, jun 2018. ISSN 1091-6490. doi: 10.1073/pnas.1722330115. URL <http://www.ncbi.nlm.nih.gov/pubmed/29915087>.
- Elphège P. Nora, Bryan R. Lajoie, Edda G. Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, Nynke L. van Berkum, Johannes Meisig, John Sedat, Joost Gribnau, Emmanuel Barillot, Nils Blüthgen, Job Dekker, and Edith Heard. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 485(7398):381–385, apr 2012. ISSN 0028-0836. doi: 10.1038/nature11049. URL <http://www.nature.com/doifinder/10.1038/nature11049>.
- Gabriel Oh, Sasha Ebrahimi, Matthew Carlucci, Aiping Zhang, Akhil Nair, Daniel E. Groot, Viviane Labrie, Peixin Jia, Edward S. Oh, Richie H. Jeremian, Miki Susic, Tenjin C. Shrestha, Martin R. Ralph, Juozas Gordevičius, Karolis Koncėvičius, and Art Petronis. Cytosine modifications exhibit circadian oscillations that are involved in epigenetic diversity and aging. *Nature Communications*, 9(1):644, dec 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-03073-7. URL <http://www.nature.com/articles/s41467-018-03073-7>.
- Y Ouyang, C R Andersson, T Kondo, S S Golden, and C H Johnson. Resonating circadian clocks enhance fitness in cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 95(15):8660–4, jul 1998. ISSN 0027-8424. doi: 10.1073/PNAS.95.15.8660. URL <http://www.ncbi.nlm.nih.gov/pubmed/9671734><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC21132>.
- Bryn M. Owen, David J. Mangelsdorf, and Steven A. Kliewer. Tissue-specific actions of the metabolic hormones FGF15/19 and FGF21. *Trends in Endocrinology & Metabolism*, 26(1):22–29, jan 2015. ISSN 1043-2760. doi: 10.1016/J.TEM.2014.10.002. URL <https://www.sciencedirect.com/science/article/pii/S104327601400174X>.
- Weihong Pan and Abba J. Kastin. The Blood-Brain Barrier. *The Neuroscientist*, 23(2):124–136, apr 2017. ISSN 1073-8584. doi: 10.1177/1073858416639005. URL <http://journals.sagepub.com/doi/10.1177/1073858416639005>.
- Carrie L Partch, Carla B Green, and Joseph S Takahashi. Molecular architecture of the mammalian circadian clock. *Trends in cell biology*, 24(2):90–9, mar 2014. ISSN 1879-3088.

- doi: 10.1016/j.tcb.2013.07.002. URL <http://www.sciencedirect.com/science/article/pii/S096289241300113X>.
- Jennifer E. Phillips-Cremins, Michael E.G. Sauria, Amartya Sanyal, Tatiana I. Gerasimova, Bryan R. Lajoie, Joshua S.K. Bell, Chin-Tong Ong, Tracy A. Hookway, Changying Guo, Yuhua Sun, Michael J. Bland, William Wagstaff, Stephen Dalton, Todd C. McDevitt, Ranjan Sen, Job Dekker, James Taylor, and Victor G. Corces. Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment. *Cell*, 153(6):1281–1295, jun 2013. URL <http://linkinghub.elsevier.com/retrieve/pii/S0092867413005291><https://www.sciencedirect.com/science/article/pii/S0092867413005291>.
- Ana Pombo and Niall Dillon. Three-dimensional genome architecture: players and mechanisms. *Nature Reviews Molecular Cell Biology*, 16(4):245–257, apr 2015. ISSN 1471-0072. doi: 10.1038/nrm3965. URL <http://www.nature.com/articles/nrm3965>.
- Sofia A. Quinodoz, Noah Ollikainen, Barbara Tabak, Ali Palla, Jan Marten Schmidt, Elizabeth Detmar, Mason M. Lai, Alexander A. Shishkin, Prashant Bhat, Yodai Takei, Vickie Trinh, Erik Aznauryan, Pamela Russell, Christine Cheng, Marko Jovanovic, Amy Chow, Long Cai, Patrick McDonel, Manuel Garber, and Mitchell Guttman. Higher-Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell*, jun 2018. ISSN 0092-8674. doi: 10.1016/J.CELL.2018.05.024. URL <https://www.sciencedirect.com/science/article/pii/S0092867418306366>{#}bib11.
- Maria S. Robles, Sean J. Humphrey, and Matthias Mann. Phosphorylation Is a Central Mechanism for Circadian Control of Metabolism and Physiology. *Cell Metab*, 25(1):118–127, jan 2017. ISSN 1550-4131. doi: 10.1016/j.cmet.2016.10.004. URL <https://www.sciencedirect.com/science/article/pii/S1550413116305356><http://www.ncbi.nlm.nih.gov/pubmed/27818261>.
- Till Roenneberg. Twitter as a means to study temporal behaviour. *Current Biology*, 27(17):R830–R832, sep 2017. ISSN 0960-9822. doi: 10.1016/J.CUB.2017.08.005. URL <https://www.sciencedirect.com/science/article/pii/S096098221731014X>.
- Till Roenneberg and Martha Merrow. The Circadian Clock and Human Health. *Current Biology*, 26(10):R432–R443, may 2016. ISSN 0960-9822. doi: 10.1016/J.CUB.2016.04.011. URL <https://www.sciencedirect.com/science/article/pii/S0960982216303335>.
- Marc D Ruben, Gang Wu, David F Smith, Robert E Schmidt, Lauren J Francey, Ron C Anafi, and John B Hogenesch. A population-based human enCYCLOPedia for circadian medicine. *bioRxiv*, page 301580, apr 2018. doi: 10.1101/301580. URL <https://www.biorxiv.org/content/early/2018/04/14/301580>.
- C Saini, D M Suter, A Liani, P Gos, and U Schibler. The mammalian circadian timing system: synchronization of peripheral clocks. *Cold Spring Harbor symposia on quantitative biology*, 76(0):39–47, jan 2011. ISSN 1943-4456. doi: 10.1101/sqb.2011.76.010918. URL <http://symposium.cshlp.org/content/76/39.long>.

Bibliography

- Christoph Scheiermann, Yuya Kunisaki, Daniel Lucas, Andrew Chow, Jung-Eun Jang, Dachuan Zhang, Daigo Hashimoto, Miriam Merad, and Paul S. Frenette. Adrenergic Nerves Govern Circadian Leukocyte Recruitment to Tissues. *Immunity*, 37(2):290–301, aug 2012. URL <https://www.sciencedirect.com/science/article/pii/S1074761312002920>.
- Christoph Scheiermann, Yuya Kunisaki, and Paul S Frenette. Circadian control of the immune system. *Nature reviews. Immunology*, 13(3):190–8, mar 2013. ISSN 1474-1741. doi: 10.1038/nri3386. URL <http://dx.doi.org/10.1038/nri3386>.
- Ueli Schibler, Ivana Gotic, Camille Saini, Pascal Gos, Thomas Curie, Yann Emmenegger, Flore Sinturel, Pauline Gosselin, Alan Gerber, Fabienne Fleury-Olela, Gianpaolo Rando, Maud Demarque, and Paul Franken. Clock-Talk: Interactions between Central and Peripheral Circadian Oscillators in Mammals. *Cold Spring Harbor symposia on quantitative biology*, 80:223–32, jan 2015. ISSN 1943-4456. doi: 10.1101/sqb.2015.80.027490. URL <http://www.ncbi.nlm.nih.gov/pubmed/26683231>.
- Adrien Senecal, Brian Munsky, Florence Proux, Nathalie Ly, Floriane E. Braye, Christophe Zimmer, Florian Mueller, and Xavier Darzacq. Transcription Factors Modulate c-Fos Transcriptional Bursts. *Cell Reports*, 8(1):75–83, jul 2014. ISSN 2211-1247. doi: 10.1016/J.CELREP.2014.05.053. URL <https://www.sciencedirect.com/science/article/pii/S2211124714004471>.
- Guangsen Shi, David Wu, Louis J Ptáček, and Ying-Hui Fu. Human genetics and sleep behavior. *Current Opinion in Neurobiology*, 44:43–49, jun 2017. ISSN 0959-4388. doi: 10.1016/J.CONB.2017.02.015. URL <https://www.sciencedirect.com/science/article/pii/S095943881630229X?via%3Dihub>.
- Marieke Simonis, Petra Klous, Erik Splinter, Yuri Moshkin, Rob Willemsen, Elzo de Wit, Bas van Steensel, and Wouter de Laat. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature genetics*, 38(11):1348–54, nov 2006. ISSN 1061-4036. doi: 10.1038/ng1896. URL <http://www.ncbi.nlm.nih.gov/pubmed/17033623>.
- Flore Sinturel, Alan Gerber, Daniel Mauvoisin, Jingkui Wang, David Gatfield, Jeremy J. Stubblefield, Carla B. Green, Frédéric Gachon, and Ueli Schibler. Diurnal Oscillations in Liver Mass and Cell Size Accompany Ribosome Assembly Cycles. *Cell*, 169(4):651–663.e14, may 2017. ISSN 00928674. doi: 10.1016/j.cell.2017.04.015. URL <http://www.sciencedirect.com/science/article/pii/S0092867417304282http://linkinghub.elsevier.com/retrieve/pii/S0092867417304282>.
- Jonathan Aryeh Sobel, Irina Krier, Teemu Andersin, Sunil Raghav, Donatella Canella, Federica Gilardi, Alexandra Styliani Kalantzi, Guillaume Rey, Benjamin Weger, Frédéric Gachon, Matteo Dal Peraro, Nouria Hernandez, Ueli Schibler, Bart Deplancke, Felix Naef, and CycliX Consortium. Transcriptional regulatory logic of the diurnal cycle in the mouse liver. *PLOS Biology*, 15(4):e2001069, apr 2017. ISSN 1545-7885. doi: 10.1371/journal.pbio.2001069. URL <http://dx.plos.org/10.1371/journal.pbio.2001069>.

- Kai-Florian Storch, Ovidiu Lipan, Igor Leykin, N Viswanathan, Fred C Davis, Wing H Wong, and Charles J Weitz. Extensive and divergent circadian gene expression in liver and heart. *Nature*, 417(6884):78–83, may 2002. ISSN 0028-0836. doi: 10.1038/nature744. URL <http://dx.doi.org/10.1038/nature744>.
- Joseph S. Takahashi. Transcriptional architecture of the mammalian circadian clock. *Nature Reviews Genetics*, 18(3):164–179, mar 2017. ISSN 1471-0056. doi: 10.1038/nrg.2016.150. URL <http://www.nature.com/articles/nrg.2016.150>.
- P. F. Thaben and P. O. Westermark. Detecting Rhythms in Time Series with RAIN. *Journal of Biological Rhythms*, 29(6):391–400, dec 2014. ISSN 0748-7304. doi: 10.1177/0748730414553029. URL <http://jbr.sagepub.com/cgi/doi/10.1177/0748730414553029>.
- Kevin Thurley, Christopher Herbst, Felix Wesener, Barbara Koller, Thomas Wallach, Bert Maier, Achim Kramer, and Pål O Westermark. Principles for circadian orchestration of metabolic pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 114(7):1572–1577, feb 2017. ISSN 1091-6490. doi: 10.1073/pnas.1613103114. URL <http://www.ncbi.nlm.nih.gov/pubmed/28159888><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5321018>.
- Alexandra J. Trott and Jerome S. Menet. Regulation of circadian clock transcriptional output by CLOCK:BMAL1. *PLOS Genetics*, 14(1):e1007156, jan 2018. ISSN 1553-7404. doi: 10.1371/journal.pgen.1007156. URL <http://dx.plos.org/10.1371/journal.pgen.1007156>.
- Bas van Steensel and Andrew S. Belmont. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell*, 169(5):780–791, may 2017. ISSN 0092-8674. doi: 10.1016/J.CELL.2017.04.022. URL <https://www.sciencedirect.com/science/article/pii/S0092867417304737><https://www.sciencedirect.com/science/article/pii/S0092867417304737#!>
- Jingkui Wang, Daniel Mauvoisin, Eva Martin, Florian Atger, Antonio Núñez N Galindo, Loïc Dayon, Federico Sizzano, Alessio Palini, Martin Kussmann, Patrice Waridel, Manfredo Quadroni, V Dulic, Felix Naef, Frédéric Gachon, Vjekoslav Dulić, Felix Naef, and Frédéric Gachon. Nuclear Proteomics Uncovers Diurnal Regulatory Landscapes in Mouse Liver. *Cell Metab*, 25(1):102–117, 2017. ISSN 15504131. doi: 10.1016/j.cmet.2016.10.003. URL <http://www.ncbi.nlm.nih.gov/pubmed/27818260>.
- Jingkui Wang, Laura Symul, Jake Yeung, Cédric Gobet, Jonathan Sobel, Sarah Lück, Pål O Westermark, Nacho Molina, and Felix Naef. Circadian clock-dependent and -independent posttranscriptional regulation underlies temporal mRNA accumulation in mouse liver. *Proceedings of the National Academy of Sciences of the United States of America*, 115(8):E1916–E1925, feb 2018. ISSN 1091-6490. doi: 10.1073/pnas.1715225115. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5828596>.

Bibliography

- Pål O. Westermarck and Hanspeter Herzl. Mechanism for 12 Hr Rhythm Generation by the Circadian Clock. *Cell Reports*, 3(4):1228–1238, apr 2013. ISSN 22111247. doi: 10.1016/j.celrep.2013.03.013. URL <https://www.sciencedirect.com/science/article/pii/S2211124713001204>.
- Nicole Wittenbrink, Bharath Ananthasubramaniam, Mirjam Münch, Barbara Koller, Bert Maier, Charlotte Weschke, Frederik Bes, Jan de Zeeuw, Claudia Nowozin, Amely Wahnschaffe, Sophia Wisniewski, Mandy Zaleska, Osnat Bartok, Reut Ashwal-Fluss, Hedwig Lammert, Hanspeter Herzl, Michael Hummel, Sebastian Kadener, Dieter Kunz, and Achim Kramer. High-accuracy determination of internal circadian time from a single blood sample. *The Journal of Clinical Investigation*, jun 2018. ISSN 0021-9738. doi: 10.1172/JCI120874. URL <https://www.jci.org/articles/view/120874>.
- Yichi Xu, Weimin Guo, Ping Li, Yan Zhang, Meng Zhao, Zenghua Fan, Zhihu Zhao, and Jun Yan. Long-Range Chromosome Interactions Mediated by Cohesin Shape Circadian Gene Expression. *PLOS Genetics*, 12(5):e1005992, may 2016. ISSN 1553-7404. doi: 10.1371/journal.pgen.1005992. URL <http://dx.plos.org/10.1371/journal.pgen.1005992>.
- Bin Yang, Jennifer B. Treweek, Rajan P. Kulkarni, Benjamin E. Deverman, Chun-Kan Chen, Eric Lubeck, Sheel Shah, Long Cai, and Viviana Gradinaru. Single-Cell Phenotyping within Transparent Intact Tissue through Whole-Body Clearing. *Cell*, 158(4):945–958, aug 2014. ISSN 0092-8674. doi: 10.1016/J.CELL.2014.07.017. URL <https://www.sciencedirect.com/science/article/pii/S0092867414009313>#!
- Y. Yang. Can the strengths of AIC and BIC be shared? A conflict between model identification and regression estimation. *Biometrika*, 92(4):937–950, dec 2005. ISSN 0006-3444. doi: 10.1093/biomet/92.4.937. URL <http://biomet.oxfordjournals.org/cgi/doi/10.1093/biomet/92.4.937>.
- Jake Yeung, Jérôme Mermet, Céline Jouffe, Julien Marquis, Aline Charpagne, Frédéric Gachon, and Felix Naef. Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs. *Genome research*, 28(2):182–191, feb 2018. ISSN 1549-5469. doi: 10.1101/gr.222430.117. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5793782>.
- Seung-Hee Yoo, Shin Yamazaki, Phillip L Lowrey, Kazuhiro Shimomura, Caroline H Ko, Ethan D Buhr, Sandra M Siepka, Hee-Kyung Hong, Won Jun Oh, Ook Joon Yoo, Michael Menaker, and Joseph S Takahashi. PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proceedings of the National Academy of Sciences of the United States of America*, 101(15):5339–46, apr 2004. ISSN 0027-8424. doi: 10.1073/pnas.0308709101. URL <http://www.pnas.org/content/101/15/5339.abstract>.
- Arnold Zellner. Bayesian Estimation and Prediction Using Asymmetric Loss Functions. *Journal of the American Statistical Association*, 81(394):446–451, jun 1986. ISSN 0162-1459. doi: 10.1080/01621459.1986.10478289. URL <http://www.tandfonline.com/doi/abs/10.1080/01621459.1986.10478289>.

- Yinxu Zhan, Luca Mariani, Iros Barozzi, Edda G Schulz, Nils Blüthgen, Michael Stadler, Guido Tiana, and Luca Giorgetti. Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. *Genome research*, 27(3):479–490, jan 2017. ISSN 1549-5469. doi: 10.1101/gr.212803.116. URL <http://www.ncbi.nlm.nih.gov/pubmed/28057745><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5340975>.
- Ray Zhang, Nicholas F Lahens, Heather I Ballance, Michael E Hughes, and John B Hogenesch. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45):16219–24, nov 2014. ISSN 1091-6490. doi: 10.1073/pnas.1408886111. URL <http://www.pnas.org/content/111/45/16219.short>.
- Yubo Zhang, Chee-Hong Wong, Ramon Y. Birnbaum, Guoliang Li, Rebecca Favaro, Chew Yee Ngan, Joanne Lim, Eunice Tai, Huay Mei Poh, Eleanor Wong, Fabianus Hendriyan Mulawadi, Wing-Kin Sung, Silvia Nicolis, Nadav Ahituv, Yijun Ruan, and Chia-Lin Wei. Chromatin connectivity maps reveal dynamic promoter–enhancer long-range associations. *Nature*, 504(7479):306–310, dec 2013. ISSN 0028-0836. doi: 10.1038/nature12716. URL <http://www.nature.com/articles/nature12716>.
- Honglei Zhao, Emmanouil G. Sifakis, Noriyuki Sumida, Lluís Millán-Ariño, Barbara A. Scholz, J. Peter Svensson, Xingqi Chen, Anna L. Ronnegren, Carolina Diettrich Mallet de Lima, Farzaneh Shahin Varnoosfaderani, Chengxi Shi, Olga Loseva, Samer Yammine, Maria Israelsson, Li-Sophie Rathje, Balázs Némethi, Erik Fredlund, Thomas Helleday, Márta P. Imreh, and Anita Göndör. PARP1- and CTCF-Mediated Interactions between Active and Repressed Chromatin at the Lamina Promote Oscillating Transcription. *Molecular Cell*, 59(6):984–997, sep 2015. ISSN 1097-2765. doi: 10.1016/J.MOLCEL.2015.07.019. URL <https://www.sciencedirect.com/science/article/pii/S1097276515005778>.
- Bokai Zhu, Qiang Zhang, Yinghong Pan, Emily M. Mace, Brian York, Athanasios C. Antoulas, Clifford C. Dacso, and Bert W. O’Malley. A Cell-Autonomous Mammalian 12 hr Clock Coordinates Metabolic and Stress Rhythms. *Cell Metabolism*, 25(6):1305–1319.e9, jun 2017. ISSN 1550-4131. doi: 10.1016/J.CMET.2017.05.004. URL <https://www.sciencedirect.com/science/article/pii/S1550413117302905>.

Jake Yeung

Education

- Ph.D. in Biotechnology & Bioengineering,** 2018 (Expected)
École Polytechnique Fédérale de Lausanne (EPFL),
Thesis: Tissue-specific circadian transcriptional regulation.
- M.Sc. in Genome Science and Technology,** 2014
University of British Columbia
GPA: 91.5/100,
Thesis: Identification of RNA binding proteins associated with differential splicing in neuroendocrine prostate cancer.
- B.A.Sc. in Chemical Engineering,** 2012
University of Waterloo
GPA: 90.3/100,
Graduated with Distinction and Dean's Honours,
Thesis: Mathematical modeling of a renewable energy hub.

Skills

- Bioinformatics:** Temporal analysis of RNA-Seq, 4C-Seq, DNase-Seq, ATAC-Seq, and smRNA FISH
- Engineering:** Machine learning, mathematical modeling, dynamical systems
- Computing:** R, Python, MATLAB, UNIX-like OS, shell scripts

Manuscripts in the Pipeline

- B. Weger, C. Gobet, J. Yeung, Eva Martin, Sonia Jimenez, Aurelie Balvay, Anne Foussier, Aline Charpagne, Chieh Jason Chou, Felix Naef, Frederic Gachon, *The gut microbiome is required for sex-specific diurnal rhythms of gene expression and metabolism*, (in revision). 2018

Manuscripts

- J. Mermet*, J. Yeung*, C. Hurni, D. Mauvoisin, K. Gustafson, C. Jouffe, D. Nicolas, Y. Emmenegger, C. Gobet, P. Franken, F. Gachon, F. Naef, *Clock-dependent chromatin topology modulates circadian transcription and behavior*, Genes & Development (accepted, click for draft). 2018
- J. Wang, L. Symul, J. Yeung, C. Gobet, J. Sobel, S. Luck, P. Westermarck, N. Molina, F. Naef, *Circadian clock –dependent and –independent post-transcriptional regulation underlies temporal mRNA accumulation in mouse liver*, PNAS. 2018
- J. Yeung*, J. Mermet*, C. Jouffe, J. Marquis, A. Charpagne, F. Gachon, F. Naef, *Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs*, Genome Research (featured on cover of February issue). 2018
- R. Shrestha, E. Hodzic, T. Sauerwald, P. Dao, K. Wang, J. Yeung, S. Anderson, F. Vandin, G. Haffari, C. Collins, S. C. Sahinalp, *HIT'nDRIVE: Patient-specific multi-driver gene prioritization for precision oncology*, Genome Research. 2017
- J. Mermet, J. Yeung, F. Naef, *Systems Chronobiology: Global Analysis of Gene Regulation in a 24-Hour Periodic World*, Cold Spring Harbor Perspectives in Biology. 2016
- A. Wyatt*, F. Mo*, ..., J. Yeung et. al. , *The inter-tumor transcriptome heterogeneity of high-risk primary prostate cancer*, Genome Biology. 2014

163

D. Lanen, J. Cocking, S. Walker, M. Fowler, R. Fraser, S. Young, L. Ahmadi, A. Thai, J. Yeung and A. Yip, *Economic Analysis of a green energy hub with energy storage under fixed and variable pricing structures*, International Conference on Smart Energy Grid Engineering. **2014**

Talks at Conferences

J. Mermet, J. Yeung, F. Naef, *Clock-dependent chromatin topology modulates circadian transcription and behavior*, Cold Spring Harbor Laboratory: Biology of Genomes, Cold Spring Harbor, 8-12 May. **2018**

J. Yeung, J. Mermet, F. Naef, *Transcription factor activity rhythms and tissue-specific chromatin interactions orchestrate circadian gene expression across organs*, Genomic Regulation 2018, Haute-Nendaz, Switzerland, 15-18 March. **2018**

J. Yeung, J. Mermet, F. Naef, *Interaction of Tissue-specific and Circadian Transcriptional Regulatory Networks*, European Biological Rhythms Society (EBRS), Amsterdam, Netherlands, 30 July - 3 Aug. **2017**

J. Yeung, J. Mermet, F. Naef, *Tissue-specific Circadian Transcriptional Regulation*, Gordon Conference Seminars (GRS), Stowe, USA, 15 July - 16 July. **2017**

J. Yeung, J. Mermet, F. Naef, *Integrated analysis reveals distinct roles of promoters underlying rhythmic gene expression*, Gene Regulation and Promoter Meeting: Eukaryotic Promoter Database 30th Anniversary Symposium, Lausanne, Switzerland, Sep 29-30. **2016**

J. Yeung, J. Mermet, F. Naef, *Integrating Functional Genomics Data Reveals Tissue-Dependent Mechanisms Underlying Circadian Gene Expression*, Society for Research on Biological Rhythms (SRBR), Tampa Bay, USA, May 21-25. **2016**
SRBR Merit Award

J. Yeung, *Identification of RNA binding proteins associated with differential splicing in neuroendocrine prostate cancer.*, 4th Annual Oncology, Bioinformatics, Genome Science Research Day, Vancouver, Canada Mar 21. **2014**
Silver award for outstanding research talk

Papers at Conferences

R. Shrestha*, E. Hodzic*, J. Yeung*, K. Wang, T. Sauerwald, P. Dao, S. Anderson, H. Beltran, M. Rubin, C. Collins, G. Haffari, S. C. Sahinalp, *HIT'nDRIVE: Multi-Driver Gene Prioritization based on Hitting Time*, Research in Computational Molecular Biology. **2013**

Awards

2017: Francophone Society of Chronobiology Travel Award (European Biological Rhythms Society in Amsterdam)

2015-2018: Natural Resources and Engineering Research Council of Canada Doctoral Fellowship

2012-2014: Natural Resources and Engineering Research Council of Canada Graduate Award

2012-2014: Genome Science and Technology Graduate Award

2013: The Sidhoo Charities Fund Travel Grant (Genome Informatics conference at Cold Spring Harbor)

2012: First Prize Walmart Green Student Challenge Competition

2011: University of Waterloo Undergraduate Research Assistantship

2010: University of Waterloo Dr. F. Hecker Foreign Exchange Scholarship

2007-2012: Dean's Honours List

2007: University of Waterloo President's Scholarship

Teaching and Mentoring Experience

Teaching Assistant, *Computational and Mathematical Modeling in Biology*, **2014-2018**
Prof. Naef,
EPFL.

Guest Lecturer, *Gene Set Analysis*,
Genome Science and Technology Laboratory Course: Genome Analysis Module,
UBC.

2013

Languages

English (native), French (conversational), Chinese (conversational, Cantonese and Mandarin), German (basic)

References

Felix Naef,

École Polytechnique Fédérale de Lausanne,

Email: felix.naef@epfl.ch.

Main PhD advisor

Paul Franken,

University of Lausanne,

Email: paul.franken@unil.ch.

Experimental collaborator and expert in sleep homeostasis. Together we work on temporal dynamics in gene expression during sleep-wake cycles and in response to acute sleep-deprivation.

Frederic Gachon,

Nestlé Institute of Health Sciences S.A.,

Email: frederic.gachon@epfl.ch.

Experimental collaborator and expert in circadian biology. His lab is involved in performing assays on most of the mouse experiments (RNA-Seq, ribosome profiling, nuclear proteomics, phosphoproteomics).

Hobbies

Eating homemade baked goods, playing Bach on the piano, and snowboarding in the Swiss alps.

