# CHARACTERIZATION OF THE MOLECULAR MECHANISMS ASSOCIATED WITH CIRCADIAN CLOCK CONTROL OF TRANSCRIPTION

A Dissertation

by

ALEXANDRA J. TROTT

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

| | |
|---|---|
| Chair of Committee, | Jerome S. Menet |
| Committee Members, | Paul E. Hardin |
| | Matthew S. Sachs |
| | Mary Bryk |
| Intercollegiate Faculty Chair, | Dorothy E. Shippen |

December 2018

Major Subject: Genetics

# ABSTRACT

The mammalian circadian clock, found in virtually every cell, coordinates rhythmic expression to control the daily regulation of biological functions. These cell autonomous circadian clocks rely on a molecular feedback loop involving the heterodimeric transcription factor CLOCK:BMAL1 to activate the transcription of repressor elements *Period* (*Per1/2*) and *Cryptochrome* (*Cry1/2*). PER and CRY feedback to inhibit CLOCK:BMAL1's transcriptional activation of this feedback loop, as well as expression of genes needed to regulate biological functions. However, recently a surprising disconnect has been found between CLOCK:BMAL1 genome-wide binding sites and CLOCK:BMAL1's target genes expression. The goal of this research was to investigate the mechanisms of how the circadian clock controls transcription and how disruption of the circadian clock leads to diseases.

To achieve this goal, we did a computational analysis of genome-wide sequencing datasets. We found that while CLOCK:BMAL1 promotes rhythmic nucleosome removal, this was not sufficient to activate these enhancers. Likely, these enhancers rely on cooperation of CLOCK:BMAL1 and the recruitment of ubiquitously expressed transcription factors, and not tissue-specific transcription factors to control transcription. These data were exemplified by analysis of how fasting effects the amplitude of rhythmically transcribed CLOCK:BMAL1 target genes, but not the expression of CLOCK:BMAL1. Altogether this suggest that CLOCK:BMAL1 creates transcriptionally permissive enhancers, priming their target genes for transcription activation and may

provide a new framework of how disruption of the circadian clock leads to pathological conditions.

To study disruption of the circadian clock we used a shift work model in rats to discover if this can emulate cardiovascular disease. Additionally, since temporal food consumption has been found to have preventive effect on shift work induced metabolic disorders, we restricted food intake to the night for a group of shift working rats. We found that five weeks of shift work in rats causes a significant increase in collagen deposition in the heart and that gene expression for shift work rats that had their food restricted, exhibited a significant up-regulation of profibrotic genes. Altogether, this suggests that five weeks of shift work in rats is capable of inducing cardiovascular disease through up-regulation of profibrotic collagen deposition in the heart.

# DEDICATION

This Dissertation is dedicated to my parents Kelvin and Deborah Trott for the constant love and support. I love you guys from here, to the moon, and beyond.

# ACKNOWLEDGEMENTS

After the longest period that I have every spent at a single school, today is the day: writing this note of thanks is the last touch on my dissertation. This period has been one of intense learning, not only scientifically, but also on a personal level. I would like to take this time to say thank you to the people who have supported and helped me so much.

First, I would like to say thank you to Dr. Jerome Menet. You took a chance on me when I needed to find a new lab and you have always been supportive of my career goals. You have been an amazing teacher and you have shown me what a good scientist (and person) should be.

I am also so grateful to have had the pleasure to work with such amazing lab members, in particular Josh Beytebiere and Ben Greenwell. Your constant support, feedback, collaborations, and of course brotherly teasing has made this time invaluable to me.

I would really like to thank my committee members, Dr. Paul Hardin, Dr. Matthew Sachs, and Dr. Mary Bryk your feedback, personal experiences, and professional guidance has provided me with the right tools to further my dissertation and become a better scientist. Additionally, I would like to thank everyone in the CJC group at TAMU for your guidance and for all the fun times at the SRBR.

I am also extremely grateful for all of my friends that I have made here at TAMU. The ability to either talk science or play soccer has always kept me on my toes and learning something new. I am especially thankful to Kelsey Schulte and Sarah Degarmo for

proofreading and giving me edits. I know that trying to read something that is not in your field can be a big challenge and I really appreciate you both taking the time to do that.

Lastly, I would like to thank my family. Especially, my parents for all their late nights helping me to learn to read or helping me to finish my homework. Your unconditional love and guidance are the only reasons I have made it this far. Finally, my husband Chris and my puppy Joey, you guys are always a supportive shoulder or cuddle partner and continuously keep my life entertaining.

CONTRIBUTORS AND FUNDING SOURCES

## Contributors

Research done in this dissertation was supervised by my Advisory Committee consisting of Dr. Jerome Menet (chair), Dr. Paul Hardin, Dr. Matt Sachs, and Dr. Mary Bryk from the departments of Biology and Biochemistry and Biophysics, and the program of Genetics at Texas A&M University.

All work described in Chapter II of this dissertation was done by myself, under the supervision and guidance of my Advisory Committee. We are grateful to Ben Greenwell for help with the bioinformatics analysis and discussion, as well as Joshua Beytebiere, Sarah Beagle and members of the Menet lab for discussion on the project. We thank Christine Merlin, Paul Hardin, Deborah Bell-Pedersen and Sebastian Kadener for helpful discussions and comments on the manuscript. We also thank the Texas A&M Institute for Genome Sciences and Society for providing access and maintenance to their high-performance computing cluster.

Analysis of work described in chapter III was conducted by myself with coding help from Ben Greenwell, as well as under the supervision and guidance of my Advisory Committee. The animal protocol was carried out by our collaborators Dr. Carolina C. Escobar, Dr. Rudolf M. Buijs and Natali Guerrero-Vargas at the Departamento de Biología Celular y Fisiología, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Distrito Federal, México.

**Funding Sources**

The studies and experiments performed in this dissertation were supported by a Pathways to the Doctorate Fellowship from Texas A&M University to Alexandra Trott, from Texas A&M University / Department of Biology Start Up Funds to Jerome S Menet, and a TAMU-CONACYT research grant program to Jerome S Menet and Ruud Buijs.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

# CHAPTER I

# INTRODUCTION AND LITERATURE REVIEW

## The autonomous mammalian circadian clock

Living organisms have many physiological activities that show rhythmic changes within a period of approximately 24 hours, even under constant environmental conditions or without any external time cues. These daily 24-hour variations are called circadian rhythms and are generated by a cell-autonomous time-measuring system called the circadian clock [1]. Additionally, the circadian clock can also be entrained by zeitgebers (German word for "time giver" or "synchronizer"), which are environmental cues that synchronize the circadian clock. It is a convention in the field of circadian rhythms to refer to light on as Zeitgeber (ZT) 0 and light off as ZT12. There are many additional examples of zeitgebers such as light, temperature, social interactions, exercise, and eating and drinking. This ability for organisms to tell time allows for an organism to anticipate environmental changes and to adapt to many parameters such as nutrient availability and predation [2]. Circadian clocks can be found in a wide variety of organisms, from cyanobacteria and fungi to higher eukaryotic plants and mammals, and dysregulation of the clock disrupts biochemical and metabolic pathways which results in diseases, such as diabetes or cancer [3,4].

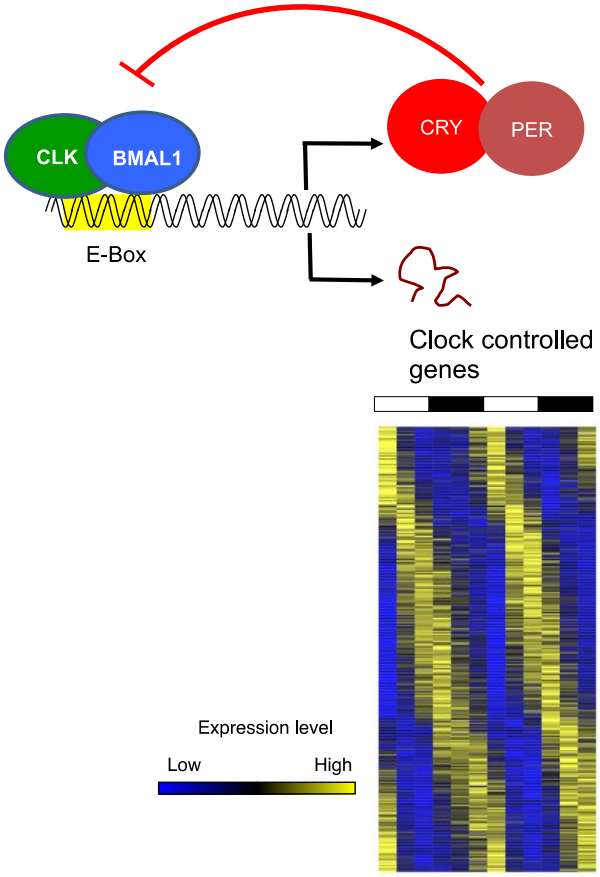
## Mammalian molecular circadian clock is controlled by a negative feedback loop

In mammals, the master clock resides in the hypothalamic suprachiasmatic nucleus (SCN) while peripheral tissues also have self-sustaining molecular clocks. These peripheral clocks are

ubiquitous and are found in almost every cell [5]. Peripheral clock cells can also respond to inputs individually depending on entrainment signals and will control different physiological outputs. However, they still receive some inputs from the SCN to coordinate biological processes as a whole [6]. These central and peripheral clocks coordinate clock genes that form a transcription/ translation-based negative-feedback loop to generate the 24-hour molecular clock [7,8]. This feeding back loop is initiated by heterodimeric basic helix-loop-helix (bHLH)-PAS positive transcription factor CLOCK:BMAL1. This heterodimeric transcription factor binds to CACGTG E-box or E-box-like sequences to promote transcription activation of negative regulatory genes, such as *Period* (*Per1* and *Per2*) and *Cryptochrome* (*Cry1* and *Cry2*) genes. These PER and CRY proteins form a complex, enter the nucleus, and bind to CLOCK:BMAL1 to inhibit CLOCK:BMAL1's transcriptional activation [9-11]. The current model in the field, illustrated in Figure I-1, assumes that CLOCK:BMAL1 binds rhythmically to negative regulatory genes, as well as other rhythmically expressed genes, and recruits transcriptional machinery to promote rhythmic transcription [7]. This model, however, has been recently challenged by the finding that the majority of CLOCK:BMAL1 target genes are not rhythmically expressed [12]. Additionally, that CLOCK:BMAL1 directly regulates their DNA binding sites by opening the chromatin and thus indirectly regulating transcriptional activation [13].

**Figure I-1 Current Model of the Circadian Clocks Positive and Negative Feedback Loops and Rhythmic Gene Expression.**

CLOCK and BMAL1 form a heterodimer and bind to E-boxes initiating transcription of their negative regulatory elements *Cry* and *Per* genes. PER and CRY from a complex and re-enter the nucleus and inhibit CLOCK:BMAL1 activation. CLOCK:BMAL1 also initiates rhythmic gene expression of other clock controlled genes. This rhythmic gene expression is illustrated in a heatmap. Highly expressed genes are shown in yellow and lower expressed genes are in blue.




## Mammalian transcription regulation

Transcription in mammals is a multistep process that starts with initiation of transcription by RNA polymerase II (RNA pol II) at specific DNA sites called promoters. Messenger RNA (mRNA) is produced after transcription elongation and termination of RNA Pol II activity and other various RNA processing events. This mRNA is then transported out of the

nucleus for translation to occur [14]. While all these steps can be specifically regulated, transcription initiation is generally the main step controlling gene expression [15].

Chromatin is a dynamic structure that is not only in charge of packaging the entire eukaryotic genome but also regulates the accessibility of the DNA for transcription. A nucleosome at the transcription start site can inhibit the ability of Pol II and transcription [16]. There are four core histone variants that make up a nucleosome, H3, H2A, H2B, and H1 [17]. These histone variants can vary depending on the needs of the cell. One of the ways the cell regulates DNA accessibility is by facilitating histone exchanges by incorporation of histone variants. Histone variants such as H3.3 and H2A.Z are modifications of H3 and H2A, and have been correlated with active transcription. Histone post-translational modifications such as acetylation, methylation, ubiquitylation, and sumoylation can also be used to regulate chromatin structure and transcription. These post-translational modifications affect nucleosome stability by interacting with the chemical structures inside the nucleosome, neighboring nucleosomes, or histone-DNA interactions [17]. For example, H3K9 acetylation is associated with active promoters and H3K27 acetylation is associated with active enhancers [18,19]. Such modifications can result in open or closed chromatin which affects the binding of transcription factors (TFs) and the machinery needed for transcription. TFs can also bind and alter the chromatin landscape, making it more permissible for transcription by competing with nucleosomes for the accessibility to the DNA [20].

For transcription to occur, TFs typically bind to enhancers and promoters found in open chromatin, and this initiates the recruitment of other TFs, the pre-initiation complex (PIC), and the Mediator [14,21]. The PIC is made up of general transcription factors (GTFs) TFIIB, TFIID, TFIIE, TFIIF and TFIIH, as well as RNA Pol II (Pol II). The GTFs and Pol II assemble at promoters and the protein interactions that occur between the PIC, the Mediator, and the other TFs regulates Pol

II activity [14]. This Pol II activity is generally paused until the Pol II C-terminal domain (CTD) serine5 is phosphorylated and the DNA in the promoter region is unwound and opened by TFIIH [22]. Meyer et al., 2010 also found that TFIIH phosphorylation of the CTD of Pol II was completely dependent on the Mediator, suggesting that the PIC, the Mediator, and TFIIH work together to control Pol II and transcription.

The Mediators basic function is to communicate regulatory signals from TFs and the PIC directly to RNA Pol II. However, the exact mechanism by which this occurs is still unknown. The core Mediator complex is generally made up of 26 proteins and the composition of the Mediator complex dictates what function the Mediator will have on transcription [21,23]. For example, the Med12 subunit has been associated with gene repression by recruiting the repressive CDK8 subcomplex to the PIC and Pol II [24] and the Med1 subunit has been associated with gene activation and recruitment of the PIC and Pol II [23]. The Mediator has also been found to be important for the organization of gene looping, which enables the coordination and regulation of gene expression, such that it connects enhancers to the promoter and thus regulates initiation, elongation, and chromatin architecture [21]. Additionally an essential component for the coordinating transcription activation of gene networks is this three-dimensional organization of the genome or gene looping [25]. Previous studies have shown the promoter of one gene interacts not only with enhancers that can be megabases away, but other gene promoters. These interactions are not random and appear to be occurring to promote coordinated gene expression for related biological functions [26-28]. The chromatin looping landscape appears to involve many proteins, but some of the core proteins are CCCTC-binding factor (CTCF), cohesion loading factor Nipbl, Cohesin, and the Mediator [29,30]. CTCF binds to the DNA and has been associated with mediating long-range interactions by facilitating chromatin remodeling proteins and organizing the chromatin into transcriptionally

active and silent domains [26,31]. Cohesion loading factor Nipbl helps load cohesion on to the DNA and helps to facilitate the regulation of gene expression by associating with the Mediator complex [23,32]. For this reason, the three-dimensional organization, the Mediator complex, the PIC, and possibly other TFs may promote interactions between gene promoters and enhancers and maintain the chromatin landscape.

**Dysregulation of the circadian clock's coordination of biological processes**

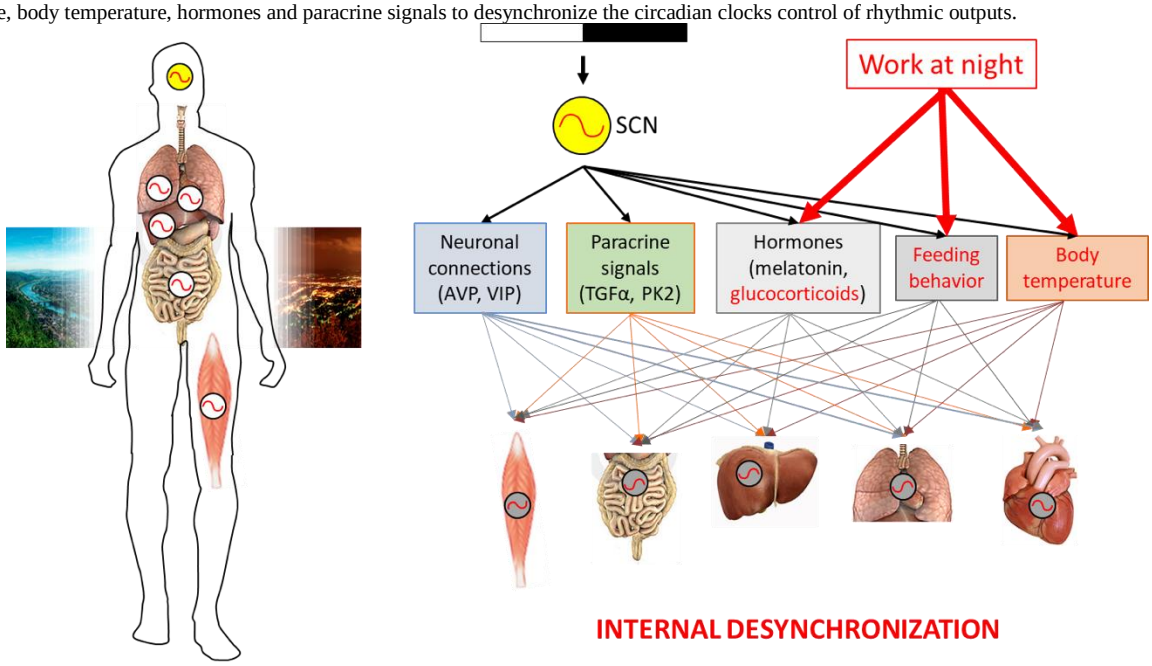
The circadian clock coordinates transcription to activate biological processes needed throughout the 24-hour day. However, many modern day lifestyles now alter the zeitgebers that are needed to coordinate the circadian clock [33]. For example, there are many nontraditional work schedules that are becoming increasingly more frequent to meet the high demands of today's industries [34]. Many of these schedules require people to work during the night when the circadian clock is coordinating inactivity.

Shift work or any nontraditional work schedules that occur outside the typical working hours of 9am to 5pm have been correlated with an increased risk of pathological diseases [35]. Approximately 15-30% of the European and American work force do some form of shift work and despite the ability for these workers to force themselves to stay awake and function on the job, a majority of shift workers do not exhibit a shift in their circadian clock functions [33,34]. Working at night has been found to alter food consumption, rhythms in body temperature, and rhythms in hormones [36,37].

The circadian clock in mammals is a hierarchical collection of cell-autonomous biological clocks that are coordinated by the SCN [5]. This SCN is directly entrained by light from photo-responsive retinal ganglion cells [38]. The SCN then sends signals to peripheral downstream

circadian clocks to control overt rhythms of body temperature, feeding behavior, and hormone release [33]. This entrainment by light means that the SCN can be reset with light at night. Furthermore, even low levels of light at night have been shown to have the ability to reset the circadian clock in the SCN, suggesting that even the use of low levels of artificial light used at home may alter the circadian clock [33]. This desynchronization between light at night and the phase of the SCN's circadian clock has been hypothesized to lead to metabolic diseases, cardiovascular disease, cancer, and an increased risk of mortality(Figure I-2) [33].

**Figure I-2 Current Model of How Disruption of the Circadian Clock Leads to Internal Desynchronization**
Cell autonomous circadian clocks, found throughout the body are coordinated by zeitgebers and the SCN. Shift work affects daily rhythms of food intake, body temperature, hormones and paracrine signals to desynchronize the circadian clocks control of rhythmic outputs.




## Shift workers have an increased risk of cardiovascular disease

A meta-analysis on 17 studies found that shift workers had a 40% increased risk of cardiovascular disease compared to day workers [35]. However, the exact mechanisms between how this desynchronization of the circadian clock and shift work and how it causes an increased risk of cardiovascular disease is unknown [35,39-41]. Cardiovascular disease includes many different types of pathologies found in the heart and in general includes any diseased blood vessels, structural problems, and blood clots [42]. Many studies have been done to try and decipher the molecular mechanisms that occur to lead to cardiovascular disease [39,42-44]. The primary risk factors that have been targeted are food consumption, smoking, and social stress [39,45]. Shift work generally leads to

a mismatch in eating time and the enzymatic activity controlled by circadian rhythms. This has been found to be correlated with an increase in cholesterol levels which can lead to hypertension and excess stress on the heart [39,45,46]. Smoking has been highly correlated with cardiovascular disease and furthermore shift workers have been noted to smoke more than day workers [39]. Smoking has also been highly correlated with coronary heart disease by damaging the lining of arteries, increasing in atheroma (fatty material) and increasing blood pressure [39,45,47]. Another known consequence to shift work is the stress of interfering with family or social interactions. Shift work generally tends to reduce the time available for social interaction and thus can lead to social isolation. This stress has been associated with increased cholesterol levels and myocardial infarction. However, the mechanisms by which this behavioral stress impacts physiology are incredibly hard to study [39,48,49].

Many different studies have tried to identify the biological mechanisms that occur to cause cardiovascular disease. It is suggested that cardiovascular disease occurs in two parts and can be identified with biomarkers. The first part is a slow buildup of plaque in the arteries, called atherosclerosis [39]. This has been identified by looking at cholesterol, triglyceride, and apolipoproteins circulation in the blood. The second part is thrombosis and has been identified by looking at coagulation and fibrinolytic proteins [39]. The last major biomarkers used to diagnose cardiovascular disease are high blood pressure and electrocardiograph which have been used in particular to diagnose hypertrophy in the heart [39,43]. Studies on shift work have had mixed results on whether the increased risk of cardiovascular disease is associated with an increase in these biomarkers, thus it is still unknown how shift work causes cardiovascular disease [34,39].

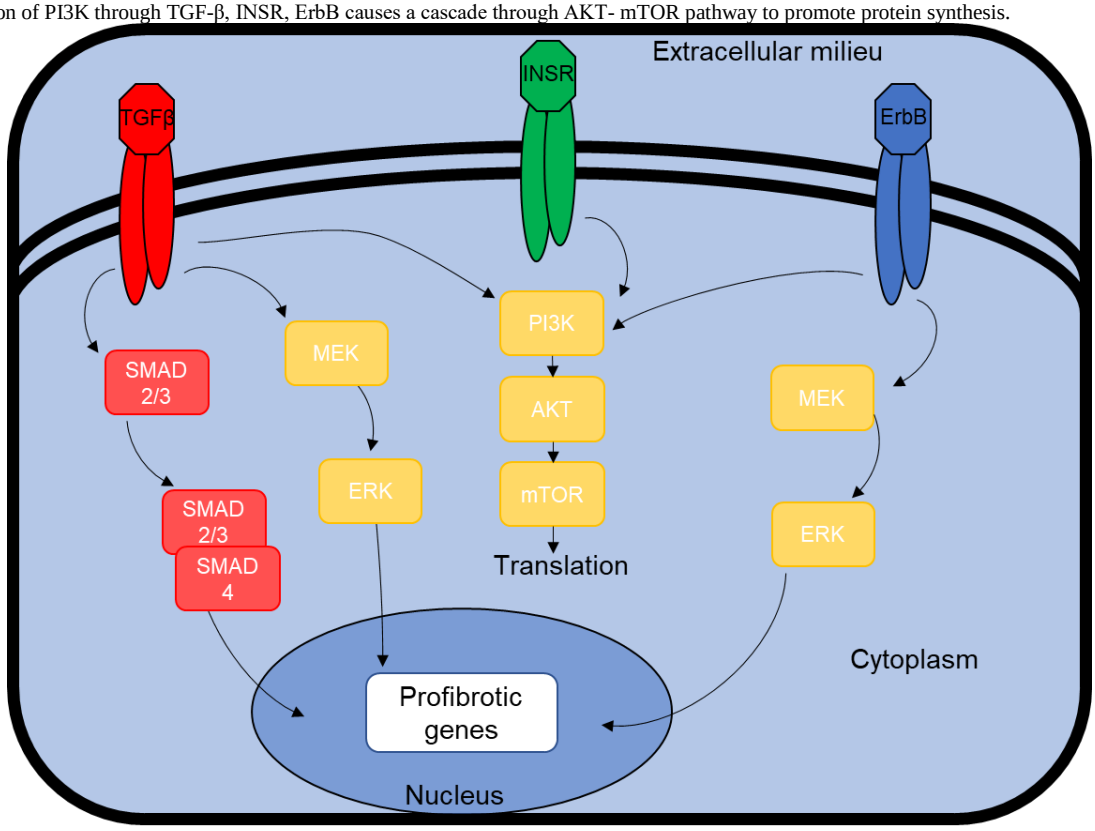
**Pathological cardiac remodeling**

After cardiovascular disease manifests it is always associated with an increase in interstitial fibrosis [44]. This increase in activation and expansion of cardiac fibroblast alters the extracellular matrix of the heart by increasing collagen levels and other extracellular matrix proteins [44,50]. One of the main pathways that is responsible for activating proliferation of fibroblasts and extracellular matrix remodeling is the transforming growth factor beta (TGF-β) pathway [51,52]. This pathway is essential in promoting extracellular matrix deposition by increasing collagen and fibronectin synthesis, by inhibiting expression of matrix metalloproteinases (MMPs) and matrix degradation through activation of tissue inhibitors of metalloproteinases (TIMPs) [52,53].

The TGF-β pathway is also considered a master switch in that it also activates and cooperates multiple other signaling pathways, such as epidermal growth factor receptor (ErbB), mammalian target of rapamycin (mTOR), Mitogen-activated protein kinase kinase (MEK), protein kinase B (AKT) and insulin receptor tyrosine kinase (INSR) pathways (Figure I-3) [51]. The TGF-β pathway does this by phosphorylating SMAD2/3 which bind to SMAD4 causing it to translocate to the nucleus to activate many profibrotic genes and ErbB ligands needed to activate the ErbB pathway [44,51]. The ErbB pathway is essential for cardiac paracrine mediator of cell to cell interactions to maintain cardiac function and activates a signaling cascade needed to adapt to stress in the myocardium [54]. SMAD independent pathways for TGF- β include interactions through phosphoinositide 3-kinase (PI3K), AKT, and mTOR which are primarily responsible for changes in translational and hypertrophy found in heart failure [55]. TGF- β also activates MEK which activates extracellular signal-regulated kinase (ERK) and regulates transcription needed for activation of collagen and fibroblast activation [44]. Additionally, the INSR pathway is also capable

of interacting with PI3K and has been shown to be hyperactive in pathological hypertrophic hearts induced by pressure overload [55].

**Figure I-3 Model of Signaling Pathways Regulating Cardiac Fibrosis.**
Activation of TGF-β receptor activates the canonical SMAD pathway which in turn activates expression of profibrotic genes. TGF-β receptor non-canonical activation results in activation of MEK and PI3K. MEK activates ERK and causes ERK translocate to the nucleus to activate transcription. Activation of PI3K through TGF-β, INSR, ErbB causes a cascade through AKT- mTOR pathway to promote protein synthesis.



**Project aims**

Virtually every mammalian cell harbors a molecular circadian clock that enables biological functions to adapt to the daily environmental variation. These mammalian clocks rely on the rhythmic recruitment of the heterodimeric transcription factor CLOCK:BMAL1 to their target genes to coordinate the rhythmic expression of 10-15% of the transcriptome. This rhythmic expression of the transcriptome promotes the activation of key biochemical and metabolic pathways needed for daily adaptation to the environment [7,56,57].

The recent characterization of CLOCK:BMAL1 DNA binding sites at the genome-wide level revealed that the majority of their target genes are not rhythmically expressed, suggesting that CLOCK:BMAL1 rhythmic DNA binding alone does not directly activate transcription and that other factors contribute to transcriptional output [12]. These other factors may include transcription activation hallmarks like nucleosome positioning, histone modifications, transcription factors and RNA polymerase II (Pol II) recruitment to DNA.

The first aim of this study is to discover the disconnect that is occurring between the homogenous binding of CLOCK:BMAL1 and its heterogeneous output. We performed a meta-analysis on genome wide datasets and found that CLOCK:BMAL1 are not sufficient to generate transcriptionally active enhancers and that ubiquitously expressed transcription factors are likely contributing to CLOCK:BMAL1 transcriptional output. This suggests that CLOCK:BMAL1 promote transcriptionally permissive chromatin landscapes allowing for the binding of other transcription factors to activate transcription when changes occur environmentally or pathological conditions arise.

The second aim is to investigate what happens when work is done out of synchronization with the circadian clock. This desynchronization of the circadian clock has been correlated with

diseases, such as cancer and metabolic disorders [3,4]. Furthermore, shift work has also been reported to be associated with many different pathologies, including metabolic disorders, cancer, and cardiovascular disease [35,39,58,59]. The mechanisms of how shift work causes these different pathologies is unknown, however many studies have suggested that it is a desynchronization of working out of sync with the circadian clock [35,40].

To address the issue of how or if shift work and desynchronization of the circadian clock is causing cardiovascular disease, we conducted RNA-Seq and Picrosirius Red staining on hearts of rats that were forced to do shift work. We found that five weeks of shift work was sufficient to increase collagen deposition. Furthermore, we found that time of eating did not inhibit collagen deposition caused by shift work. Surprisingly, we did find that rats forced to do shift work but only allowed to eat at night had a significant change in gene expression compared to rats forced to do shift work but allowed to eat *ad libitum* and control rats. Our analysis suggests that five weeks of shift work is capable of significantly increasing collagen production in the heart and suggests a mechanism of how shift work may be causing cardiovascular disease.

# CHAPTER II

# REGULATION OF CIRCADIAN CLOCK TRANSCRIPTIONAL OUTPUT BY CLOCK:BMAL1*

## Overview

The mammalian circadian clock relies on the transcription factor CLOCK:BMAL1 to coordinate the rhythmic expression of 15% of the transcriptome and control the daily regulation of biological functions. The recent characterization of CLOCK:BMAL1 cistrome revealed that although CLOCK:BMAL1 binds synchronously to all of its target genes, its transcriptional output is highly heterogeneous. By performing a meta-analysis of several independent genome-wide datasets, we found that the binding of other transcription factors at CLOCK:BMAL1 enhancers likely contribute to the heterogeneity of CLOCK:BMAL1 transcriptional output. While CLOCK:BMAL1 rhythmic DNA binding promotes rhythmic nucleosome removal, it is not sufficient to generate transcriptionally active enhancers as assessed by H3K27ac signal, RNA Polymerase II recruitment, and eRNA expression. Instead, the transcriptional activity of CLOCK:BMAL1 enhancers appears to rely on the activity of ubiquitously expressed transcription factors, and not tissue-specific transcription factors, recruited at nearby binding sites. The contribution of other transcription factors is exemplified by how fasting, which effects several transcription factors but not CLOCK:BMAL1, either decreases or increases the amplitude of many rhythmically expressed CLOCK:BMAL1 target genes.

---

Together, our analysis suggests that CLOCK:BMAL1 promotes a transcriptionally permissive chromatin landscape that primes its target genes for transcription activation rather than directly activating transcription, and provides a new framework to explain how environmental or pathological conditions can reprogram the rhythmic expression of clock-controlled genes.

## Introduction

Virtually every mammalian cell harbors a circadian clock that regulates rhythmic gene expression to enable biological functions to occur at the most appropriate time of day. Circadian clocks rely on transcriptional feedback loops which are initiated in mammals by the heterodimeric transcription factor CLOCK:BMAL1 for review, [60]. CLOCK:BMAL1 rhythmically binds to DNA to activate the rhythmic transcription of the core clock genes *Period* (*Per1*, *Per2*, *Per3*), *Cryptochrome* (*Cry1* and *Cry2*), *Rev-erb* (*Rev-erbα* and *Rev-erbβ*) and *Ror* (*Rorα*, *Rorβ* and *Rorγ*). Upon expression and maturation, PERs and CRYs form a repressive complex that rhythmically inhibits CLOCK:BMAL1-mediated transcription first on-DNA and then off-DNA [7,61-63]. Furthermore, REV-ERBs and RORs rhythmically regulate *Bmal1* expression by repressing or activating its transcription, which promotes robustness of circadian oscillations [64,65]. In addition to activating the rhythmic transcription of core clock components, CLOCK:BMAL1 also regulates rhythmic expression of thousands of clock-controlled genes to generate oscillations in biochemistry, physiology and behavior, and thus control the rhythmic organization of most biological functions [56,57,66].

Characterizing the mechanisms through which CLOCK:BMAL1 regulates expression of its target genes has largely been carried out by determining how CLOCK:BMAL1 regulates the transcription of core clock genes (*Per*, *Cry* and *Rev-erb*), and target genes (e.g., *Dbp,* named for *D*

*site of albumin promoter binding protein*). Results from many laboratories show that the rhythmic binding of CLOCK:BMAL1 to e-boxes located in core clock gene promoters is necessary and sufficient for rhythmic transcription [62,63,67-69]. Upon DNA binding during the light phase, CLOCK:BMAL1 promotes chromatin modifications by recruiting histone-modifying enzymes to core clock gene promoters and enhancers. These enzymes include the histone acetyltransferases p300 and CBP, which mediate the acetylation of H3K9 and H3K27, and the histone methyltransferases MLL1 and MLL3 (Myeloid/Lymphoid Or Mixed-Lineage Leukemia 1 and 3), which promote the tri-methylation of H3K4 [7,70-75]. CLOCK:BMAL1 rhythmic DNA binding was also recently shown to promote rhythmic nucleosome removal, thereby generating a chromatin landscape that is favorable for the binding of other transcription factors at its enhancers [13]. Finally, CLOCK:BMAL1 recruits transcriptional co-activators, including components of the mediator complex and RNA Polymerase II (Pol II) to initiate core clock gene transcription [7,76,77]. During the repression phase in the early night, binding of the PER/CRY complex to DNA-bound CLOCK:BMAL1 is accompanied by the co-recruitment of histone deacetylases and demethylases and the removal of the H3K9ac, H3K27ac and H3K4me3 marks [75,78-82]. While these mechanisms are required for CLOCK:BMAL1-mediated transcription activation of core clock genes, it still remains unclear if the same mechanisms regulate the rhythmic expression of clock-controlled genes.

The recent characterization of CLOCK and BMAL1 mouse liver cistromes revealed that although CLOCK:BMAL1 binds synchronously during the middle of the day to thousands of enhancers and promoters, the transcription of its target genes is highly heterogeneous [7,12,83,84]. Indeed, not all CLOCK:BMAL1 target genes are rhythmically expressed, and a large fraction of the rhythmically expressed target genes are transcribed at night, in antiphase to maximal

CLOCK:BMAL1 DNA binding [12]. These data therefore suggest that the mechanisms by which CLOCK:BMAL1 regulates transcription of core clock genes differs from the regulation of other clock-controlled genes, and that additional mechanisms account for the activation of rhythmic gene expression by the circadian clock.

To uncover these mechanisms and to delineate the transcriptional logic underlying CLOCK:BMAL1 heterogeneous transcriptional output, we performed a meta-analysis of genome-wide datasets investigating the molecular events occurring at CLOCK:BMAL1 DNA binding sites, including CLOCK:BMAL1 rhythmic DNA binding, epigenetic modifications and transcription activation. Our analysis reveals that while CLOCK:BMAL1 DNA binding is sufficient to decondense the chromatin and prime its enhancers for transcriptional activation, it is not sufficient to generate transcriptionally active enhancers. Our results also indicate that many transcription factors bind to CLOCK:BMAL1 enhancers, and their recruitment likely contributes to CLOCK:BMAL1 clock-controlled transcriptional output. Altogether, our data support that CLOCK:BMAL1 regulation of clock-controlled gene expression relies on the cooperation between CLOCK:BMAL1 and other transcription factors. Furthermore, our data also suggest that a major role of CLOCK:BMAL1 is to generate a permissive chromatin landscape to rhythmically prime its enhancers for the recruitment of other transcription factors, rather than directly promoting transcription activation.

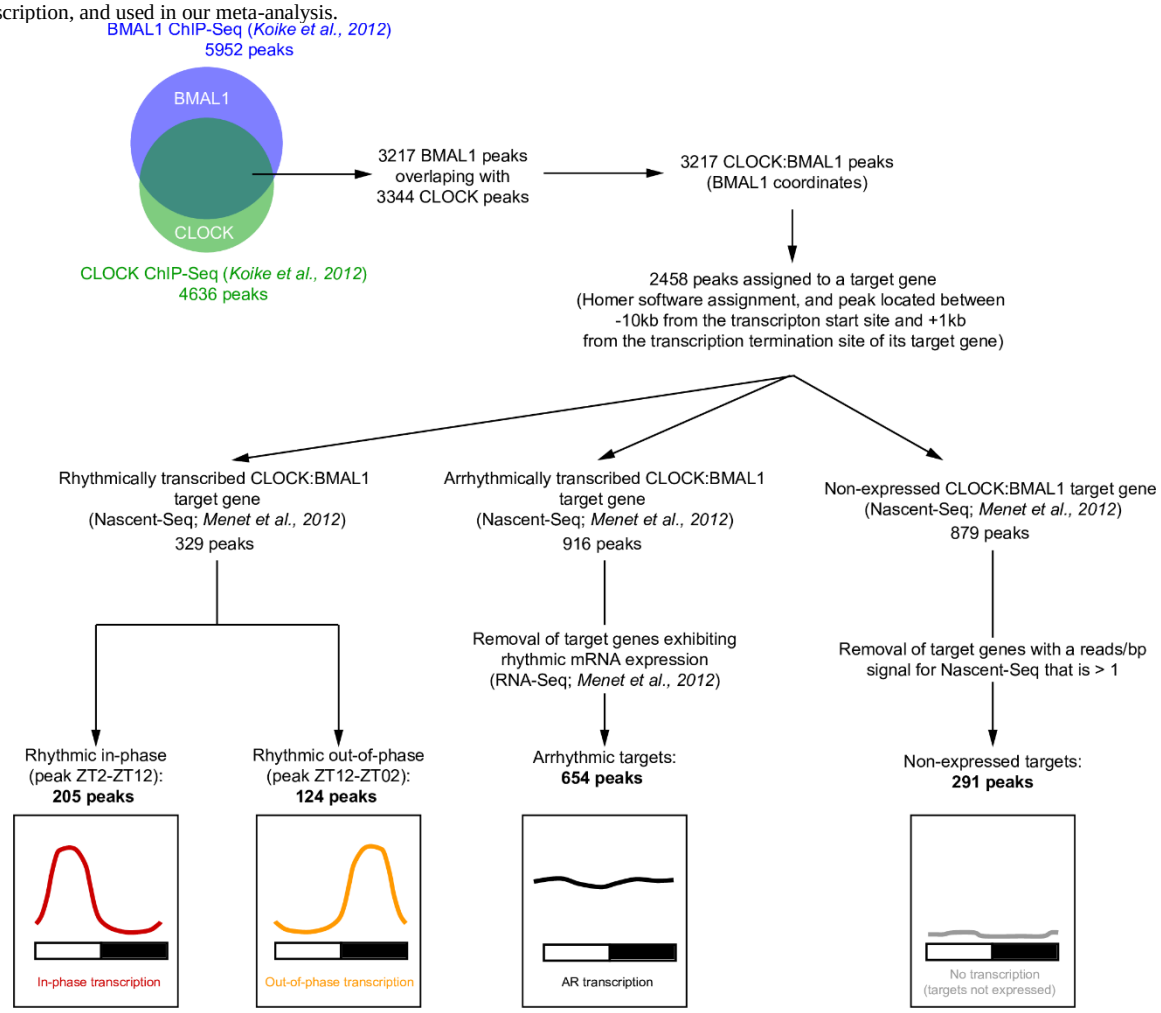
## Results

### *CLOCK:BMAL1 transcriptional output is heterogeneous*

To characterize the mechanisms by which CLOCK:BMAL1 regulates the transcriptional activity of its target genes at the genome-wide level in the mouse liver, we first generated a list of high-confidence CLOCK:BMAL1 DNA binding sites by determining the overlap between CLOCK and BMAL1 ChIP-Seq peaks in the mouse liver [7]. This analysis resulted in a list of 3217 CLOCK:BMAL1 binding sites, of which 2458 peaks can be assigned to a direct target gene (i.e., a CLOCK:BMAL1 peak located between -10kb of a target gene transcription start site and +1kb of a target gene transcription termination site; see Figure II-1 and Methods section for details).

**Figure II-1 Assignment of CLOCK:BMAL1 DNA Binding Sites to Their Target Gene Transcriptional Output in the Mouse Liver.**

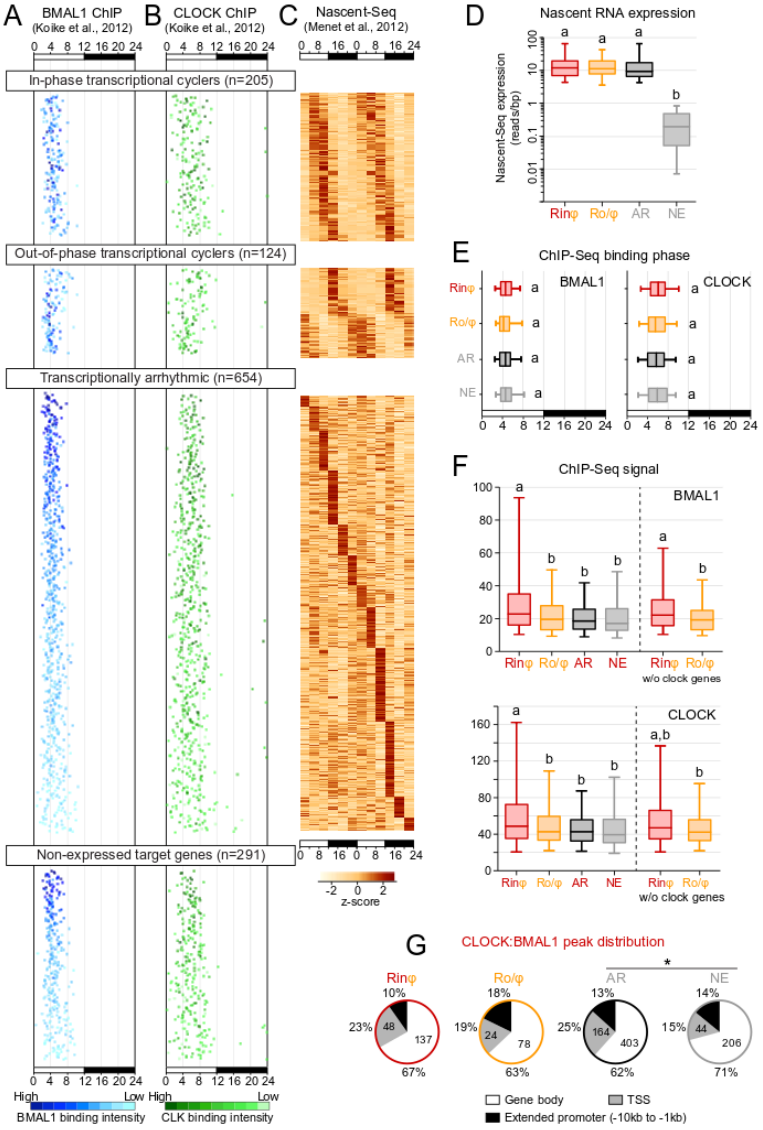
Flowchart illustrating the procedure used to identify CLOCK:BMAL1 target genes and to determine their transcriptional output in the mouse liver. See Methods section for details. Briefly, publicly available lists of CLOCK and BMAL1 DNA binding sites from [7] were compared and the overlapping CLOCK and BMAL1 peaks were identified as CLOCK:BMAL1 DNA binding sites (BMAL1 peak coordinates were kept for downstream analysis). Of the 3217 identified CLOCK:BMAL1 peaks, 2458 were assigned to a target gene (peak located by HOMER software between -10kb of a gene transcription start site and +1kb of a gene transcription termination site). The remaining 759 peaks were listed as intergenic. The list of 2458 CLOCK:BMAL1 peaks was then parsed based on their target genes transcriptional output using our publicly available Nascent-Seq analysis of rhythmic transcription in the mouse liver [12]. 329 CLOCK:BMAL1 peaks were found to target rhythmically transcribed genes in the mouse liver. Of these, 205 peaks were found to target rhythmically transcribed genes with a peak of transcription coinciding with CLOCK:BMAL1 rhythmic DNA binding (from ZT02 to ZT12; in-phase rhythmic transcriptional cyclers or Rinφ), whereas 124 peaks were targeting genes with a peak of rhythmic transcription out-of-phase with CLOCK:BMAL1 DNA binding (from ZT12 to ZT02; out-of-phase transcription cyclers or Ro/φ). A total of 916 CLOCK:BMAL1 peaks were assigned to genes exhibiting an arrhythmic nascent RNA profile. To ensure that these target genes are "true" arrhythmically expressed target genes, the list was further filtered by removing those exhibiting rhythmic mRNA expression using the dataset from [12], resulting in a final list of 654 CLOCK:BMAL1 peaks targeting arrhythmically transcribed genes. Finally, the remaining CLOCK:BMAL1 peaks were assigned to genes expressed below the expression threshold set to determine rhythmic gene expression. Because this threshold is set to call rhythmically expressed genes with high confidence rather than calling "true" non-expressed genes, we further filtered this list of peaks by removing genes exhibiting an averaged signal greater than 1 read/bp for the 12 time points of the Nascent-Seq dataset. This filtering resulted in a list of 291 CLOCK:BMAL1 peaks targeting non-expressed genes. The list of the 3217 CLOCK:BMAL1 peaks parsed based on their target gene transcription, and used in our meta-analysis.



To determine the extent to which rhythmic CLOCK:BMAL1 DNA binding contributes to rhythmic transcription activation at the genome-wide level, we used a public mouse liver Nascent-Seq dataset that characterized the levels of nascent RNA expression over the course of a 24-hr day [12]. A Nascent-Seq dataset was preferred over RNA-Seq because nascent RNA expression directly reflects transcription activation, and is unaffected by the post-transcriptional regulations that contribute to rhythmic mRNA expression in the mouse liver [7,12,85]. We found that only a small fraction of CLOCK:BMAL1 target genes are rhythmically transcribed (~26%; Figure II-1). Noticeably, not all rhythmic target genes are transcribed during the day, i.e., coincidently with rhythmic CLOCK:BMAL1 DNA binding (ZT02-ZT12). Indeed, 38% of the rhythmic CLOCK:BMAL1 target genes exhibit a peak of transcription between ZT12 and ZT02, out-of-phase with the rhythmic DNA binding of CLOCK:BMAL1 (n = 124 CLOCK:BMAL1 peaks) (Figure II-2 A-C; Figure II-1). Importantly, our analysis also reveals that the majority of CLOCK:BMAL1 direct target genes are either arrhythmically transcribed (AR; n = 654 CLOCK:BMAL1 peaks) or not expressed (NE; n = 291 CLOCK:BMAL1 peaks) (Figure II-2 A-D; Figure II-1).

**Figure II-2 Mouse Liver CLOCK:BMAL1: Transcriptional Output is Heterogeneous.**

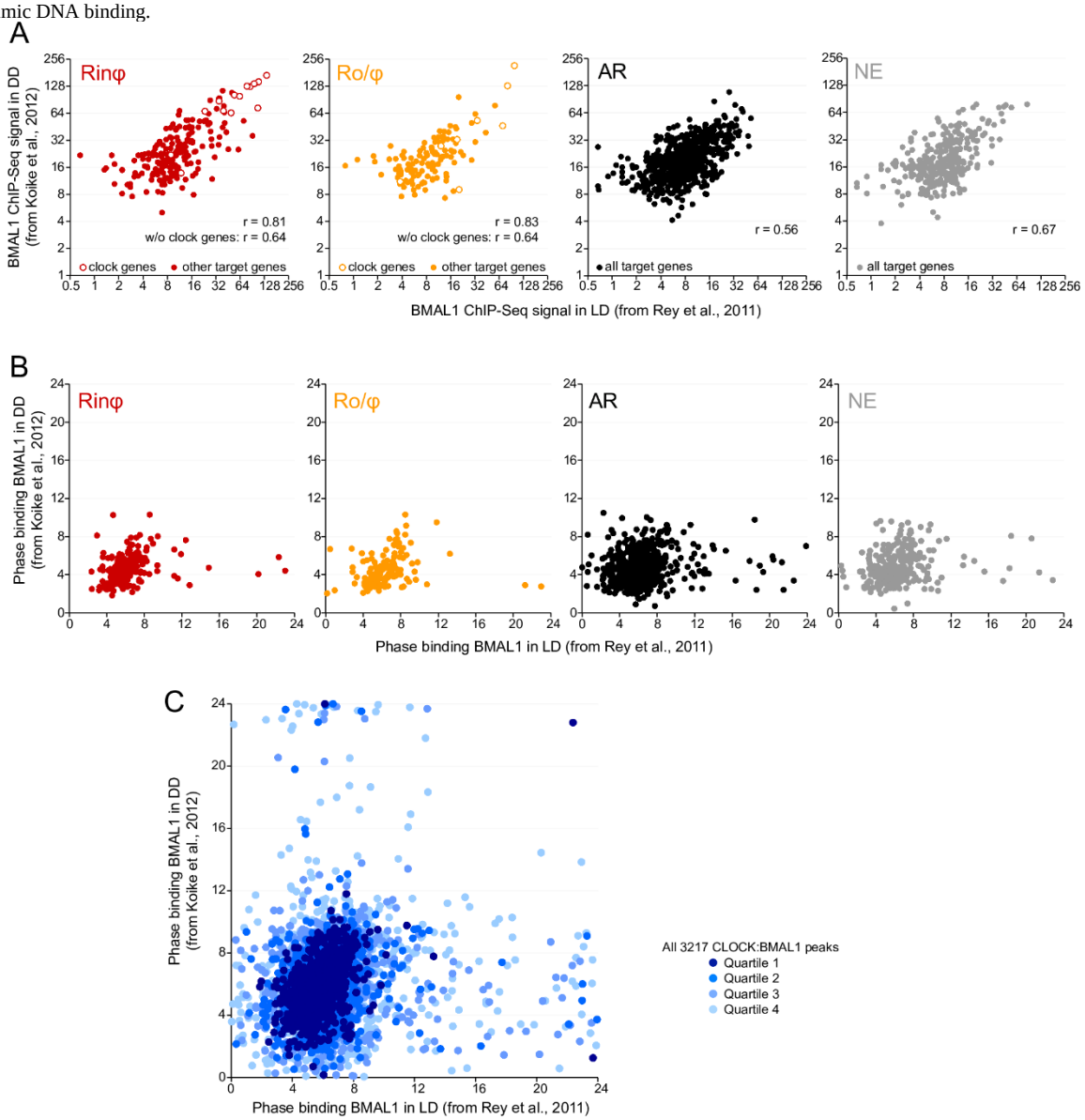
A, B. Mouse liver BMAL1 (blue, A) and CLOCK (green, B) ChIP-Seq peaks from Koike et al., 2012 were mapped to their target genes and parsed based on their transcriptional output (Nascent-Seq from Menet et al., 2012, C). Each dot represents the phase of maximal DNA binding, and the ChIP-Seq signal is displayed using different shades of color to illustrate differences in binding intensity. C. Heatmap representation of the Nascent-Seq signal of direct CLOCK:BMAL1 target genes classified based on their transcriptional output in the mouse liver. Each lane represents the Nascent-Seq signal of a gene corresponding to the CLOCK and BMAL1 peaks in A and B. Nascent-Seq signal was ordered based on the phase of nascent RNA oscillations for the in-phase and out-of-phase transcriptional cyclers, and based on BMAL1 ChIP-Seq signal for arrhythmically transcribed genes. No heatmap could be generated for the non-expressed genes because of the lack of nascent RNA expression. D. Nascent RNA expression, calculated as reads/bp, for each of the 4 CLOCK:BMAL1 transcriptional output group (Rinφ: in-phase transcriptional cyclers; Ro/φ: out-of-phase transcriptional cyclers; AR: arrhythmically transcribed target genes; NE: non-expressed target genes) . Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$). E. Phase of maximal BMAL1 (left) and CLOCK (right) rhythmic DNA binding for each of the 4 CLOCK:BMAL1 transcriptional output categories. Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$). F. BMAL1 (top) and CLOCK (bottom) ChIP-Seq signal for each of the 4 CLOCK:BMAL1 transcriptional output groups. Signal is also displayed for the Rinφ and Ro/φ groups after removal of the ChIP-Seq signal at peaks targeting core clock genes. Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$). G. Location of CLOCK:BMAL1 peaks within gene loci for each of the 4 CLOCK:BMAL1 transcriptional output groups. TSS (transcriptional start site; +/- 1kb from annotated TSS); Gene body: + 1kb from TSS to + 1kb from transcription termination site; Extended promoter: - 10 kb to - 1 kb from the annotated TSS. Numbers correspond to the percentage and numbers of peaks (outside and inside the pie chart, respectively) within each location for each group. * denotes a significant difference in the distribution of peaks between the AR and NE groups (chi square test; $p < 0.05$).



To determine if this result may be due to comparing samples collected in constant darkness (ChIP-Seq) and in a light:dark (LD) cycle (Nascent-Seq), we also analyzed a mouse liver BMAL1 ChIP-Seq rhythm performed under LD condition [83]. BMAL1 binding phase and ChIP-Seq signal under LD condition both exhibit a remarkably high level of similarity to those under DD conditions, and this even for the AR or NE target genes (Figure II-3). This therefore suggests that the large number of arrhythmically transcribed or not expressed CLOCK:BMAL1 target genes is not a consequence of using datasets generated under different lighting conditions. Taken together, these results indicate that the mechanisms underlying CLOCK:BMAL1-mediated rhythmic transcription of core clock genes (i.e., *Per1*, *Per2*, *Per3*, *Cry1*, *Cry2*, *Rev-erbα*, *Rev-erbβ* and *Dbp*) are not prevalent at the genome-wide level. They also suggest that the rhythmic recruitment of CLOCK:BMAL1 at its target gene promoters and enhancers is not sufficient to activate transcription for the majority of its target genes.

**Figure II-3 Effect of LD vs. DD Lighting Conditions on BMAL1 Rhythmic DNA Binding.**

Mouse liver BMAL1 ChIP-Seq datasets performed in mouse exposed to LD12:12 (Rey et al., 2011) or constant darkness (DD, Koike et al., 2012) were compared to determine if the lighting conditions (LD *vs.* DD) impact BMAL1 rhythmic DNA binding phase and signal. A. Correlation between BMAL1 ChIP-Seq signal in LD and DD for each of the 4 CLOCK:BMAL1 transcriptional output categories (rhythmic-in-phase (Rinφ, red); rhythmic out-of-phase (Ro/φ, orange); arrhythmic (AR, black); and non expressed (NE, grey) target genes). Peaks targeting core clock genes are depicted with an open circle. B. Correlation between the phase of BMAL1 DNA binding in LD and DD for each of the 4 CLOCK:BMAL1 transcriptional output categories. C. Correlation between the phase of BMAL1 DNA binding in LD and DD for all 3217 CLOCK:BMAL1 ChIP-Seq peaks from the Koike et al., 2012 dataset (see methods section for details). ChIP-Seq peaks were classified based on BMAL1 ChIP-Seq signal from Koike et al., 2012, and divided into 4 equal size quartiles. Peaks with higher ChIP-Seq signal display a better phase correlation in BMAL1 rhythmic DNA binding.



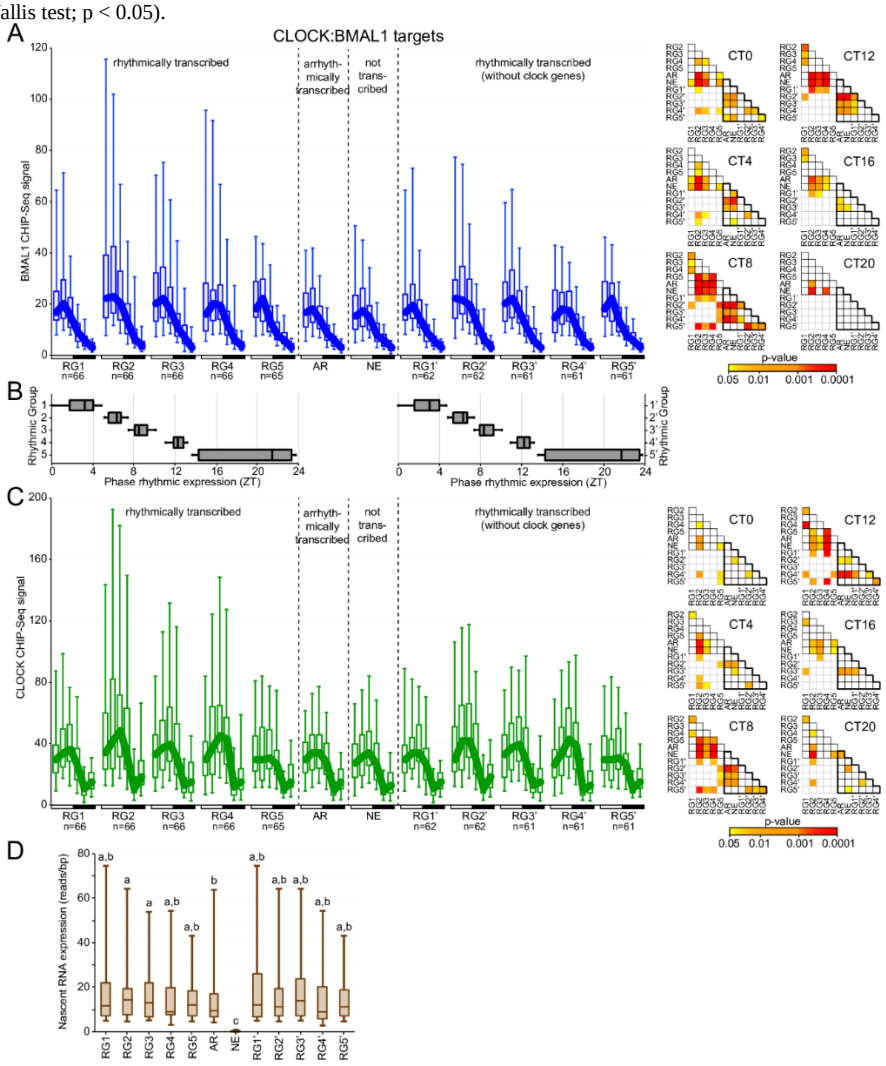
***CLOCK:BMAL1 heterogeneous transcriptional output is not mediated by differences in CLOCK:BMAL1 DNA binding***

To investigate the mechanisms underlying CLOCK:BMAL1 heterogeneous transcriptional output, we first examined if differences in the phase, intensity or location of CLOCK:BMAL1 DNA binding might explain the differences in transcription activation. The phase of CLOCK:BMAL1 DNA binding was found to be indistinguishable between all four transcriptional output categories, as both CLOCK and BMAL1 rhythmically bind to DNA with a peak between ZT3 and ZT9 for almost all target genes (Figure II-2 A, B, E). We then used CLOCK and BMAL1 ChIP-Seq signal as a readout to determine DNA binding intensity, and found that both CLOCK and BMAL1 ChIP-Seq signals are significantly higher at DNA binding sites targeting the in-phase transcriptional cyclers (Rinφ) when compared to peaks targeting the 3 other groups (out-of phase cyclers, arrhythmically expressed and non-expressed target genes) (Figure II-2 A, B, F; Kruskal-Wallis test, $p < 0.05$).

Remarkably, the binding intensity of CLOCK and BMAL1 at non-expressed target genes (NE) is similar to the binding intensity observed at the out-of-phase transcriptional cyclers (Ro/φ) and arrhythmically transcribed (AR) target genes, suggesting that CLOCK:BMAL1 DNA binding alone does not directly activate transcription at most of its target genes (e.g., comparisons between Figure II-2 D and F). To verify that these results are not due to the cut-offs we used to partition CLOCK:BMAL1 transcriptional output, we performed similar analyses using direct correlations between BMAL1 or CLOCK ChIP-Seq signal and the phase of rhythmic transcription, as well as by partitioning rhythmic target genes in five groups of equal sizes. These analyses confirmed our results (Figure II-4 and Figure II-5). While rhythmically transcribed target genes peaking from ZT5 to ZT13 exhibit higher BMAL1 and CLOCK ChIP-Seq signal, no differences in DNA binding signal were observed between the rhythmically expressed targets peaking from ZT13 to ZT5 and the AR and NE groups (Figure II-4).

**Figure II-4 Analysis of BMAL1 and CLOCK ChIP-Seq Signal Based on the Phase of Target Gene Transcription.**

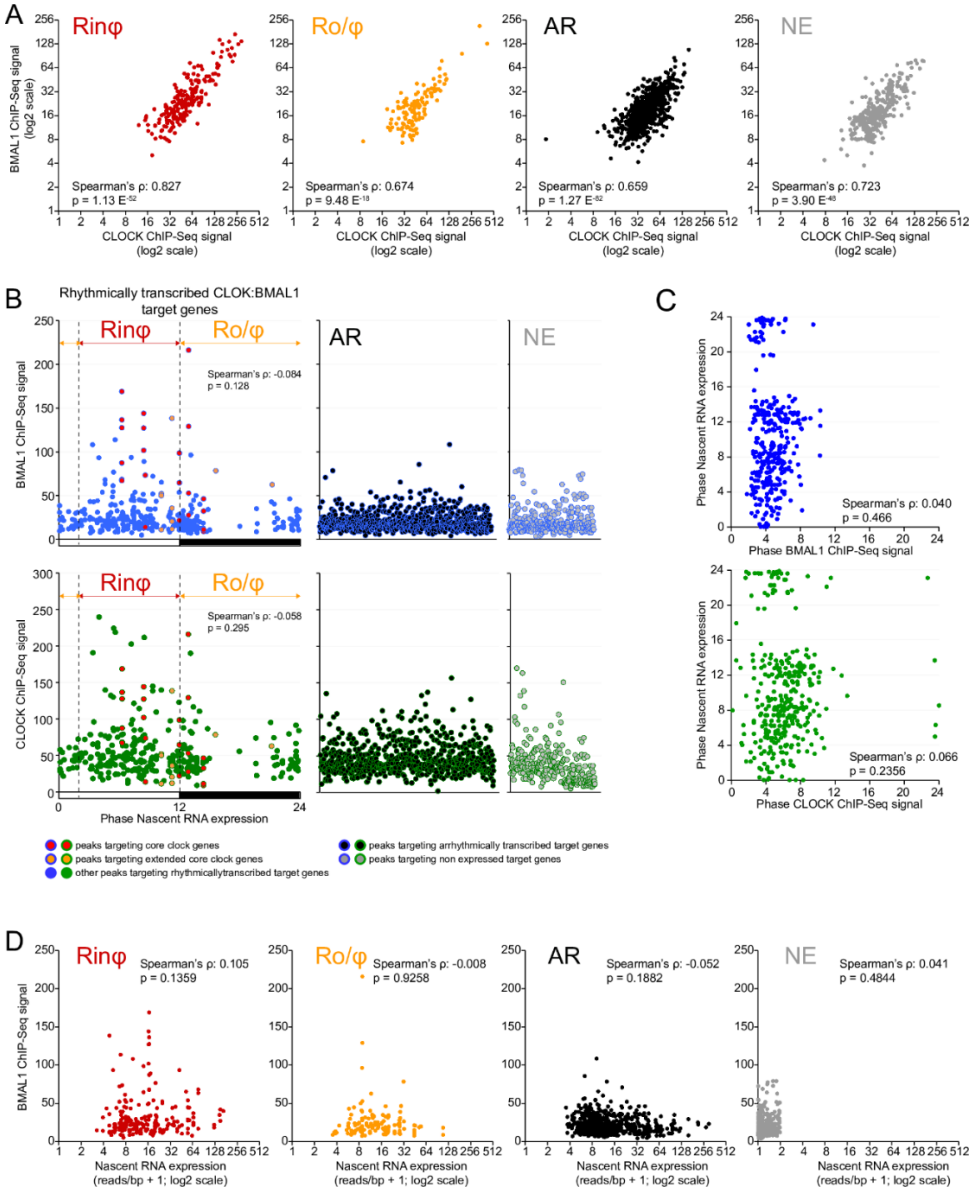
A. Analysis of BMAL1 ChIP-Seq signal from Koike et al. (2012) at CLOCK:BMAL1 peaks targeting rhythmically transcribed genes (RG), arrhythmically transcribed genes (AR) or not transcribed genes (NE). Peaks targeting rhythmic targets are binned in five groups of equal size for either all peaks (n = 329; groups RG1 to RG5), or those targeting non-core clock genes only (n = 307; groups RG1' to RG5'). Data are represented as boxplots for each group and time points, and the thick line displays CLOCK:BMAL1 DNA binding rhythm based on the median of ChIP-Seq signal. Statistical analysis was performed by Kruskal-Wallis non-parametric test, and pair-wise post-hoc analyses are displayed for each of the six time points using color-coding of the p-values. B. Phases of nascent RNA expression of rhythmically transcribed CLOCK:BMAL1 target genes are displayed for either all rhythmic target genes (left, groups RG1 to RG5), or only non-core clock rhythmic target genes (right, groups RG1' to RG5'). Nascent RNA expression was retrieved from Menet et al., 2012. C. Analysis of CLOCK ChIP-Seq signal from Koike et al. (2012) at CLOCK:BMAL1 peaks was performed as for BMAL1 ChIP-Seq signal in A. D. Nascent RNA expression of rhythmically transcribed CLOCK:BMAL1 is displayed for either all rhythmic targets (groups RG1 to RG5), or for non-core clock target genes (groups RG1' to RG5'), as well was for arrhythmically transcribed target genes (AR), or non-expressed target genes (NE). Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$).



In addition, we did not find any significant correlation between either CLOCK or BMAL1 ChIP-Seq signals and the phase of DNA binding or the phase of rhythmic transcription (Figure II-5). Because CLOCK:BMAL1 peaks targeting core clock genes are enriched in the Rinφ and Ro/φ groups and exhibit higher ChIP-Seq signal than clock-controlled (output) genes, we also compared CLOCK and BMAL1 ChIP-Seq signals between groups after removing peaks targeting the core clock genes (i.e., comparing clock-controlled genes only). Whereas BMAL1 ChIP-Seq signal intensity was still significantly higher at the Rinφ target genes compared to the three other groups, CLOCK DNA binding intensity was similar between all 4 groups (Figure II-2 F). Our data therefore indicate that while higher BMAL1 DNA binding signal may contribute to Rinφ transcription, the different transcriptional output of CLOCK:BMAL1 target genes cannot be explained solely by differences in CLOCK:BMAL1 DNA binding intensity.

**Figure II-5 Correlation Analysis between CLOCK:BMAL1 ChIP-Seq Signal and CLOCK:BMAL1 Target Genes Nascent RNA Expression.**

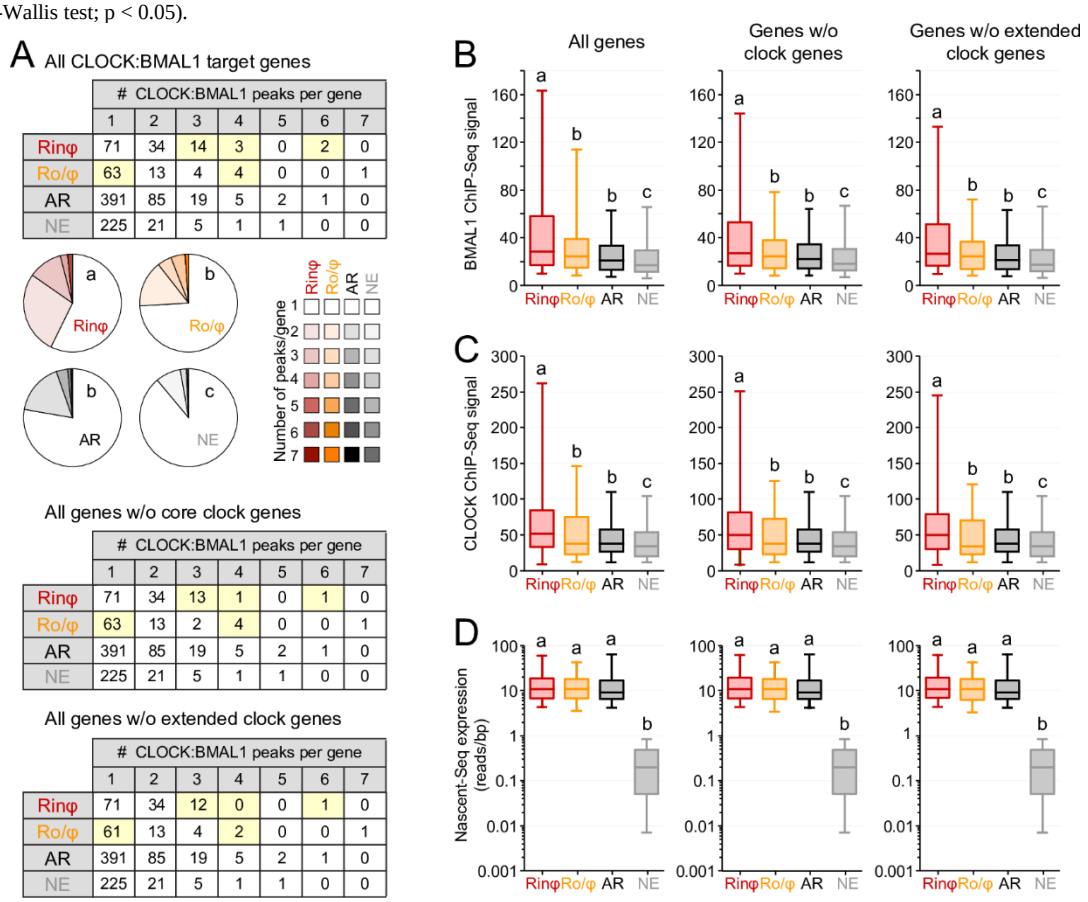
A. Correlation between BMAL1 and CLOCK ChIP-Seq signal at CLOCK:BMAL1 ChIP-Seq peaks in the mouse liver from Koike et al., 2012 datasets, parsed based on the transcriptional output of CLOCK:BMAL1 target genes (in-phase transcriptional cyclers, Rinφ, red; out-of-phase transcriptional cyclers, Ro/φ, orange, arrhythmically transcribed target genes, AR, black; not transcribed target genes, NE, grey; see text for details). B. Correlation between BMAL1 (top) and CLOCK (bottom) ChIP-Seq signal and the phase of nascent RNA expression of rhythmic CLOCK:BMAL1 target genes in the mouse liver (Nascent-Seq data from Menet et al., 2012). The dash lines depict the cut-offs used to partition the in-phase cyclers (Rinφ; from ZT02 to ZT12) to the out-of-phase cyclers (Ro/φ; from ZT12 to ZT02). Distinction is made between CLOCK:BMAL1 peaks targeting core clock genes (*Per1*, *Per2*, *Cry2*, *Dbp*, *Rev-erbα*, and *Rev-erbβ*; circles filled in red), extended core clock genes (*Tef*, *Hlf*, *Gm129*, and *Rorγ*; circles filled in orange), to those targeting clock-controlled genes (filled in blue and green for BMAL1 and CLOCK, respectively). CLOCK:BMAL1 peaks targeting arrhythmically transcribed genes (circles filled in black) and non expressed genes (circles filled in grey) are shown for comparison. C. Correlation between the phase of BMAL1 (top) or CLOCK (bottom) DNA binding and the phase of transcription of rhythmically transcribed CLOCK:BMAL1 target genes in the mouse liver. D. Correlation between BMAL1 ChIP-Seq signal and nascent RNA expression levels of CLOCK:BMAL1 target genes in the mouse liver, parsed based on the transcriptional output of CLOCK:BMAL1 target genes.



We also examined if differences in the location of CLOCK:BMAL1 DNA binding sites are associated with differences in transcriptional output by mapping CLOCK:BMAL1 peaks to either the transcription start site (TSS), gene body or extended promoter (-10 kb to -1 kb from the TSS) of their target genes. While the AR and NE groups were found to be statistically different (chi square test; $p < 0.05$), we did not observe any differences between the rhythmic target groups (Rinφ and Ro/φ) and the arrhythmically or not expressed groups (Figure II-2 G). The vast majority of CLOCK:BMAL1 peaks were located within enhancers (i.e., gene body or extended promoter), and only ~10-19% of CLOCK:BMAL1 peaks were mapped to TSS. Finally, we examined if differences in the number of genes targeted by multiple CLOCK:BMAL1 peaks were associated with differences in transcriptional output. We found that in-phase transcriptional cyclers were more frequently targeted by multiple CLOCK:BMAL1 peaks, and that conversely, non-expressed target genes were less frequently targeted by multiple peaks (Figure II-6). However, the lack of differences between the Ro/φ and AR groups indicates that the presence of multiple ChIP-Seq peaks does not directly influence the rhythmicity of CLOCK:BMAL1 target genes.

**Figure II-6 Contribution of CLOCK:BMAL1 Peaks Targeting the Same Genes to CLOCK:BMAL1 ChIP-Seq Signal and CLOCK:BMAL1 Target Gene Nascent RNA Expression.**

A. The number of CLOCK:BMAL1 target genes is displayed based on the number of CLOCK:BMAL1 ChIP-Seq peaks for each of the 4 categories of CLOCK:BMAL1 transcriptional output (in-phase transcriptional cyclers, Rinφ, red; out-of-phase transcriptional cyclers, Ro/φ, orange, arrhythmically transcribed target genes, AR, black; not transcribed target genes, NE, grey; see text for details). Top table: all CLOCK:BMAL1 target genes; Middle table: target genes without core clock genes (*Per1*, *Per2*, *Cry2*, *Dbp*, *Rev-erbα*, and *Rev-erbβ*); Bottom table: CLOCK:BMAL1 target genes without core clock genes (*Per1*, *Per2*, *Cry2*, *Dbp*, *Rev-erbα*, and *Rev-erbβ*) and other associated clock genes (*Tef*, *Hlf*, *Gm129*, and *Rorγ*). Yellow boxes indicate the location of clock genes within the table. The distribution of the number of CLOCK:BMAL1 ChIP-Seq peaks per gene is also displayed as a pie chart for all CLOCK:BMAL1 peaks. Groups with different letters are significantly different (Fischer's exact test (two-sided test); $p < 0.05$). B, C. BMAL1 (B) and CLOCK (C) ChIP-Seq signal at CLOCK:BMAL1 ChIP-Seq peaks (from Koike et al., 2012) is displayed for each of CLOCK:BMAL1 transcriptional output category. In this analysis, ChIP-Seq signal at CLOCK:BMAL1 peaks targeting the same gene was summed up (see panel A for the number of genes with multiple peaks for each category). Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$). D. Nascent RNA expression of CLOCK:BMAL1 target genes parsed based on CLOCK:BMAL1 target gene transcription, for all CLOCK:BMAL1 targets (left), target genes without core clock genes (middle), and target genes without core clock genes and other associated clock genes (*Tef*, *Hlf*, *Gm129*, and *Rorγ*). Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$).



Taken together, our analysis indicates that CLOCK:BMAL1 heterogeneous transcriptional output cannot be simply attributed to differences in the phase, intensity or location of CLOCK and BMAL1 binding to the DNA. While stronger DNA binding intensity may contribute to rhythmic transcription during the light phase, additional mechanisms are likely to contribute to CLOCK:BMAL1 transcriptional output heterogeneity.

***Recruitment of PERs and CRYs at CLOCK:BMAL1 DNA binding sites does not contribute to the heterogeneous CLOCK:BMAL1 transcriptional output***

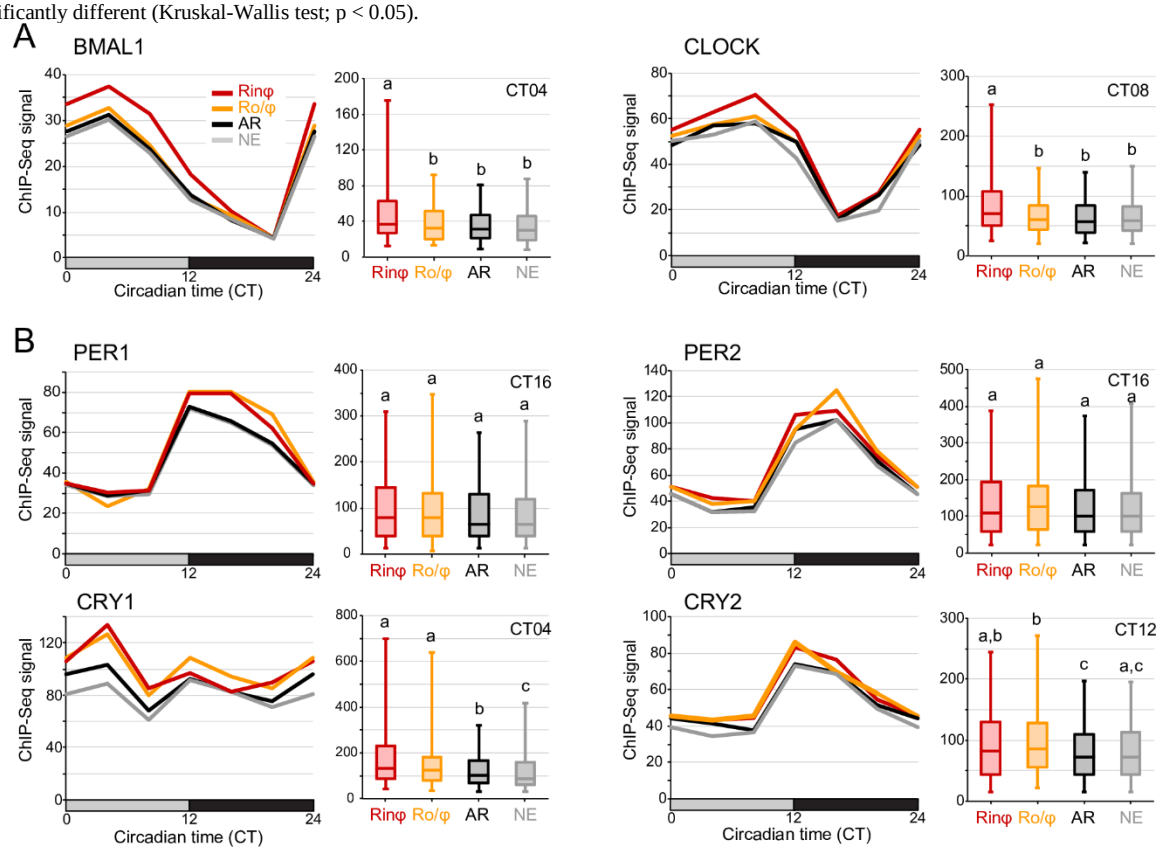
Circadian repression in mammals is initiated at the beginning of the night by the recruitment of the PER/CRY repressive complex and its associated histone deacetylases and methyltransferases to CLOCK:BMAL1 on DNA [78-80,82,86]. Because a differential recruitment of PERs and CRYs at CLOCK:BMAL1 DNA binding sites could lead to differences in CLOCK:BMAL1-mediated transcriptional output (e.g., decreased recruitment at arrhythmically transcribed target genes, delayed recruitment of out-of-phase transcriptional cyclers, *etc.*), we investigated the DNA binding profile of PER1, PER2, CRY1 and CRY2 at CLOCK:BMAL1 DNA binding sites for each of the four transcriptional output groups using publicly available ChIP-Seq datasets [7].

Our analysis shows that PER1, PER2 and CRY2 are rhythmically recruited at CLOCK:BMAL1 DNA binding sites with little difference between the four transcriptional output groups (Figure II-7). Maximal DNA binding for PER1, PER2 and CRY2 occur at CT12-16 for all groups, and differences were mostly observed for CRY2, where higher ChIP-Seq signal was found for rhythmically expressed target genes (Figure II-7). On the other hand, analysis of CRY1 recruitment to CLOCK:BMAL1-bound enhancers revealed more pronounced differences between

all four groups. CRY1 is a potent circadian repressor that is preferentially recruited at the beginning of the light phase just prior CLOCK:BMAL1 transcription activation (i.e., CT0-4), a mechanism proposed to poise CLOCK:BMAL1 for transcription activation [7]. We found that CRY1 recruitment at CT4 is significantly higher for rhythmically transcribed target genes (both Rinφ and Ro/φ) than for arrhythmically transcribed and non-expressed genes (Figure II-7). In addition, CRY1 recruitment was significantly decreased in non-expressed CLOCK:BMAL1 target genes than arrhythmic genes at CT4. These data thus suggest that CRY1 recruitment to CLOCK:BMAL1 DNA binding sites is, in addition to its well-characterized repressive effect, linked to rhythmic transcription activation. Consistent with this hypothesis are the higher levels for Ro/φ at CT12 compared to Rinφ (Figure II-7).

**Figure II-7 Recruitment of PERs and CRYs to CLOCK:BMAL1 Peaks Does Not Correlate with the Heterogeneous CLOCK:BMAL1 Transcriptional Output.**

A. (*Left*) Circadian rhythm of BMAL1 and CLOCK ChIP-Seq signal in the mouse liver at CLOCK:BMAL1 DNA binding sites for each of the 4 CLOCK:BMAL1 transcriptional output groups. (*Right*) Distribution of BMAL1 and CLOCK ChIP-Seq signal for each of the 4 CLOCK:BMAL1 transcriptional output groups at the time of maximal DNA binding (CT04 for BMAL1 and CT08 for CLOCK). B. (*Left*) Circadian rhythm of PER1, PER2, CRY1, and CRY2 ChIP-Seq signal in the mouse liver at CLOCK:BMAL1 DNA binding sites for each of the 4 CLOCK:BMAL1 transcriptional output groups. (*Right*) Distribution of PER1, PER2, CRY1, and CRY2 ChIP-Seq signal for each of the 4 CLOCK:BMAL1 transcriptional output group at the time of maximal DNA binding (CT16 for PER1 and PER2, CT04 for CRY1 and CT12 for CRY2. For both A and B panels, datasets were retrieved from Koike et al., 2012 and re-analyzed (see Methods section for more details). Values correspond to the ChIP-Seq signal median for each group. To improve visualization, CT0 ChIP-Seq values were repeated at CT24. Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$).



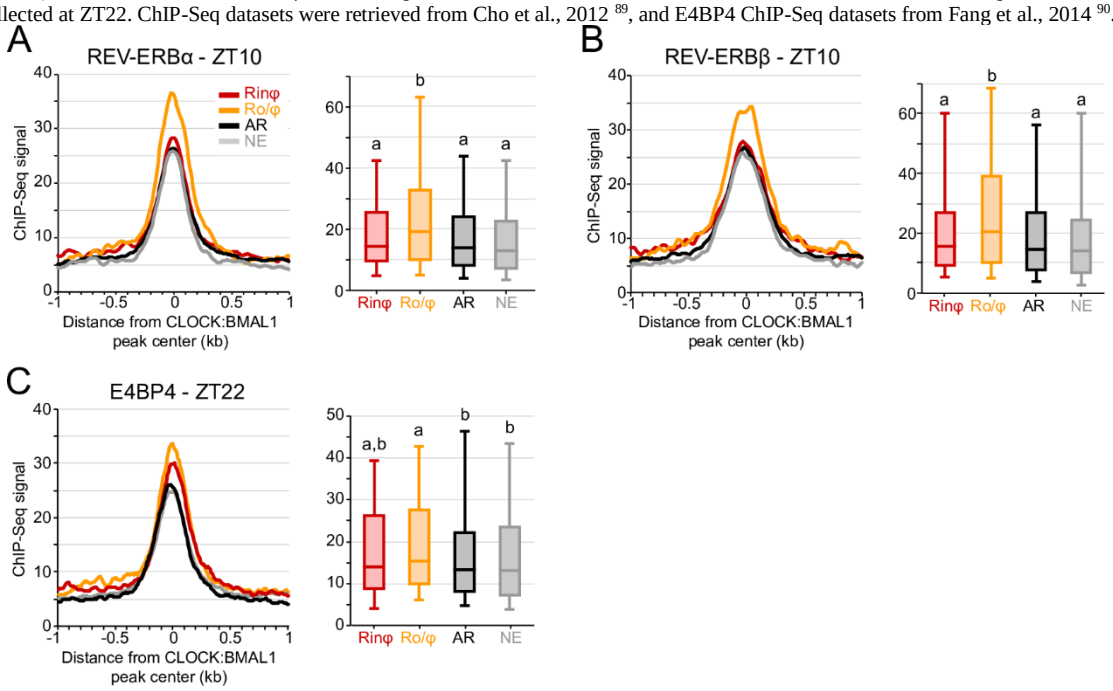
***REV-ERBα and REV-ERBβ ChIP-Seq signal is higher at CLOCK:BMAL1 peaks targeting genes transcribed at night***

Based on the mechanisms mediating the delayed transcription of the CLOCK:BMAL1 target gene *Cry1* [87], a model incorporating the nuclear receptors *Rev-erb* (repressor) and *Ror* (activator), and the D-box transcriptional factors *E4bp4* (also called *Nfil3*; repressor), *Dbp*, *Hlf* and *Tef* (activators) has been proposed to explain the different phases of rhythmic gene expression in the mouse liver [87,88]. In this model, co-binding of D-box transcription factors with CLOCK:BMAL1 is proposed to delay the phase of CLOCK:BMAL1 target genes from the morning to the afternoon (i.e., from ~ZT6 to ~ZT12), and additional binding of REV-ERBs and RORs would further delay the phase of transcription to the night (e.g., ~ZT18). To test if the binding of REV-ERBs and D-box transcription factors contribute to the delay of the out-of-phase CLOCK:BMAL1 target genes, we used publicly available ChIP-Seq datasets to determine REV-ERBα, REV-ERBβ [89], and E4BP4 [90] DNA binding intensity at CLOCK:BMAL1 enhancers. We found that REV-ERBα and REV-ERBβ DNA binding, which peaks at ZT10 for all target genes [89], is significantly higher at CLOCK:BMAL1 peaks targeting genes transcribed during the night consistent with the model proposed based on Cry1 expression; [87,88], and no differences were observed between Rinφ, AR and NE target genes (Figure II-8 A, B; Kruskal-Wallis test, $p < 0.05$). The binding of E4BP4, which is maximal at ZT22 [90], was also enriched at CLOCK:BMAL1 enhancers targeting the Ro/φ genes, but to a lesser extent than what was observed for the REV-ERBs (Figure II-8 C). In particular, no significant difference in enrichment was observed between the Rinφ and the Ro/φ groups, perhaps because co-binding of both CLOCK:BMAL1 and D-box transcription factors drives rhythmic transcription in the afternoon around ZT12, a time used for our cut-off to differentiate the in-phase from out-of-phase transcription cyclers. In summary, our

analysis indicates that the binding of REV-ERBα and REV-ERBβ (and eventually E4BP4) at CLOCK:BMAL1 enhancers may, as suggested by others [87,88], contribute to the delayed transcription of rhythmically expressed CLOCK:BMAL1 target genes.

**Figure II-8 REV-ERBα and REV-ERBβ ChIP-Seq Signal is High at CLOCK:BMAL1 DNA Binding Sites Targeting Genes Transcribed at Night.**

A-C *(Left).* Average ChIP-Seq signal for REV-ERBα (A), REV-ERBβ (B) and E4BP4 (C) at CLOCK:BMAL1 DNA binding sites (center ± 1kb) for each of the 4 CLOCK:BMAL1 transcriptional output group. A-C (*Right).* Distribution of REV-ERBα (A), REV-ERBβ (B) and E4BP4 (C) ChIP-Seq signal at CLOCK:BMAL1 peaks for each of the 4 CLOCK:BMAL1 transcriptional output groups (signal for each peak was averaged at CLOCK:BMAL1 peak center ± 250bp). Groups are labeled as in Figure II-2. Those with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$). REV-ERBα and REV-ERBβ ChIP were performed from mice liver collected at ZT10, while E4BP4 ChIP was performed from mice liver collected at ZT22. ChIP-Seq datasets were retrieved from Cho et al., 2012 [89], and E4BP4 ChIP-Seq datasets from Fang et al., 2014 [90].



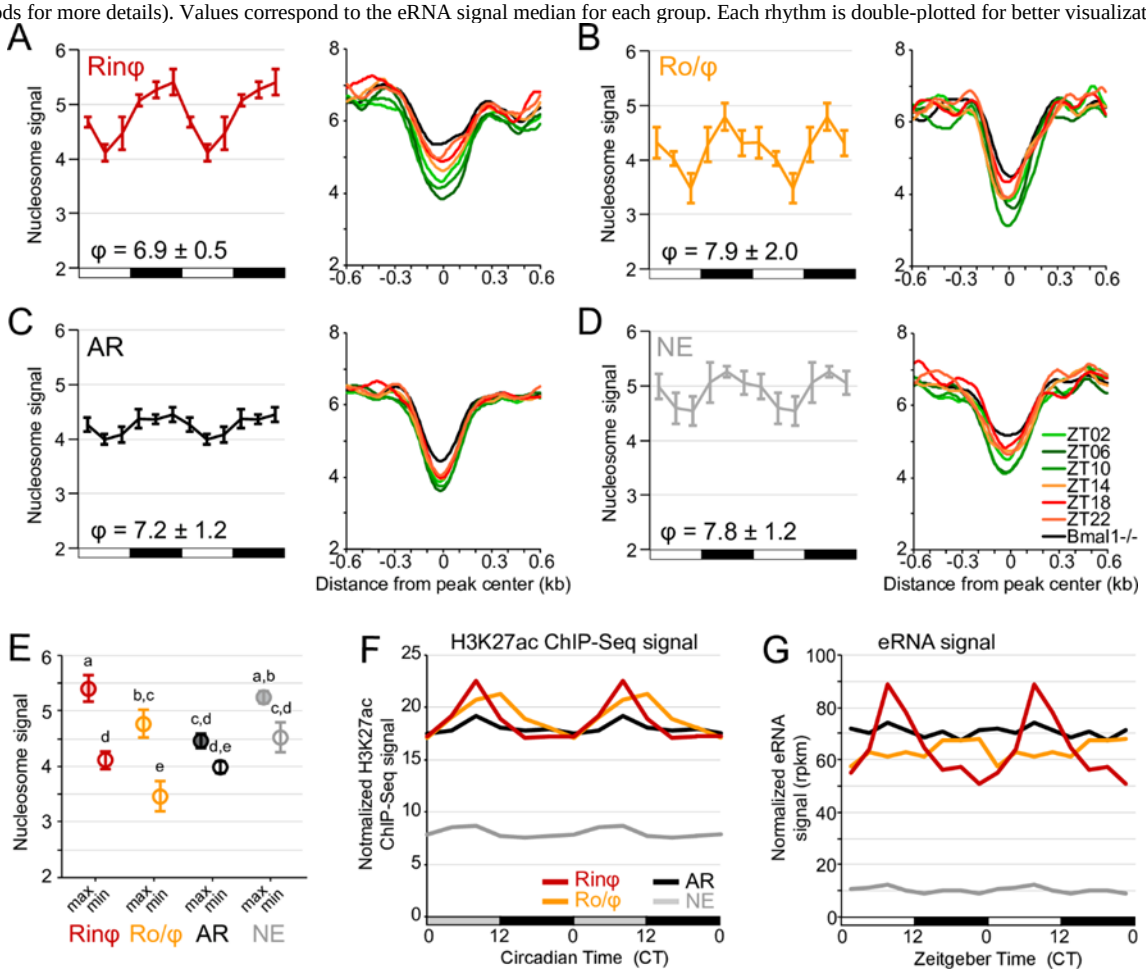
***CLOCK:BMAL1 promotes rhythmic nucleosome removal independently of its transcriptional output***

Our inability to detect substantial differences in CLOCK:BMAL1 DNA binding that would explain the heterogeneity of CLOCK:BMAL1 transcriptional output suggests that mechanisms other than the recruitment of core clock proteins to target gene promoters control CLOCK:BMAL1-mediated transcription. The recent finding that CLOCK:BMAL1 promotes the removal of nucleosomes when bound to DNA may represent one of these mechanisms [13]. Indeed, by mediating the removal of nucleosomes, CLOCK:BMAL1 would enable other transcription factors to access CLOCK:BMAL1 enhancers (most transcription factors bind better to naked DNA than DNA wrapped around nucleosomes).

To test if CLOCK:BMAL1-mediated nucleosome removal can contribute to the heterogeneity of CLOCK:BMAL1 transcriptional output, we examined mouse liver nucleosome signal over the 24-hr day at CLOCK:BMAL1 DNA binding sites for each of the transcriptional output groups, using a public MNase-Seq dataset (micrococcal nuclease digestion of mouse liver chromatin at 6-time points and high-throughput sequencing of mononucleosomes [13]). Our analysis reveals that nucleosome signal is rhythmic at CLOCK:BMAL1 DNA binding sites for each of the transcriptional output categories, i.e., even at CLOCK:BMAL1 DNA binding sites targeting arrhythmically transcribed and non-expressed genes (Figure II-9 A-D; Figure II-10).

**Figure II-9 CLOCK:BMAL1 Rhythmic DNA Binding is Associated with Rhythmic Nucleosome Signal, but Not with Rhythmic Histone Post-translational Modification and eRNA Transcription.**

A-D. Rhythmic nucleosome signal at CLOCK:BMAL1 DNA binding sites for each of the 4 CLOCK:BMAL1 transcriptional output groups: (A) Rhythmic-in-phase (Rinφ, red); (B) Rhythmic out-of-phase (Ro/φ, orange); (C) arrhythmic (AR, black); (D) non expressed (NE, grey) target genes). Nucleosome signal was retrieved from mouse liver MNase-Seq datasets [13], which consists of 6 time points each separated by 4 hours with n = 4 mice for each time point. (*Left*): 6-time points rhythm of nucleosome signal at CLOCK:BMAL1 binding sites (calculated at CLOCK:BMAL1 peak center ± 75 bp for each peak), displayed as the average ± s.e.m. of the signal (n = 4) calculated for each mouse and for each transcriptional output category. The phase of rhythm (average ± s.e.m. from 4 independent rhythm) is indicated in the bottom right. Each rhythm is double-plotted for better visualization. (*Right*): average nucleosome signal for each transcriptional output group at CLOCK:BMAL1 DNA-binding sites (±0.6 kb) during the light phase (ZT2, ZT6, and ZT10; green) and dark phase (ZT14, ZT18, and ZT22; red/orange) of wild-type mice and in Bmal1-/- mice (average signal for six time points; black). E. maximal and minimal nucleosome signal from the 6-time points rhythms for each of the CLOCK:BMAL1 transcriptional output groups. Groups with different letters are significantly different (2-way ANOVA; $p < 0.05$). F. Circadian rhythm of H3K27ac ChIP-Seq signal in the mouse liver at CLOCK:BMAL1 DNA binding sites (calculated at CLOCK:BMAL1 peak center ± 1 kb) for each of the 4 CLOCK:BMAL1 transcriptional output group. Datasets were retrieved from Koike et al., 2012 [7] and re-analyzed (see methods for more details). Values correspond to the ChIP-Seq signal median for each group. Each rhythm is double-plotted for better visualization. G. Rhythm of enhancer RNA (eRNA) signal in the mouse liver at CLOCK:BMAL1 DNA binding sites (calculated at CLOCK:BMAL1 peak center ± 500 bp) for each of the 4 CLOCK:BMAL1 transcriptional output groups. Datasets were retrieved from Fang et al., 2014 [90] and re-analyzed (see methods for more details). Values correspond to the eRNA signal median for each group. Each rhythm is double-plotted for better visualization.

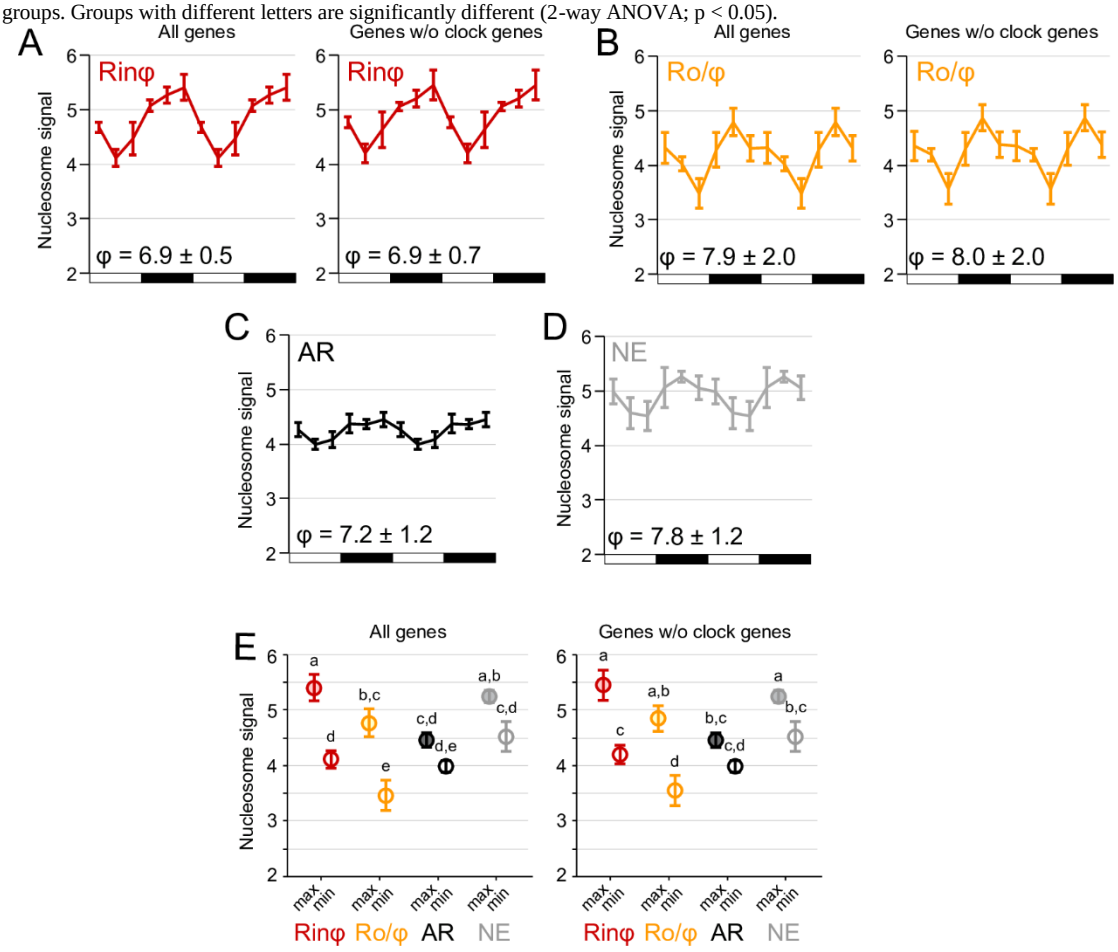


Importantly, the phase of the rhythms is similar for all groups and minimal nucleosome signal coincides with maximal CLOCK:BMAL1 DNA binding during the light phase. Closer inspection of the levels of nucleosome signal and rhythm amplitude reveals important differences between each of the four transcriptional output categories (Figure II-9 E). First, the amplitude of the rhythms is significantly decreased for arrhythmically transcribed target genes. While minimal levels of nucleosome signal during the day are similar between the AR and Rinφ groups, nucleosome signal remains low during the night (i.e., when CLOCK:BMAL1 is not bound to DNA) at CLOCK:BMAL1 peaks targeting AR genes (Figure II-9 E). This suggests that some transcription factors may be still bound to DNA during the night in the AR group (when CLOCK:BMAL1 is not bound to DNA), thereby preventing the reformation of nucleosomes. This may promote transcription at night and thus lead to arrhythmic transcription. Second, the overall nucleosome signal is significantly lower at CLOCK:BMAL1 peaks targeting Ro/φ genes than for Rinφ genes, without any significant effect on the amplitude of the rhythm (Figure II-9 E). In addition, the time of minimal nucleosome signal is delayed by 4 hours between Rinφ and Ro/φ: while it coincides with the time of maximal CLOCK:BMAL1 DNA binding for Rinφ genes (ZT06), minimal nucleosome signal is observed at ZT10 for Ro/φ genes. This delayed nucleosome signal for the out-of-phase transcriptional cyclers may be explained by the significant recruitment of REV-ERBα and REV-ERBβ (Figure II-8 A, B). Indeed, CLOCK:BMAL1 has been recently proposed to facilitate circadian repression by promoting the recruitment of REV-ERBα through chromatin decondensation [91]. Thus, the increased binding of REV-ERBs at CLOCK:BMAL1 enhancers at ZT10 may promote a further decrease in nucleosome signal. Furthermore, anti-phase binding of the RORs on ROREs during the night would prevent a full nucleosome re-compaction, thereby promoting lower levels of nucleosome signal at CLOCK:BMAL1 peaks targeting Ro/φ

target genes. Finally, there are no significant differences of nucleosome signal between CLOCK:BMAL1 DNA binding sites targeting in-phase transcriptional cyclers than those targeting non-expressed target genes (Figure II-9 E). This intriguing result suggests that although CLOCK:BMAL1 is unable to promote transcription activation at NE target genes, its rhythmic DNA binding still mediates a rhythm in nucleosome signal. One possible explanation for this result is that CLOCK:BMAL1 decondenses the chromatin to facilitate the binding of other transcription factors, but those would not be recruited at NE target genes except under specific conditions (e.g. environmental stressors), thereby preventing activation of transcription under standard conditions.

**Figure II-10 Rhythmic Nucleosome Signal at CLOCK:BMAL1 DNA Binding Sites.**

A-D: Nucleosome signal was retrieved from mouse liver MNase-Seq datasets (Menet et al., 2014), which consists of 6 time points each separated by 4 hours with n = 4 mice for each time point. Each graph displays a 6-time points rhythm of nucleosome signal at CLOCK:BMAL1 binding sites (calculated at CLOCK:BMAL1 peak center ± 75 bp for each peak), displayed as the average ± s.e.m. of the signal (n = 4) calculated for each mouse and for each transcriptional output category: (A) Rhythmic-in-phase (Rinφ, red); (B) Rhythmic out-of-phase (Ro/φ, orange); (C) arrhythmic (AR, black); (D) non expressed (NE, grey) target genes). The phase of rhythm (average ± s.e.m. from 4 independent rhythm, calculated by Fourier transform) is indicated in the bottom right. Each rhythm is double-plotted for better visualization. For both Rinφ and Ro/φ groups, the nucleosome rhythm is calculated at all CLOCK:BMAL1 peaks targeting rhythmic target genes (left), or only at peaks targeting rhythmic non core clock genes (removal of nucleosome signal at peaks targeting *Cry2*, *Dbp*, *Rev-erbα*, and *Rev-erbβ* for the Rinφ group, and of the peaks targeting *Per1* and *Per2* for the Ro/φ group. E: maximal and minimal nucleosome signal from the 6-time points rhythms for each of the CLOCK:BMAL1 transcriptional output groups. Groups with different letters are significantly different (2-way ANOVA; $p < 0.05$).



***CLOCK:BMAL1 does not directly contribute to the transcriptional activity of its enhancers***

Our data indicate that CLOCK:BMAL1 rhythmic DNA binding promotes the rhythmic removal of nucleosomes at all four transcriptional output categories. We then asked if CLOCK:BMAL1 can also promote the formation of transcriptionally active enhancers. To address this question, we used public datasets [7,19,90] to examine the rhythmic pattern of two independent markers of enhancer activity at CLOCK:BMAL1 DNA binding sites: the post-translational modification H3K27 acetylation (H3K27ac), which positively correlates with enhancer activity at almost all enhancers and TSS [92], and the expression levels of enhancer RNA (eRNA), which are relatively short non-coding RNA molecules (50-2000 nucleotides) transcribed at active enhancer regions [93].

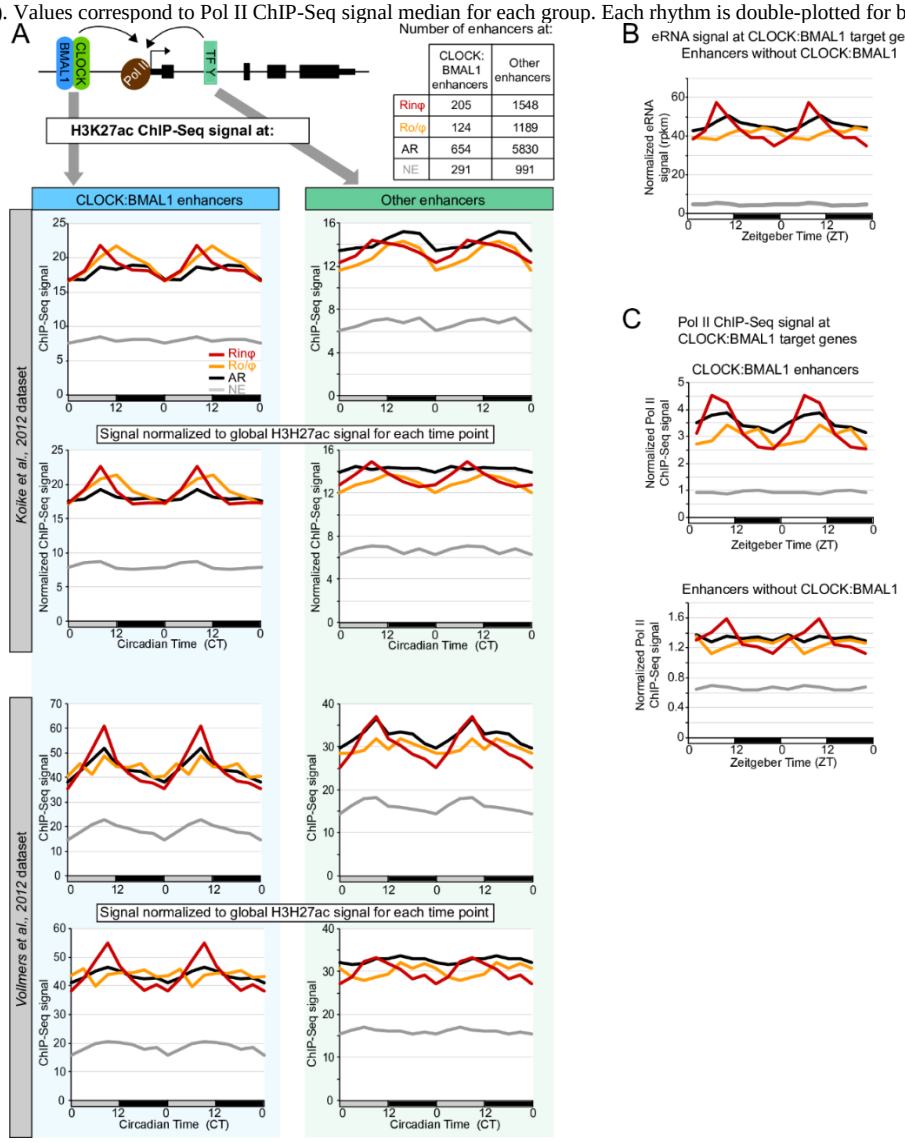
While mouse liver H3K27ac ChIP-Seq signal is rhythmic and high during the light phase at CLOCK:BMAL1 DNA binding sites targeting in-phase transcriptional cyclers (consistent with CLOCK:BMAL1 directly facilitating the acetylation of H3K27; Figure II-9 F and Figure II-11), significant differences were observed at CLOCK:BMAL1 DNA binding sites targeting the other 3 transcriptional output categories. Rhythmic H3K27ac rhythm is delayed for the out-of-phase transcriptional cyclers, and the amplitude of H3K27ac rhythm is significantly dampened at CLOCK:BMAL1 DNA binding sites targeting arrhythmically transcribed genes (Figure II-9 F and Figure II-11 A). Remarkably, levels of H3K27ac are close to background levels at CLOCK:BMAL1 peaks targeting non-expressed genes. Given that CLOCK:BMAL1 rhythmically binds to relatively similar levels for all four transcriptional output categories, our analysis suggests that CLOCK:BMAL1 DNA binding does not directly contribute to the acetylation of H3K27.

To extend on this observation, we then examined another marker of enhancer transcriptional activity by assessing eRNA expression at CLOCK:BMAL1 DNA binding sites

using a publicly available GRO-Seq dataset [90]. The analysis confirmed the results obtained with H3K27ac (Figure II-9 G). Rhythmic eRNA expression is only observed at CLOCK:BMAL1 enhancers targeting rhythmically transcribed genes, and eRNA expression at enhancers targeting non-expressed genes is dramatically decreased to levels close to background (Figure II-9 G). Importantly, these differences in eRNA expression between the four CLOCK:BMAL1 transcriptional output categories are further corroborated by similar variations in RNA Polymerase II (Pol II) ChIP-Seq signal at CLOCK:BMAL1 enhancers (Figure II-11 C). Altogether, our analysis therefore demonstrates that, contrary to what has been typically found for core clock genes, CLOCK:BMAL1 DNA binding is not sufficient to promote the activation of its enhancers. Instead, our results suggest that CLOCK:BMAL1 rhythmically opens the chromatin to facilitate the binding of other transcription factors at its enhancers, and that the nature of these transcription factors (e.g., activators, repressors) significantly contributes to CLOCK:BMAL1 transcriptional output.

**Figure II-11 CLOCK:BMAL1 Does Not Directly Promote H3K27ac Post-Translational Modification.**

A. Circadian rhythm of H3K27ac ChIP-Seq signal in the mouse liver at CLOCK:BMAL1 DNA binding sites (*Left; blue background*) or at non-CLOCK:BMAL1 enhancers located in CLOCK:BMAL1 target genes (*Right; green background*) for each of the 4 CLOCK:BMAL1 transcriptional output groups: rhythmic-in-phase (Rinφ, red); rhythmic out-of-phase (Roφ, orange); arrhythmic (AR, black); and non expressed (NE, grey) target genes. Datasets were retrieved from Koike et al., 2012 (top) or Vollmers et al., 2012 (bottom) and re-analyzed (see methods section for more details). Values correspond to the ChIP-Seq signal median for each group, and were calculated for each CLOCK:BMAL1 peak as the average of reads/bp at CLOCK:BMAL1 DNA binding sites center ± 1 kb normalized to one million sequencing reads. For each dataset, H3K27ac ChIP-Seq signal was further normalized by mean normalization to account for the differences in ChIP-Seq efficiency between each sequencing sample (bottom graphs for each datasets). This normalization assumes that the overall genome-wide levels of H3K27ac are constant at any time in the mouse liver. To this end, we normalized H3K27ac ChIP-Seq signal for each peak to the averaged H3K27ac signal calculated at the top 40,000 DNase hypersensitive sites for each time point  Graphs are double-plotted to improve visualization. B. Rhythm of enhancer RNA (eRNA) signal in the mouse liver at enhancers targeting a CLOCK:BMAL1 target gene and that do not harbor a CLOCK:BMAL1 DNA binding site (calculated at CLOCK:BMAL1 peak center ± 500 bp) for each of the 4 CLOCK:BMAL1 transcriptional output group. Datasets were retrieved from Fang et al., 2014 [90] and re-analyzed (see methods section for details). Values correspond to the eRNA signal median for each group. Each rhythm is double-plotted for better visualization. C. Rhythm of RNA Polymerase II ChIP-Seq signal in the mouse liver at CLOCK:BMAL1 enhancers (top) and enhancers targeting a CLOCK:BMAL1 target gene but without a CLOCK:BMAL1 DNA binding site for each of the 4 CLOCK:BMAL1 transcriptional output group. Datasets were retrieved from Le Martelot et al., 2012 [77], re-analyzed and normalized by a ranking analysis (see methods section for details). Values correspond to Pol II ChIP-Seq signal median for each group. Each rhythm is double-plotted for better visualization.



***Differential recruitment of transcription factors at CLOCK:BMAL1 enhancers***

To test our hypothesis that transcription factors bind at CLOCK:BMAL1 enhancers to contribute to their transcriptional activity and thereby impact on CLOCK:BMAL1-mediated transcription, we assessed if transcription factors were differentially recruited at CLOCK:BMAL1 DNA binding sites within each transcriptional output group. To this end, we performed a DNA binding motif analysis using HOMER Software Suite that we further validated using mouse liver transcription factor ChIP-Seq datasets. As expected, the motif analysis revealed that CLOCK:BMAL1 DNA binding motif (e-box of the sequence CACGTG) is highly enriched at CLOCK:BMAL1 enhancers for all transcriptional output categories (Figure II-12 A). Surprisingly however, we found that motifs for liver-specific transcription factors (e.g., *Cebp*, *Hnf1*, *Hnf4* and *Hnf6*) were also enriched for all four transcriptional output categories, and thus even at CLOCK:BMAL1 enhancers targeting non-expressed genes (Figure II-12 B, Figure II-13). On the contrary, motifs for ubiquitous transcription factors (u-TFs; broadly expressed transcription factors with a transcriptional activity regulated by external factors) were almost always enriched for specific CLOCK:BMAL1 transcriptional output groups (Figure II-12 C, Figure II-13). For example, CRE motif was enriched at all CLOCK:BMAL1 enhancers except those targeting out-of-phase transcriptional cyclers, and FXR motif was enriched at all CLOCK:BMAL1 enhancers except those targeting out-of-phase transcriptional cyclers. Noticeably, the motifs for NF-κB which binds to DNA and becomes transcriptionally active upon infection and inflammation; [94,95], and CTCF which establishes discrete functional chromatin domains by promoting DNA looping; [96,97,98] were enriched at enhancers targeting on-expressed genes.

**Figure II-12 Tissue-Specific and Ubiquitous Transcription Factors Are Differentially Recruited at CLOCK:BMAL1 Enhancers.**

A-C. Enrichment for the DNA binding motif of CLOCK:BMAL1 (A), tissue-specific transcription factors (B) and ubiquitous transcription factors (C) at CLOCK:BMAL1 DNA binding sites for each of the four CLOCK:BMAL1 transcriptional output categories. Enrichment was calculated using HOMER and is reported as the ratio between the calculated enrichment over the calculated background. * $q < 0.05$ (Benjamini-Hochberg procedure). D-F. ChIP-Seq signal of tissue-specific transcription factors (D), ubiquitous transcription factors (E), and transcriptional co-activators / RNA Polymerase II at ZT6 (F) at CLOCK:BMAL1 DNA binding sites (peak center ± 250bp) for each of the transcriptional output categories. Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$). G. Transcription factor DNA binding variability index at CLOCK:BMAL1 DNA binding sites. The TF DNA binding variability index reflects differential TF DNA binding by calculating the variance of TF ChIP-Seq signal between the four CLOCK:BMAL1 transcriptional output groups (see Methods for details). The variability index is displayed as a dot for each TF: CLOCK, BMAL1, PER1, PER2, CRY1, CRY2 (blue), seven ts-TFs (CEBPA, CEBPB, FOXA1, FOXA2, HNF1, HNF4A, and HNF6; red), thirteen u-TFs (REV-ERBα, REV-ERVβ, RORα, E4BP4, RXR, LXR, PPARα, GR-ZT12, E2F4, STAT5, BCL6, ERα, and GABPA; green), as well as for p300, CBP, and Pol II at seven time points (ZT02 to ZT26) (black). The horizontal lines represent the variability index median for the first 3 groups of TF. ChIP-Seq datasets used in this analysis are described in the method section. The variability index was calculated using all CLOCK:BMAL1 peaks analyzed in Fig. 1 (left), or CLOCK:BMAL1 peaks that do not target a clock gene (removal of TF ChIP-Seq signal at peaks targeting *Per1*, *Per2*, *Cry2*, *Dbp*, *Rev-erbα*, and *Rev-erbβ*, *Tef*, *Hlf*, *Gm129*, and *Rorγ*; right).

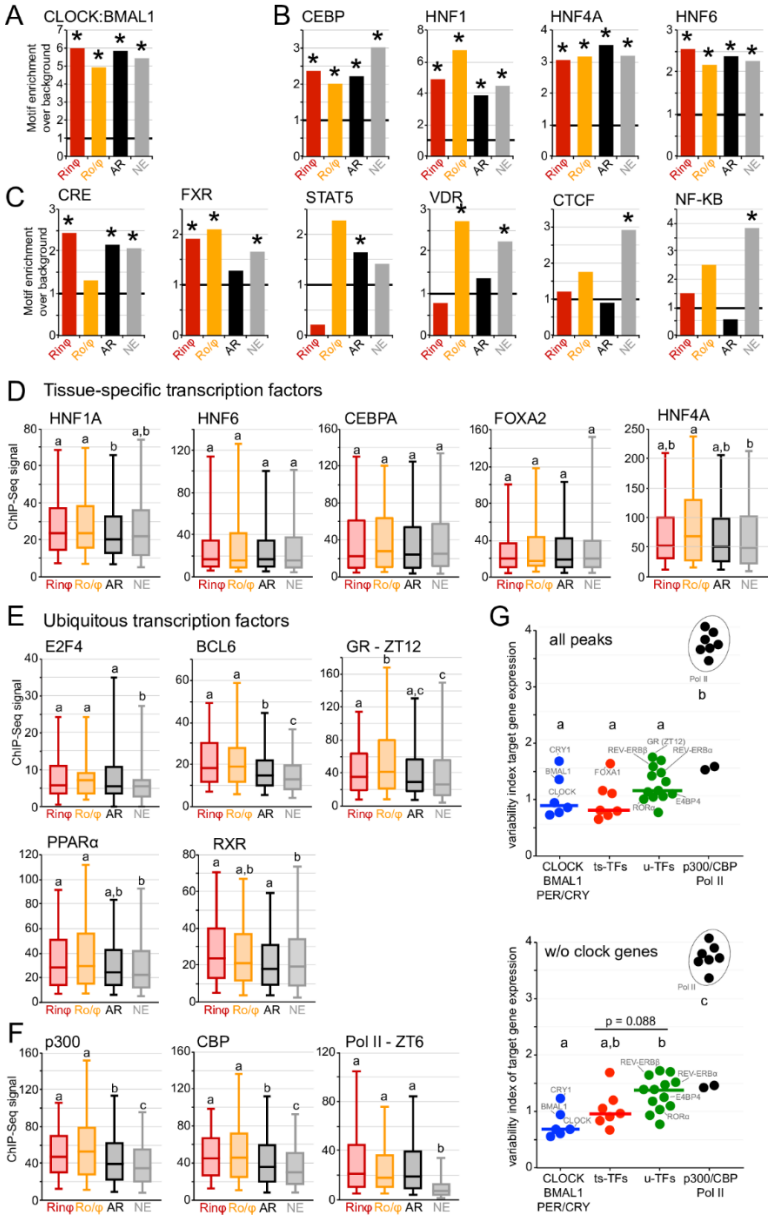
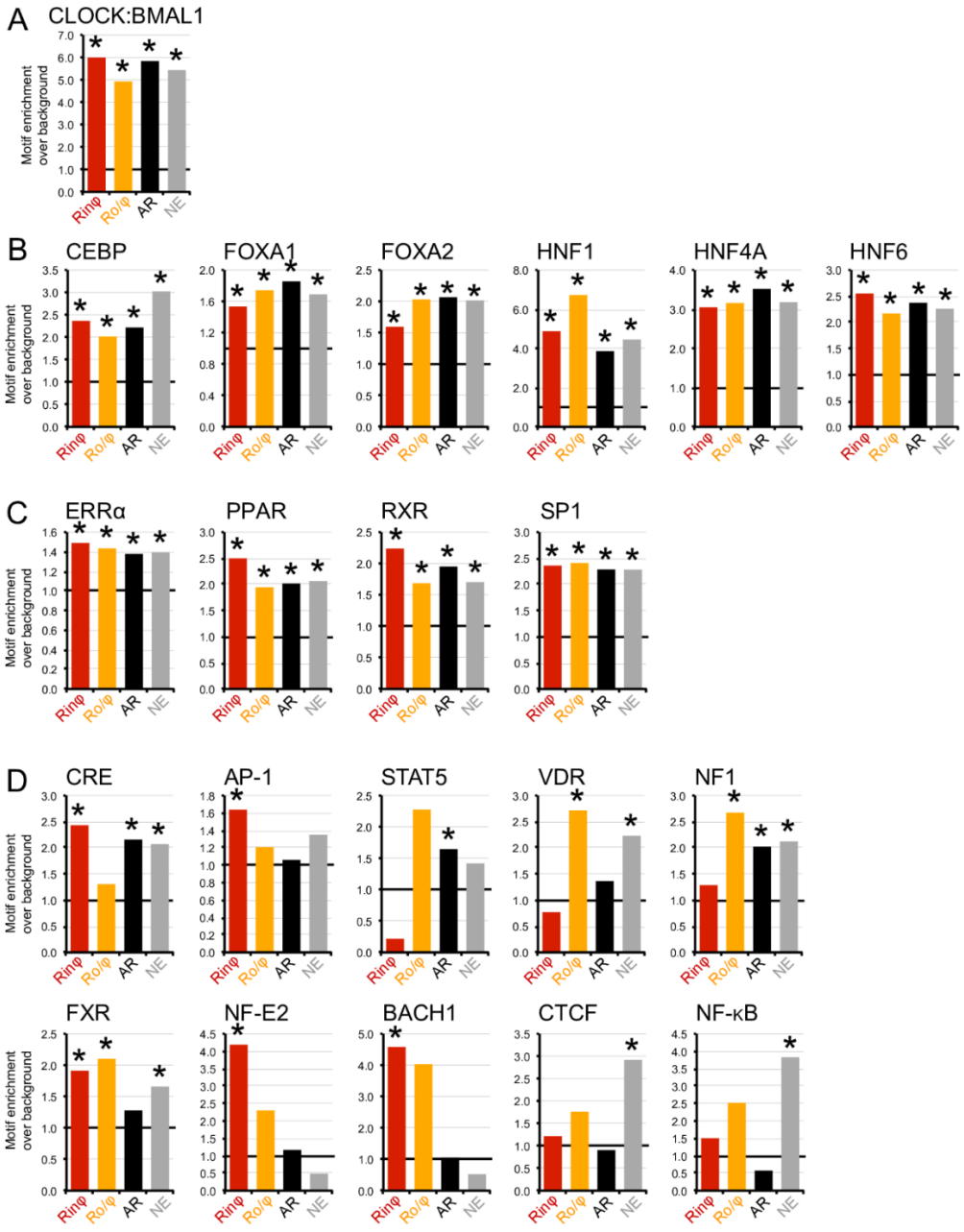


**Figure II-13 Transcription Factor DNA Binding Motif Analysis at CLOCK:BMAL1 Enhancers and Based on CLOCK:BMAL1 Transcriptional Output.**

Enrichment for transcription factor DNA binding motifs (calculated using the HOMER software suite) at CLOCK:BMAL1 DNA binding sites for each of the four CLOCK:BMAL1 transcriptional output categories: rhythmic-in-phase (Rinφ, red); rhythmic out-of-phase (Ro/φ, orange); arrhythmic (AR, black); and non expressed (NE, grey) target genes. Enrichment is reported as the ratio between the calculated enrichment over the calculated background. * q < 0.05 (Benjamini-Hochberg procedure). Motif enrichment is shown for: CLOCK:BMAL1 (A); tissue-specific transcription factors (B); as well as ubiquitous transcription factors for which the motif is enriched for all 4 output groups (C) or specific group(s) (D).

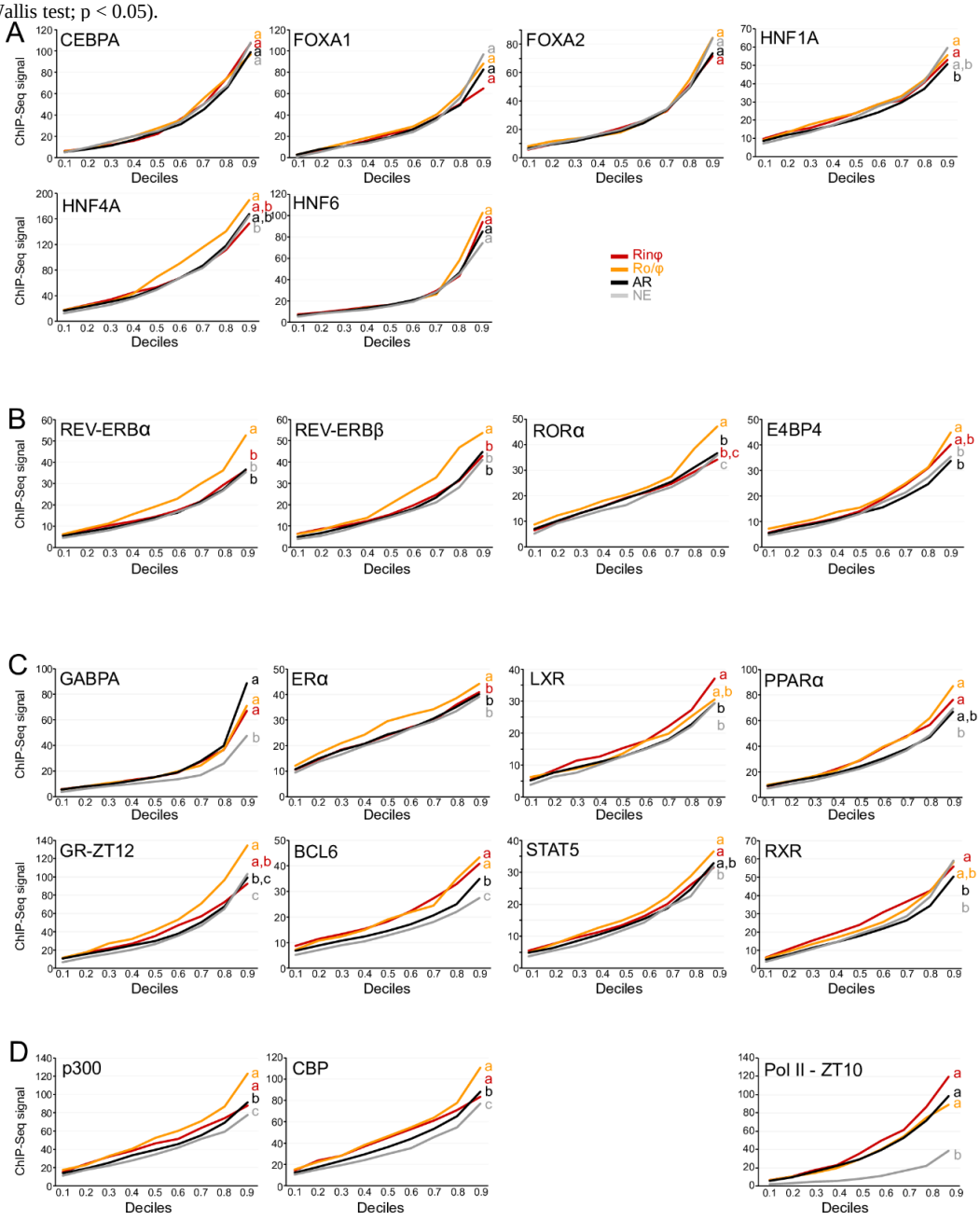


To assess the relevance of this difference of motif enrichments between tissue-specific (ts-TFs) and u-TFs, we determined the DNA binding pattern of several transcription factors at CLOCK:BMAL1 enhancers in the mouse liver using publicly available transcription factors ChIP-Seq datasets [99-104]. This *in vivo* analysis largely confirmed the computational motif analysis: most liver-specific TFs were found to bind at CLOCK:BMAL1 DNA binding sites independently of the transcriptional output, whereas u-TFs were more specifically enriched in specific CLOCK:BMAL1 transcriptional output groups (Figure II-12 D, E, Figure II-14). For example, *Hnf4a* and *Hnf1* are the only liver-specific TF to exhibit a differential binding between CLOCK:BMAL1 transcriptional output groups of the six TFs tested (*Foxa1*, *Foxa2*, *Hnf1*, *Hnf4A*, *Hnf6* and *Cepba*). Conversely, all twelve u-TFs investigated exhibit DNA binding differences at CLOCK:BMAL1 enhancers between categories of transcriptional output (Figure II-12 D, E, Figure II-14). Although each u-TF bound to different subsets of CLOCK:BMAL1 enhancers, u-TF recruitment was generally higher in rhythmically expressed target genes and lower in non-expressed target genes compared to the arrhythmic CLOCK:BMAL1 target group. (Figure II-12 D, E, Figure II-14).

To further characterize the differences in TF DNA binding between CLOCK:BMAL1, u-TFs and ts-TFs, we computed a TF DNA binding variability index by calculating the standard deviation of the ChIP-Seq signal between the 4 CLOCK:BMAL1 transcriptional output groups (see Methods for details). We found that the DNA binding variability at CLOCK:BMAL1 peaks is comparable between CLOCK:BMAL1 and ts-TFs, whereas there is significantly more variability for u-TFs than for CLOCK and BMAL1 when peaks targeting core clock genes are removed from the analysis (Figure II-12 G). While there are variability index differences among ts-TFs and u-TFs, this analysis further supports our finding that u-TF recruitment at CLOCK:BMAL1 peaks is globally more variable than for ts-TF (Figure II-12 G).

**Figure II-14 Transcription Factor ChIP-Seq Signal at CLOCK:BMAL1 Enhancers and Based on CLOCK:BMAL1 Transcriptional Output.**

Mouse liver ChIP-Seq signal of tissue-specific transcription factors (A), ubiquitous transcription factors (B, C), and transcriptional co-activators / RNA Polymerase II at ZT10 (D) at CLOCK:BMAL1 DNA binding sites (peak center ± 250bp) for each of the transcriptional output categories. ChIP-Seq signal is represented for each output group based on its distribution (every decile). Groups with different letters are significantly different (Kruskal-Wallis test; $p < 0.05$).

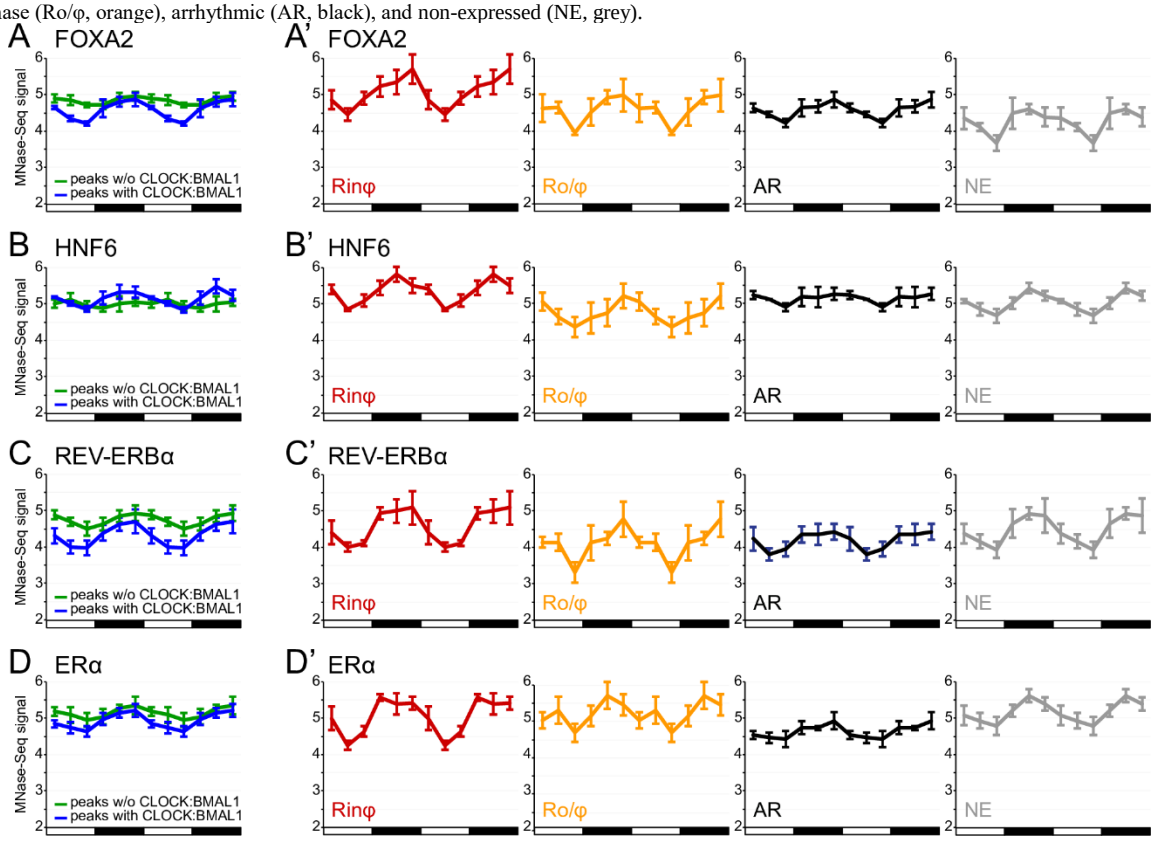


***CLOCK:BMAL1 likely cooperates with other transcription factors to regulate the transcription of its direct target genes***

Altogether, our data indicate that the mechanisms by which CLOCK:BMAL1 regulates transcription of clock-controlled genes differ from the well-characterized CLOCK:BMAL1-mediated regulation of core clock gene expression. Specifically, our data show that although CLOCK:BMAL1 mediates rhythmic nucleosome removal at its enhancers, it is not sufficient to generate an active enhancer or drive rhythmic transcription. We thus propose a model whereby CLOCK:BMAL1 regulates transcription of clock-controlled genes by rhythmically opening chromatin to facilitate the binding of other transcription factors at its enhancers (Figure II-16 A). This possibility is supported by results showing that nucleosome signal is rhythmic at the DNA binding sites of several TFs when those sites are located close to a CLOCK:BMAL1 peak, and not rhythmic when CLOCK:BMAL1 binding is absent (Figure II-15).

**Figure II-15 Analysis of Nucleosome Signal Over the 24-hr Day at the DNA Binding Sites of Two Tissue-specific TF and Two Ubiquitous TF in the Mouse Liver.**

Nucleosome signal at four TF DNA binding sites was retrieved from mouse liver MNase-Seq datasets (Menet et al., 2014), which consists of 6 time points each separated by 4 hours with n = 4 mice for each time point. MNase-Seq data are displayed for two tissue-specific TFs: FOXA2 (A) and HNF6 (B), and two ubiquitous TFs: REV-ERBα (C) and ERα (D). Nucleosome signal was calculated at TF peak center ± 75 bp for each peak and at each TF ChIP-Seq peak coordinate, and is displayed as the average ± s.e.m of the signal (n = 4) calculated for each mouse. Each rhythm is double-plotted for better visualization. *(Left, A-D)* Nucleosome signal at TF ChIP-Seq peaks harboring a CLOCK:BMAL1 peak (blue), or at the top 10,000 TF ChIP-Seq peaks that do not harbor a CLOCK:BMAL1 peak (green). *(Right, A'-D')* Nucleosome signal at TF ChIP-Seq peaks harboring a CLOCK:BMAL1 peak, and parsed based on CLOCK:BMAL1 transcriptional output: Rhythmic-in-phase (Rinφ, red), Rhythmic out-of-phase (Ro/φ, orange), arrhythmic (AR, black), and non-expressed (NE, grey).



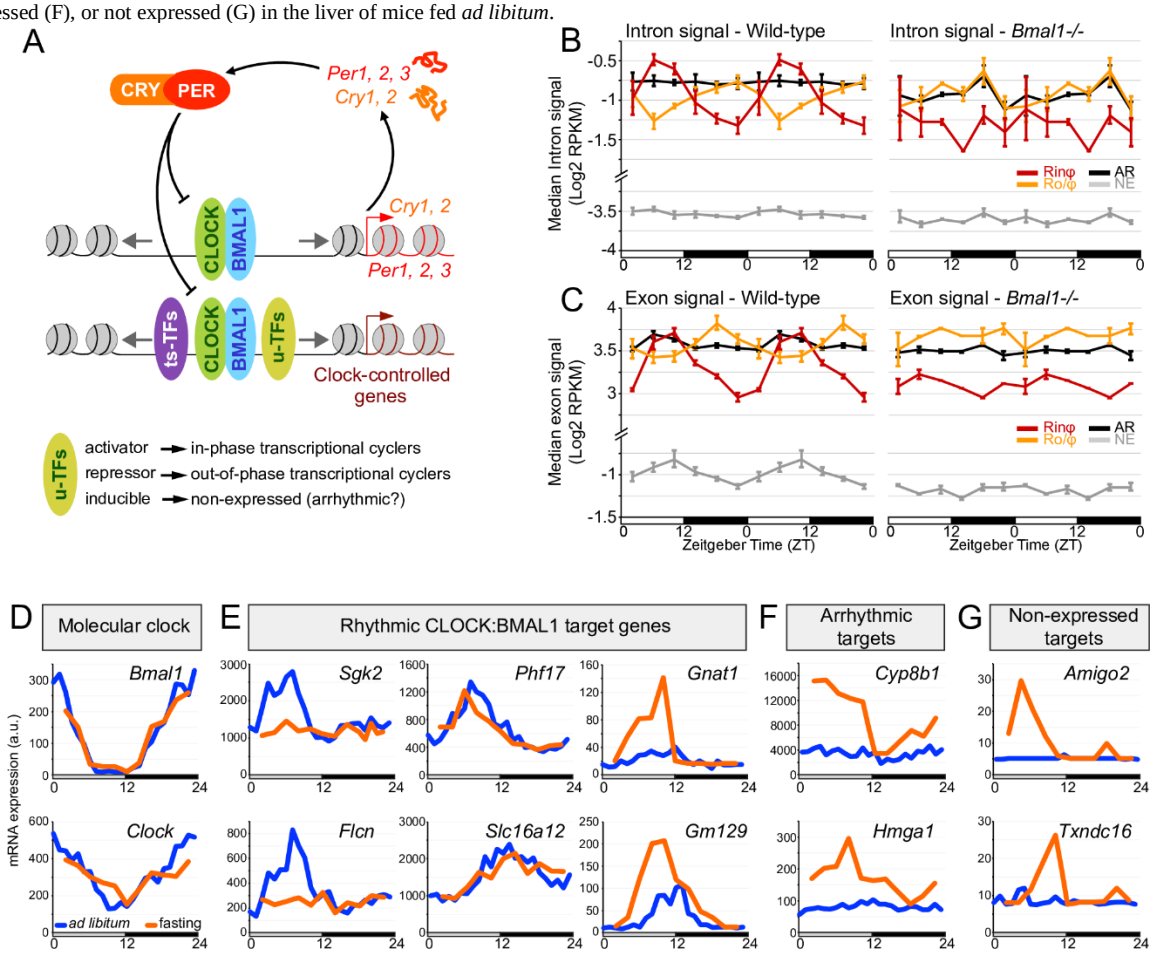
Consequently, the transcriptional activities of these transcription factors would dictate the transcriptional outcome of clock-controlled genes rather than CLOCK:BMAL1 (Figure II-16 A). For example, binding of positive transcription factors along with CLOCK:BMAL1 would activate enhancers and lead to transcription activation during the day, whereas binding of transcriptional repressors (e.g., REV-ERBα/β) would inhibit CLOCK:BMAL1 enhancer activity and thereby contribute to rhythmic transcription peaking during the night, in anti-phase with CLOCK:BMAL1 DNA binding (Figure II-16 A). If no transcription factors are recruited (e.g., inducible transcription factors), CLOCK:BMAL1 enhancers remain silent and target genes are not expressed or are arrhythmically expressed (Figure II-16 A). Arrhythmically expressed genes at CLOCK:BMAL1 enhancers may also have positive transcription factors bound at all times overriding the absence of CLOCK:BMAL1 DNA binding at night (see result section about rhythmic nucleosome signal and Figure II-9 E). Our results also suggest that u-TFs regulate CLOCK:BMAL1 transcriptional output more prevalently than ts-TFs. It may be that ts-TFs facilitate the binding of CLOCK:BMAL1 at tissue-specific enhancers rather than contributing to CLOCK:BMAL1 transcriptional output (see discussion). To validate this model experimentally, we investigated how i) *Bmal1* knockout, and ii) changes in environmental conditions (that alter u-TFs transcriptional activities) affect CLOCK:BMAL1 transcriptional output.

If the activity of u-TFs contributes to CLOCK:BMAL1 regulation of clock-controlled gene transcription, then, a knockout of *Bmal1* (which eliminates CLOCK:BMAL1-mediated transcription [105]) should differentially affect the expression of CLOCK:BMAL1 target genes. Specifically, target gene expression levels in *Bmal1-/-* mouse should be arrhythmic and low for the in-phase transcriptional cyclers (no recruitment of positive transcription factors at CLOCK:BMAL1 enhancers), while they should be arrhythmic and high for the out-phase

transcription cyclers (no recruitment of repressors at CLOCK:BMAL1 enhancers). These effects should also be more obvious during the light phase when CLOCK:BMAL1 binds to DNA. In addition, *Bmal1* knockout should have a reduced effect on arrhythmically and non-expressed target genes. These predictions were confirmed by analyzing a public dataset that characterized the genome-wide effect of *Bmal1* knockout in the mouse liver using RNA-Seq of rRNA-depleted total RNA (Figure II-16 B, C) [106]. For both intronic and exonic RNA-Seq signal, the expression of Rinφ genes in *Bmal1-/-* mouse liver is at the trough level of wild-type mice, and at peak levels in Ro/φ genes. Moreover, *Bmal1* knockout does not significantly affect the expression levels of arrhythmic and non-expressed CLOCK:BMAL1 target genes (Figure II-16 B, C).

**Figure II-16 CLOCK:BMAL1 Regulation of Clock-controlled Gene Expression Likely Relies on the Cooperation of CLOCK:BMAL1 with Other Transcription Factors.**

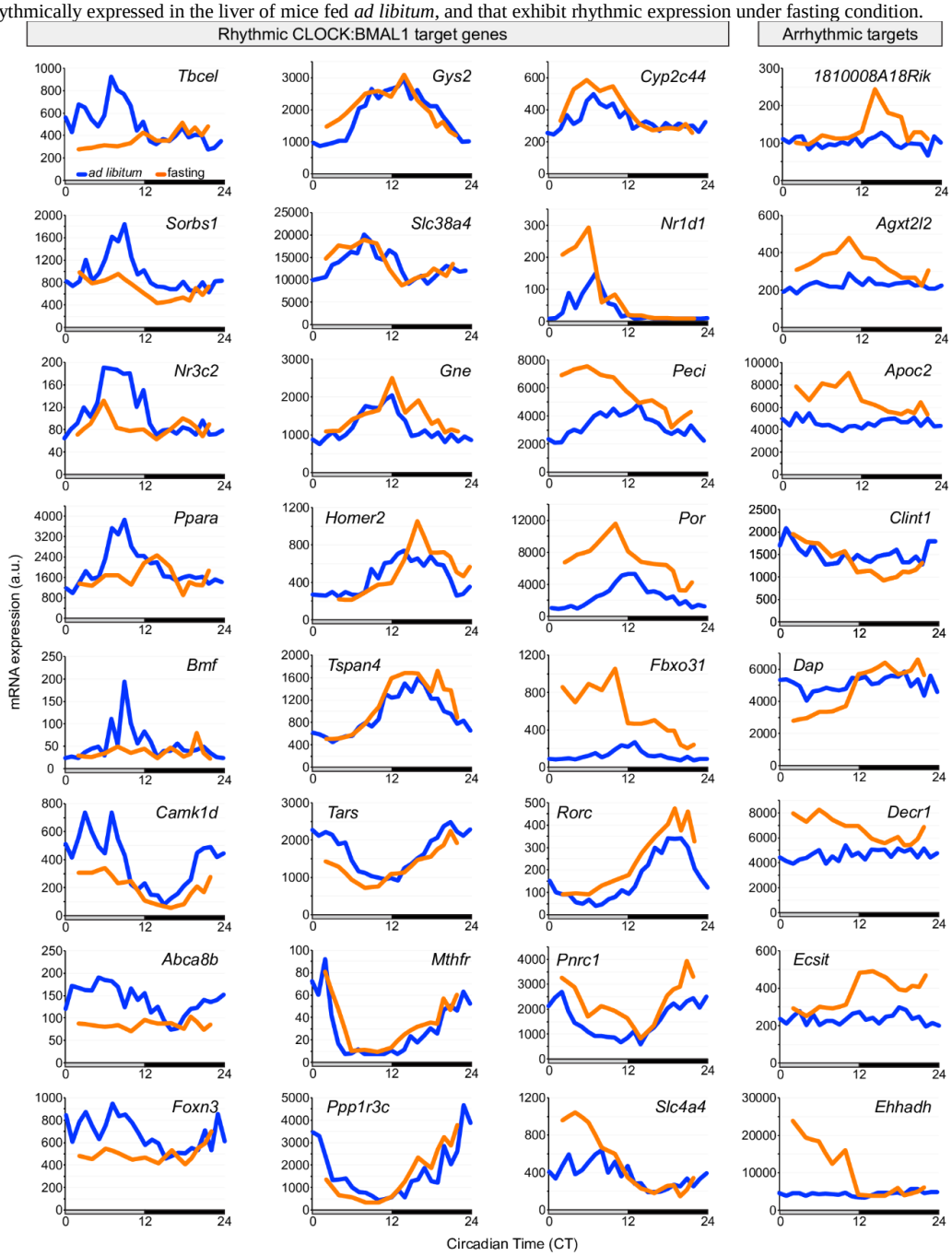
A. Proposed model incorporating tissue-specific (ts-TFs) and ubiquitous (u-TFs) transcription factors into CLOCK:BMAL1 regulation of clock-controlled gene transcription. See text for details. B, C. Intron (B) and exon (C) signals of direct CLOCK:BMAL1 target genes classified based on their transcriptional output (Rinφ: in-phase transcriptional cyclers; Ro/φ: out-of-phase transcriptional cyclers; AR: arrhythmically transcribed target genes; NE: non-expressed target genes), in wild-type (left) and *Bmal1-/-* (right) mouse liver. Values correspond to the median RPKM for each transcriptional output group, and are displayed as the average ± s.e.m. of four (wild-type) or two (*Bmal1-/-*) independent samples for each time point. Data were retrieved from public RNA-Seq datasets [106], and are double-plotted for better visualization. D-G. Rhythm of mRNA expression in the liver of mice fed *ad libitum* (blue) or fasted for at least 22 hours (orange). Data were retrieved from a public dataset [107]. Mouse liver mRNA expression is displayed for *Clock* and *Bmal1* (D), as well as CLOCK:BMAL1 target genes that are rhythmically expressed (E), arrhythmically expressed (F), or not expressed (G) in the liver of mice fed *ad libitum*.



Our model also predicts that the transcriptional output of CLOCK:BMAL1 target genes can be altered by environmental changes that affect u-TF DNA binding capacity. External signals that activate or repress the binding of u-TFs are predicted to impact CLOCK:BMAL1 cooperation with other transcription factors, and thereby change the transcriptional output of CLOCK:BMAL1 target genes. For example, signals that enable the recruitment of positive transcription factors at CLOCK:BMAL1 enhancers could increase the amplitude of rhythmic transcription and/or initiate the rhythmic expression of target genes that are arrhythmic under control conditions. Conversely, signals that inhibit the binding of transcription factors that normally cooperate with CLOCK:BMAL1 could blunt the rhythmic expression of some CLOCK:BMAL1 target genes. To test this hypothesis, we analyzed how fasting, which is known to affect the transcriptional activity of many u-TFs [108,109], alters CLOCK:BMAL1 target gene expression in the mouse liver using a public dataset [107]. Strikingly, while the expression of *Clock*, *Bmal1* and several direct rhythmic target genes (e.g., *Phf17*, *Slc16a2*) are unaffected by fasting, some other targets exhibit a significantly altered gene expression profile (Figure II-16 D-G, Figure II-17 for additional examples). For example, some rhythmic target genes become arrhythmically expressed under fasting (e.g., *Sgk2, Flcn*) while other targets exhibit an increased amplitude of expression (e.g., *Gnat1, Gm129*) (Figure II-16 E).

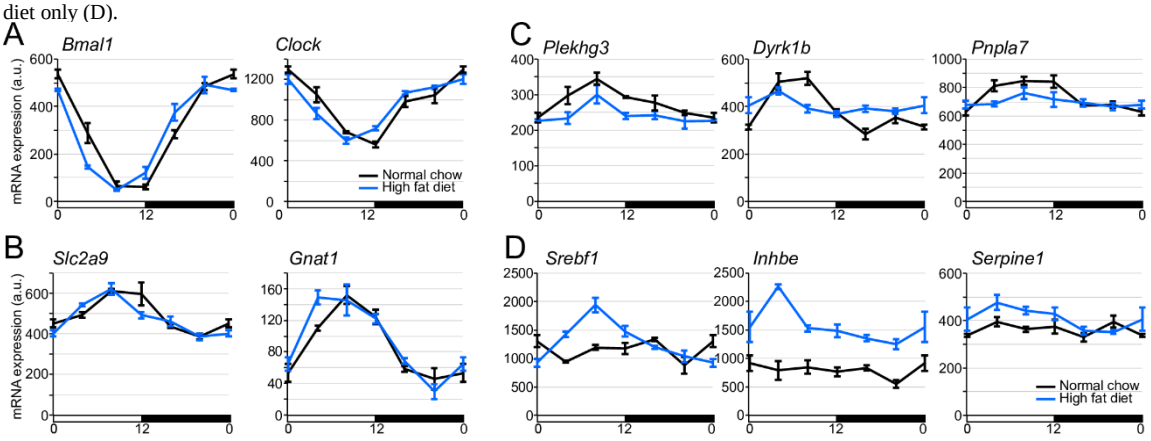
**Figure II-17 Additional Examples of CLOCK:BMAL1 Target Genes Exhibiting Changes in Expression Under Fasting Condition in the Mouse Liver.**

Rhythm of mRNA expression in the liver of mice fed ad libitum (blue) or fasted for at least 22 hours (orange). Data were retrieved from a public dataset (Vollmers et al., 2009). Mouse liver mRNA expression is displayed for CLOCK:BMAL1 target genes that are rhythmically expressed in the liver of mice fed *ad libitum*, and that exhibit under fasting condition a decrease in the rhythm amplitude (first column), no change (second column) or an increase in the rhythm amplitude (third column). Mouse liver mRNA expression is also displayed for CLOCK:BMAL1 target genes that are arrhythmically expressed in the liver of mice fed *ad libitum*, and that exhibit rhythmic expression under fasting condition.



Remarkably, some direct CLOCK:BMAL1 target genes that are arrhythmically or not expressed under *ad libitum* condition become rhythmically expressed under fasting condition (Figure II-16 F, G). Because not all CLOCK:BMAL1 target genes are equally affected by fasting, these results cannot simply be explained by a global change in CLOCK:BMAL1 transcriptional activity under fasting condition. One possibility is that fasting enhances or represses the transcriptional capabilities of several u-TFs that cooperate with CLOCK:BMAL1, thereby altering the transcriptional output of many direct CLOCK:BMAL1 target genes. Similar results were found by analyzing a public dataset investigating the effect of high-fat diet on rhythmic gene expression in the mouse liver (Figure II-18) [110].

**Figure II-18 Effect of High-fat Diet on CLOCK:BMAL1 Target Gene Expression.**
A-D. Six-time point rhythm of liver mRNA expression in mice fed with normal chow (black) or high fat diet (blue). Values were retrieved from a public dataset [110] and correspond to the average ± s.e.m. of 3 independent samples. Mouse liver mRNA expression is displayed for *Clock* and *Bmal1* (A), and some CLOCK:BMAL1 target genes that are rhythmic under both normal chow and high fat diet (B); normal chow only (C); and high fat diet only (D).

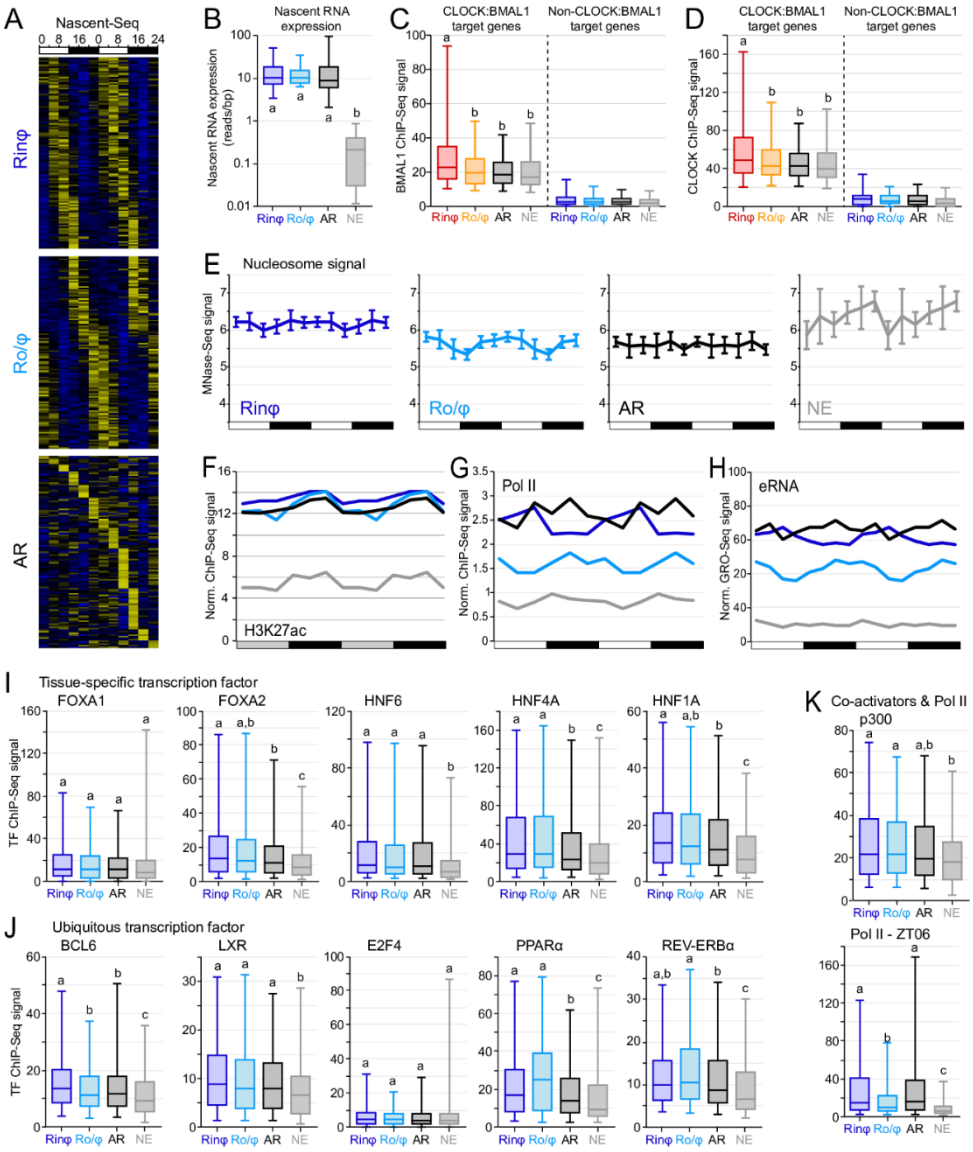


## Discussion

Based on the mechanisms by which CLOCK:BMAL1 regulates the expression of several core clock genes, it is commonly assumed that the rhythmic binding of CLOCK:BMAL1 to DNA is necessary and sufficient to drive the rhythmic transcription of its target genes. However, the recent characterization of CLOCK and BMAL1 cistromes in the mouse liver revealed that CLOCK:BMAL1 target gene transcription is highly heterogeneous, thereby suggesting that CLOCK:BMAL1 regulation of clock-controlled gene expression relies on more complex mechanisms than those underlying core clock gene rhythmic transcription [7,12,83,84]. We report here that CLOCK:BMAL1 heterogeneous transcriptional output does not stem from differences in the DNA binding profiles of CLOCK and BMAL1, or the PER/CRY circadian repressive complex. Instead, we found that while CLOCK:BMAL1 rhythmically promotes chromatin decondensation at its enhancers, it is not sufficient to promote transcription activation. Based on these data and the characterization of transcription factor DNA binding profiles at CLOCK:BMAL1 enhancers, we propose that CLOCK:BMAL1 regulates the expression of clock-controlled genes by generating a permissive chromatin landscape that facilitates the binding of other transcription factors at its enhancers rather than directly promoting rhythmic transcription. Interestingly, analysis of a random set of genes not directly targeted by CLOCK:BMAL1 but exhibiting similar profiles of expression of the four CLOCK:BMAL1 transcriptional output categories suggests that this mechanism is largely specific to CLOCK:BMAL1 (Figure II-19).

**Figure II-19 Analysis of Nucleosome Signal, Enhancer Activity, and TF DNA Binding at Cis-regulatory Regions Targeting Non-CLOCK:BMAL1 Target Genes.**

Analysis of a random set of genes not targeted by CLOCK:BMAL1 and transcribed similarly to the 4 CLOCK:BMAL1 transcriptional output groups (n = 125 genes for each group) was performed to determine the extent to which the findings reported in the manuscript are specific to CLOCK:BMAL1. The same criteria as those used for the characterization of CLOCK:BMAL1 transcriptional output were used, and groups are thus similarly referred as Rhythmic-in-phase (Rinφ, dark blue), Rhythmic out-of-phase (Ro/φ, light blue), arrhythmic (AR, black), and non expressed (NE, grey). Cis-regulatory regions targeting the randomly selected control genes are defined as DNase I hypersensitive sites (DHS) located within -10kb of the TSS to +1kb of the TTS (similarly to what was done for CLOCK:BMAL1 target genes). Analysis of these DHS suggests that many of our findings are specific to CLOCK:BMAL1 A. Heatmap displaying nascent RNA expression of the random set of genes and parsed based on the transcriptional output. Nascent-Seq signal was ordered based on the phase of nascent RNA oscillations for the in-phase and out-of-phase transcriptional cyclers. Ordering of arrhythmically transcribed genes is based on the peak time of maximal expression; the lack of a distinctive 24-hr rhythm profile of nascent RNA expression over the 48-hr time-scale is indicative of arrhythmic transcription. NE genes are not displayed due to the lack of expression. B. Average nascent RNA expression level for the 4 control groups. C-D. BMAL1 (C) and CLOCK (D) ChIP-Seq signal at DNase I hypersensitive sites (DHS) targeting the randomly selected control genes. ChIP-Seq signal for CLOCK:BMAL1 target genes is provided for comparison. E. Nucleosome rhythm at DHS targeting the randomly selected control genes (similar to Figure II-9 A-D). F-H. H3K27ac, Pol II and eRNA expression at DHS targeting the randomly selected control genes (similar to Figure II-9 F, Figure II-11 C and Figure II-4 G, respectively). I-K. ts-TF (I), u-TF (J), and p300 and Pol II (K) ChIP-Seq signal at DHS targeting the randomly selected control genes (similar to Figure II-12 D-F).



The current models describing the regulation of rhythmic gene expression by circadian clocks in other eukaryotes are also based on how core clock components regulate their own transcription via transcriptional feedback loops. For example, the mechanisms underlying transcriptional regulation by CLOCK:BMAL1 orthologs in *Neurospora* (WCC for White Collar Complex) and *Drosophila* (CLK:CYC heterodimer) are based largely on how they regulate the expression of the core clock genes *frequency* (in *Neurospora*), and *period* and *timeless* (in *Drosophila*) [61,111-113]. Given that the circadian clock mechanisms are highly conserved in eukaryotes, it is likely that both WCC and CLK:CYC also regulate their target gene expression by remodeling the chromatin and facilitating the binding of other transcription factors. Consistent with this hypothesis, WCC and CLK:CYC transcriptional outputs are also heterogeneous [114,115], and both recruit chromatin remodelers to promote nucleosome eviction at their binding sites [116-120].

The recent characterization of many transcription factor cistromes revealed that the number of transcription factor DNA binding sites often exceeds the number of anticipated target genes, suggesting that many of these DNA binding sites are non-functional [121,122]. Although many CLOCK:BMAL1 DNA binding sites could be considered as non-functional because they target arrhythmically or not expressed genes, the observation that CLOCK:BMAL1 rhythmically promotes nucleosome eviction at enhancers targeting both arrhythmically expressed (albeit with a decreased amplitude) and non-expressed genes instead indicates that CLOCK:BMAL1 rhythmic DNA binding is not "silent". More specifically, our data suggest that the majority of CLOCK:BMAL1 DNA binding events are functional, in that they rhythmically shape the chromatin landscape, and that transcription activation requires additional downstream events to be initiated (e.g., recruitment of other transcription factors). This hypothesis is further supported by

our finding that CLOCK:BMAL1 does not directly generate a transcriptionally active enhancer. Indeed, both H3K27ac ChIP-Seq signal and eRNA transcription are minimal at CLOCK:BMAL1 enhancers targeting non-expressed genes, and are delayed at CLOCK:BMAL1 enhancers targeting out-of-phase transcriptional cyclers (Figure II-9F, G). The observation that H3K27ac ChIP-Seq signal at CLOCK:BMAL1 enhancers correlates with CLOCK:BMAL1 transcription output rather than CLOCK:BMAL1 DNA binding phase/intensity seems inconsistent with the well-described interactions between core clock proteins and histone modifiers [7,70-75,81], and thus raises the question on whether or not CLOCK:BMAL1 DNA binding occurs with enzymatically activate histone modifiers. Interestingly, instances of enhancers bound by p300/CBP but lacking H3K27ac (and transcriptional activity) have been described at enhancers targeting developmental genes in human ES cells [123,124]. Those enhancers, which are termed poised enhancers, share most of the properties of active enhancers, including similar levels of nucleosome depletion, p300, and chromatin remodelers binding. However, these poised enhancers are unable to drive gene expression in ES cells until they acquire H3K27ac signal during differentiation [124]. Here we found that the binding of CBP and p300 at non-expressed target genes is above background levels, and that the differences in CBP and p300 DNA binding between non-expressed and expressed target genes are smaller than those observed for H3K27ac and Pol II ChIP-Seq signal (Figure II-9 F, Figure II-12 F, Figure II-14). It is thus tempting to speculate that the concept of poised enhancers extends to the circadian field, with CLOCK:BMAL1 rhythmically priming the chromatin landscape of "circadian poised enhancers". While those circadian poised enhancers would share properties of active enhancers (similar CLOCK:BMAL1 DNA binding, nucleosome eviction rhythm, *etc.*), they would be transcriptionally inactive and require the binding of other transcription-associated factors needed to trigger H3K27ac and rhythmic transcription.

Investigation of the transcription factors that are recruited at CLOCK:BMAL1 enhancers revealed a surprising difference between u-TFs and ts-TFs. In particular, ts-TFs are recruited at similar levels between expressed and non-expressed CLOCK:BMAL1 target genes, suggesting that they do not significantly contribute to the heterogeneity of CLOCK:BMAL1 transcriptional output. Because ts-TFs are known to establish tissue-specific enhancers and enable the binding of u-TFs in a tissue-specific manner [125-128], it is likely that ts-TFs contribute primarily to the binding of CLOCK:BMAL1 at tissue-specific enhancers and thus enable the generation of a tissue-specific circadian transcriptional program [56,57,66,129]. Contrary to ts-TFs, u-TFs appear to bind at CLOCK:BMAL1 enhancers targeting specific transcriptional output categories, suggesting that their nature (i.e., activator or repressor, constitutively active or inducible), as well as mode of cooperation with CLOCK:BMAL1, likely contributes to the heterogeneity of CLOCK:BMAL1 target gene transcription. For example, the transcriptional repressors REV-ERBα and REV-ERBβ are enriched at CLOCK:BMAL1 enhancers targeting out-of-phase transcriptional cyclers, agreeing with the recently proposed model of facilitated repression whereby CLOCK:BMAL1 remodels its enhancer chromatin to facilitate the recruitment of REV-ERBs and delay the transcriptional output of some of its target genes [91]. Since rhythmically expressed genes tend to exhibit higher u-TF ChIP-Seq signal than arrhythmic and non-expressed genes (Figure II-14), and given the low expression of in-phase transcriptional cyclers in *Bmal1-/-* mice, we propose that a major function of CLOCK:BMAL1 is to facilitate the recruitment of both positive and negative transcription factors to drive the rhythmic transcription of clock-controlled genes (i.e., not just to facilitate the binding of the circadian repressors REV-ERBα/β). Although the mechanisms underlying of this cooperation between CLOCK:BMAL1 and other transcription factors are still unknown, nucleosome-mediated cooperation between transcription factors is not unprecedented

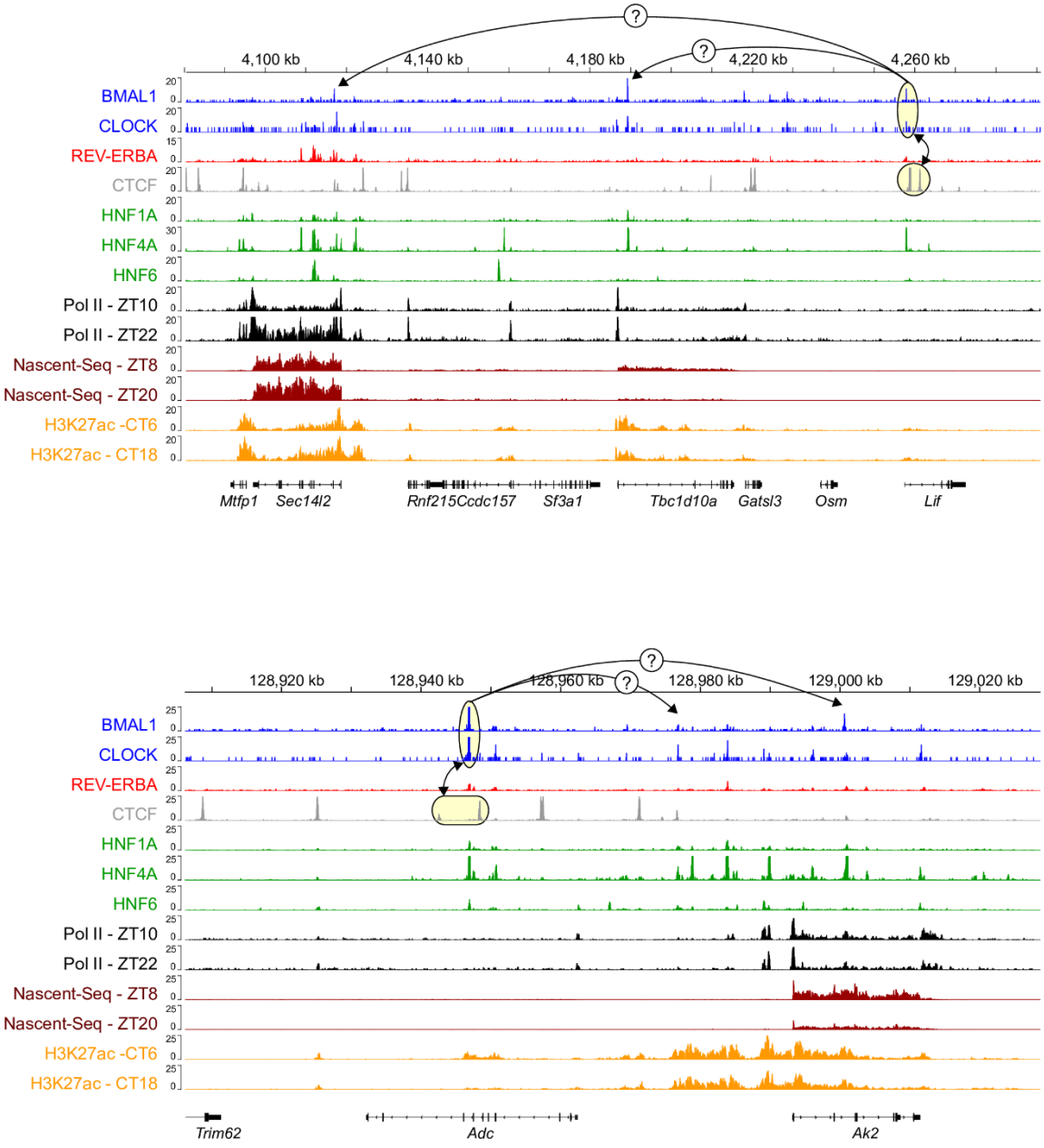
[130-135], and several papers have shown that two non-interacting TFs can synergistically bind to DNA through a mechanism whereby the first TF leads to partial unwrapping of nucleosomal DNA, thus making the site of the second TF more accessible and thereby increasing DNA binding.

This cooperation between CLOCK:BMAL1 and other TFs may explain why a large fraction of CLOCK:BMAL1 target genes are not expressed: u-TF recruitment is not sufficient to activate CLOCK:BMAL1 enhancers and promote transcription. In support of this idea, CLOCK:BMAL1 enhancers targeting non-expressed target genes are enriched for the NF-κB transcription factor motif, which is known to mediate transcriptional response to immune and inflammatory responses [94]. Because the genome-wide characterization of circadian clock mechanisms has mostly been carried out in healthy mice raised in standard laboratory conditions, NF-κB is likely inactive, sequestered in the cytosol and its target genes are not expressed. CLOCK:BMAL1 may thus prime NF-κB DNA binding upon inflammation or immune response, thereby triggering a rhythmic response to acute infection. Interestingly, such a mechanism may explain, at least in part, why the immune host response oscillates based on the time-of-day bacterial infection [136-139].

We also found that CLOCK:BMAL1 enhancers at non-expressed target genes are enriched for the transcription factor CTCF (CCCTC-binding factor; Figure II-12 A). CTCF is known to promote long-range interactions between two or more genomic sequences, and thus bring sequences that are far apart in the linear genome into close proximity [97]. This may suggest that some CLOCK:BMAL1 DNA binding sites situated in non-expressed gene loci actually target other clock-controlled genes located hundreds of kilobases apart through long-range interactions, as recently described for one CLOCK:BMAL1 DNA binding site in the mouse liver [140]. Although it is impossible to assess the prevalence of CLOCK:BMAL1 binding sites mediating long-range chromatin interactions without the appropriate experiments, we found a few examples suggesting that this is a likely possibility (Figure II-20).

**Figure II-20 The Enrichment of CTCF (CCCTC-binding Factor) at CLOCK:BMAL1 Enhancers Targeting Non-expressed Genes May Underscore the Role of Long-range Chromatin Interactions between CLOCK:BMAL1 Enhancers and its Target Genes.**

Visualization file of ChIP-Seq signal at two CLOCK:BMAL1 enhancers (yellow boxes) targeting non expressed genes (Top: *Lif* and Bottom: *Adc*), and exhibiting significant CTCF ChIP-Seq signal. Arrows with question marks indicate a potential CTCF-mediated long-range chromatin interaction that would enable CLOCK:BMAL1 to regulate the rhythmic transcription of genes located more than 50kb away from its enhancer.



Transcription regulation in higher eukaryotes relies on the activity of multiple enhancers [141,142]. It is thus likely that CLOCK:BMAL1 target gene expression results from a complex integration between CLOCK:BMAL1 enhancers and other enhancers. Our results indicate that enhancers targeting the same gene typically share the same transcriptional activity profiles (H3K27ac signal, eRNA levels, and Pol II ChIP-Seq signal; Figure II-11). Based on these observations, we cannot exclude that other enhancers targeting arrhythmically expressed CLOCK:BMAL1 target genes outcompete CLOCK:BMAL1 enhancers, to produce constitutive expression. Further experiments aimed at revealing hierarchical influences of enhancers on the regulation of gene expression at the genome-wide level will be required to directly test this hypothesis.

It was recently proposed that altering the environmental conditions can reprogram circadian transcriptional programs (e.g., high-fat diet and antibiotics treatment in the liver, LPS treatment in the lung [110,143-145]). Our model that CLOCK:BMAL1 regulates the expression of clock controlled genes by facilitating the binding of other TFs represents a mechanistic framework for explaining how environmental signals can mediate this transcriptional reprogramming. Indeed, activation of new signaling pathways by environmental changes is likely to modulate multiple transcriptional programs, thereby altering how CLOCK:BMAL1 cooperate with those programs to drive rhythmic gene expression. Importantly, this mechanism may also explain, at least in part, why the number and nature of rhythmically expressed genes vary between datasets and laboratories [146-148]. Indeed, differences in diet, light environment and housing may all lead to changes in u-TF transcriptional activity, which may in turn affect clock-controlled gene expression.

In conclusion, our data indicate that the mechanisms by which CLOCK:BMAL1 regulates the transcription of core clock genes do not apply to clock-controlled genes, and suggest that the

primary function of CLOCK:BMAL1 is to regulate the chromatin landscape at its enhancers to facilitate the binding of other transcription factors. Our results therefore highlight the emerging role of other transcription factors in regulating the ~15% of genes that are rhythmically expressed in a given mammalian tissue, and suggests that clock-controlled gene expression relies more on the interplay between the circadian clock and other signaling pathways. Given that the majority of CLOCK:BMAL1 target genes are either arrhythmically or not expressed under standard conditions, our data also suggest that these non-oscillating genes may become rhythmically expressed under other environmental and/or pathological conditions, and thus expand the total number of genes under circadian control to more than 50% in mammals [66]. Finally, because the clockwork mechanisms are highly conserved between eukaryotes (e.g., heterogeneous transcriptional output, poor reproducibility between datasets characterizing circadian gene expression, regulation of chromatin landscape by core clock components), it is likely that the mechanisms we uncovered largely apply to all eukaryotic circadian clocks.

## Materials and Methods

### *Sequencing datasets and alignment to the mouse genome*

Unless notified below, publicly available datasets used in this paper were downloaded from the NCBI or EBI websites in either sra or fastq formats (see Table II-1 for accession numbers). Files in sra format were converted to fastq files using the sratoolkit (version 2.3.5-2). Fastq files were mapped to the mouse genome (version mm10) using bowtie2 [149] or tophat2 [150]. For all datasets, we only considered reads that mapped uniquely to the mouse genome (i.e., one unique genomic location). Datasets were further filtered to remove duplicated reads using samtools

(rmdup function) or a custom-made script. Additional information is provided for each dataset as Supplementary Materials and Methods from Trott and Menet, 2018.

**Table II-1 Public Datasets Used in this Study.**

| Reference | SRA/GEO dataset | Data | Remapped |
|---|---|---|---|
| Koike et al., 2012 [7] | GSE39977 | ChIP-Seq: BMAL1, CLOCK, PER1, PER2, CRY1, CRY2, H3K27ac | Yes[1] |
| Menet et al., 2012 [12] | GSE36916 | Nascent-Seq<br>RNA-Seq | Yes[2] |
| Rey et al., 2011 [83] | GSE26602 | ChIP-Seq: BMAL1 | Yes |
| Cho et al., 2012 [89] | GSE34020 | ChIP-Seq: Reverba, Reverbb | Yes |
| Menet et al., 2014 [13] | GSE47145 | MNase-Seq | Yes[3] |
| Le Martelot et al., 2012 [77] | GSE35790 | ChIP-Seq: RNA polymerase II (Pol II) | Yes |
| Vollmers et al., 2012 [19] | SRA025656 | ChIP-Seq: H3K9ac, H3K27ac | Yes |
| Fang et al., 2014 [90] | GSE59486 | ChIP-Seq: E4BP4, Roralpha<br>GRO-Seq | Yes[4] |
| Ling et al., 2010 [151] | GSE21777 | DNase-Seq | No[5] |
| Faure et al., 2012 [99] | https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-941/ | ChIP-Seq: GABPA, HNF1, HNF4A, E2F4, CEBPA, HNF6, p300, CBP, CTCF | Yes |
| Lim et al., 2015 [100] | GSE59752 | ChIP-Seq: CEBPB, GR | Yes |
| Boergesen et al., 2012 [101] | GSE35262 | ChIP-Seq: LXR, PPARa, RXR | Yes |
| Gordon et al., 2014 [103] | GSE52351 | ChIP-Seq: ERalpha | Yes |
| MacIsaac et al., 2010 [102] | GSE17067 | ChIP-Seq: FOXA1, FOXA2 | Yes |
| Zhang et al., 2012 [104] | GSE31578 | ChIP-Seq: BCL6, STAT5 | Yes |
| Atger et al., 2015 [106] | GSE73554 | RNA-Seq: Wild-type and Bmal1-/- mouse liver | No[6] |
| Eckel-Mahan et al., 2015 [110] | GSE52333 | Mouse liver RNA expression (microarray data) from wild-type mice fed normal chow or high fat diet | No[7] |
| Vollmers et al., 2009 [107] | GSE13093, GSE13064 | Mouse liver RNA expression (microarray data) from wild-type mice fed *ad libitum* or subjected to fasting for 24 hours. | Yes |

[1] Datasets from Koike et al., 2012 were downloaded as sra files and remapped to the mouse genome (mm10 version) to analyze ChIP-Seq signal at various enhancers. However, we used the CLOCK and BMAL1 peak coordinates provided in the supplemental table of the original paper to generate our list of high confidence CLOCK:BMAL1 DNA binding sites (see details below), and also determine the phase and intensity of CLOCK and BMAL1 DNA binding.

[2] Nascent-Seq and RNA-Seq dataset were remapped to the mouse genome version mm10 (e.g., to generate the Figure II-19). However, Nascent-Seq and RNA-Seq signals were retrieved from the original paper (Figure II-8—source data 1 and Figure II-9—source data 2) and used to generate the data presented in Figure II-2.

[3] While fastq files were remapped to the mouse genome version mm10, further analysis was performed as described in the original paper [13]. For example, the 50nt reads were extended to 147nt to match the length of a mononucleosome.

[4] Analysis of the GRO-Seq dataset has been performed similarly to the original paper and reads were also extended to 150bp as in Fang et al., 2014 [90].

[5] No reanalysis of the DNase-Seq was performed. The file GSE21777_M-CM_peaks.txt, which contain the list of DNaseI hypersensitive sites (DHS) in the mouse liver, was downloaded from the ncbi website (GSE21777) and the DHS peak coordinates (mouse genome version mm9) were converted to mouse genome version mm10 using a liftOver tool downloaded from the UCSC genome browser website (conversion resulted in a list of 104,556 DHS peaks).

[6] The processed files with intron and exon RNA-Seq signal (fpkm) from wild-type and *Bmal1-/-* mouse liver were downloaded from the ncbi website and directly used to generate the Figure II-16 B and Figure II-16 C (files GSE73554_KO_RF_Intron_Exon_RFP.txt and
GSE73554_WT_RF_Intron_Exon_RFP.txt).

[7] Processed microarray data were downloaded from the ncbi website and used to generate the Figure II-16 D-G. The original statistical analysis performed by the authors from the journal Cell website was used in our study to assess rhythmic gene expression.

### *Identification of CLOCK:BMAL1 DNA binding sites*

Genomic locations of CLOCK and BMAL1 DNA binding sites in the mouse liver provided in the original paper [7] were used to generate our list of high confidence CLOCK:BMAL1 DNA binding sites. Genomic locations were converted to the mm10 version of the mouse genome using UCSC genome browser liftOver tools, and processed as indicated in Figure II-1 to generate our list of high confidence CLOCK:BMAL1 DNA binding sites. Overlap between CLOCK and BMAL1 ChIP-Seq peaks was determined using bedtools (intersectBed) and coordinates from BMAL1 ChIP-Seq datatsets were further kept to generate a list of 3217 CLOCK:BMAL1 peaks. We also used the original data provided by the authors to assign CLOCK:BMAL1 peaks to their putative target genes (original analysis performed using HOMER tools). In particular, we defined a CLOCK:BMAL1 target gene as a gene with at least one CLOCK:BMAL1 peak located between -10kb of the transcription start site and +1kb from the transcription termination site. Using this

criteria, 2458 CLOCK:BMAL1 peaks were assigned to a target gene, and the remaining 759 peaks were assigned as an intergenic CLOCK:BMAL1 DNA binding site.

The 2458 CLOCK:BMAL1 peaks assigned to a target gene were then parsed based on the transcription profile of their target genes using the Nascent-Seq datasets from Menet et al., 2012 [12]. We directly used the original Nascent-Seq expression values and the assessment of their rhythmic expression from the original paper without performing new analysis. Details on how genes were determined to be rhythmically transcribed are provided in Supplementary Materials and Methods. Using these data, CLOCK:BMAL1 peaks were parsed into 4 different categories of transcriptional output (see also Figure II-1):

(i) Rhythmic transcription in phase with CLOCK:BMAL1 binding (peak of transcription between ZT02 and ZT12), 205 CLOCK:BMAL1 peaks,

(ii) Rhythmic transcription out-of-phase with CLOCK:BMAL1 binding (peak of transcription between ZT12 and ZT02), 124 CLOCK:BMAL1 peaks,

(iii) Arrhythmic transcription, 654 CLOCK:BMAL1 peaks (criteria used to defined arrhythmic transcription are detailed in Supplementary Materials and Methods). Note that to decrease the number of false positive in the list of arrhythmically expressed genes, we removed genes that exhibit arrhythmic nascent RNA expression, but exhibit rhythmic mRNA expression (using the RNA-Seq dataset from Menet et al., 2012 [12],

(iv) Not transcribed (average reads/bp between the 12 sample < 1), 291 CLOCK:BMAL1 peaks.

The remaining peaks, which were not analyzed in this study, were categorized as:

(v) Post-transcriptional cyclers, 262 CLOCK:BMAL1 peaks. These peaks target genes with an arrhythmic Nascent-Seq signal, but rhythmic RNA-Seq signal (based on the reads/bp values published in Menet et al., 2012 datasets [12]).

(vi) Low expression levels, 588 CLOCK:BMAL1 peaks. These CLOCK:BMAL1 peaks target genes with Nascent-Seq and/or RNA-Seq signals that are below threshold for the analysis of rhythmic expression, but above the threshold of 1 read/bp set to define the "not-expressed" CLOCK:BMAL1 target genes. These thresholds were defined in Menet et al., 2012 [12].

(vii) No signal value, 334 CLOCK:BMAL1 peaks. These peaks were assigned to a target gene (defined as described above with the HOMER software and localization criteria), but no information was found in the Nascent-Seq or RNA-Seq. Several of these peaks target genes encoding for a non-coding RNA, as well as genes with alternative gene symbol.

The list of the 3217 CLOCK:BMAL1 peaks parsed into the different transcriptional output categories can be found in S1 Table from Trott and Menet, 2018. The phase of rhythmic CLOCK:BMAL1 DNA binding, ChIP-Seq signal, and genomic location of CLOCK:BMAL1 DNA binding sites were retrieved from the Koike et al., 2012 original paper supplementary [7] and processed to generate the analysis presented in Figure II-2.

***Analysis of ChIP-Seq, MNase-Seq and GRO-Seq signal at CLOCK:BMAL1 peaks***

ChIP-Seq, MNase-Seq and GRO-Seq signal was retrieved from bam files containing uniquely mapped reads (and duplicated reads removed) at CLOCK:BMAL1 enhancers using custom-made scripts [13]. Specifically, signal was retrieved at:

- CLOCK:BMAL1 peak center ± 250 bp for transcription factors, CBP, 300 and Pol II,
- CLOCK:BMAL1 peak center ± 1kb for histone modifications,
- CLOCK:BMAL1 peak center ± 500 bp for eRNA,
- CLOCK:BMAL1 peak center ± 75bp for nucleosome signal,

and normalized to the sequencing depth. Differences in the window size were calculated based on the width of the ChIP-Seq signal at CLOCK:BMAL1 DNA binding sites (e.g., H3K27ac ChIP-Seq signal is significantly wider than any transcription factor ChIP-Seq signal). Because we aimed at assessing the role of CLOCK:BMAL1 in removing a nucleosome at its DNA binding site, we chose a narrower window size of 150bp (see Figure II-9 A-D). All analyses were performed at individual CLOCK:BMAL1 ChIP-Seq peaks, and this even for peaks targeting the same gene. Data presented in Figure II-5 examined the role of multiple peaks targeting the same gene on BMAL1 and CLOCK ChIP-Seq signals. For all datasets, ChIP-Seq signal is displayed as the number of reads/bp per 100,000,000 reads.

***Analysis of ChIP-Seq and GRO-Seq signal at other enhancers targeting CLOCK:BMAL1 target genes***

Enhancers lying into CLOCK:BMAL1 target gene loci (-10kb from the transcription start site to +1kb from the transcription termination site) were identified using a public mouse liver DNAse-Seq dataset (see above) [151] and bedtools (intersectBed function). Enhancers were then parsed based on the presence or not of a CLOCK:BMAL1 ChIP-Seq peak (3155 out of the 3217 CLOCK:BMAL1 ChIP-Seq peaks are located into a DNaseI hypersensitive site). Because a majority of the 104,556 DHS peaks only displayed low levels of ChIP-Seq (transcription factors, Pol II, H3K27ac) and GRO-Seq signals as shown in the ENCODE project, [152,153], we filtered the number of DHS lying into a CLOCK:BMAL1 target gene by only considering those being into the top 40,000 DHS list (based on DNase-Seq signal), obtaining the following number of DHS peaks:

(i) In-phase transcription cyclers: 1548 DHS peaks,

(ii) Out-of-phase transcription cyclers: 1189 DHS peaks,

(iii) Arrhythmically expressed target genes, 5830 DHS peaks,

(iv) Not expressed target genes, 991 DHS peaks.

H3K27ac and Pol II ChIP-Seq signals, as well as GRO-Seq (eRNA) signal, were retrieved at those DHS sites (as well as those overlapping with a CLOCK:BMAL1 peak) using the DHS peak coordinate and normalized to 100,000,000 reads. Signal was then normalized to the coordinate length (in bp) to obtain the signal displayed as reads/bp per 100,000,000 reads. The coordinates used were, for the same reason as above for CLOCK:BMAL1 DNA binding sites:

- H3K27ac ChIP-Seq: DHS genomic coordinate center ± 1 kb,
- Pol II ChIP-Seq: DHS "real" genomic coordinates,
- GRO-Seq (eRNA signal): DHS genomic coordinate center ± 500 bp.

Because our analysis revealed the existence of small but significant overall variations of H3K27ac and Pol II ChIP-Seq signal between time points (see Figure II-11), we further normalized the datasets by performing either a mean normalization (H3K27ac) or a ranking analysis (Pol II). For H3K27ac ChIP-Seq datasets [7,19], averaged H3K27ac signal was calculated at the top 40,000 DHS peaks (the top 40,000 DHS peaks concentrate the majority of TFs DNA binding sites; peak center ± 1 kb; total of 104,556 total DHS peaks; dataset from Ling et al., 2010 [151]) for each time point. This averaged signal was then used to normalize the raw H3K27ac ChIP-Seq signal, by calculating for each time point the ratio between H3K27ac signal for each peak and this averaged signal (see Figure II-11). Pol II ChIP-Seq dataset [77] were normalized by performing a ranking normalization (method similar to a quantile normalization). To this end, Pol II ChIP-Seq signal was calculated at all 104,556 DHS peaks (peaks mapped in Ling et al., 2010 paper [151]), and sorted based on the ChIP-Seq values. The raw values for each DHS peak were then normalized using the sorted averaged ChIP-Seq signal at each of the 104,556 ranks for all time points.

### *Motif analysis*

Motif analysis was performed at CLOCK:BMAL1 enhancers (original peak coordinates) for each of the transcriptional categories using the findsMotifGenome.pl script from the HOMER suite. Parameters were as the following: -size given –len8. The resulting table was sorted by the q-value and a q-value less than 0.05 was considered significant. Percent enrichment (percent of target sequences with motifs / percent of background sequences with motif) was then calculated for motifs found to be significant in at least one of the CLOCK:BMAL1 transcriptional output category.

### *Determination of the TF DNA binding variability index*

To determine the variance of each TF DNA binding (CLOCK, BMAL1, ts-TFs and u-TFs) between the four CLOCK:BMAL1 transcriptional output categories, we computed a TF DNA binding "variability index" based on the analysis performed in Figure II-14. The variability index was calculated by summing up the standard deviation of the ChIP-Seq signal between the 4 transcriptional output groups, which was calculated for each decile (0.1 to 0.9) and normalized to the averaged signal for each decile (the standard deviation is higher for upper deciles because ChIP-Seq signals are higher). This index reflects differential DNA binding strength between groups, as similar binding between the 4 groups results in small standard deviation values for each decile, and thus a small variability index. Conversely, differences in DNA binding signal between groups result in larger standard deviation values and thus a larger variability index.

***Generation of a list of control genes not targeted by CLOCK:BMAL1***

To determine if the results described in this paper are specific to CLOCK:BMAL1, we also performed an analysis on genes not targeted by CLOCK:BMAL1, but exhibiting similar profiles of expression to the 4 CLOCK:BMAL1 transcriptional output categories (Rinφ, Ro/φ, AR and NE). To this end, 125 genes were randomly selected for each of the 4 groups, using criteria similar to those used to define CLOCK:BMAL1 transcriptional output (see above). Levels of expression for each group were not significantly different to those of CLOCK:BMAL1 target genes (Kruskal-Wallis test). Nucleosome signal, H3K27ac ChIP-Seq signal, Pol II ChIP-Seq signal, eRNA expression, tissue specific and ubiquitous transcription factor ChIP-Seq signal were all calculated as described above for CLOCK:BMAL1 target genes. Statistical analysis was also performed similarly to CLOCK:BMAL1 transcriptional output.

***Statistical analysis***

Statistical analysis was done using JMP®, Version *Pro 12.0.1*. SAS Institute Inc., Cary, NC, 1989-2007. Differences in sequencing signal, represented in the boxplot graphs, were analyzed for statistical enrichment using the nonparametric Kruskal-Wallis test. Rhythmic analysis of nucleosome signal and ChIP-Seq signal was performed using a Fourier analysis (Figure II-9 A-D) (see Supplementary Materials and Methods for details). Differences in the amplitude of nucleosome signal rhythm (Figure II-9 E) were analyzed using a 2-way ANOVA. Differences in CLOCK:BMAL1 ChIP-Seq peaks genomic location were analyzed using a chi-square test (Figure II-2 G), and differences in the number of CLOCK:BMAL1 peaks per target genes (Figure II-6) were analyzed by a Fisher's exact test. Differences were considered significant when $p < 0.05$.

# CHAPTER III

# MOLECULAR CHARACTERIZATION OF THE EFFECTS OF SHIFT WORK AND FOOD CONSUMPTION ON CARDIOVASCULAR DISEASE IN THE RAT

## Overview

Shift work misaligns the circadian clock and leads to an increased risk of cancer, cardiovascular diseases, and metabolic disorders. Furthermore, food consumed during the rest phase is a major contributor to this misalignment, as food access restricted to the endogenous active phase, at night, prevents against the adverse effects of shift work on obesity and diabetes in rats. In this study, we aimed to investigate the molecular mechanisms by which shift work and food consumption contribute to an increased risk of cardiovascular disorders. To this end, we used a model of shift work in rats, whereby animals are exposed to a 12:12 light:dark cycle and are forced to be active for eight-hours during their natural rest phase during the day, Monday to Friday for five consecutive weeks. Given the preventive effects of temporal restriction of food intake, a group of shift worker rats with food access restricted to the night was included in addition to controls and shift workers. To gain insight into the molecular underpinnings of shift-work associated cardiovascular pathologies, we performed Picrosirius Red staining and RNA-Seq analysis in rat hearts. Our results show that shift work rats, regardless of the time of food consumption, have an increase in collagen deposition in the heart. In addition, the expression of many genes encoding key fibrotic pathways were found to be up-regulated in shift worker rats that had their food restricted to the active phase. Altogether, our results suggest that five weeks of shift work in rats is capable of inducing cardiovascular disease through an up-regulation of collagen deposition in the heart.

## Introduction

Shift work is a work schedule that has been defined many times as being any work done outside the general working hours during the day [35]. This type of work schedule is becoming increasingly popular among employers to meet the high demand of productivity in modern society. Many of these industries including manufacturing, agriculture, transportation, healthcare and emergency services require some form of shift work. This has been reported to be approximately 15-30% of workers in America and Europe, and this work force is rapidly increasing [34]. Shift work has been reported to disrupt many aspects of one's lifestyle including job strain and social stress. Furthermore, shift work has been correlated with an increased risk of pathological diseases such as cancer, cardiovascular disease, and metabolic disorders like obesity and diabetes [34-36,39].

Numerous studies have been conducted that have correlated shift work with cardiovascular disease. A meta-analysis was done in 1999 on 17 studies observing the risk of shift work and cardiovascular disease, concluded that shift workers have a 40% increased risk of cardiovascular disease compared to day workers [35]. However, significant heterogeneity has been found among these results and thus it is still unknown how the quantity and duration of shift work affect cardiovascular disease [34].

The exact mechanisms on how shift work causes cardiovascular disease is unknown, however many mechanisms have been hypothesized [35]. Additionally, shift work is known to alter the circadian clock and food consumption, all of which have been correlated to cause metabolic disruption and diabetes [36]. Furthermore, both of these have been known to contribute to cardiovascular disease by increasing the load that the heart must work under which can lead to an accumulation of myocardial fibrosis [154,155]. Almost all cardiovascular disease is associated with activation and expansion of cardiac fibroblasts and an increase in collagen deposition in the

extracellular matrix of the heart [155,156]. Persistent presence of stress conditions will result in a continual deposition of collagen and extracellular matrix proteins resulting in fibrosis. The TGF-β signaling pathway has been recognized as one of most potent profibrotic cytokines known to stimulate fibrosis, via myofibroblast differentiation and extracellular matrix remodeling [51,52,157,158].

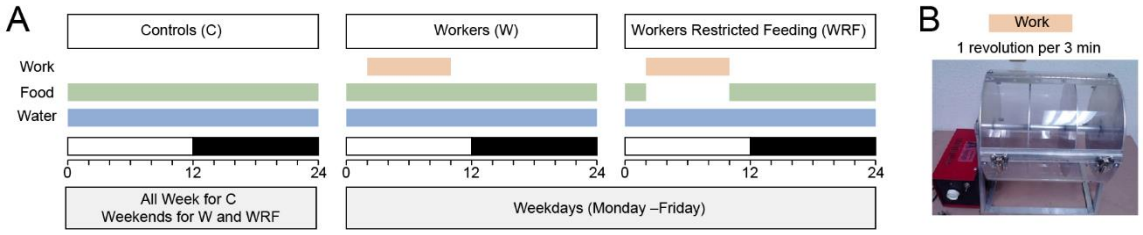
To uncover the mechanisms on how shift work causes cardiovascular disease, we used a shift work model developed in rats to mimic the conditions of a human night shift worker. Our data revealed that five weeks of shift work induces an increase in collagen deposition in the rat heart. This induction of collagen deposition is irrelevant to time of day eating and thus suggests that the act of shift work alone is sufficient to increase collagen deposition in the heart. Surprisingly, we observed very little differential gene expression between the Worker and Control rats. However, we observed a significant change in differential gene expression for rats forced to do shift work during the day but had their food consumption restricted to the night. Furthermore, we observed that many of these differentially expressed genes were found in pathways associated with fibrosis and hypertrophy. Altogether, our results suggest that shift work is capable of inducing cardiovascular disease through up-regulation of collagen deposition in the heart.

## Results

### *Shift work induces collagen deposition in the heart*

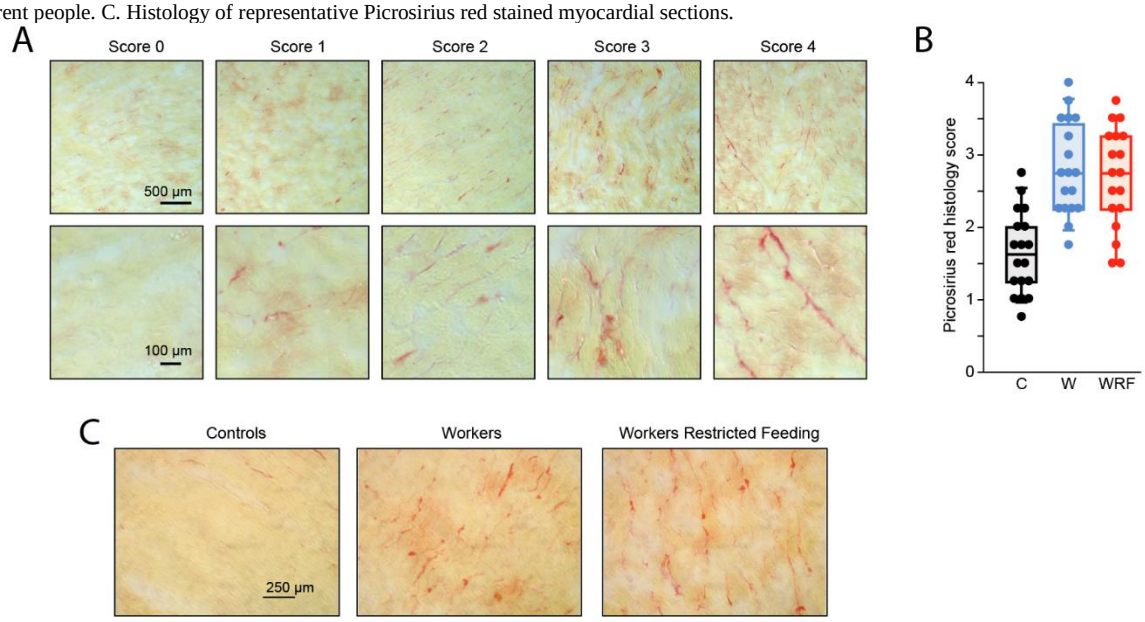
To determine the molecular mechanisms by which shift work and food consumption contribute to cardiovascular disease we used a rat shift work model (Figure III-1 A, B). A total of 54 rats were maintained on a 12:12 light dark cycle (7am to 7pm) and then randomly assigned to Control (C), Worker (W), or Worker Restricted Feeding (WRF) groups. Shift work consisted of rats being placed in a slowly rotating drum (1rev/3min) as published previously by Salgado-Delgado et al., 2010. Control rats had access to *ad libitum* food and water throughout the experiment. Worker rats had access to *ad libitum* food and water but were forced to do shift work. Worker Restricted Feeding rats were also forced to do shift work, had *ad libitum* access to water, but only had access to food when not doing shift work. W and WRF rats were forced to do shiftwork Monday-Friday and were given the weekend off where they were fed and watered like C. Rats maintained this paradigm for five weeks before rats were euthanized and hearts were collected for analysis.

**Figure III-1 Shift Work Paradigm.**
A, B. Rats individually housed and maintained on a 12 hour light: 12 hour dark schedule were randomly assigned to one of the three groups: Control (C), Worker (W), or Worker Restricted Feeding (WRF). All rats had *ad libitum* access to water. C rats had *ad libitum* access to food and were not moved from their cages. W and WRF rats were placed in a rotating drum (1 revolution/3min) for eight hours Monday- Friday to simulate shift work, and were left in their cages over the weekend. W rats had *ad libitum* access to food at all times, including in the wheel, while the WRF rats only had access to food at night when not simulating shift work.




Increases in collagen deposition have been well established as a marker of fibrosis and cardiovascular disease [159-161]. To investigate if shift work increased levels of fibrotic tissue in the heart, we conducted Picrosirius Red staining on cryostat sections of our C, W, WRF rat hearts. Following the staining of these 54 samples, we conducted a blind study to determine if there were differences in collagen deposition between the different paradigms (Figure III-2 A). A score of zero was given for no collagen deposition and maximum collagen deposition was given a score of four. The blind study resulted in significantly higher score for W and WRF compared to C, indicating that five weeks of shift work results in more collagen deposition in the rat heart (Figure III-2 B, C).

**Figure III-2 Shift Work in Rats Causes Fibrosis in the Heart.**
A. Representative examples of Picrosirius red stained rat 10 µm myocardial sections for scoring of the blind study. A score of 0-4 was given for collagen deposition. Yellow =viable myocardium, and red = collagen. B. Boxplot of blind scoring of 18 samples per group conducted by four different people. C. Histology of representative Picrosirius red stained myocardial sections.



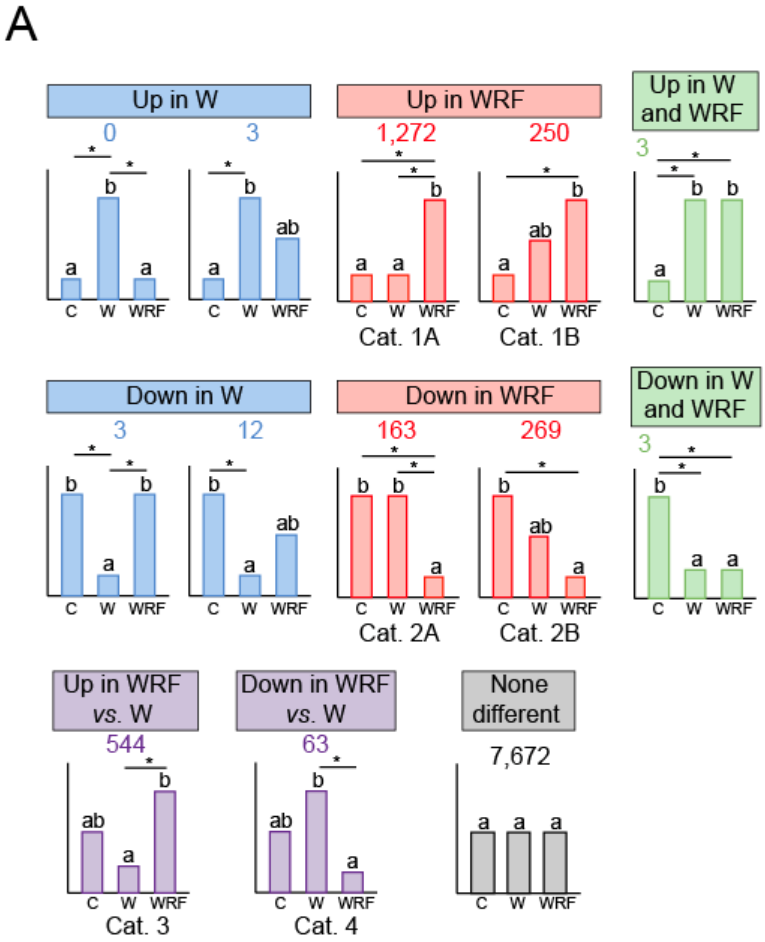

These results suggest that five weeks of shift work is enough time to induce larger amounts of collagen deposition in the heart. Additionally, they suggest that eating during the natural active phase at night does not inhibit shiftwork induced collagen deposition in the rat heart.

### *Gene expression is significantly altered in WRF*

To further assess if we were observing cardiovascular disease in W and WRF rats, we sequenced the transcriptome of C, W, WRF rat hearts. This consisted of 18 biological replicates per group with a median seq depth of 25 million reads per sample. Gene expression was mapped to RN6 using Tophat2 and further analyzed using Cufflinks [162,163]. As previously published, a single gene copy is equivalent to 1-3 reads per kilobase of exon model per million mapped reads

(RPKM) [164]. To remove any false positives, we chose to keep genes that where expressed in 18 of the 54 biological samples with a log2 (fpkm+1) value over 1. This left us with 10,257 total genes expressed in C, W, WRF hearts. To determine differential gene expression, we did a one-way ANOVA with post-hoc Tukey's HSD (p-value <0.05) analysis on C, W, WRF. Categories of significantly differentially expressed genes are illustrated by bar graphs with the total number of genes written on top of the graph in Figure III-3 A. Majority of the genes found to be differentially expressed are up-regulated or down-regulated in WRF compared to both W and C groups, which we have labeled as Categories 1A, 1B, 2A, and 2B (Figure III-3 A). Strikingly, only six genes were found to be significantly up or down regulated in both W and WRF compared to C. Although, approximately 600 genes were found to be significantly differentially expressed between W and WRF (Categories 3 & 4) these were not significantly different than C (Figure III-3 A).

**Figure III-3 Shift Work and Restricted Feeding Significantly Alters Gene Expression in the Rat Heart.**
A. Differential gene expression was analyzed with a one-way ANOVA on paradigms with post hoc analysis using Tukey's HSD. Genes were considered as differentially expressed between paradigms with a p-value < 0.05. Categories of statistically significant differential gene expression are illustrated by bar graphs with the total number of genes written above for each category. Groups with different letters are significantly different with a Tukey's HSD p-value <0.05. The majority of the genes were either up-regulated (Categories 1A, 1B, 3) or down-regulated (Categories 2A, 2B, 4) in WRF compared to C and W.

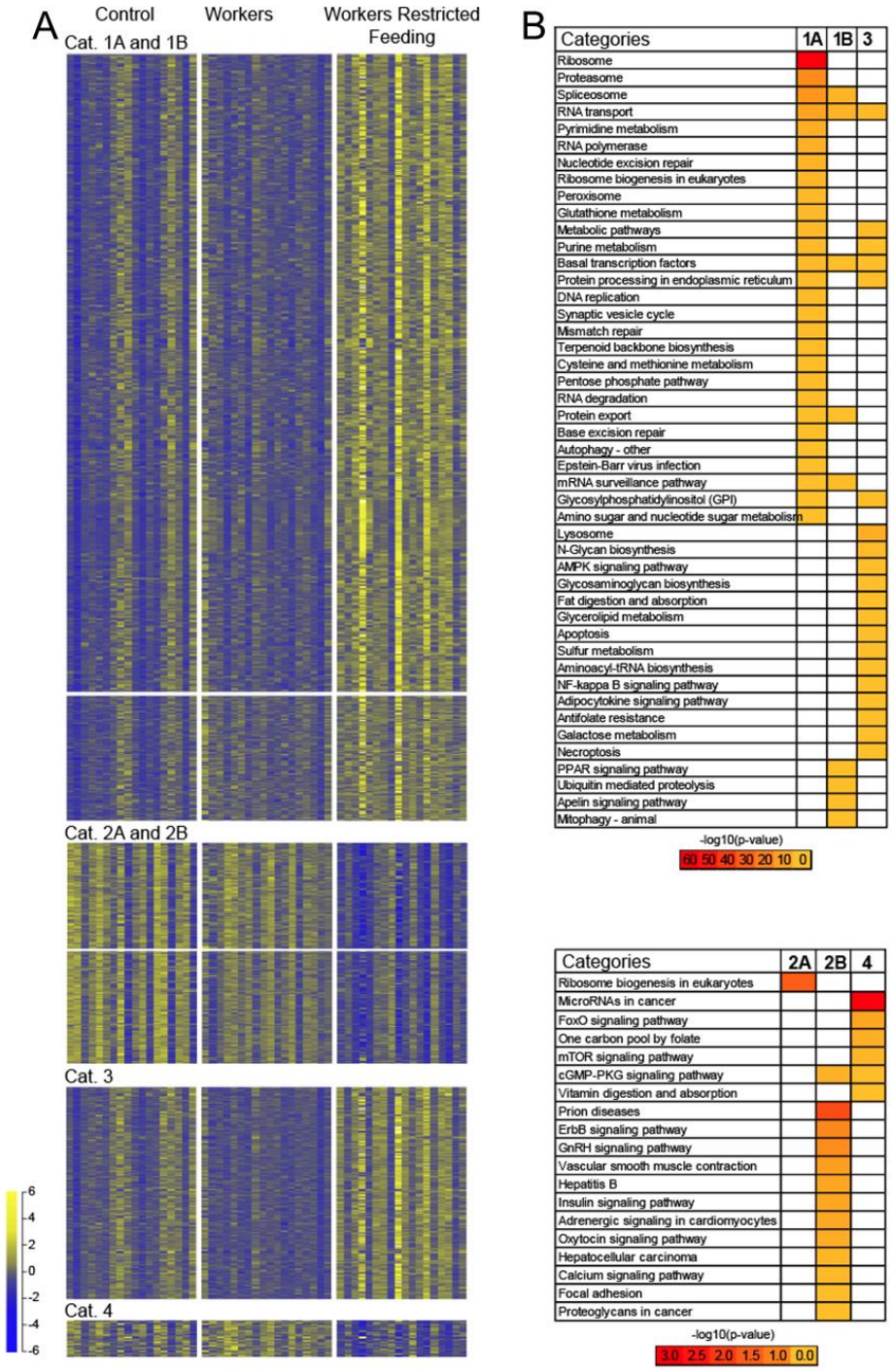



Heatmap visualization of the differences in standardized gene expression in these categories illustrates the individual biological variation in gene expression; however, the over-all trend of up or down regulation of gene expression for each group is easily observed between the categories (Figure III-4 B). KEGG Pathway analysis [165] indicates a significant up-regulation of the

ribosome and proteasome pathways in category 1A, as well as an enrichment of NF-kappa B signaling pathway, AMPK signaling pathway, and PPAR signaling pathway in categories 1B and 3 (Figure III-4 C). Furthermore, categories 2A, 2B, and 4 were enriched for the FoxO signaling pathway, mTOR signaling pathway, ErbB signaling pathway all which have been suggested to be involved in remodeling of the extracellular matrix of the heart leading to fibrosis and cardiovascular disease [51,166-168].

**Figure III-4 WRF Significantly Upregulates Gene Expression for Ribosomes, Proteasomes, and Metabolic Pathways.**

A. Heatmap of standardized gene expression for categories 1A, 1B, 2A, 2B, 3, and 4 genes, n=18 animals per paradigm sorted by time of euthanasia. Yellow indicates up-regulation, while blue indicates down-regulation. B. KEGG pathway analysis of the differentially expressed Categories 1A ,1B, 2A, 2B, 3, 4. KEGG pathways with a p-value < 0.05 are colored from yellow to red based on p-value, with red being the most significant.



***Collagen and fibroblast marker gene expression is significantly up regulated in WRF***

To understand the large differences in collagen deposition in W and WRF, we looked at collagen gene expression. Interestingly, the only genes that were significantly differentially expressed were collagen type 1 alpha 1 chain (*Col1a1)* and collagen type 1 alpha 2 chain (*Col1a2*), which were expressed at a greater level in WRF (Figure III-5 A). Additionally, an increased expression of these genes has been linked to remodeling the cardiac matrix, leading to fibrosis and cardiac disease [169].

To investigate if the increase in collagen deposition observed in W and WRF is because of an increase in fibroblasts, we looked at fibroblast marker gene expression. Again, we observed that WRF was the only group with a significant differential gene expression of fibroblast markers, fibroblast-specific protein-1 (*s100a4*) and Thy-1 cell surface antigen (*Thy1*), were significantly up-regulated in WRF compared to C and W (Figure III-6 A). Surprisingly, W rats did not show up-regulation for any markers of fibroblasts, which are predominantly the manufactures of collagen in the heart [50].

**Figure III-5 Some Collagen Genes are Up-Regulated in WRF Rats.**
A. Boxplots of gene expression for collagen related genes in the heart. Groups with different letters are significantly different with a Tukey's HSD p-value <0.05.

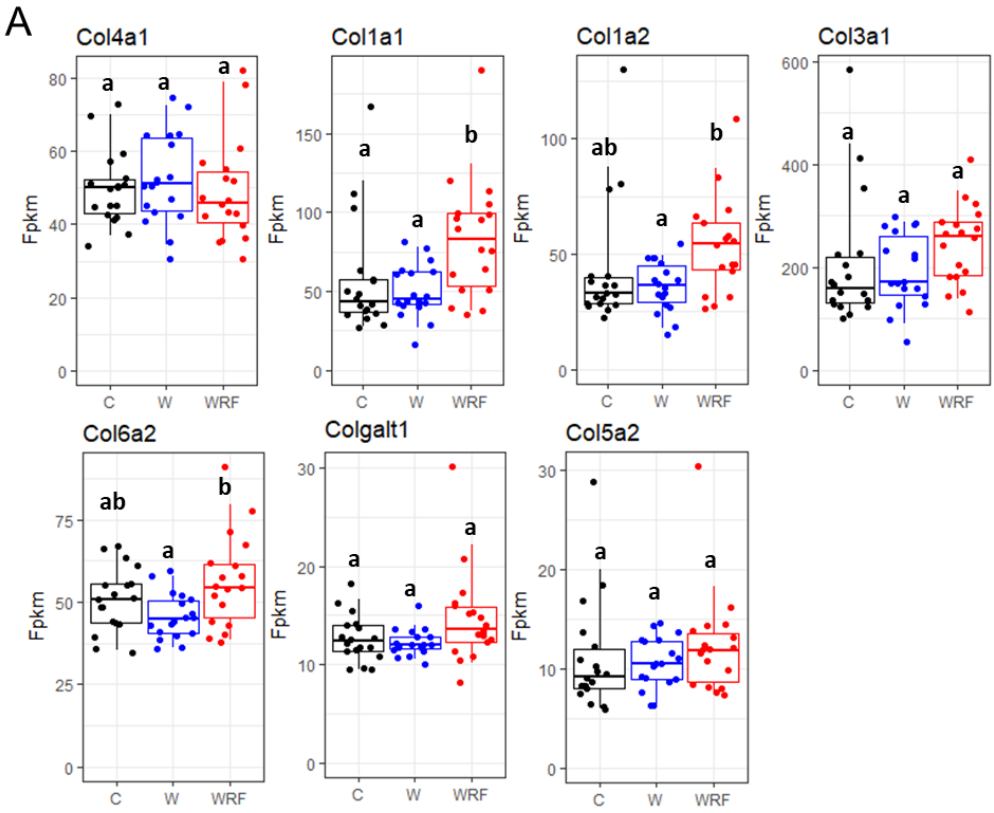
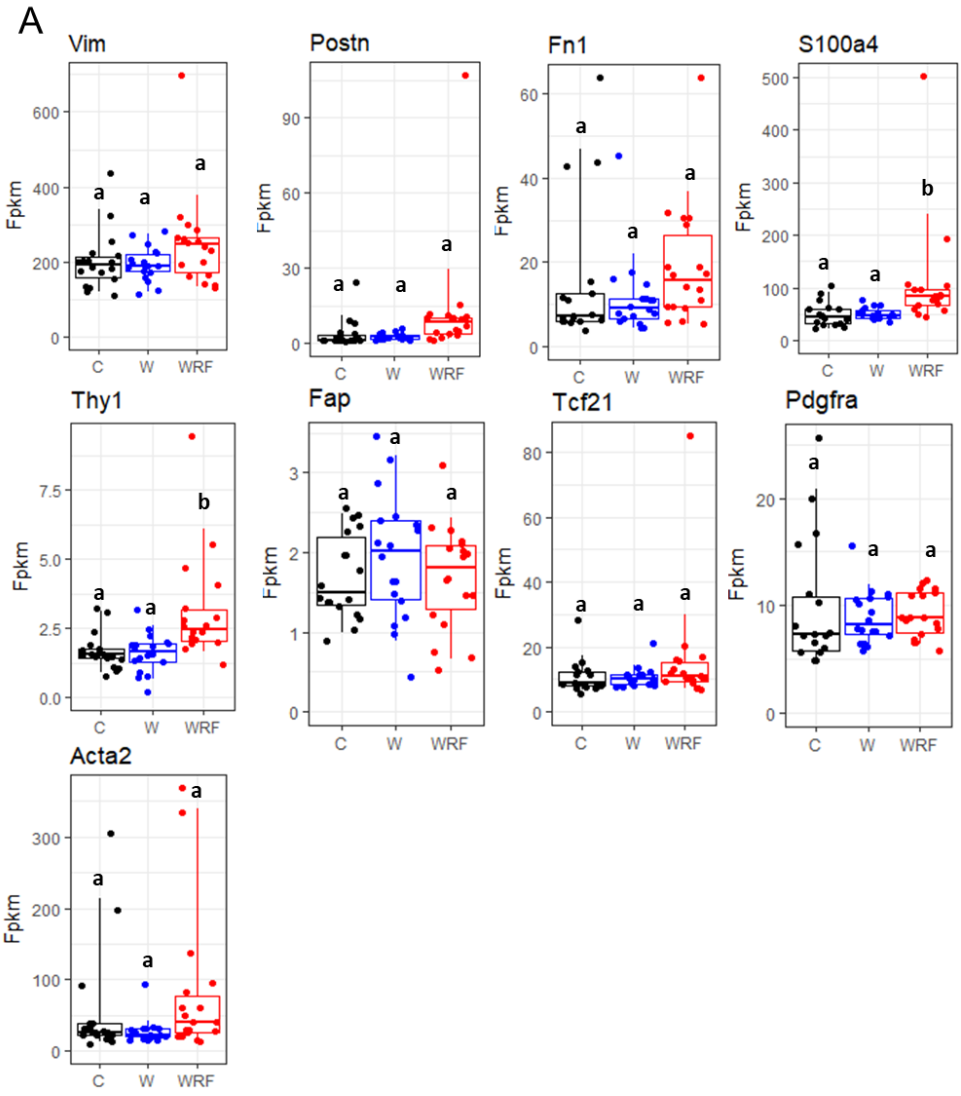


**Figure III-6 Only Two Fibroblast Markers are Up-Regulated in WRF Rats.**
A. Boxplots of gene expression for fibroblast related markers. Groups with different letters are significantly different with a Tukey's HSD p-value <0.05.




As shown previously, W and WRF rats have an increased deposition of collagen between myocardial cells, yet we do not observe a strong increase in gene expression for both W and WRF.

This lack of a significant increase in gene expression for our W group brings to question what is up-regulating collagen deposition in our W rat hearts.

***Open chromatin regions reveal fibrotic transcription factor motifs***

ENCODE consortium revealed that ~95% of transcription factors bind in open chromatin regions [170]. To understand the mechanisms behind the differential gene expression observed between the different paradigms, we set out to find the transcription factors responsible for activating or repressing this transcription. We conducted DNase-Seq in the rat heart to discover where the open chromatin areas were and what transcription factors may be binding to alter this gene expression. To validate our DNase-Seq in the rat heart, we overlapped our peaks with those of ENCODE'S mouse DNase-Seq peaks [170]. Overall, with strict peak calling, we had a 76% overlap of our rat rn6 enhancers to the mouse mm10 enhancers (Figure III-7 A). Figure III-7 B is an IGV browser of the comparison of DNase-Seq mapped reads to Nr1d1 and Col1a1 genes, and is an example of the synteny and the differences of enhancers found between rat and mouse. The yellow boxed peaks illustrated in Figure III-7 B are examples of peaks that were further analyzed using HOMER motif analysis to discover transcription factors that might be altered by shift work [171-173].

We conducted our motif analysis using HOMER on open chromatin regions observed 10kb up-stream of transcription start sites (TSS) and 1kb down-stream of transcription termination sites (TTS) of the genes found in categories 1A, 1B, 2A, 2B, 3, & 4 [173]. We found a significant enrichment of transcription factor motifs associated with transcription factors that have been correlated with fibrosis and cardiovascular disease in the heart for differentially expressed genes in WRF rats [157,174-178] (Figure III-8 A).

**Figure III-7 Shared Synteny Between DHSs in the Mouse and Rat.**

(A) Rat DNase-Seq peaks were converted to mouse mm10 genome with LiftOver and then overlapped with mouse DNase hypersensitive sites (DHS) from the ENCODE project to observe synteny between genomes. 76% of the lifted-over rat DHS overlapped with the mouse ENCODE dataset. (B) IGV browser visualization of rat DNase-Seq reads and mouse DNase-Seq reads for Nr1d1 and Col1a1. DNase-Seq peaks are boxed in yellow and illustrate the differences between in the synteny of DHS between mouse and rat.

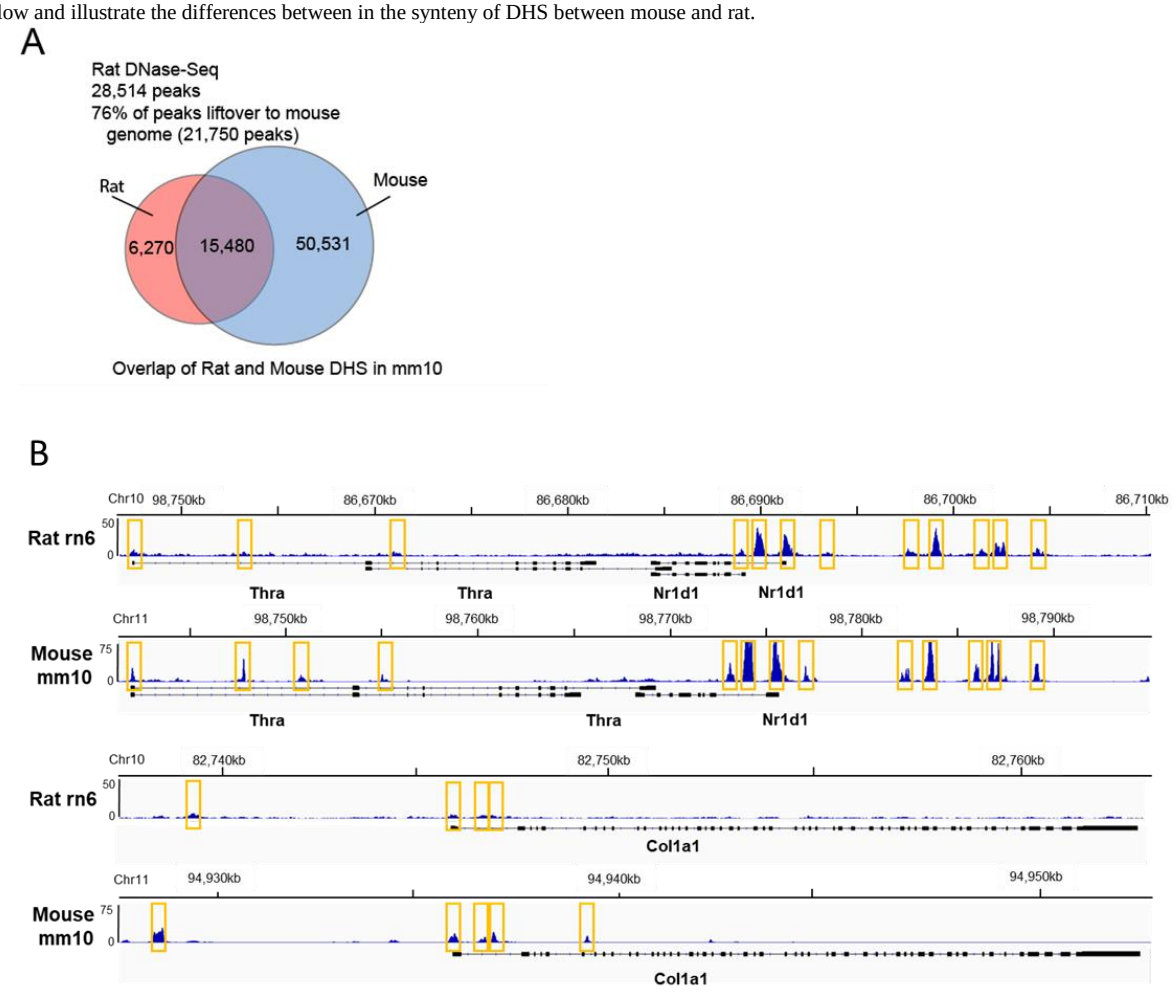


**Figure III-8 Motif Analysis of DHS in Genes Found in Differentially Expressed Categories.**
HOMER motif analysis of DHS within genes found in categories 1A_1B, 2A_2B, 3, and 4. Motifs with a p-value < 0.05 are colored from yellow to red based on p-value, with red being the most significant.

A

| Motif Name | Consensus | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| ETS | AACCGGAAGT | | | | |
| ELF1 | AVCCGGAAGT | | | | |
| Elk4 | NRYTTCCGGY | | | | |
| Elk1 | HACTTCCGGY | | | | |
| GABPA | RACCGGAAGT | | | | |
| Sp1 | GGCCCCGCCCCC | | | | |
| Fli1 | NRYTTCCGGH | | | | |
| ETS1 | ACAGGAAGTG | | | | |
| ETV1 | AACCGGAAGT | | | | |
| KLF3 | NRGCCCCRCCCHBNN | | | | |
| Etv2 | NNAYTTCCTGHN | | | | |
| EWS:FLI1-fusion | VACAGGAAAT | | | | |
| NRF | STGCGCATGCGC | | | | |
| NRF1 | CTGCGCATGCGC | | | | |
| Ronin | RACTACAACTCCCAGVAKGC | | | | |
| ERG | ACAGGAAGTG | | | | |
| Sp5 | RGKGGGCGGAGC | | | | |
| NFY | RGCCAATSRG | | | | |
| GFY-Staf | RACTACAATTCCCAGAAKGC | | | | |
| Klf4 | GCCACACCCA | | | | |
| EHF | AVCAGGAAGT | | | | |
| GFY | ACTACAATTCCC | | | | |
| bZIP50 | GATGACGTCA | | | | |
| ELF5 | ACVAGGAAGT | | | | |
| CRE | CSGTGACGTCAC | | | | |
| Klf9 | GCCACRCCCACY | | | | |
| TFE3 | GTCACGTGACYV | | | | |
| HY5 | RRTSACGTSD | | | | |
| YY1 | CAAGATGGCGGC | | | | |
| KLF5 | DGGGYGKGGC | | | | |
| NAP | ARGTTACGTRTN | | | | |
| Atf1 | GATGACGTCA | | | | |
| Atf2 | NRRTGACGTCAT | | | | |
| Atf7 | NGRTGACGTCAY | | | | |
| BIM1 | NNNNNNVTCACGTGM | | | | |
| BIM3 | TWVTCACGTGAB | | | | |
| BORIS | CNNBRGCGCCCCCTGSTGGC | | | | |
| Cbf1 | TCACGTGAYH | | | | |
| c-Jun-CRE | ATGACGTCATCY | | | | |
| E-box | SSGGTCACGTGA | | | | |
| Ets1-distal | MACAGGAAGT | | | | |
| EWS:ERG-fusion | ATTTCCTGTN | | | | |
| KLF6 | MKGGGYGTGGCC | | | | |
| Lhx3 | ADBTAATTAR | | | | |
| Rap210 | TGTCGGCA | | | | |
| Usf2 | GTCACGTGGT | | | | |

-log10(p-value)

185 150 115 80 45 10 2

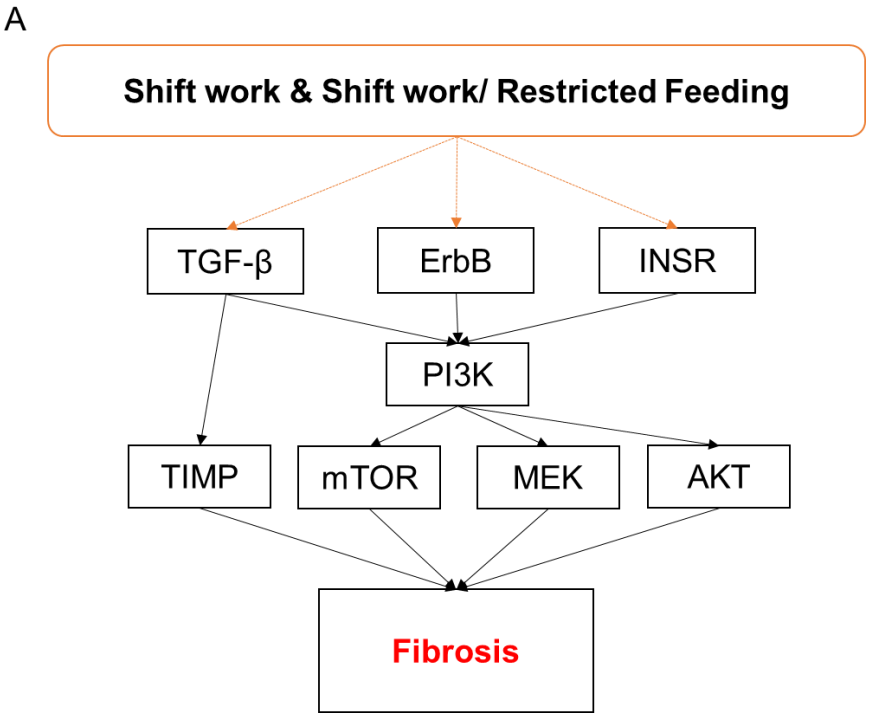
In summary, our rat DNase-Seq open chromatin regions overlap well with the ENCODE'S mouse DNase-Seq open chromatin regions. Furthermore, our motif analysis identified transcription factors known to cause fibrosis and suggests that they may be altering our differential gene expression observed between our paradigms. Importantly, our open chromatin regions and RNA-seq show an enrichment for the same pathways and transcription factors that may be being altered to initiate cardiovascular diseases in shift workers.

***Fibrosis mechanisms cannot be fully explained with gene expression***

Several mechanisms may be contributing to the collagen deposition we are observing in W and WRF. Previous studies have demonstrated that the transforming growth factor beta (TGF-β), epidermal growth factor (ErbB), insulin signaling (INSR), and solute carrier family 2 facilitated glucose transporter (GLUT) pathways can be responsible for activating fibroblasts leading to fibrosis and cardiac failure in the heart (Figure III-9 A) [51,154,168,169].

**Figure III-9 Proposed Model of How Shift Work and Shift Work Restricted Feeding May Increase Fibrosis through Multiple Signaling Pathways.**

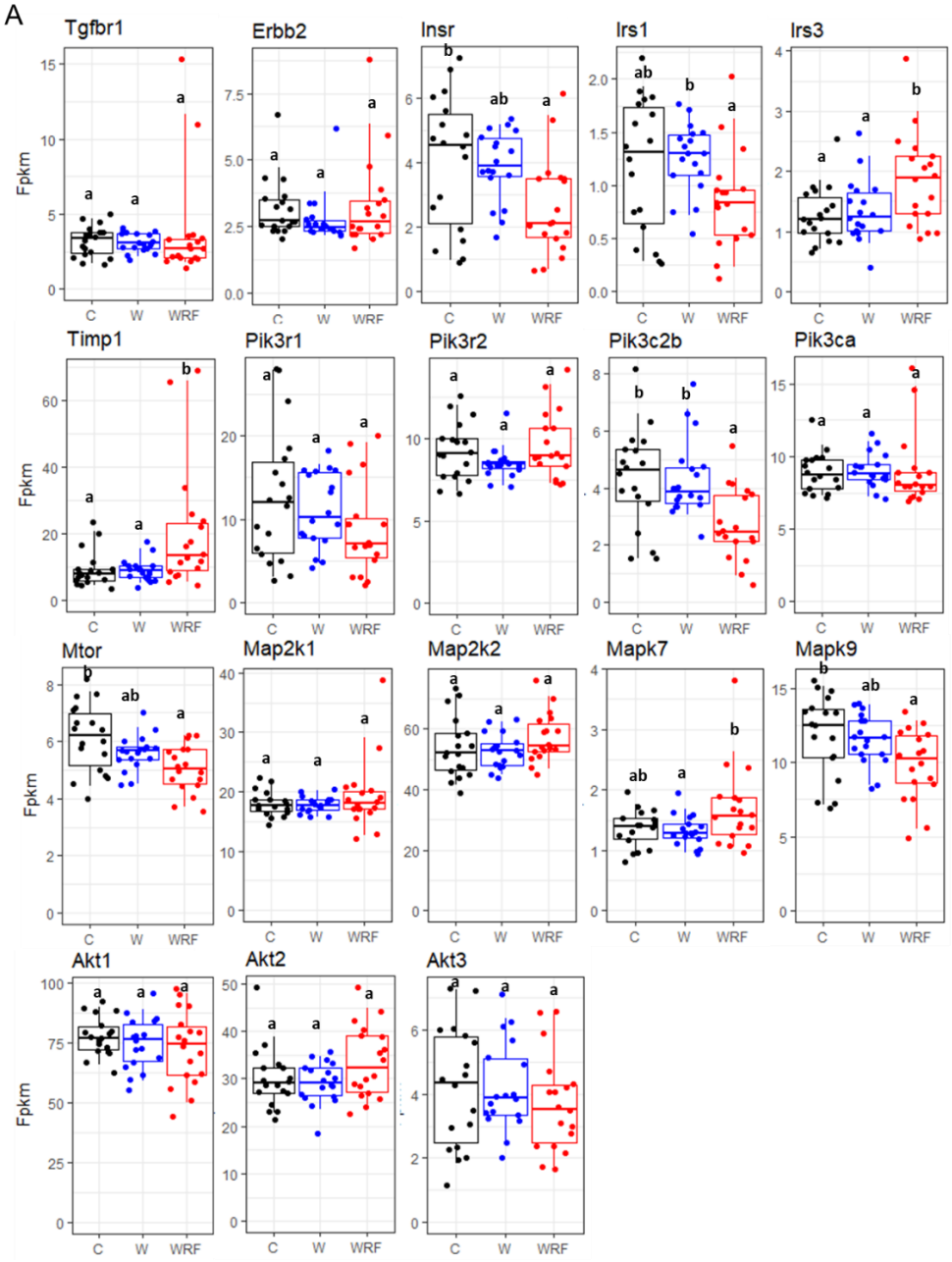
A. Pathway schematic of how shift work leads to increased collagen and fibrosis in the heart. Shift work promotes the activation of the TGF-β, ErbB, and INSR pathways. TGF-β pathway initiates the activation of TIMPs which stimulates collagen deposition. TGF-β, ErbB, and INSR pathways all activate PI3K which in-turn activates mTOR, MEK, and AKT. mTOR and AKT are involved in up-regulation of translation and MEK initiates a cascade that activates transcription of profibrotic genes.




One of the most important factors for regulating the transition of epicardial cells to fibroblasts is the TGF- β pathway [44]. Strangely enough, we do not observe a significant change in gene expression of transforming growth factor beta receptor (*Tgfr1),* but we do observe a significant up-regulation of gene expression of TIMP metallopeptidase inhibitor1 (*Timp1*) in WRF (Figure III-10 A). Timp1 is up-regulated when the TGF- β pathway is up-regulated and is responsible for inhibiting metalloproteinases and in the process stimulates collagen deposition and fibrosis, suggesting that this maybe one of the pathways that is altered in WRF to increase collagen

deposition [169]. Up-regulation of the TGF- β pathway is considered a master switch in its ability to induce other downstream profibrotic mediators such as ErbB, INSR, and GLUT [51]. However, regulation of the other pathways has been shown to independently mediate enhanced collagen production in fibroblasts [51,179]. We observed that expression of Erb-B2 receptor tyrosine kinase 2 (*Erbb2*) is not significantly differentially expressed between the three paradigms. However, we do observe differences in expression of phosphatidylinositol-4,5-bisphosphate 3 kinase (PI3K)'s catalytic and regulator subunits, as well as, insulin receptor (*Insr*), insulin receptor substrate1(*Irs1*), and insulin receptor substrate 3 (*Irs3*). Surprisingly, we do not see differences in gene expression of AKT serine/threonine kinases (*Akt1*), (*Akt2*), (*Akt3*) (Figure III-10 A).

**Figure III-10 WRF Alters the Expression for Some of the Genes Involved in Activation of Fibrosis.**
A. Boxplots of gene expression for genes involved in multiple signaling pathways in the Figure III-9A. Groups with different letters are significantly different with a Tukey's HSD p-value <0.05.



Together these results suggest that shift work induces collagen deposition in the heart, even when eating has been shifted to the natural active phase. Additionally, these pathways may be playing a role in the collagen deposition in the W and WRF. However, gene expression alone may not be sufficient to explain differences in collagen deposition between the paradigms, which suggests that shift work mediated regulation of these pathways may be post-transcriptional instead of transcriptional.

**Discussion**

Shift work has been strongly correlated with cardiovascular disease and it has been suggested that shift workers have a 40% increase in cardiovascular disease compared to day workers. Furthermore, the mechanism underlying how shift work is causing cardiovascular disease is unknown [41,180]. In our experimental model, using rats that are forced to be active during the day (W) and rats that are forced to be active during the day but only allowed to eat during the night (WRF), we show that shift work causes an increase in collagen deposition in the heart for both W and WRF groups. This is surprising because we provide the first evidence of how shift work may be causing cardiovascular disease and, furthermore, that time of eating does not seem to inhibit this effect in the heart. These findings are contradictory with what has been previously published for shift work and metabolic disorders, such that eating during the natural active phase while doing shift work is sufficient to inhibit metabolic disorders, such as obesity and diabetes [36,181].

To address the mechanism behind how shift work is inducing an increase in collagen deposition in both W and WRF, we characterized, for the first time, the transcriptome of C, W, and WRF rat hearts. Surprisingly, we found a significant difference in gene expression for WRF rats compared to C and W. Interestingly, Song et al. 2012 also found a significant difference in

gene expression of mice that had pathological hypertrophy compared to mice that had physiological hypertrophy, suggesting that more pathways are significantly altered in hearts undergoing heart failure. Additionally, genes that were up regulated in our WRF rats were significantly enriched for ribosome and proteasome KEGG Pathways. This suggests that there may be an up-regulation in proteins, which previous publications have found to be an early marker of heart disease and is needed for increased proteins found in pathological hypertrophy [182,183]. In this regard, it is of great interest and importance to further study these changes that we observed for gene expression, and how they translate to changes in protein expression. Furthermore, we found NF-kappa B signaling, AMPK signaling, PPAR, FoxO, mTOR, and ErbB signaling pathways to be differentially expressed in our W and WRF rats. These pathways when altered have been attributed to cause fibrosis, hypertrophy, and heart failure [54,184,185].

Investigation of transcription factors that may be recruited to these differentially expressed genes revealed an enrichment of motifs for transcription factors that regulate connective tissue growth factor (CCN2) which is essential in regulating angiogenesis, chondrogenesis, and wound healing [176]. Friend leukemia integration 1 (FLI1) and ETS proto-oncogene 1 (ETS1) were both found in enhancers surrounding genes that are up-regulated in WRF, and both have been found to regulate extracellular matrix genes such as Collagen type 1. Additionally, we found Kruppel-like factors (KLF) motifs surrounding many of the differentially expressed genes and these have been implicated in many cellular processes such as glucose homeostasis and cardiac hypertrophy [177]. In particular, KLF5 has been reported to play a role in cardiovascular remodeling through TGF-β pathway [186].

On a cellular level the heart is composed of cardiac myocytes, fibroblasts, endothelial cells, mast cells, and vascular smooth muscle cells surrounded by extracellular matrix [187]. Under certain

types of stresses, cardiac myocytes increase their functional load, activating signaling pathways by changing gene expression, increasing protein synthesis, and increasing collagen deposition [188]. In response to this increase in functional load, the heart mass will increase to try and normalize wall stress. This enlargement of cardiac muscle cells is referred to as cardiac hypertrophy. Additionally, this increase in heart mass typically occurs in the ventricles. If equilibrium cannot be reached the hypertrophied heart can dilate and contraction of the myocytes fails occur progressing in heart failure [154]. We observe an increase in ribosome subunit expression as well as an enrichment of the KLF motif for genes that were differentially expressed in W and WRF rats. This data implicates that cardiac hypotrophy is occurring; however, further research needs to be done to see if there is an increase in heart or cardiac myocyte size for W and WRF rats compared to C.

Many animal studies have been done to try and discover the molecular mechanisms that occur in pathological hypertrophy, however not all of these have been done without contention. Collagen 1 and collagen 3 are always found with pathological hypertrophy, and this is what we found in our W and WRF rat hearts. This deposition of collagen is produced by fibroblasts to replace myocytes that undergo apoptosis [53,159,184,189]. Furthermore, many signaling pathways such as TGF-β, ErbB, and INSR have been attributed to lead to fibrosis and hypertrophy [51,155]. The TGF- β signaling pathway, in particular, is responsible for inhibiting expression of matrix metalloproteinases (MMPs), needed to degrade extracellular matrix proteins, and activating of tissue inhibitors of metalloproteinases (TIMPs), which inhibit MMPs. It is this disturbance in the balance between the MMPs and TIMPs that allow for accumulation of collagens in the extracellular matrix [53]. More specifically, our data also shows an increase in Timp1 gene expression, eluding to this pathway being altered by WRF. TGF- β signaling pathway is also a

multifunctional cytokine that plays an important role in proliferation, differentiation, and apoptosis by signaling through PI3K to mTOR, MEK, AKT[155]. Furthermore, signaling through ErbB and INSR pathways also feed into PI3K and are essential for myofilament structure, glucose uptake, and angiogenesis [54,55]. Surprisingly, our gene expression data is unclear on whether these pathways are being altered or not by shift work. Further analysis is needed to observe if post-transcriptional mechanisms are occurring in W and WRF rats to affect these pathways and to cause the increase in collagen deposition in W and WRF rats.

In conclusion our data suggests that five weeks of shift work, independent to time of eating, induces an increase in collagen deposition in the rat heart. Furthermore, eating at night during the natural active phase is not sufficient to inhibit collagen deposition. Astonishingly, we observe significant differential gene expression in WRF compared to W and C. Genes that are up-regulated in WRF are enriched for ribosome, proteasome, NF-kappa B, AMPK, and PPAR signaling pathways. Open chromatin regions surrounding genes found to be differentially expressed were found to be enriched for transcription factors found to be involved in fibrosis and cardiovascular disease. Surprisingly, we only observe an increase in expression of collagen genes and fibroblast markers in WRF and not W when compared to C. Our results indicate that gene expression alone is not sufficient to explain collagen deposition and that post transcriptional mechanisms must be occurring. Further studies will be needed to precisely verify the molecular pathways in which shift work is increasing collagen levels in the rat heart.

## Methods

### *Animals*

Male Wistar rats around the age of 7-8 weeks underwent a shiftwork protocol for five weeks. Each animal was housed individually in a 12 L:12 D cycle (7am to 7pm) throughout the protocol. Protocol consisted of 54 animals randomly assigned to three groups: 18 Controls 'C' (undisturbed and no shiftwork protocol, *ad libitum* food and water), 18 Workers 'W' (shiftwork protocol, *ad libitum* food and water), 18 Worker Restricted Feeding 'WRF' (shiftwork protocol, *ad libitum* water, access to food only at night). Shiftwork protocol consisted of five days of scheduled 'work' (workday: Monday to Friday, 9am to 5pm) followed by two days without 'work' and given *ad libitum* food and water (weekend: Saturday and Sunday). Experiments were approved by Universidad Nacional Autonoma de Mexico, in accordance with animal handling, Norma Oficial Mexicana NOM-062-ZOO-1999.

### *Shiftwork protocol*

In order to simulate shiftwork, rats were placed in a slowly rotating drum that were designed to keep the animal awake during their natural rest phase or when lights were turned on. Drums rotated with a speed of one revolution/three minutes, this was done to allow animals to move, eat, and drink freely. However, animals were not able fall asleep; forcing the animals to stay awake but not requiring them to exhibit laborious movements. Food and water were hung from a concentric middle tube and available depending on which group the rats were placed in. Outside of work periods rats were housed in their individual cages with *ad libitum* food and water and kept under 12:12 LD cycles [190].

***Heart Tissue Collection***

Rats entrained on 12:12 LD and after five weeks of shiftwork protocol were sacrificed every four hours starting one hour after lights turned on at 7am. This was done to account for any time of day effects. Heart tissue was collected and flash frozen and then shipped to Texas A&M University.

***Whole transcriptome sequencing (RNA-Seq)***

RNA was isolated from crushed frozen heart samples using Trizol reagent according to the manufacturer's instructions (Life Technologies). Total RNA was diluted to 4µg and ribosomal RNA was depleted using Dyna oligo dT beads following manufacturer's instructions (Invitrogen). Sequencing libraries were constructed using NEB Utra Directional RNA Library Prep Kit from Illumina (New England BioLabs) following manufacturer's instructions (n=54). Total intact RNA was fragmented for 15 mins at 94$^{\circ}$C before first and second strand cDNA synthesis. Purification of double stranded cDNA was done using AMPure XP beads following NEB Ultra Directional RNA Library Prep Kit instructions. PCR library enrichment was amplified for 13 cycles and following clean up were quantified by Quantus Fluorometer (# E6150, Promega) and qPCR with TRUseq library standards. Libraries were generated with multiple bar-coded adaptors, mixed in equimolar ratio and sequenced on three lanes of an Illumina NextSeq with a sequencing length of 76bps.

***DNase-sequencing***

Nuclei was isolated from crushed frozen heart samples (n=2) using a dounce homogenizer in lysis buffer (0.32M sucrose, 10mM Tris-HCl ph8, 5mM CaCl2, 5mM MgCl2, 2.0mM EDTA, 0.5mM EGTA, 1mM DTT). Samples were dounced 6x with pestle A and 4x with pestle B.

Homogenate was mixed with sucrose lysis buffer (2.2M sucrose, 10mM tris HCl ph8, 5mM CaCl2, 5mM MgCl2, 2.0mM EDTA, 0.5mM EGTA, 1mM DTT) and then layered on top of sucrose cushion (2.05M sucrose, 10% glycerol, 125mM glycine, 10 mM HEPES pH7.6, 15 mM KCl, 2 mM EDTA, 1mM PMSF, 0.15mM spermine, 0.5mM spermidine, 0.5mM DTT) in an ultracentrifuge tube. Samples were centrifuged for 45 min at 24,000 rpm (100,000g) at 2°C in a Beckmann SW28 rotor. Supernatant was removed and nuclei were washed in resuspension buffer (10 mM Tris, pH 7.5, 150 mM NaCl, 2 mM EDTA, 1mM PMSF) 3x and then counted and aliquoted to tubes of five million nuclei and flash frozen. Flash frozen nuclei were thawed and nuclei were pelleted and supernatant was removed before tubes were digested with DNase I digestion buffer (6mM CaCl2, 75mM NaCl2, 13.5mM Tris-HCl ph8, 13.5mM NaCl, 54mM KCl, 0.9mM EDTA, 0.45mM EGTA, 0.45 spermidine) and 80units/ml of DNase I enzyme. Following digestion, equal volumes of stop buffer (50mM Tris-HCl ph8, 100mM NaCl, 0.1% SDS, 100mM EDTA, 1mM spermidine, 0.3mM spermine) were added to each tube and then incubated at 55°C for 1 hour. RNA was removed by adding 10mg/ml of RNase A and samples were incubated at 37°C for 30 minutes. Samples were purified by phenol/chloroform extraction and then resuspended in water. DNA size fractionation was done using a sucrose gradient and centrifuged for 30,000 rpm for 20 hours at 20°C in a SW-41 rotor with no brake and minimum acceleration. DNA fractions with a size less than 500bps were kept and pooled together for purification and concentrated using sodium acetate and ethanol and were resuspended in water. DNase-Seq libraries were generated using NEBNext® ChIP-Seq Library Prep Master Mix Set (# E6240, NEB) as per the manufacturer's instructions. DNA collected from the DNase I digestion was quantified using the Quantus Fluorometer (# E6150, Promega), and 10 ng was used to generate the libraries. Libraries were amplified 14 cycles using adaptor ligates DNA using NEBNext Multiplex

oligonucleotides and Phusion Taq (M0530S). Any adaptor dimers were removed using a gel purification and phenol chloroform and ethanol precipitation. Libraries were quantified using TRUseq library standards and qPCR before paired-end sequencing was done with an Illumina NextSeq at a read length of 75bps.

***Alignment of RNA-Seq to the RAT (rn6) and differential gene expression analysis***

Libraries were sequenced to a median depth of ~25 million reads /sample and mapped to rn6 with an average Tophat2 alignment of ~90% using the following criteria: --read-realign-edit-dist 2 -g 1 --b2-sensitive [162]. Gene expression was analyzed using Cufflinks with the following criteria: --library-type fr-firststrand [163]. Visualization of the forward and reverse strands were done using a custom shell script with total signal normalized to 10 million reads and viewed with the Integrated Genome Viewer [171]. Genes where 18 of the 54 biological samples had a log2 (fpkm+1) value over 1 were kept for further processing (75% of the total number of genes). Differential gene expression was done using a one-way ANOVA on shiftwork groups. Post hoc analysis was performed using Tukey's HSD with a p-value $<0.05$.

***Alignment of DNase-Seq to the RAT (rn6) and motif analysis***

Libraries were sequenced to a median depth of ~ 36 million and mapped to rn6 with an average bowtie2 alignment of ~ 85% using the following criteria: -t –phred33 -X 650 [191]. Sam files were converted to bam files and merged using SAMtools view, sort, and merge [192]. Duplicates were removed with Picard 2.8.2 MarkDuplicatesWithMateCigar created by the Broad Institute. A custom python script was used to remove unpaired reads, unmapped reads, and filled in sequences between paired-end reads. Sorted coordinates were then made into BWfiles using Bedtools, normalized to 10 million reads and viewed with the Integrated Genome Viewer [171,193]. Peak calling was done using findPeaks from the HOMER suite using the following commands: -size 45 -ntagThreshold 2 -region [194]. Open chromatin regions for genes of interest were determined by using HOMER gene annotation and annotation for these genes were extended 10kb from transcription start sites and 1kb from transcription termination sites. These gene windows were overlapped with DNase-Seq reads using intersectBed from Bedtools suite[193]. Motif analysis was done on these overlapped files using the perl script findMotifsGenome.pl from HOMER using the following command: -size given [194].

# CHAPTER IV

# DISCUSSION, SUMMARY, AND FUTURE DIRECTIONS

## Mechanisms controlling rhythmic transcription

In mammals, CLOCK: BMAL1 regulates gene expression rhythmically and it is assumed that this regulation is necessary and sufficient to drive rhythmic expression of genes needed for biological processes. However, gene expression associated with CLOCK:BMAL1 cistromes is highly heterogenous [7,12,83]. We suggest that this heterogenous genes expression is not due to nonfunctionally binding of CLOCK:BMAL1, but that the regulation of gene expression of CLOCK:BMAL1 targets relies on much more complex mechanism. This hypothesis is further supported by the finding that 82.2% of protein coding genes found in the Baboon were found to be rhythmic in at least one tissue, again suggesting that regulation of rhythmic expression relies on more than just CLOCK:BMAL1 binding [195]. This issue lead to the question of whether CLOCK:BMAL1 binds in a tissue specific manner to control tissue specific rhythmic gene expression. This tissue specificity of regulation of rhythmic gene expression suggest that the binding of CLOCK:BMAL1 is not the same in every tissue and that regulation of gene expression is more complicated than CLOCK:BMAL1 activation. Furthermore, full body knockouts of BMAL1 exhibit a distinct change in expression of genes; genes are either found to be up-regulated or down-regulated. This suggests that CLOCK:BMAL1 do not activate expression, but recruit other transcription factors to regulate expression of target genes possibly in a tissue specific manner.

We proposed in chapter II that CLOCK:BMAL1 rhythmically promotes permissive chromatin landscaping that facilitates the binding of transcription factors needed to activate or

repress transcription. Additionally, that this poised enhancer is what is need for the body to adapt to environmental changes and is how the circadian clock controls rhythmic biological processes needed for particular times of the day. Furthermore, this suggest a mechanism of how disruption of the circadian clock leads to many different biological diseases because without these primed enhancers the cell cannot activate transcription fast enough to respond to a stimulus.

Additional work has been done in our lab to further confirm that CLOCK:BMAL1 promote a transcriptionally permissive landscape by characterizing BMAL1 cistrome in three different tissues [196]. This data illustrated that BMAL1 binds tissue specifically with other ts-TFs and that tissue specific expression of target genes relies on interactions with other TFs. Again suggesting that the binding of CLOCK:BMAL1 is to promote a permissive enhancer and not to activate transcription. Furthermore, many of the areas where BMAL1 was binding in a tissue specific manner were considered to be open chromatin in all three tissues [196]. This data suggests that BMAL1 is not a pioneer transcription factor and instead prevents DNA from rewrapping around nucleosomes with the help of ts-TFs [13,196]. Furthermore, this suggests a way in which the cell autonomous circadian clock regulates biological process in a tissue specific way.

Transcription is regulated in many different ways by many different proteins. Traditionally, transcription factors bind to enhancers and recruit additional TFs, PIC, and the Mediator to activate transcription [14,21]. We have observed that CLOCK:BMAL1 appears to bind with other ts-TFs and recruit in ub-TFs to activate or repress transcription. However, little is understood of how these TFs interact with the PIC and the Mediator. One study has been able to link CLOCK with the Mediator (MED1) through the interaction of TRAP150 (thyroid hormone receptor-associated protein-150) [197]. They observed that TRAP150 bridged the gap between CLOCK:BMAL1 and the transcriptional machinery and that when they depleted TRAP150 that this interaction was reduced.

Furthermore, they suggest that this interaction with TRAP150, which is rhythmically regulated, is critical to coordinate rhythmic transcription by linking CLOCK:BMAL1 to basic transcriptional machinery at genes driven predominantly by CLOCK:BMAL1 [197]. These findings suggest a way in which CLOCK:BMAL1 coordinates core clock gene expression, but still leaves to question if CLOCK:BMAL1 interact with the PIC and Mediator on a genome wide level.

Additionally in chapter II we suggest that it may be possible for CLOCK:BMAL1 to mediate long-range chromatin interactions. We observed an enrichment of the CTCF transcription factor motif at genes that were non-expressed and found that enhancers targeting the same gene exhibited similar transcriptional activity profiles. To address whether it was possible for CLOCK:BMAL1 to mediate these long-range chromatin interactions, a PolII ChIA-PET (Chromatin Interaction Analysis by Paired-End Tag Sequencing) experiment was conducted at ZT6 and ZT18 to identify interactions of DHSs for genes undergoing active transcription [196]. This experiment identified that many CLOCK:BMAL1 sites interacted with other DHSs and that these interactions were higher at ZT6, the time of maximal CLOCK:BMAL1 DNA binding. Furthermore, many of these interactions at ZT6 were for genes that were found to be expressed rhythmically in-phase (Rinφ) with CLOCK:BMAL1 DNA binding. This suggests that transcription of these genes relies on CLOCK:BMAL1 coordinating a permissive chromatin landscaping of their enhancers. However, this experiment needs to be done in a BMAL1 knockout to verify this hypothesis. Additionally, many of the AR genes showed CLOCK:BMAL1 DHSs interacting with many other DHSs, suggesting that temporal control of transcription may rely more strongly on the activity of the TFs binding in DHSs without CLOCK:BMAL1. Furthermore, Ro/φ genes showed an enrichment for Rev-erbα, a core clock repressor transcription factor. A recent study has found Rev-erbα to be involved in rhythmically modulating chromatin looping associated

with circadian gene expression [198]. This study found that Rev-erbα bound enhancers recruited in histone deacetylases, evicted elongation factors BRD4 and MED1, and opposed chromatin looping [198]. This data Altogether further suggests that chromatin looping is occurring to help coordinate rhythmic gene expression and that regulation maybe occurring around CLOCK:BMAL1 enhancers.

Altogether this suggests that the circadian clocks transcriptional mechanisms are complex and multilayered. The pervasive circadian regulation of biological processes are intimately linked to regulation of physiologies, requiring gene expression to coincide with metabolic demands to optimize energy utilization [7,199]. Our analysis and those done by others, suggests ways in which the dynamic circadian clock may adapt to an ever-changing environment.

**Disruption of the circadian clock leads to disease**

In mammals, the circadian system is a collection of cell autonomous biological clocks that are all synchronized by the SCN [5]. This master circadian clock is synchronized by light:dark cycles and sends signals to the rest of the peripheral clocks by controlling body temperature, feeding behavior, and hormone releases [5,33]. Disruptions of light:dark cycles or any of the signals sent by the SCN to the rest of the body has been correlated to many different diseases [40,200,201].

Shift work requires people to work outside the normal 9am to 5pm schedule and causes people to be exposed to many alter light:dark cycles, changes in body temperature, altered feeding behaviors, and changes in hormone releases [35]. Additionally, many of diseases such as metabolic disease, cardiovascular disease, and cancer have been correlated with shift work and it has been suggested that this is caused by disrupting the coordination of the circadian clock [33].

Many studies have been done to try and understand the mechanism of how shift work causes disease. Population-based studies are problematic because humans have to many outcomes that are long-term, as well as too much variation in lifestyle between each individual [35,39,202]. Therefore, an animal model is preferred and has been used to try and investigate the mechanisms that underly the diseases correlated with shift work.

Multiple studies done by our collaborators aimed at addressing how shift work causes metabolic disorders. They did this by forcing male Wistar rats to stay active during their rest phase and had different groups altering time of food availability, because food consumption is known as being a very important synchronizer of the peripheral clocks and that changes in time of food intake has led to metabolic disorders [143,190,203]. Additionally, shifting the time of food intake is very characteristic feature of shift workers [35]. In this study they found that even the Control rats, who did no amount of shift work but had food restricted to their rest phase, during the day, exhibited an increase in bodyweight and fat accumulation compared to *ad libitum* or night fed Controls [181,190]. Furthermore, they observed that the phase of the molecular circadian clock altered in the liver changed according to when the rats were fed, and that obesity and diabetic symptoms could be inhibited by time of feeding [36,204]. These findings suggest that it is the time of food intake and not shift work that leads to metabolic disorders. These findings raise the question of whether the act of shift work causes health disorders or if it is the changes in signals that desynchronizes the peripheral clocks.

In chapter III we aimed to investigate the molecular mechanisms by which shift work and or food consumption contribute to cardiovascular disorders. We found that five consecutive weeks of shift work Monday-Friday was enough time to increase collagen deposition in the hearts of both Worker rats and Worker Restricted Feeding rats. We also discovered a significant difference in

gene expression for Worker Restricted Feeding rats compared to Controls and Workers. This was surprising because eating during the normal active phase was enough to prevent metabolic disorders and a shift in the circadian clock in the liver [36]. Therefore, since we observed an increase in collagen deposition in both the Workers and Worker Restricted Feeding rats, this suggests that the desynchrony caused by food consumption is not causing an increase collagen deposition.

Furthermore, we did not observe a shift in the expression of molecular circadian clock in the heart for Worker Restricted Feeding rats compared to Control rats (data not shown), however we still observed an increase in collagen and significant changes in gene expression. This data suggests that desynchronization of the molecular circadian clock in the heart is not needed to cause an increase in collagen deposition or up-regulation of profibrotic gene expression. This suggests that these changes may actually be caused by doing shift work and not a desynchronization of the molecular circadian clock in the rat heart.

Food consumption is not the only signal altered when these animals do shift work. These rats are also forced to stay awake during the day and are thus exposed to light and the stress of being forced to stay awake out of synchronization with their natural wake and sleep cycle. Many studies have indicated a positive correlation between stressors and corticosterone levels [37]. Furthermore, corticosterone observed in rats forced to do shift work has a biphasic peak, one when the rats are forced to work and the second just as lights turn off [181,190]. This peak just as the rats are forced to do shift work suggests that these rats do experience stress and that this may be playing a role in the collagen deposition observed in Worker and Worker Restricted Feeding rat hearts. Additionally, the stress of shift work may be causing a heightened activation of the sympathetic nervous system enhancing the adrenergic signaling pathways via catecholamines, norepinephrine, and epinephrine [55,205]. This heightened activation of the adrenergic signaling pathway has also

been associated with triggering cardiovascular disease and unless treated has been known to progress to heart failure [205,206]. However, this hypothesis that shift work is causing a heightened activation of the adrenergic signaling pathway in shift working rats will need to be validated.

Additionally, in Chapter III we discussed that the TGF-β, ErbB, INSR and GLUT pathways activate fibroblast that can lead to fibrosis and cardiac failure in the heart [51,154,168,169]. However, even though we observed an increase in collagen deposition, we observed very little changes in gene expression associated with these pathways in Worker and Worker Restricted Feeding rats. These results suggest that regulation of these pathways occurs at the protein level and that future work such as proteomics should be done to address possible changes caused by shift work.

Our analysis, is the first to correlate shift work, regardless of the time of eating, with an increase in collagen deposition in the heart. Altogether this suggests that shift work can cause desynchronization at many different levels and that gene expression alone is not sufficient to explain how shift work causes cardiovascular disease.

REFERENCES


1 Balsalobre, A., Damiola, F. & Schibler, U. A serum shock induces circadian gene expression in mammalian tissue culture cells. *Cell* **93**, 929-937 (1998).

2 Paranjpe, D. A. & Sharma, V. K. Evolution of temporal order in living organisms. *Journal of Circadian Rhythms*, doi:http://doi.org/10.1186/1740-3391-3-7 (2005).

3 Fu, L., Pelicano, H., Liu, J., Huang, P. & Lee, C. The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. *Cell* **111**, 41-50 (2002).

4 Marcheva, B. *et al.* Disruption of the clock components CLOCK and BMAL1 leads to hypoinsulinaemia and diabetes. *Nature* **466**, 627-631, doi:10.1038/nature09253 (2010).

5 Mohawk, J. A., Green, C. B. & Takahashi, J. S. Central and peripheral circadian clocks in mammals. *Annual review of neuroscience* **35**, 445-462, doi:10.1146/annurev-neuro-060909-153128 (2012).

6 Hogenesch, J. B. & Ueda, H. R. Understanding systems-level properties: timely stories from the study of clocks. *Nature reviews. Genetics* **12**, 407-416, doi:10.1038/nrg2972 (2011).

7 Koike, N. *et al.* Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338**, 349-354, doi:10.1126/science.1226339 (2012).

8 Yoshitane, H. *et al.* Roles of CLOCK phosphorylation in suppression of E-box-dependent transcription. *Mol Cell Biol* **29**, 3675-3686, doi:10.1128/MCB.01864-08 (2009).

9 Kume, K. *et al.* mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell* **98**, 193-205 (1999).

10 Yoo, S. H. *et al.* A noncanonical E-box enhancer drives mouse Period2 circadian oscillations in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2608-2613, doi:10.1073/pnas.0409763102 (2005).

11 Glossop, N. R. & Hardin, P. E. Central and peripheral circadian oscillator mechanisms in flies and mammals. *Journal of cell science* **115**, 3369-3377 (2002).

12 Menet, J. S., Rodriguez, J., Abruzzi, K. C. & Rosbash, M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *eLife* **1**, e00011, doi:10.7554/eLife.00011 (2012).

13 Menet, J. S., Pescatore, S. & Rosbash, M. CLOCK:BMAL1 is a pioneer-like transcription factor. *Genes & development* **28**, 8-13, doi:10.1101/gad.228536.113 (2014).

14 Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 129-143, doi:10.1038/nrm3952 (2015).

15 Ogata, K., Sato, K. & Tahirov, T. H. Eukaryotic transcriptional regulatory complexes: cooperativity from near and afar. *Curr Opin Struct Biol* **13**, 40-48 (2003).

16 Soboleva, T. A., Nekrasov, M., Ryan, D. P. & Tremethick, D. J. Histone variants at the transcription start-site. *Trends in genetics : TIG* **30**, 199-209, doi:10.1016/j.tig.2014.03.002 (2014).

17 Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nature reviews. Molecular cell biology* **16**, 178-189, doi:10.1038/nrm3941 (2015).

18 Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annual review of biochemistry* **78**, 273-304, doi:10.1146/annurev.biochem.77.062706.153223 (2009).

19 Vollmers, C. *et al.* Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell metabolism* **16**, 833-845, doi:10.1016/j.cmet.2012.11.004 (2012).

20 Wang, L. *et al.* INO80 facilitates pluripotency gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. *Cell stem cell* **14**, 575-591, doi:10.1016/j.stem.2014.02.013 (2014).

21 Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* **16**, 155-166, doi:10.1038/nrm3951 (2015).

22 Meyer, K. D., Lin, S. C., Bernecky, C., Gao, Y. & Taatjes, D. J. p53 activates transcription by directing structural shifts in Mediator. *Nature structural & molecular biology* **17**, 753-760, doi:10.1038/nsmb.1816 (2010).

23 Kagey, M. H. *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-435, doi:10.1038/nature09380 (2010).

24 Knuesel, M. T., Meyer, K. D., Donner, A. J., Espinosa, J. M. & Taatjes, D. J. The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independently of mediator. *Mol Cell Biol* **29**, 650-661, doi:10.1128/MCB.00993-08 (2009).

25 Levine, M., Cattoglio, C. & Tjian, R. Looping Back to Leap Forward: Transcription Enters a New Era. *Cell* **157**, 13-25, doi:10.1016/j.cell.2014.02.009 (2014).

26 Li, G. *et al.* Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* **148**, 84-98, doi:10.1016/j.cell.2011.12.014 (2012).

27 Zhang, Y. *et al.* Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature* **504**, 306-310, doi:10.1038/nature12716 (2013).

28 Sandhu, K. S. *et al.* Large-scale functional organization of long-range chromatin interaction networks. *Cell reports* **2**, 1207-1219, doi:10.1016/j.celrep.2012.09.022 (2012).

29 Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature reviews. Genetics* **14**, 390-403, doi:10.1038/nrg3454 (2013).

30 Kieffer-Kwon, K. R. *et al.* Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* **155**, 1507-1520, doi:10.1016/j.cell.2013.11.039 (2013).

31 Hou, T. *et al.* Characterization of domain-peptide interaction interface: a generic structure-based model to decipher the binding specificity of SH3 domains. *Molecular & cellular proteomics : MCP* **8**, 639-649, doi:10.1074/mcp.M800450-MCP200 (2009).

32 Seitan, V. C. *et al.* Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome research* **23**, 2066-2077, doi:10.1101/gr.161620.113 (2013).

33 Evans, J. A. & Davidson, A. J. Health consequences of circadian disruption in humans and animal models. *Progress in molecular biology and translational science* **119**, 283-323, doi:10.1016/B978-0-12-396971-2.00010-5 (2013).

34 Wang, D., Ruan, W., Chen, Z., Peng, Y. & Li, W. Shift work and risk of cardiovascular disease morbidity and mortality: A dose-response meta-analysis of cohort studies. *European journal of preventive cardiology* **25**, 1293-1302, doi:10.1177/2047487318783892 (2018).

35 Knutsson, A. Health disorders of shift workers. *Occupational medicine (Oxford, England)* **53**, 103-108 (2003).

36 Salgado-Delgado, R. C. *et al.* Shift Work or Food Intake during the Rest Phase Promotes Metabolic Disruption and Desynchrony of Liver Genes in Male Rats. *PLOS ONE* **8**, e60052, doi:10.1371/journal.pone.0060052 (2013).

37 Ulhôa, M. A., Marqueze, E. C., Burgos, L. G. A. & Moreno, C. R. C. Shift Work and Endocrine Disorders. *International Journal of Endocrinology* **2015**, 826249, doi:10.1155/2015/826249 (2015).

38 Berson, D. M. Strange vision: ganglion cells as circadian photoreceptors. *Trends in Neurosciences* **26**, 314-320, doi:https://doi.org/10.1016/S0166-2236(03)00130-9 (2003).

39 Boggild, H. & Knutsson, A. Shift work, risk factors and cardiovascular disease. *Scandinavian journal of work, environment & health* **25**, 85-99 (1999).

40 Morris, C. J., Yang, J. N. & Scheer, F. A. The impact of the circadian timing system on cardiovascular and metabolic function. *Progress in brain research* **199**, 337-358, doi:10.1016/B978-0-444-59427-3.00019-8 (2012).

41 Mosendane, T., Mosendane, T. & Raal, F. J. Shift work and its effects on the cardiovascular system. *Cardiovascular Journal of Africa* **19**, 210-215 (2008).

42 Stewart, J., Manmathan G Fau - Wilkinson, P. & Wilkinson, P. Primary prevention of cardiovascular disease: A review of contemporary guidance and literature. (2017).

43 Dimmeler, S. Cardiovascular disease review series. *EMBO Molecular Medicine* **3**, 697-697, doi:10.1002/emmm.201100182 (2011).

44 Travers, J. G., Kamal, F. A., Robbins, J., Yutzey, K. E. & Blaxall, B. C. Cardiac Fibrosis: The Fibroblast Awakens. *Circulation research* **118**, 1021-1040, doi:10.1161/CIRCRESAHA.115.306565 (2016).

45 Tenkanen, L., Sjoblom, T. & Harma, M. Joint effect of shift work and adverse life-style factors on the risk of coronary heart disease. *Scandinavian journal of work, environment & health* **24**, 351-357 (1998).

46 Han, S. S. *et al.* Circadian control of bile acid synthesis by a KLF15-Fgf15 axis. *Nature communications* **6**, 7231, doi:10.1038/ncomms8231 (2015).

47 Niiranen, T. J. & Vasan, R. S. Epidemiology of cardiovascular disease: recent novel outlooks on risk factors and clinical approaches. *Expert Rev Cardiovasc Ther* **14**, 855-869, doi:10.1080/14779072.2016.1176528 (2016).

48 Akerstedt, T. & Wright, K. P., Jr. Sleep Loss and Fatigue in Shift Work and Shift Work Disorder. *Sleep medicine clinics* **4**, 257-271, doi:10.1016/j.jsmc.2009.03.001 (2009).

49 Mendlewicz, J. Disruption of the circadian timing systems: molecular mechanisms in mood disorders. *CNS drugs* **23 Suppl 2**, 15-26, doi:10.2165/11318630-000000000-00000 (2009).

50 Ivey, M. J. & Tallquist, M. D. Defining the Cardiac Fibroblast: A New Hope. *Circulation journal : official journal of the Japanese Circulation Society* **80**, 2269-2276, doi:10.1253/circj.CJ-16-1003 (2016).

51 Andrianifahanana, M. *et al.* Profibrotic up-regulation of glucose transporter 1 by TGF-beta involves activation of MEK and mammalian target of rapamycin complex 2 pathways. (2016).

52 Bujak, M. & Frangogiannis, N. G. The role of TGF-beta signaling in myocardial infarction and cardiac remodeling. (2007).

53 Kwak, H.-B. Aging, exercise, and extracellular matrix in the heart. *Journal of exercise rehabilitation* **9**, 338-347, doi:10.12965/jer.130049 (2013).

54 Pentassuglia, L. & Sawyer, D. B. The role of Neuregulin-1beta/ErbB signaling in the heart. *Experimental cell research* **315**, 627-637, doi:10.1016/j.yexcr.2008.08.015 (2009).

55 Fu, Q., Wang, Q. & Xiang, Y. K. Insulin and β Adrenergic Receptor Signaling: Crosstalk in Heart. *Trends in Endocrinology & Metabolism* **28**, 416-427, doi:https://doi.org/10.1016/j.tem.2017.02.002 (2017).

56 Panda, S. *et al.* Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* **109**, 307-320 (2002).

57 Storch, K. F. *et al.* Extensive and divergent circadian gene expression in liver and heart. *Nature* **417**, 78-83, doi:10.1038/nature744 (2002).

58 Bass, J. & Takahashi, J. S. Circadian integration of metabolism and energetics. *Science* **330**, 1349-1354, doi:10.1126/science.1195027 (2010).

59 Kerin, A. & Aguirre, A. Improving health, safety, and profits in extended hours operations (shiftwork). *Industrial health* **43**, 201-208 (2005).

60 Partch, C. L., Green, C. B. & Takahashi, J. S. Molecular architecture of the mammalian circadian clock. *Trends in cell biology* **24**, 90-99, doi:10.1016/j.tcb.2013.07.002 (2014).

61 Menet, J. S., Abruzzi, K. C., Desrochers, J., Rodriguez, J. & Rosbash, M. Dynamic PER repression mechanisms in the Drosophila circadian clock: from on-DNA to off-DNA. *Genes & development* **24**, 358-367, doi:10.1101/gad.1883910 (2010).

62 Ye, R. *et al.* Dual modes of CLOCK:BMAL1 inhibition mediated by Cryptochrome and Period proteins in the mammalian circadian clock. *Genes & development* **28**, 1989-1998, doi:10.1101/gad.249417.114 (2014).

63 Chiou, Y. Y. *et al.* Mammalian Period represses and de-represses transcription by displacing CLOCK-BMAL1 from promoters in a Cryptochrome-dependent manner. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E6072-E6079, doi:10.1073/pnas.1612917113 (2016).

64 Guillaumond, F., Dardente, H., Giguere, V. & Cermakian, N. Differential control of Bmal1 circadian transcription by REV-ERB and ROR nuclear receptors. *Journal of biological rhythms* **20**, 391-403, doi:10.1177/0748730405277232 (2005).

65 Preitner, N. *et al.* The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251-260 (2002).

66 Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 16219-16224, doi:10.1073/pnas.1408886111 (2014).

67 Gekakis, N. *et al.* Role of the CLOCK protein in the mammalian circadian mechanism. *Science* **280**, 1564-1569 (1998).

68 Hogenesch, J. B., Gu, Y. Z., Jain, S. & Bradfield, C. A. The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5474-5479 (1998).

69 Ripperger, J. A. & Schibler, U. Rhythmic CLOCK-BMAL1 binding to multiple E-box motifs drives circadian Dbp transcription and chromatin transitions. *Nature genetics* **38**, 369-374, doi:10.1038/ng1738 (2006).

70 Valekunja, U. K. *et al.* Histone methyltransferase MLL3 contributes to genome-scale circadian transcription. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1554-1559, doi:10.1073/pnas.1214168110 (2013).

71 Katada, S. & Sassone-Corsi, P. The histone methyltransferase MLL1 permits the oscillation of circadian gene expression. *Nature structural & molecular biology* **17**, 1414-1421, doi:10.1038/nsmb.1961 (2010).

72 Xu, H. *et al.* Cryptochrome 1 regulates the circadian clock through dynamic interactions with the BMAL1 C terminus. *Nature structural & molecular biology* **22**, 476-484, doi:10.1038/nsmb.3018 (2015).

73 Etchegaray, J. P., Lee, C., Wade, P. A. & Reppert, S. M. Rhythmic histone acetylation underlies transcription in the mammalian circadian clock. *Nature* **421**, 177-182, doi:10.1038/nature01314 (2003).

74 Curtis, A. M. *et al.* Histone acetyltransferase-dependent chromatin remodeling and the vascular clock. *The Journal of biological chemistry* **279**, 7091-7097, doi:10.1074/jbc.M311973200 (2004).

75 DiTacchio, L. *et al.* Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock. *Science* **333**, 1881-1885, doi:10.1126/science.1206022 (2011).

76 Lande-Diner, L., Boyault, C., Kim, J. Y. & Weitz, C. J. A positive feedback loop links circadian clock factor CLOCK-BMAL1 to the basic transcriptional machinery. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 16021-16026, doi:10.1073/pnas.1305980110 (2013).

77 Le Martelot, G. *et al.* Genome-wide RNA polymerase II profiles and RNA accumulation reveal kinetics of transcription and associated epigenetic changes during diurnal cycles. *PLoS biology* **10**, e1001442, doi:10.1371/journal.pbio.1001442 (2012).

78 Duong, H. A., Robles, M. S., Knutti, D. & Weitz, C. J. A molecular mechanism for circadian clock negative feedback. *Science* **332**, 1436-1439, doi:10.1126/science.1196766 (2011).

79 Duong, H. A. & Weitz, C. J. Temporal orchestration of repressive chromatin modifiers by circadian clock Period complexes. *Nature structural & molecular biology* **21**, 126-132, doi:10.1038/nsmb.2746 (2014).

80 Masri, S. *et al.* Partitioning circadian transcription by SIRT6 leads to segregated control of cellular metabolism. *Cell* **158**, 659-672, doi:10.1016/j.cell.2014.06.050 (2014).

81 Nakahata, Y. *et al.* The NAD+-dependent deacetylase SIRT1 modulates CLOCK-mediated chromatin remodeling and circadian control. *Cell* **134**, 329-340, doi:10.1016/j.cell.2008.07.002 (2008).

82 Asher, G. *et al.* SIRT1 regulates circadian clock gene expression through PER2 deacetylation. *Cell* **134**, 317-328, doi:10.1016/j.cell.2008.06.050 (2008).

83 Rey, G. *et al.* Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. *PLoS biology* **9**, e1000595, doi:10.1371/journal.pbio.1000595 (2011).

84 Yoshitane, H. *et al.* CLOCK-controlled polyphonic regulation of circadian rhythms through canonical and noncanonical E-boxes. *Molecular and cellular biology* **34**, 1776-1787, doi:10.1128/MCB.01465-13 (2014).

85 Kojima, S., Sher-Chen, E. L. & Green, C. B. Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression. *Genes & development* **26**, 2724-2736, doi:10.1101/gad.208306.112 (2012).

86 Kim, J. Y., Kwak, P. B. & Weitz, C. J. Specificity in Circadian Clock Feedback from Targeted Reconstitution of the NuRD Corepressor. *Molecular cell* **56**, 738-748, doi:10.1016/j.molcel.2014.10.017 (2014).

87 Ukai-Tadenuma, M. *et al.* Delay in feedback repression by cryptochrome 1 is required for circadian clock function. *Cell* **144**, 268-281, doi:10.1016/j.cell.2010.12.019 (2011).

88 DeBruyne, J. P. & Hogenesch, J. B. A CRY in the Night. *Developmental cell* **20**, 144-145, doi:10.1016/j.devcel.2011.01.014 (2011).

89 Cho, H. *et al.* Regulation of circadian behaviour and metabolism by REV-ERB-alpha and REV-ERB-beta. *Nature* **485**, 123-127, doi:10.1038/nature11048 (2012).

90 Fang, B. *et al.* Circadian enhancers coordinate multiple phases of rhythmic gene transcription in vivo. *Cell* **159**, 1140-1152, doi:10.1016/j.cell.2014.10.022 (2014).

91 Zhu, B. *et al.* Coactivator-Dependent Oscillation of Chromatin Accessibility Dictates Circadian Gene Amplitude via REV-ERB Loading. *Molecular cell* **60**, 769-783, doi:10.1016/j.molcel.2015.10.024 (2015).

92 Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21931-21936, doi:10.1073/pnas.1016071107 (2010).

93 Kim, T. K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182-187, doi:10.1038/nature09033 (2010).

94 Hayden, M. S. & Ghosh, S. NF-kappaB, the first quarter-century: remarkable progress and outstanding questions. *Genes & development* **26**, 203-234, doi:10.1101/gad.183434.111 (2012).

95 Smale, S. T. Hierarchies of NF-kappaB target-gene regulation. *Nature immunology* **12**, 689-694, doi:10.1038/ni.2070 (2011).

96 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

97 Ong, C. T. & Corces, V. G. CTCF: an architectural protein bridging genome topology and function. *Nature reviews. Genetics* **15**, 234-246, doi:10.1038/nrg3663 (2014).

98 Vietri Rudan, M. *et al.* Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell reports* **10**, 1297-1309, doi:10.1016/j.celrep.2015.02.004 (2015).

99 Faure, A. J. *et al.* Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome research* **22**, 2163-2175, doi:10.1101/gr.136507.111 (2012).

100 Lim, H. W. *et al.* Genomic redistribution of GR monomers and dimers mediates transcriptional response to exogenous glucocorticoid in vivo. *Genome research* **25**, 836-844, doi:10.1101/gr.188581.114 (2015).

101 Boergesen, M. *et al.* Genome-wide profiling of liver X receptor, retinoid X receptor, and peroxisome proliferator-activated receptor alpha in mouse liver reveals extensive sharing of binding sites. *Mol Cell Biol* **32**, 852-867, doi:10.1128/MCB.06175-11 (2012).

102 MacIsaac, K. D. *et al.* A quantitative model of transcriptional regulation reveals the influence of binding location on expression. *PLoS computational biology* **6**, e1000773, doi:10.1371/journal.pcbi.1000773 (2010).

103 Gordon, F. K. *et al.* Research resource: Aorta- and liver-specific ERalpha-binding patterns and gene regulation by estrogen. *Mol Endocrinol* **28**, 1337-1351, doi:10.1210/me.2013-1395 (2014).

104 Zhang, Y., Laz, E. V. & Waxman, D. J. Dynamic, sex-differential STAT5 and BCL6 binding to sex-biased, growth hormone-regulated genes in adult mouse liver. *Mol Cell Biol* **32**, 880-896, doi:10.1128/MCB.06312-11 (2012).

105 Bunger, M. K. *et al.* Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* **103**, 1009-1017 (2000).

106 Atger, F. *et al.* Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E6579-6588, doi:10.1073/pnas.1515308112 (2015).

107 Vollmers, C. *et al.* Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 21453-21458, doi:10.1073/pnas.0909591106 (2009).

108 Goldstein, I. & Hager, G. L. Transcriptional and Chromatin Regulation during Fasting - The Genomic Era. *Trends in endocrinology and metabolism: TEM* **26**, 699-710, doi:10.1016/j.tem.2015.09.005 (2015).

109 Desvergne, B., Michalik, L. & Wahli, W. Transcriptional regulation of metabolism. *Physiological reviews* **86**, 465-514, doi:10.1152/physrev.00025.2005 (2006).

110 Eckel-Mahan, K. L. *et al.* Reprogramming of the circadian clock by nutritional challenge. *Cell* **155**, 1464-1478, doi:10.1016/j.cell.2013.11.034 (2013).

111 Froehlich, A. C., Liu, Y., Loros, J. J. & Dunlap, J. C. White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter. *Science* **297**, 815-819, doi:10.1126/science.1073681 (2002).

112 Froehlich, A. C., Loros, J. J. & Dunlap, J. C. Rhythmic binding of a WHITE COLLAR-containing complex to the frequency promoter is inhibited by FREQUENCY. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5914-5919, doi:10.1073/pnas.1030057100 (2003).

113 Allada, R., White, N. E., So, W. V., Hall, J. C. & Rosbash, M. A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless. *Cell* **93**, 791-804 (1998).

114 Hurley, J. M. *et al.* Analysis of clock-regulated genes in Neurospora reveals widespread posttranscriptional control of metabolic potential. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 16995-17002, doi:10.1073/pnas.1418963111 (2014).

115 Abruzzi, K. C. *et al.* Drosophila CLOCK target gene characterization: implications for circadian tissue-specific gene expression. *Genes & development* **25**, 2374-2386, doi:10.1101/gad.178079.111 (2011).

116 Cha, J., Zhou, M. & Liu, Y. CATP is a critical component of the Neurospora circadian clock by regulating the nucleosome occupancy rhythm at the frequency locus. *EMBO reports* **14**, 923-930, doi:10.1038/embor.2013.131 (2013).

117 Wang, B., Kettenbach, A. N., Gerber, S. A., Loros, J. J. & Dunlap, J. C. Neurospora WC-1 recruits SWI/SNF to remodel frequency and initiate a circadian cycle. *PLoS genetics* **10**, e1004599, doi:10.1371/journal.pgen.1004599 (2014).

118 Sancar, C. *et al.* Combinatorial control of light induced chromatin remodeling and gene activation in Neurospora. *PLoS genetics* **11**, e1005105, doi:10.1371/journal.pgen.1005105 (2015).

119 Kwok, R. S., Li, Y. H., Lei, A. J., Edery, I. & Chiu, J. C. The Catalytic and Non-catalytic Functions of the Brahma Chromatin-Remodeling Protein Collaborate to Fine-Tune Circadian Transcription in Drosophila. *PLoS genetics* **11**, e1005307, doi:10.1371/journal.pgen.1005307 (2015).

120 Kwok, R. S., Lam, V. H. & Chiu, J. C. Understanding the role of chromatin remodeling in the regulation of circadian transcription in Drosophila. *Fly (Austin)* **9**, 145-154, doi:10.1080/19336934.2016.1143993 (2015).

121 MacQuarrie, K. L., Fong, A. P., Morse, R. H. & Tapscott, S. J. Genome-wide transcription factor binding: beyond direct target regulation. *Trends in genetics : TIG* **27**, 141-148, doi:10.1016/j.tig.2011.01.001 (2011).

122 Banks, C. J., Joshi, A. & Michoel, T. Functional transcription factor target discovery via compendia of binding and expression profiles. *Scientific reports* **6**, 20649, doi:10.1038/srep20649 (2016).

123 Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Molecular cell* **49**, 825-837, doi:10.1016/j.molcel.2013.01.038 (2013).

124 Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279-283, doi:10.1038/nature09692 (2011).

125 Glass, C. K. & Natoli, G. Molecular control of activation and priming in macrophages. *Nature immunology* **17**, 26-33, doi:10.1038/ni.3306 (2016).

126 Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Nature reviews. Molecular cell biology* **16**, 144-154, doi:10.1038/nrm3949 (2015).

127 Grontved, L. *et al.* C/EBP maintains chromatin accessibility in liver and facilitates glucocorticoid receptor recruitment to steroid response elements. *The EMBO journal* **32**, 1568-1583, doi:10.1038/emboj.2013.106 (2013).

128 Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D. & Carroll, J. S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nature genetics* **43**, 27-33, doi:10.1038/ng.730 (2011).

129 Perelis, M. *et al.* Pancreatic beta cell enhancers regulate rhythmic transcription of genes controlling insulin secretion. *Science* **350**, aac4250, doi:10.1126/science.aac4250 (2015).

130 Mirny, L. A. Nucleosome-mediated cooperativity between transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 22534-22539, doi:10.1073/pnas.0913805107 (2010).

131 Miller, J. A. & Widom, J. Collaborative competition mechanism for gene activation in vivo. *Molecular and cellular biology* **23**, 1623-1632 (2003).

132 Moyle-Heyrman, G., Tims, H. S. & Widom, J. Structural constraints in collaborative competition of transcription factors against the nucleosome. *Journal of molecular biology* **412**, 634-646, doi:10.1016/j.jmb.2011.07.032 (2011).

133 Polach, K. J. & Widom, J. A model for the cooperative binding of eukaryotic regulatory proteins to nucleosomal target sites. *Journal of molecular biology* **258**, 800-812, doi:10.1006/jmbi.1996.0288 (1996).

134 Vashee, S., Melcher, K., Ding, W. V., Johnston, S. A. & Kodadek, T. Evidence for two modes of cooperative DNA binding in vivo that do not involve direct protein-protein interactions. *Current biology : CB* **8**, 452-458 (1998).

135 Adams, C. C. & Workman, J. L. Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Molecular and cellular biology* **15**, 1405-1421 (1995).

136 Nguyen, K. D. *et al.* Circadian gene Bmal1 regulates diurnal oscillations of Ly6C(hi) inflammatory monocytes. *Science* **341**, 1483-1488, doi:10.1126/science.1240636 (2013).

137 Xu, H. *et al.* Myeloid cell-specific disruption of Period1 and Period2 exacerbates diet-induced inflammation and insulin resistance. *The Journal of biological chemistry* **289**, 16374-16388, doi:10.1074/jbc.M113.539601 (2014).

138 Gibbs, J. *et al.* An epithelial circadian clock controls pulmonary inflammation and glucocorticoid action. *Nature medicine* **20**, 919-926, doi:10.1038/nm.3599 (2014).

139 Bellet, M. M. *et al.* Circadian clock regulates the host response to Salmonella. *Proc Natl Acad Sci U S A* **110**, 9897-9902, doi:10.1073/pnas.1120636110 (2013).

140 Aguilar-Arnal, L. *et al.* Cycles in spatial and temporal chromosomal organization driven by the circadian clock. *Nature structural & molecular biology* **20**, 1206-1213, doi:10.1038/nsmb.2667 (2013).

141 Lelli, K. M., Slattery, M. & Mann, R. S. Disentangling the many layers of eukaryotic transcriptional regulation. *Annual review of genetics* **46**, 43-68, doi:10.1146/annurev-genet-110711-155437 (2012).

142 de Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499-506, doi:10.1038/nature12753 (2013).

143 Eckel-Mahan, K. & Sassone-Corsi, P. Metabolism and the circadian clock converge. *Physiological reviews* **93**, 107-135, doi:10.1152/physrev.00016.2012 (2013).

144 Murakami, M. *et al.* Gut microbiota directs PPARgamma-driven reprogramming of the liver circadian clock by nutritional challenge. *EMBO reports* **17**, 1292-1303, doi:10.15252/embr.201642463 (2016).

145 Haspel, J. A. *et al.* Circadian rhythm reprogramming during lung inflammation. *Nature communications* **5**, 4753, doi:10.1038/ncomms5753 (2014).

146 Keegan, K. P., Pradhan, S., Wang, J. P. & Allada, R. Meta-analysis of Drosophila circadian microarray studies identifies a novel set of rhythmically expressed genes. *PLoS computational biology* **3**, e208, doi:10.1371/journal.pcbi.0030208 (2007).

147 Covington, M. F., Maloof, J. N., Straume, M., Kay, S. A. & Harmer, S. L. Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome biology* **9**, R130, doi:10.1186/gb-2008-9-8-r130 (2008).

148 Doherty, C. J. & Kay, S. A. Circadian control of global gene expression patterns. *Annual review of genetics* **44**, 419-444, doi:10.1146/annurev-genet-102209-163432 (2010).

149 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).

150 Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)* **25**, 1105-1111, doi:10.1093/bioinformatics/btp120 (2009).

151 Ling, G., Sugathan, A., Mazor, T., Fraenkel, E. & Waxman, D. J. Unbiased, genome-wide in vivo mapping of transcriptional regulatory elements reveals sex differences in chromatin structure associated with sex-specific liver gene expression. *Mol Cell Biol* **30**, 5531-5544, doi:10.1128/MCB.00601-10 (2010).

152 Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75-82, doi:10.1038/nature11232 (2012).

153 Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).

154 Bernardo, B. C., Weeks, K. L., Pretorius, L. & McMullen, J. R. Molecular distinction between physiological and pathological cardiac hypertrophy: Experimental findings and therapeutic strategies. *Pharmacology & Therapeutics* **128**, 191-227, doi:https://doi.org/10.1016/j.pharmthera.2010.04.005 (2010).

155 Guo, Y. *et al.* Entanglement of GSK-3beta, beta-catenin and TGF-beta1 signaling network to regulate myocardial fibrosis. (2017).

156 Liu, G., Ma, C., Yang, H. & Zhang, P. Y. Transforming growth factor beta and its role in heart disease. (2017).

157 Verrecchia, F. & Mauviel, A. Transforming growth factor-beta and fibrosis. (2007).

158 Pan, X., Chen Z Fau - Huang, R., Huang R Fau - Yao, Y., Yao Y Fau - Ma, G. & Ma, G. Transforming growth factor beta1 induces the expression of collagen type I by DNA methylation in cardiac fibroblasts. (2013).

159 de Jong, S., van Veen, T. A. B., de Bakker, J. M. T. & van Rijen, H. V. M. Monitoring cardiac fibrosis: a technical challenge. *Netherlands Heart Journal* **20**, 44-48, doi:10.1007/s12471-011-0226-x (2012).

160 Hadi, A. M. *et al.* Rapid quantification of myocardial fibrosis: a new macro-based automated analysis. *Cellular Oncology (Dordrecht)* **34**, 343-354, doi:10.1007/s13402-011-0035-7 (2011).

161 Vogel, B., Siebert, H., Hofmann, U. & Frantz, S. Determination of collagen content within picrosirius red stained paraffin-embedded tissue sections using fluorescence microscopy. *MethodsX* **2**, 124-134, doi:10.1016/j.mex.2015.02.007 (2015).

162 Kim D Fau - Pertea, G. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. (2013).

163 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. (2012).

164 Mortazavi, A., Williams Ba Fau - McCue, K., McCue K Fau - Schaeffer, L., Schaeffer L Fau - Wold, B. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. (2008).

165 Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research* **45**, D353-D361, doi:10.1093/nar/gkw1092 (2017).

166 Song, H. K., Hong, S.-E., Kim, T. & Kim, D. H. Deep RNA Sequencing Reveals Novel Cardiac Transcriptomic Signatures for Physiological and Pathological Hypertrophy. *PLOS ONE* **7**, e35552, doi:10.1371/journal.pone.0035552 (2012).

167 Yu-Wai-Man, C. *et al.* Genome-wide RNA-Sequencing analysis identifies a distinct fibrosis gene signature in the conjunctiva after glaucoma surgery. *Scientific Reports* **7**, 5644, doi:10.1038/s41598-017-05780-5 (2017).

168 Xu, Y., Li X Fau - Liu, X., Liu X Fau - Zhou, M. & Zhou, M. Neuregulin-1/ErbB signaling and chronic heart failure. (2010).

169 Bonnans, C., Chou, J. & Werb, Z. Remodelling the extracellular matrix in development and disease. *Nature reviews. Molecular cell biology* **15**, 786-801, doi:10.1038/nrm3904 (2014).

170 Stamatoyannopoulos, J. A. *et al.* An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome Biology* **13**, 418, doi:10.1186/gb-2012-13-8-418 (2012).

171 Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**, 178-192, doi:10.1093/bib/bbs017 (2013).

172 Robinson, J. T. *et al.* Integrative genomics viewer. *Nature Biotechnology* **29**, 24, doi:10.1038/nbt.1754

https://www.nature.com/articles/nbt.1754#supplementary-information (2011).

173 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. (2010).

174 Sato, M. *et al.* Identification of transcription factor E3 (TFE3) as a receptor-independent activator of Galpha16: gene regulation by nuclear Galpha subunit and its activator. *The*

*Journal of biological chemistry* **286**, 17766-17776, doi:10.1074/jbc.M111.219816 (2011).

175 Fujita, J. *et al.* Ronin Governs Early Heart Development by Controlling Core Gene Expression Programs. *Cell reports* **21**, 1562-1573, doi:10.1016/j.celrep.2017.10.036 (2017).

176 Nakerakanti, S. S., Kapanadze, B., Yamasaki, M., Markiewicz, M. & Trojanowska, M. Fli1 and Ets1 have distinct roles in connective tissue growth factor/CCN2 gene regulation and induction of the profibrotic gene program. *The Journal of biological chemistry* **281**, 25259-25269, doi:10.1074/jbc.M600466200 (2006).

177 Chin, M. T. KLF15 and cardiac fibrosis: the heart thickens. *J Mol Cell Cardiol* **45**, 165-167, doi:10.1016/j.yjmcc.2008.05.022 (2008).

178 Rommel, C. *et al.* The Transcription Factor ETV1 Induces Atrial Remodeling and Arrhythmia. *Circ Res*, doi:10.1161/circresaha.118.313036 (2018).

179 Asbun, J. & Villarreal, F. J. The Pathogenesis of Myocardial Fibrosis in the Setting of Diabetic Cardiomyopathy. *Journal of the American College of Cardiology* **47**, 693-700, doi:https://doi.org/10.1016/j.jacc.2005.09.050 (2006).

180 Knutsson, A. Health disorders of shift workers.

181 Salgado-Delgado, R., Angeles-Castellanos, M., Saderi, N., Buijs, R. M. & Escobar, C. Food Intake during the Normal Activity Phase Prevents Obesity and Circadian Desynchrony in a Rat Model of Night Work. *Endocrinology* **151**, 1019-1029, doi:10.1210/en.2009-0864 (2010).

182 Brandenburger, Y. *et al.* Cardiac hypertrophy in vivo is associated with increased expression of the ribosomal gene transcription factor UBF. *FEBS Letters* **548**, 79-84, doi:https://doi.org/10.1016/S0014-5793(03)00744-0 (2003).

183 Siehl D Fau - Chua, B. H., Chua Bh Fau - Lautensack-Belser, N., Lautensack-Belser N Fau - Morgan, H. E. & Morgan, H. E. Faster protein and ribosome synthesis in thyroxine-induced hypertrophy of rat heart. (1985).

184 Gurtl, B. *et al.* Apoptosis and fibrosis are early features of heart failure in an animal model of metabolic cardiomyopathy. (2009).

185 Gallo, S. *et al.* Agonist antibodies activating the Met receptor protect cardiomyoblasts from cobalt chloride-induced apoptosis and autophagy. *Cell Death Dis* **5**, e1185, doi:10.1038/cddis.2014.155 (2014).

186 Shindo, T. *et al.* Kruppel-like zinc-finger transcription factor KLF5/BTEB2 is a target for angiotensin II signaling and an essential regulator of cardiovascular remodeling. *Nature medicine* **8**, 856-863, doi:10.1038/nm738 (2002).

187 Nag, A. C. Study of non-muscle cells of the adult mammalian heart: a fine structural analysis and distribution. *Cytobios* **28**, 41-61 (1980).

188 Hort, W. Growth of the heart in health and disease edited by Radovan Zak Raven Press, New York (1984) 480 pages, illustrated, $79.00 ISBN: 0-89004-734-0. *Clinical Cardiology* **8**, 186-186, doi:doi:10.1002/clc.4960080313 (1985).

189 Schneiders, D., Heger J Fau - Best, P., Best P Fau - Michael Piper, H., Michael Piper H Fau - Taimor, G. & Taimor, G. SMAD proteins are involved in apoptosis induction in ventricular cardiomyocytes. (2005).

190 Salgado-Delgado, R., Angeles-Castellanos, M., Buijs, M. R. & Escobar, C. Internal desynchronization in a model of night-work by forced activity in rats. *Neuroscience* **154**, 922-931, doi:10.1016/j.neuroscience.2008.03.066 (2008).

191 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. (2012).

192 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

193 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

194 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).

195 Mure, L. S. *et al.* Diurnal transcriptome atlas of a primate across major neural and peripheral tissues. *Science* **359**, doi:10.1126/science.aao0318 (2018).

196 Menet, J. S. *et al.* Tissue-specific BMAL1 cistromes reveal that enhancer-enhancer interactions regulate rhythmic transcription. *bioRxiv* (2018).

197 Lande-Diner, L., Boyault, C., Kim, J. Y. & Weitz, C. J. A positive feedback loop links circadian clock factor CLOCK-BMAL1 to the basic transcriptional machinery. *Proceedings of the National Academy of Sciences* **110**, 16021-16026 (2013).

198 Kim, Y. H. *et al.* Rev-erbalpha dynamically modulates chromatin looping to control circadian gene transcription. *Science* **359**, 1274-1277, doi:10.1126/science.aao6891 (2018).

199 Takahashi, J. S. Transcriptional architecture of the mammalian circadian clock. *Nature reviews. Genetics* **18**, 164-179, doi:10.1038/nrg.2016.150 (2017).

200 Menet, J. S. & Rosbash, M. When brain clocks lose track of time: cause or consequence of neuropsychiatric disorders. *Current opinion in neurobiology*, doi:10.1016/j.conb.2011.06.008 (2011).

201 Lowrey, P. L. & Takahashi, J. S. Mammalian circadian biology: elucidating genome-wide levels of temporal organization. *Annual review of genomics and human genetics* **5**, 407-441, doi:10.1146/annurev.genom.5.061903.175925 (2004).

202 Karlsson, B., Alfredsson, L., Knutsson, A., Andersson, E. & Toren, K. Total mortality and cause-specific mortality of Swedish shift- and dayworkers in the pulp and paper industry in 1952-2001. *Scandinavian journal of work, environment & health* **31**, 30-35 (2005).

203 Takahashi, J. S., Hong, H. K., Ko, C. H. & McDearmon, E. L. The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nature reviews. Genetics* **9**, 764-775, doi:10.1038/nrg2430 (2008).

204 Salgado-Delgado, R., Angeles-Castellanos, M., Saderi, N., Buijs, R. M. & Escobar, C. Food intake during the normal activity phase prevents obesity and circadian desynchrony in a rat model of night work. *Endocrinology* **151**, 1019-1029, doi:10.1210/en.2009-0864 (2010).

205 Hamdani, N. & Linke, W. A. Alteration of the beta-adrenergic signaling pathway in human heart failure. *Current pharmaceutical biotechnology* **13**, 2522-2531 (2012).

206 Liu, L., Zhao, M., Yu, X. & Zang, W. Pharmacological Modulation of Vagal Nerve Activity in Cardiovascular Diseases. *Neuroscience bulletin*, doi:10.1007/s12264-018-0286-7 (2018).