REGULATION OF RHYTHMIC GENE EXPRESSION BY RHYTHMIC FOOD INTAKE AND ALTERNATIVE POLYADENYLATION

A Dissertation

by

BEN JEFFERSON GREENWELL

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,	Deborah Bell-Pedersen
	David W. Threadgill
	Jun-yuan Ji
Head of Department,	David W. Threadgill

May 2020

Major Subject: Genetics

ABSTRACT

Nearly every mammalian cell exhibits daily rhythms in gene expression, which guide the activation of tissue-specific processes across the day. Rhythms in peripheral tissues are synchronized by neuronal and hormonal rhythmic signals as well as rhythms of body temperature and food intake, all of which originate from the master circadian clock located in the suprachiasmatic nucleus (SCN) of the hypothalamus. However, how these rhythmic signals contribute to the oscillations of gene expression and biological functions remains unknown. Increasing evidence suggests that systemic signals, and more specifically rhythmic food intake (RFI), can regulate rhythmic gene expression independently of the circadian clock. Additionally, experiments have shown that anywhere from 15-30% of the mouse hepatic transcriptome is rhythmic. However, these experiments only look at expression on an overall gene level, and so the importance of rhythmic gene expression may be understated as certain isoforms of arrhythmic genes may be rhythmic.

To determine the relative contribution of cell autonomous clocks versus RFI in the regulation of rhythmic gene expression, we developed a system that allows long-term manipulation of the daily rhythm of food intake in the mouse, and analyzed liver gene expression by RNA-Seq in mice fed *ad libitum*, only at night, or arrhythmically (mouse eating 1/8th of their daily food intake every 3 hours). We show that 70% of the cycling mouse liver transcriptome loses rhythmicity under

arrhythmic feeding. Remarkably, this loss of rhythmic gene expression under arrhythmic feeding is independent of the liver circadian clock, which continues to exhibit normal oscillations in core clock gene expression. Together, these results demonstrate that systemic signals driven by rhythmic food intake play a more important role than the cell-autonomous circadian clock in driving rhythms in liver gene expression.

Next, to determine if alternative polyadenylation results in differential rhythmic expression of transcripts, we performed 3'-end mRNA-Seq and found that 15% of mouse hepatic transcriptome with more than one PAS exhibits differential rhythmic expression. Importantly, the major cause for this differential expression was found to be strongly related to co-transcriptional processes *vs*. transcriptional or post-transcriptional, and indicates that gene isoforms are independently regulated.

DEDICATION

To my friends and family, who have made life worth continuing.

ACKNOWLEDGEMENTS

A great many people have helped and supported me throughout my journey through graduate school. Without the support, challenges, praise, discussion, and many, many other interactions I had, I would be a far different person today. Some days I only got through due to previous praise urging me forward.

First and foremost, I would like to thank my PI, Dr. Jerome Menet. He is a brilliant researcher who will go far as a future PI. We didn't always see eye-to-eye, but he was incredibly consistent in only wanting the best for his students. When I first entered the lab, I had no idea how to respond to questions he would ask me or how he could jump to conclusions so quickly. As I progressed, his tutelage molded me as a researcher, showing me the right questions to ask and answer. Here, near the end of my time in his lab, I'm finally able to hold long conversations with him about my research, what my results mean, and possible avenues for the future. I still have no idea how he jumps to conclusions so quickly, though.

Next, I would like to thank my labmates, who were the largest factor in keeping me going. Alexandra Trott, who I could talk to at any time about research, life, or anything in between, and was beyond a phenomenal motivator. Joshua Beytebiere, whose camaraderie kept me laughing on my worst days and focused on my best. Aishwarya Sahasrabudhe, or Ash, who joined our lab as a very lost first-year and even within the first year has gained so much knowledge as Josh and I (to a lesser extent) taught her how to survive graduate school. The many undergrads I

worked closely with: Patrick Finegan, Shanny Pao, Alexander Bosley, and Erin Beach, who helped me learn how to mentor students and guide their progress through experiments for my projects. My close friends from my cohort in fellow circadian labs, Jennifer Jung and Samantha Iiams, to whom I could always confide in and share everything, be it exciting or disappointing. Furthermore, the many, many friends I have made around College Station/Bryan, who have kept me busy and given me so many opportunities to grow as a person.

I would like to thank my committee members, Drs. Deborah Bell-Pedersen, David W. Threadgill, and Jun-yuan Ji, for their helpful guidance and encouragement throughout the entire process.

Next I would like to thank the journal club I attended year-round, that gave me the much-needed experience in both presenting both my work and that of others frequently, and in reading papers in-depth to answer any questions the audience may have.

Thanks also go to my undergraduate PI, Rodolfo Aramayo, who was the progenitor of my decision to enter graduate school. Without what I learned from taking his classes, working in his lab, talking to him, and his support and pressure to push myself harder, I would've never thought of joining graduate school in the first place.

Finally, I would like to thank my family and pets, who provided me with support throughout every step I've taken in graduate school.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

Research done in this dissertation was supervised by Dr. Jerome Menet (chair) from the Program of Genetics and Department of Biology, as well Dr. Deborah Bell-Pedersen from the Program of Genetics and Department of Biology, and Dr. David W. Threadgill and Dr. Jun-yuan Ji from the Program of Genetics and Department of Molecular and Cellular Medicine at Texas A&M University.

The research performed in Chapter II was conceived of and designed by Dr. Jerome Menet and myself. I performed all bioinformatics work. I, Dr. Jerome Menet, Dr. Christine Merlin, Joshua Beytebiere, and Aldrin Lugena built the special cages required for rhythmic feeding. Alexandra Trott, Joshua Beytebiere, Dr. Jerome Menet, Shanny Pao, Alexander Bosley, Christopher Hernandez, Amin Heravi, Shaikh Islam, Shawn Alex, and Patrick Finegan assisted me with feeding mice daily. Patrick Finegan and Christopher Hernandez assisted with RNA extraction and purification. I created all RNA-Seq libraries. Shanny Pao, Alexander Bosley, and Erin Beach performed the majority of the Western Blot experiments. Dr. Chaodong Wu, Dr. Honggui Li, and Dr. David Threadgill helped with physiological assays.

The work performed in Chapter III was conceived of and designed by myself and Dr. Jerome Menet. Teresa Lamb and Dr. Deborah Bell-Pedersen contributed and performed all of the polysomal fractionation of mouse liver. Joshua Beytebiere performed the RNA isolation and purification of all samples, i.e., total RNA, nuclear

RNA, and polysomal RNA. I created all RNA-Seq libraries, performed all the bioinformatics work, and data analysis.

Funding Sources

Funding was provided by Department of Biology Start Up funds provided to Dr. Jerome Menet. No other source of funding was used in conduction of my research.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

The circadian clock

Circadian clocks have been observed throughout a broad range of organisms including insects, plants, fungi, and cyanobacteria[1]. Nearly all of the circadian clocks through the different phyla contain a core set of genes that act in a similar manner in a transcriptional / translational feedback loop, where a positive element activates transcription of the negative elements that eventually turn off the positive element and close the loop. The positive elements also activate the expression of clock-controlled genes, resulting in circadian expression of those genes. They govern a wide variety of biological processes that result in peaks and troughs in patterns of behavior, metabolism, physiology, gene processing, and so on. In most circumstances, the circadian clock is entrained to the daily 24-hour rhythm in light exposure experienced on Earth. In order to standardize timekeeping measurements across phyla, the concept of Zeitgeber Time (ZT) arose, where (for example) in a system with 12 hours of light followed by 12 hours of dark (LD), ZT0 would indicate time at the start of lights-on and ZT12 would indicate the start of lights-off. In constant darkness (DD), the timekeeping mechanisms continue in a free-running manner in its last entrained phase, and time is indicated in Circadian Time (CT).

The mammalian circadian clock

Nearly every cell in mammals contains the molecular components of the circadian clock. These cells are synchronized by signals coming from the master circadian clock located in the suprachiasmatic nucleus (SCN) in the hypothalamus region of the brain, which is entrained primarily by light:dark cycles received through the optic nerve[1-3]. The SCN then regulates multiple pathways and processes, such as body temperature, hormones, feeding behavior, and autonomic nervous signaling in order to entrain peripheral clocks in other tissues, ensuring that rhythmic gene and protein expression is synchronized across all tissues throughout the mammalian body[2,4]. These specific pathways and events that interact and result in setting the phase of the circadian clock in every tissue are known as Zeitgebers (German: time-givers).

On a molecular level, the circadian clock is composed of multiple proteins that form a transcriptional-translational negative feedback loop[5]. The positive arm in mammals is composed of the two genes *Clock* and *Bmal1*, whose proteins heterodimerize to form CLOCK:BMAL1, a bHLH transcription factor[6]. CLOCK:BMAL1 binds at canonical e-boxes composed of the nucleotide motif CACGTG during the day in order to help activate transcription of its target genes[7] (Figure I-1). Two such target genes are the *Period* family (*Per1, Per2, Per3)* and the *Cryptochrome* family (*Cry1, Cry2*). After the *Per* and *Cry* mRNA is exported outside the nucleus and translated, PER and CRY proteins form a heterodimer that translocates back inside the nucleus during the end of the day, bind to

CLOCK:BMAL1 and repress its ability to bind DNA, resulting in removal of CLOCK:BMAL1 from the DNA[8]. As more and more of CLOCK:BMAL1 is bound by the PER:CRY complex, eventually all daytime-phased transcription is turned off including the transcription of *Per* and *Cry*, which are slowly degraded. By the end of the night, most of the PER:CRY complexes have been degraded, leading to CLOCK:BMAL1 once again being able to bind DNA and re-initiating daytime-phased transcription.

Two other transcription factors, the *Ror* family (*Rorα*, *Rorβ*, *Rorγ*) and *Rev-erb* family (*Rev-erbα, Rev-erbβ*) are also part of the molecular circadian clock, and act as transcriptional activators and repressors, respectively. They both bind to ROR/REV-ERB response elements (ROREs) near their own target genes, which include *Bmal1* itself as well as many CLOCK:BMAL1 target genes, in order to modulate gene expression (Figure I-1). Double knockout (DKO) of both *Rev-erb*s strongly affects both circadian gene expression as well as lipid homeostasis[9].

Much of the control of the circadian clock is exerted at the transcriptional level, as the core clock TF CLOCK:BMAL1 as well as the secondary TFs RORα and REV-ERBα rhythmically activate transcription of clock-controlled genes through both direct and indirect methods. Activation of biochemical pathways at certain times of the day through transcription of the corresponding genes separates opposing processes such that cellular energy is utilized in an efficient manner[10].

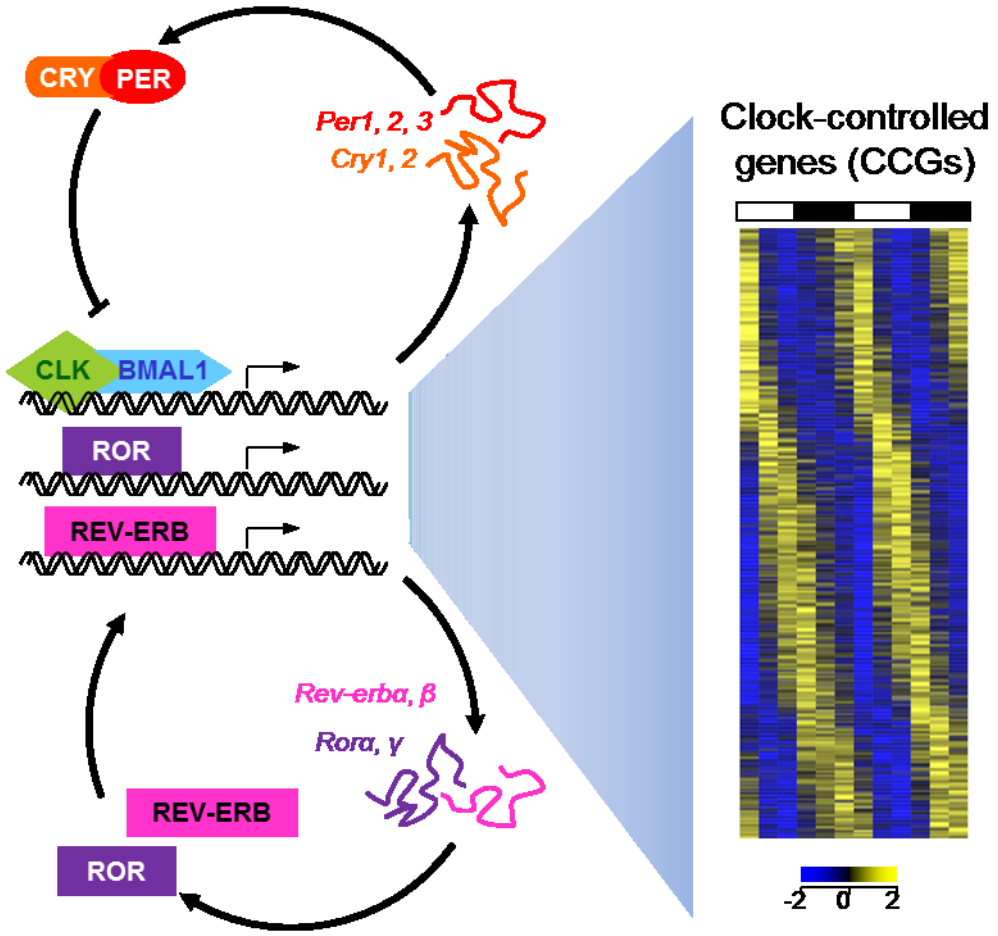

Figure I-1 The core mammalian clock genes

Regulation of gene expression and alternative polyadenylation in eukaryotes

In eukaryotes, transcription begins with the enzyme RNA Polymerase II binding to a promoter and starting transcribing at the transcription start site (TSS), which continues through the end of the gene and past the transcription termination site (TTS) until it is removed, releasing the nascent RNA transcript[11]. Cleavage and polyadenylation specificity factors (CPSF) loaded onto the C-terminal domain (CTD)

of Pol II detect the motif AAUAAA in the 3' UTR and cleave off any extraneous nucleotides approximately 15-35nt downstream at the polyadenylation site (PAS), which sometimes overlaps with the TTS[12,13]. Multiple potential PAS may be present in a gene, leading to the phenomenon known as alternative polyadenylation (APA), detailed below[14,15]. The enzyme poly(A) polymerase, part of the cleave/polyadenylation complex, then begins to synthesize a poly(A) tail onto the end of the transcript[16]. Finally, the 5' guanosine cap is added, introns are spliced out, and the now-mature mRNA is exported outside the nucleus in order to be translated or degraded[17].

Initiation of transcription is a complicated process involving many different proteins and several DNA elements, including promoters and enhancers. Enhancers are short DNA elements of approximately 50-1500 base pairs[18] that increase the likelihood of transcription occurring of their target genes. They serve as binding locations for at least 98.5% of all transcription factors[19], and so are major targets to study in order to understand transcription initiation. During the early steps of transcription initiation, the Mediator complex binds the gene promoter, growing PIC complex, and enhancers related to the target gene, bringing them into close proximity and allowing the enhancers to direct transcription. However, it is not currently well understood exactly how enhancers instruct Pol II through the Mediator complex, or potentially other genes or complexes, that results in changes in the transcription process[20].

Most mammalian genes harbor multiple polyadenylation sites, whose alternative usage increases the number of possible transcript isoforms[15,21,22]. The position of these alternative polyadenylation sites (APAS) can have large effects on 3' UTR length and/or protein sequence, leading to transcripts being truncated or containing additional exonic or intronic sequences[14,23]. The length of the 3' UTR tail has implications on stability, as it is targeted by RNA binding proteins (RBP), miRNA, and lncRNA in order to be regulated[24-27]. Moreover, recent evidence proposes that differences in the RNA transcript 3' UTR lengths can lead to profound differences in the resulting protein localization[28]. APA has also been shown to regulate processes like pluripotency[29,30], and defects in APAS can lead to a multitude of health effects including cancer[31-34], indicating that is it not an insignificant phenomenon.

The process by which APAS transcripts are chosen to be expressed both within a single tissue, as well as between different tissues, is yet to be well understood. Certain genes may choose to always express the longer isoform in one tissue, and other genes the shorter isoform, with few observable reasons as to why[35]. One emerging model on the tissue-specificity front involves using differing amounts of polyadenylation proteins during transcription, however this has yet to be shown experimentally[36]. Previous studies have suggested that APAS isoforms within tissues are co-regulated, where an increase in proximal PAS usage will lead to an equivalent decrease in distal PAS usage[33,37]. Furthermore, the binding of 3' UTRs by RBPs plays a large role in how a transcript is expressed[26,30,38,39]. Overall,

the process by which genes are individually regulated is an extensive, multifactorial issue that remains poorly understood. A number of different mechanisms mentioned here interact in order to activate transcription, all of which contain the potential for modifications leading to alternative regulation.

Circadian and tissue-specific expression

Rhythmic gene expression is a hallmark of circadian rhythms, and is found in any organisms having a circadian clock. In mice, genes are rhythmically expressed across a 24-hour period in almost every tissue, with the liver containing the highest number of rhythmically expressed genes (REG)[40]. Estimates for the number of REG in the liver under *ad libitum* conditions range from 10-20% of the transcriptome, with parts of the brain being the least rhythmic. Interestingly, this trend appears to reverse for primates, where the liver has very few REG, and the brain systems are among the most rhythmic[41,42]. While approximately 50% of the transcriptome is rhythmically expressed in at least one tissue, most genes are only rhythmically expressed in only a few (or just one) tissue.

The molecular circadian clock has multiple transcription factors (TFs) that have both unique and overlapping target genes[9,43,44]. Additionally, these TFs all have different peak and trough binding times, leading to modulation of gene expression in a circadian manner. The prevailing model in the field has been that once entrained, each peripheral clock then regulates its own tissue-specific rhythmic gene expression in a cell-autonomous manner; for example, the hepatic

liver clock should be driving most if not all of the rhythmic gene expression in the liver. However, recent papers are starting to challenge this model. One study discovered that, despite CLOCK:BMAL1 always binding to all of its target genes in the same small window of time every day, its target genes have a large amount of heterogeneity in their rhythmic expression[45]. The same circadian machinery leads to differing outputs within the liver, indicating that the clock in its transcriptional capacity cannot be the sole determinant of rhythmic gene expression[46].

Tissues may express large numbers of tissue-specific genes that are responsible for the unique characteristics and processes of each tissue[47]. In addition, not only does each tissue have shared *vs.* uniquely expressed genes, it also has shared *vs.* uniquely rhythmically expressed genes. A gene that is not expressed in one tissue may be expressed and arrhythmic in another tissue, and expressed and rhythmic in yet another tissue. As the molecular circadian components are the same between tissues[8], how the clock drives rhythmic gene expression differently between tissues is not yet well understood. However, it appears that chromatin accessibility and enhancer interactions appear to play a large role[48]. Thus, recent studies have made an effort to distinguish how the same molecular clock mechanism can result in both differential circadian expression within tissues[45] as well as between tissues[48] and have found a large number of correlations, but the true mechanisms have yet to be discovered.

Disruptions in the circadian clock have been linked to a wide variety of pathological disorders[49-54]. Additionally, a large number of biological processes

have been shown to interact with the circadian clock[55-60], where the resulting phenotype may not be immediately obvious, yet many pathways would be interrupted. Finally, the degree to which certain drugs may be metabolized or present risks of toxicity across the 24-hour day is an aspect of circadian metabolism and the control of which is an emerging field (chronotherapeutics)[40,61,62]. Therefore, disruptions in the functioning of the circadian machinery can have large downstream implications on health.

Role of rhythmic food intake in the regulation of rhythmic gene expression

The circadian clock plays an extensive role in metabolism, where it temporally regulates catabolic and anabolic processes in a manner that is believed to maximize the potential energy output from food intake and metabolism[63-65]. Multiple pathways and genes have been shown that implicate the circadian clock in the liver and its response to current nutrient status[66-69]. Additionally, desynchronization between the SCN and peripheral clocks is shown to be significantly related to the development of metabolic disorders[10,70].

Feeding in the form of feeding/fasting cycles has been extensively studied for its ability to act as a Zeitgeber and entrain the phase of the circadian clock in peripheral tissues by somehow bypassing the control of the SCN. Under *ad libitum* conditions, the SCN through its multiple pathways entices mice to eat at roughly the same time every day, resulting in the typical 75%:25% night:day breakdown of food intake as well as the SCN and liver maintaining the same phase of rhythmic gene

expression[4,71]. Under a time-restricted feeding schedule during the day (tRF-day), core clock genes in the liver, but not the SCN, will shift by up to 12 hours[72]. Further studies extended this by performing a genome-wide approach and showing that tRF-day synchronizes the rhythmic expression of many genes to be day-phased, mimicking what is seen in the core clock genes[73]. This also occurs in a clock-deficient *cry1*$^{-/-}$;*cry2*$^{-/-}$ double knockout, indicating that both the clock and feeding can act as synchronizing cues. Importantly, under fasting the expression of many genes that are rhythmic under *ad libitum* conditions is lost, raising the question of whether feeding status or rhythmic food intake is required for rhythmic expression of these genes. Further studies have shown that there appear to be a divide in gene expression in the liver, and potentially other tissues, where rhythmic gene expression of some genes is driven more by systemic signaling coming from the SCN, and for other genes is driven by a combination of feeding and systemic signals[74-76]. Finally, both circadian expression and RFI modify transcription and translation efficiency, lending support to the separation of these two effects in driving gene expression[77,78].

Some aspects of rhythmic food intake have been shown to have a large effect on the health of the organism. Restricting time of feeding to the night (tRF-night), the active phase for mice, improves health both in wild-type mice and in those with a dysfunctional circadian clock[79]. Furthermore, a nutritional challenge such as high-fat diet (HFD) or a ketogenic diet under *ad libitum* conditions results in results in obesity, diabetes, and metabolic syndromes, as well as affecting clock

gene expression[80-82]. However, under tRF-night, many of the negative effects are either negated or dampened, suggesting that tRF may be a potential tool for translational health[83].

Project aims

As described above, the role of feeding rhythms in the regulation of rhythmic gene expression was poorly understood. In particular, it was unknown whether the expression of genes controlled by feeding was due to a simple feeding/fasting switch, or if it was actually the rhythmic component of rhythmic food intake (RFI) that was driving rhythmic gene expression. Therefore, **the first aim of my Ph.D. (Chapter II)** was to determine if amplitude of RFI can regulate rhythmic gene expression without affecting oscillations in the molecular circadian clock. Previous studies examining feeding/fasting cycles had already determined that food intake is necessary for promoting rhythmic expression of many genes, even before the point where mice would die from starvation. However, little was known about the actual effect of a highly rhythmic (night-restricted) feeding schedule versus that of an arrhythmic feeding schedule, where mice had to eat an equal amount across 24 hour periods. To that end, we designed a system by which we could feed mice on custom feeding schedules and maintained them on arrhythmic, ad libitum, and night-restricted feeding schedules for 5 consecutive weeks. We discovered that RFI was positively correlated to the number of rhythmic genes in the mouse liver, which we designated as feeding-controlled genes. Furthermore, there was a class of

genes that does not respond at all to feeding status, which we termed clock-controlled. We concluded that approximately 70% of the cycling hepatic transcriptome depends at least partly on RFI for dictating gene expression.

The second aim of my Ph.D. (Chapter III) was to examine the rhythmicity of different 3' isoforms, as measured by polyadenylation site (PAS) usage, and determine if there was differential rhythmic expression between isoforms of the same genes. Studies which investigate gene expression often do so at a whole-gene level, where expression across all isoforms is summed and a single value is reported for each gene. However, how each specific isoform is regulated is not yet well understood, and thus the potential exists for many of these projects to leave key information undiscovered if isoforms are differentially regulated, instead of regulated as a whole unit as is otherwise expected. We investigated three avenues that could potentially explain the differences in rhythmic expression between the isoforms of these genes, and overall found that approximately 10% of the entire hepatic mouse transcriptome contains PAS that are differentially regulated in a rhythmic manner.

References

1 Dunlap, J. C. Molecular Bases for Circadian Clocks. *Cell* **96**, 271-290, doi:10.1016/s0092-8674(00)80566-8 (1999).

2 Akhtar, R. A. *et al.* Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus. *Current biology : CB* **12**, 540-550 (2002).

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

4 Shibata, S. Neural regulation of the hepatic circadian rhythm. *Anat Rec A Discov Mol Cell Evol Biol* **280**, 901-909, doi:10.1002/ar.a.20095 (2004).

5 Reppert, S. M. & Weaver, D. R. Coordination of circadian timing in mammals. *Nature* **418**, 935-941, doi:10.1038/nature00965 (2002).

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

7 Oishi, K. *et al.* Genome-wide expression analysis of mouse liver reveals CLOCK-regulated circadian output genes. *J Biol Chem* **278**, 41519-41527, doi:10.1074/jbc.M304564200 (2003).

8 Ko, C. H. & Takahashi, J. S. Molecular components of the mammalian circadian clock. *Hum Mol Genet* **15 Spec No 2**, R271-277, doi:10.1093/hmg/ddl207 (2006).

9 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).

10 Panda, S. Circadian physiology of metabolism. *Science* **354**, 1008-1015, doi:10.1126/science.aah4967 (2016).

11 Hsin, J. P. & Manley, J. L. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**, 2119-2137, doi:10.1101/gad.200303.112 (2012).

12 Murthy, K. G. & Manley, J. L. The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev* **9**, 2672-2683, doi:10.1101/gad.9.21.2672 (1995).

13 Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev* **25**, 1770-1782, doi:10.1101/gad.17268411 (2011).

14 Tian, B. & Manley, J. L. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* **38**, 312-320, doi:10.1016/j.tibs.2013.03.005 (2013).

15 Reyes, A. & Huber, W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Res* **46**, 582-592, doi:10.1093/nar/gkx1165 (2018).

16 Balbo, P. B. & Bohm, A. Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. *Structure* **15**, 1117-1131, doi:10.1016/j.str.2007.07.010 (2007).

17 Proudfoot, N. J. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* **352**, aad9926, doi:10.1126/science.aad9926 (2016).

18 Chen, H. *et al.* Dynamic interplay between enhancer-promoter topology and gene activity. *Nat Genet* **50**, 1296-1303, doi:10.1038/s41588-018-0175-z (2018).

19 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).

20 Vernimmen, D. & Bickmore, W. A. The Hierarchy of Transcriptional Activation: From Enhancer to Promoter. *Trends Genet* **31**, 696-708, doi:10.1016/j.tig.2015.10.004 (2015).

21 Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470-476, doi:10.1038/nature07509 (2008).

22 Tian, B., Hu, J., Zhang, H. & Lutz, C. S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* **33**, 201-212, doi:10.1093/nar/gki158 (2005).

23 Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* **18**, 18-30, doi:10.1038/nrm.2016.116 (2017).

24 Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A. & Burge, C. B. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**, 1643-1647, doi:10.1126/science.1155390 (2008).

25 Gong, C. & Maquat, L. E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284-288, doi:10.1038/nature09701 (2011).

26 Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464-469, doi:10.1038/nature07488 (2008).

27 Liu, Y. *et al.* Cold-induced RNA-binding proteins regulate circadian gene expression by controlling alternative polyadenylation. *Sci Rep* **3**, 2054, doi:10.1038/srep02054 (2013).

28 Berkovits, B. D. & Mayr, C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* **522**, 363-367, doi:10.1038/nature14321 (2015).

29 Modic, M. *et al.* Cross-Regulation between TDP-43 and Paraspeckles Promotes Pluripotency-Differentiation Transition. *Mol Cell* **74**, 951-965 e913, doi:10.1016/j.molcel.2019.03.041 (2019).

30 Ye, J. & Blelloch, R. Regulation of pluripotency by RNA binding proteins. *Cell Stem Cell* **15**, 271-280, doi:10.1016/j.stem.2014.08.010 (2014).

31 Weng, T. *et al.* Cleavage factor 25 deregulation contributes to pulmonary fibrosis through alternative polyadenylation. *J Clin Invest* **129**, 1984-1999, doi:10.1172/JCI122106 (2019).

32 Rehfeld, A., Plass, M., Krogh, A. & Friis-Hansen, L. Alterations in polyadenylation and its implications for endocrine disease. *Front Endocrinol (Lausanne)* **4**, 53, doi:10.3389/fendo.2013.00053 (2013).

33 Gruber, A. J. & Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet* **20**, 599-614, doi:10.1038/s41576-019-0145-z (2019).

34 Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673-684, doi:10.1016/j.cell.2009.06.016 (2009).

35 Creemers, E. E. *et al.* Genome-Wide Polyadenylation Maps Reveal Dynamic mRNA 3'-End Formation in the Failing Human Heart. *Circ Res* **118**, 433-438, doi:10.1161/CIRCRESAHA.115.307082 (2016).

36 MacDonald, C. C. Tissue-specific mechanisms of alternative polyadenylation: Testis, brain, and beyond (2018 update). *Wiley Interdiscip Rev RNA* **10**, e1526, doi:10.1002/wrna.1526 (2019).

37 Xu, C. & Zhang, J. Alternative Polyadenylation of Mammalian Transcripts Is Generally Deleterious, Not Adaptive. *Cell Syst* **6**, 734-742 e734, doi:10.1016/j.cels.2018.05.007 (2018).

38 Morf, J. *et al.* Cold-inducible RNA-binding protein modulates circadian gene expression posttranscriptionally. *Science* **338**, 379-383, doi:10.1126/science.1217726 (2012).

39 Benegiamo, G. *et al.* The RNA-Binding Protein NONO Coordinates Hepatic Adaptation to Feeding. *Cell Metab* **27**, 404-418 e407, doi:10.1016/j.cmet.2017.12.010 (2018).

40 Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111**, 16219-16224, doi:10.1073/pnas.1408886111 (2014).

41 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).

42 Ruben, M. D. *et al.* A database of tissue-specific rhythmically expressed human genes has potential applications in circadian medicine. *Sci Transl Med* **10**, doi:10.1126/scitranslmed.aat8806 (2018).

43 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).

44 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).

45 Trott, A. J. & Menet, J. S. Regulation of circadian clock transcriptional output by CLOCK:BMAL1. *PLoS Genet* **14**, e1007156, doi:10.1371/journal.pgen.1007156 (2018).

46 Hurley, J. M. *et al.* Analysis of clock-regulated genes in Neurospora reveals widespread posttranscriptional control of metabolic potential. *Proc Natl Acad Sci U S A* **111**, 16995-17002, doi:10.1073/pnas.1418963111 (2014).

47 Ramskold, D., Wang, E. T., Burge, C. B. & Sandberg, R. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLoS Comput Biol* **5**, e1000598, doi:10.1371/journal.pcbi.1000598 (2009).

48 Beytebiere, J. R. *et al.* Tissue-specific BMAL1 cistromes reveal that rhythmic transcription is associated with rhythmic enhancer-enhancer interactions. *Genes Dev* **33**, 294-309, doi:10.1101/gad.322198.118 (2019).

49 Harfmann, B. D. *et al.* Muscle-specific loss of Bmal1 leads to disrupted tissue glucose metabolism and systemic glucose homeostasis. *Skelet Muscle* **6**, 12, doi:10.1186/s13395-016-0082-x (2016).

50 Birky, T. L. & Bray, M. S. Understanding circadian gene function: animal models of tissue-specific circadian disruption. *IUBMB Life* **66**, 34-41, doi:10.1002/iub.1241 (2014).

51 Sato, S. *et al.* Circadian Reprogramming in the Liver Identifies Metabolic Pathways of Aging. *Cell* **170**, 664-677 e611, doi:10.1016/j.cell.2017.07.042 (2017).

52 Rudic, R. D. *et al.* BMAL1 and CLOCK, two essential components of the circadian clock, are involved in glucose homeostasis. *PLoS Biol* **2**, e377, doi:10.1371/journal.pbio.0020377 (2004).

53 Shimba, S. *et al.* Brain and muscle Arnt-like protein-1 (BMAL1), a component of the molecular clock, regulates adipogenesis. *Proc Natl Acad Sci U S A* **102**, 12071-12076, doi:10.1073/pnas.0502383102 (2005).

54 Lamia, K. A., Storch, K. F. & Weitz, C. J. Physiological significance of a peripheral tissue circadian clock. *Proc Natl Acad Sci U S A* **105**, 15172-15177, doi:10.1073/pnas.0806717105 (2008).

55 Kaasik, K. & Lee, C. C. Reciprocal regulation of haem biosynthesis and the circadian clock in mammals. *Nature* **430**, 467-471, doi:10.1038/nature02724 (2004).

56 Yang, G. *et al.* Timing of expression of the core clock gene Bmal1 influences its effects on aging and survival. *Sci Transl Med* **8**, 324ra316, doi:10.1126/scitranslmed.aad3305 (2016).

57 Toledo, M. *et al.* Autophagy Regulates the Liver Clock and Glucose Metabolism by Degrading CRY1. *Cell Metab* **28**, 268-281 e264, doi:10.1016/j.cmet.2018.05.023 (2018).

58 Thaiss, C. A. *et al.* Microbiota Diurnal Rhythmicity Programs Host Transcriptome Oscillations. *Cell* **167**, 1495-1510 e1412, doi:10.1016/j.cell.2016.11.003 (2016).

59 Le Martelot, G. *et al.* REV-ERBalpha participates in circadian SREBP signaling and bile acid homeostasis. *PLoS Biol* **7**, e1000181, doi:10.1371/journal.pbio.1000181 (2009).

60 Doi, R., Oishi, K. & Ishida, N. CLOCK regulates circadian rhythms of hepatic glycogen synthesis through transcriptional activation of Gys2. *J Biol Chem* **285**, 22114-22121, doi:10.1074/jbc.M110.110361 (2010).

61 Johnson, B. P. *et al.* Hepatocyte circadian clock controls acetaminophen bioactivation through NADPH-cytochrome P450 oxidoreductase. *Proc Natl Acad Sci U S A* **111**, 18757-18762, doi:10.1073/pnas.1421708111 (2014).

62 Gorbacheva, V. Y. *et al.* Circadian sensitivity to the chemotherapeutic agent cyclophosphamide depends on the functional status of the CLOCK/BMAL1 transactivation complex. *Proc Natl Acad Sci U S A* **102**, 3407-3412, doi:10.1073/pnas.0409897102 (2005).

63 Challet, E. Circadian clocks, food intake, and metabolism. *Prog Mol Biol Transl Sci* **119**, 105-135, doi:10.1016/B978-0-12-396971-2.00005-1 (2013).

64 Gooley, J. J. & Chua, E. C. Diurnal regulation of lipid metabolism and applications of circadian lipidomics. *J Genet Genomics* **41**, 231-250, doi:10.1016/j.jgg.2014.04.001 (2014).

65 Panda, S. *et al.* Coordinated Transcription of Key Pathways in the Mouse by the Circadian Clock. *Cell* **109**, 307-320, doi:10.1016/s0092-8674(02)00722-5 (2002).

66 Yang, X. *et al.* Nuclear receptor expression links the circadian clock to metabolism. *Cell* **126**, 801-810, doi:10.1016/j.cell.2006.06.050 (2006).

67 Sun, X. *et al.* Glucagon-CREB/CRTC2 signaling cascade regulates hepatic BMAL1 protein. *J Biol Chem* **290**, 2189-2197, doi:10.1074/jbc.M114.612358 (2015).

68 Ribas-Latre, A. & Eckel-Mahan, K. Interdependence of nutrient metabolism and the circadian clock system: Importance for metabolic health. *Mol Metab* **5**, 133-152, doi:10.1016/j.molmet.2015.12.006 (2016).

69 Zhang, Y. *et al.* The hepatic circadian clock fine-tunes the lipogenic response to feeding through RORalpha/gamma. *Genes Dev*, doi:10.1101/gad.302323.117 (2017).

70 Oosterman, J. E., Kalsbeek, A., la Fleur, S. E. & Belsham, D. D. Impact of nutrients on circadian rhythmicity. *Am J Physiol Regul Integr Comp Physiol* **308**, R337-350, doi:10.1152/ajpregu.00322.2014 (2015).

71 Noble, E. E. *et al.* Control of Feeding Behavior by Cerebral Ventricular Volume Transmission of Melanin-Concentrating Hormone. *Cell Metab* **28**, 55-68 e57, doi:10.1016/j.cmet.2018.05.001 (2018).

72 Damiola, F. *et al.* Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes Dev* **14**, 2950-2961 (2000).

73 Vollmers, C. *et al.* Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proc Natl Acad Sci U S A* **106**, 21453-21458, doi:10.1073/pnas.0909591106 (2009).

74 Kornmann, B., Schaad, O., Bujard, H., Takahashi, J. S. & Schibler, U. System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. *PLoS Biol* **5**, e34, doi:10.1371/journal.pbio.0050034 (2007).

75 Koronowski, K. B. *et al.* Defining the Independence of the Liver Circadian Clock. *Cell* **177**, 1448-1462 e1414, doi:10.1016/j.cell.2019.04.025 (2019).

76 Anzulovich, A. *et al.* Elovl3: a model gene to dissect homeostatic links between the circadian clock and nutritional status. *J Lipid Res* **47**, 2690-2700, doi:10.1194/jlr.M600230-JLR200 (2006).

77 Atger, F. *et al.* Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc Natl Acad Sci U S A* **112**, E6579-6588, doi:10.1073/pnas.1515308112 (2015).

78 Mange, F. *et al.* Diurnal regulation of RNA polymerase III transcription is under the control of both the feeding-fasting response and the circadian clock. *Genome Res* **27**, 973-984, doi:10.1101/gr.217521.116 (2017).

79 Chaix, A., Lin, T., Le, H. D., Chang, M. W. & Panda, S. Time-Restricted Feeding Prevents Obesity and Metabolic Syndrome in Mice Lacking a Circadian Clock. *Cell Metab*, doi:10.1016/j.cmet.2018.08.004 (2018).

80 Kohsaka, A. *et al.* High-fat diet disrupts behavioral and molecular circadian rhythms in mice. *Cell Metab* **6**, 414-421, doi:10.1016/j.cmet.2007.09.006 (2007).

81 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).

82 Tognini, P. *et al.* Distinct Circadian Signatures in Liver and Gut Clocks Revealed by Ketogenic Diet. *Cell Metab* **26**, 523-538 e525, doi:10.1016/j.cmet.2017.08.015 (2017).

83 Hatori, M. *et al.* Time-restricted feeding without reducing caloric intake prevents metabolic diseases in mice fed a high-fat diet. *Cell Metab* **15**, 848-860, doi:10.1016/j.cmet.2012.04.019 (2012).

CHAPTER II

RHYTHMIC FOOD INTAKE DRIVES RHYTHMIC GENE EXPRESSION MORE POTENTLY THAN THE HEPATIC CIRCADIAN CLOCK IN MICE*

Overview

Every mammalian tissue exhibits daily rhythms in gene expression to control the activation of tissue-specific processes at the most appropriate time of the day. Much of this rhythmic expression is thought to be driven cell-autonomously by molecular circadian clocks present throughout the body. By manipulating the daily rhythm of food intake in the mouse, we here show that more than 70% of the cycling mouse liver transcriptome loses rhythmicity under arrhythmic feeding. Remarkably, core clock genes are not among the 70% of genes losing rhythmic expression, and their expression continues to exhibit normal oscillations in arrhythmically-fed mice. Manipulation of rhythmic food intake also alters the timing of key signaling and metabolic pathways without altering the hepatic clock oscillations. Our findings thus demonstrate that systemic signals driven by rhythmic food intake significantly contribute to driving rhythms in liver gene expression and metabolic functions independently of the cell-autonomous hepatic clock.

---

*Reprinted with permission from “Rhythmic Food Intake Drives Rhythmic Gene Expression More Potently than the Hepatic Circadian Clock in Mice”, Greenwell, B J *et al.* 2019. *Cell Reports,* 27, 649-657.

Introduction

Nearly every mammalian cell harbors a molecular circadian clock that drives rhythmic gene expression to coordinate daily cycles in metabolism, physiology and behavior. These clocks are synchronized to the daily environmental variation by the master circadian pacemaker located in the suprachiasmatic nucleus (SCN) of the hypothalamus, which is itself entrained to the light:dark cycle via direct retinal innervation [1,2]. The SCN utilizes multiple cues to synchronize peripheral clocks, including rhythms in neuronal signaling, hormones secretion, body temperature, and food intake [2]. The hierarchical organization of the circadian system positions the SCN as the master coordinator of all peripheral clocks, ensuring that they are all properly entrained to the environment and synchronized throughout the body. Entrained peripheral clocks are thought to then regulate rhythmic gene expression in a cell-autonomous manner to initiate tissue-specific circadian transcriptional programs that control the rhythmicity of biological processes [3-5]. Experiments using temporal-restricted feeding paradigms demonstrated that the daily rhythm of food intake is a major synchronizing cue for the circadian clock and circadian transcriptional programs in the liver and other peripheral tissues [6-13]. However, recent evidence suggests that SCN-driven cues, in particular the rhythm of food intake, can also drive rhythmic gene expression in peripheral tissues without involving cell-autonomous molecular clocks [9,14-17]. In this study, we investigated this possibility by analyzing the role of the daily rhythm of food intake in driving circadian hepatic functions in the mouse. We found that, contrary to current models, rhythmic

food intake drives the majority of rhythmic gene expression independently of the cell-autonomous hepatic clock.

Results and discussion

To characterize the contribution of rhythmic food intake (RFI) to circadian biology and rhythmic gene expression, we developed a feeding system that allows for the long-term manipulation of RFI in the mouse (Figure II-A). This system exposes each mouse to a new feeding compartment every 3 hours. We fed mice under 3 feeding paradigms: $1/8^{th}$ of the daily food intake every 3 h (arrhythmic feeding, AR), only at night (*i.e*., night-restricted feeding, NR), or *ad libitum* (LB) (Figure II-A, Figure II-2A-B). As previously shown, mice fed *ad libitum* in a 12:12 light:dark (LD12:12) cycle exhibit robust rhythms of food intake and eat 75% of their daily food intake at night (Figure II- B-C, Figure II-2B). Change to the NR or AR paradigms profoundly affects the daily profile of RFI. While all mice still eat the majority of their food at night, the amount of chow eaten during the light phase varied considerably from 48.9% in AR-fed mice to 0% in NR-fed mice (and 23.8% in LB-fed mice; Figures 1B, 1C). Because AR-fed mice displayed difficulties in adapting fully to the AR-feeding paradigm and ate less from ZT6 to ZT12, we doubled the number of mice and separated them after 5 weeks into an AR group (robust dampening of RFI) and DR group (less robust dampening of RFI) (Figure II- B-C).

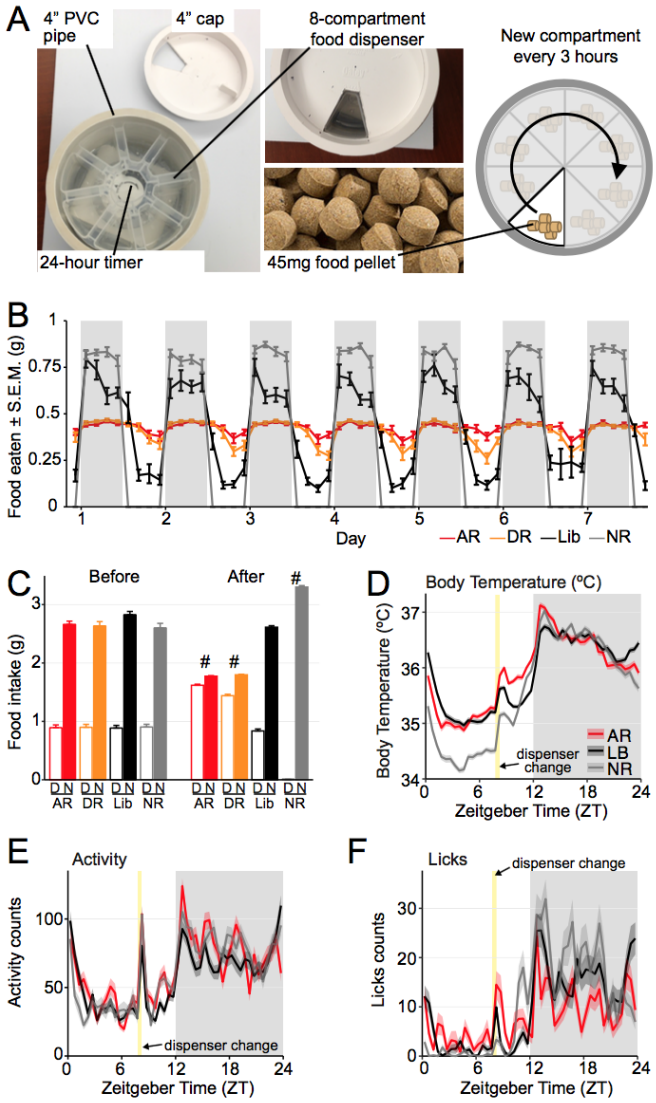

Figure II-1 Mice fed arrhythmically remain behaviorally rhythmic

(A) Overview of the feeding system. An 8-compartment food dispenser is placed on a 24 h timer and capped by a lid such that only one compartment is accessible at a time. **(B)** Average food eaten from each of the eight compartments for seven consecutive days in mice acclimated to their feeding paradigm for 4 weeks (mean ± s.e.m; n = 18 per paradigm). AR: arrhythmic feeding (red); DR: dampened feeding rhythm (orange); LB: *ad libitum* feeding (black); NR: night-restricted feeding (gray) **(C)** Quantification of the food eaten during the day and night. # indicates a significant difference in food eaten during the day *vs*. night before and after adaptation to feeding paradigms (p-value < 0.05, repeated measures two-way ANOVA). **(D-F)** Rhythms of body temperature (D), physical activity (E), and lick counts (drinking behavior) (F) for seven consecutive days (n=7-8 per feeding paradigm). The yellow bar indicates time of dispenser change-out at ZT8. Shading represents the s.e.m.

To determine if RFI manipulation alters other physiological and behavioral rhythms, we implanted mice with telemeters and tracked the body temperature, physical activity, and number of water bottle licks (interpreted as drinking behavior) (Figure II- D-F, Figure II-2 C-E). Body temperature continued to exhibit normal daily oscillations in AR mice and LB mice, and was significantly decreased by ~1ºC during the day in NR mice, potentially due to the lack of feeding during the rest phase as shown in other studies[6,13] (Figure II-D, Figure II-2C). Physical activity and drinking behavior were not affected by changes in RFI, and continued to exhibit robust oscillations across the 24-hour day (Figure II- E-F, Figure II-2 D-E). Thus, mice fed arrhythmically remain behaviorally rhythmic and still exhibit a rhythmic drive to feed when active at night, indicating that we uncoupled the rhythm of food intake from other rhythmic behaviors in AR-fed mice.

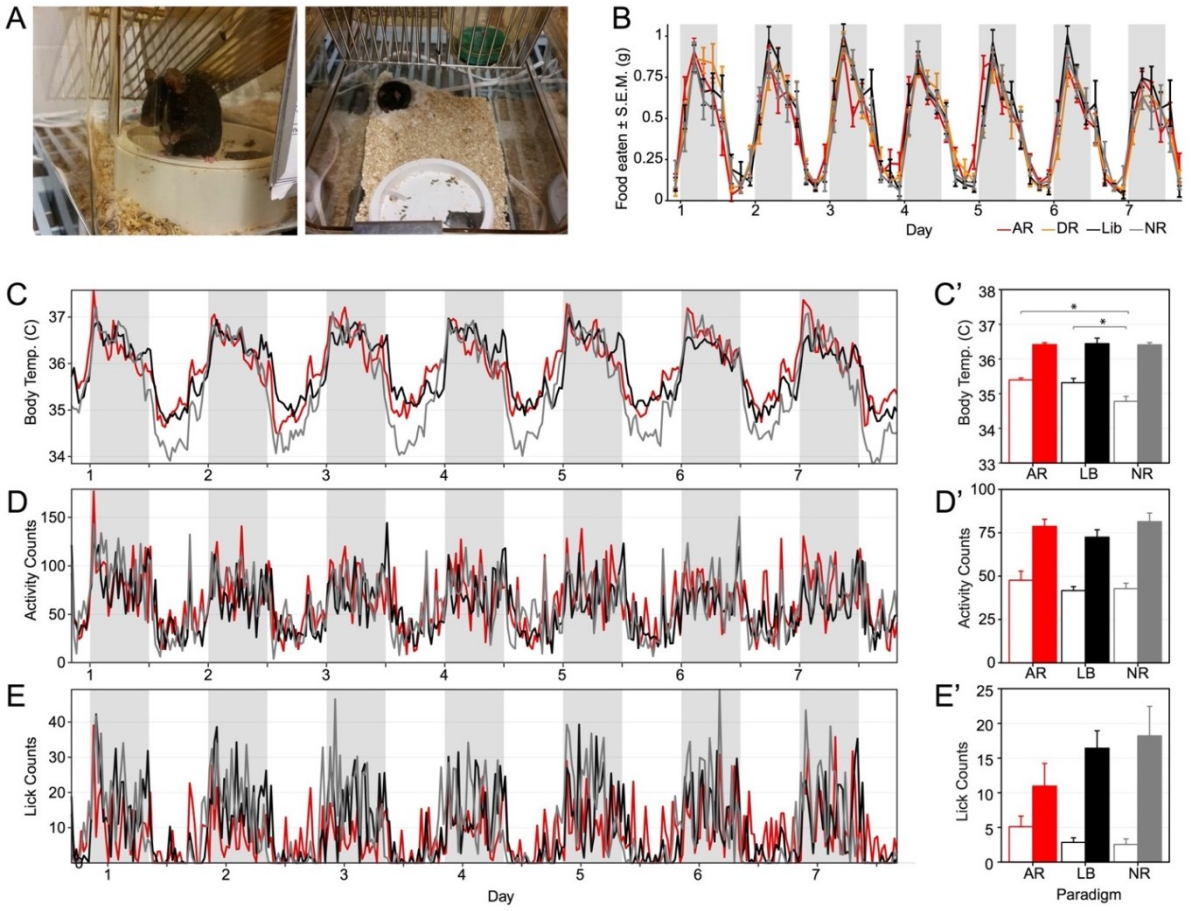

Figure II-2 (Supplement) Mice fed arrhythmically remain behaviorally rhythmic
(A) Pictures of the feeding system. Note that the lid blocks the access to 7 out of the 8 food compartments such that only 3 hours of food is available at one time. **(B)** Food eaten from each of the eight compartments for the first seven consecutive days of feeding while mice are all on *ad libitum* feeding, n=18 per feeding paradigm. AR: arrhythmic feeding (red); DR: dampened feeding rhythm (orange); LB: *ad libitum* feeding (black); NR: night-restricted feeding (gray) **(C-E)** Seven consecutive days of body temperature, physical activity, and lick counts that are averaged in Figure II- D-F (n=7-8 per feeding paradigm). (C'-E') Quantification of the seven consecutive days represented in (C-E), with empty bars corresponding to the data recorded during daytime, and solid bars to data recorded during nighttime.

RFI is a potent cue for synchronizing circadian rhythms in peripheral tissues[4,6,9,12,14]. To determine if manipulation of RFI alters the rhythmic hepatic transcriptome, we collected the livers of mice fed for 5 weeks under the three different feeding paradigms in LD12:12 every 4 hours for 24 hours, and sequenced 3'-mRNA (n = 3 per paradigm and timepoint). Genome-wide analysis of rhythmic gene expression, performed using four independent statistical programs (see Methods for details), revealed that the number of rhythmically expressed genes under each feeding paradigm correlates with the amplitude of RFI (Figure II-3A, Table II-1, Table II-2). However, rhythmically expressed genes exhibit a relatively poor overlap between each feeding paradigm, with approximately 1600 and 500 genes found to be uniquely rhythmically expressed under NR and AR feeding, respectively (Figure II-3B).

Program	AR	LB	NR
Harmonic regression	390	800	1814
Metacycle	1345	1630	3103
F24	417	880	1874
RAIN	1527	2287	3344

Table II-1 Statistical analysis of rhythmic gene expression

A: Number of rhythmically expressed genes in the mouse liver based on the feeding paradigm and the statistical program (q-value≤ 0.05). AR: arrhythmic feeding; LB: *ad libitum* feeding; NR: night-restricted feeding.

Number of programs	AR	LB	NR
1	1811	2473	3767
2	1061	1454	2718
3	448	912	1926
4	359	758	1724

Table II-2 Number of rhythmically expressed genes based on the feeding paradigm

Number of rhythmically expressed genes based on the feeding paradigm (AR: arrhythmic feeding; LB: *ad libitum* feeding; NR: night-restricted feeding) and the number of programs that classify a gene as rhythmically expressed (q-value≤ 0.05). The four programs used were Harmonic regression[18], Metacycle[19], F24[20], and RAIN[21].

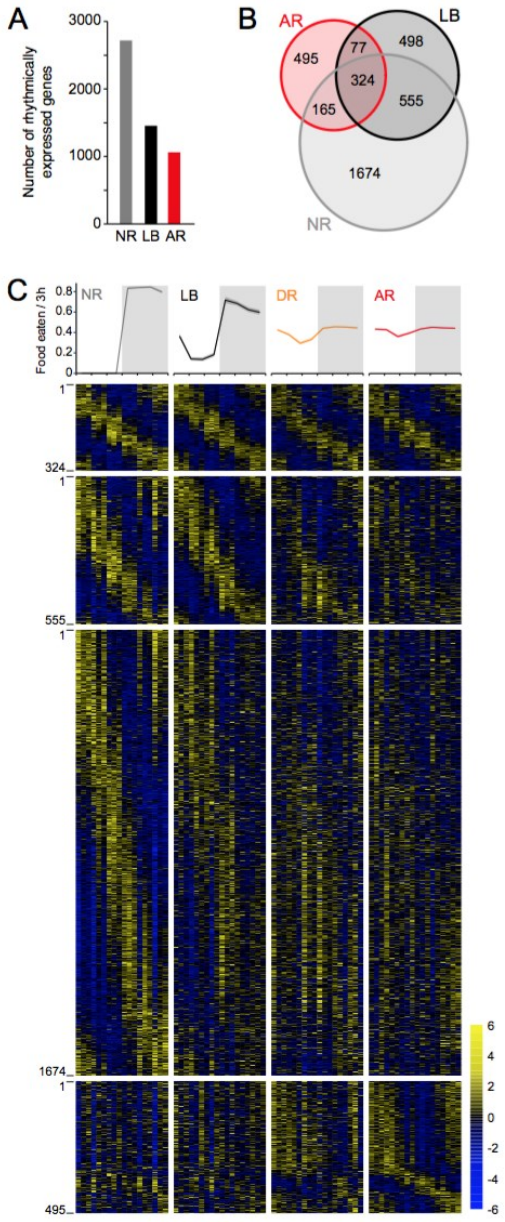

Figure II-3 Manipulation of rhythmic food intake impairs rhythmic gene expression in the mouse liver

(A) Number of rhythmically expressed genes in the liver of mice fed arrhythmically (AR; red), *ad libitum* (LB; black) or only at night (NR; gray). See STAR methods for details about the statistical analysis. **(B)** Overlap of rhythmically expressed genes between the three feeding paradigms. **(C)** Above: averaged food eaten profile. Shading represents the s.e.m. Below: heatmap of standardized expression for four categories: RRR (genes rhythmic in NR-, LB-, and AR-fed mice; 324 genes), RRA (genes rhythmic in NR- and LB-fed mice, and arrhythmic in AR-fed mice; 555 genes), RAA (genes rhythmic in NR-fed mice only; 1674 genes), and AAR (genes rhythmic in AR-fed mice only; 495 genes).

Figure II-3 Continued

Data for each column are grouped by timepoint (n = 3 per timepoint) and plotted from left to right by increasing timepoint starting at ZT2. Rows are sorted according to the peak phase in LB (RRR, RRA), NR (RAA), or AR (AAR). Expression and averaged feeding data for the DR mice (dampened-rhythm of food intake; orange) are shown for comparison, but were not considered for analysis.

To characterize the contribution of RFI to the regulation of rhythmic gene expression in the mouse liver, we focused our analysis on the three major feeding paradigms (NR, LB, and AR; see methods for details). Genes were categorized into several groups: genes rhythmic in all three feeding paradigms (named RRR for rhythmic in NR, LB, and AR), genes whose decreased rhythmicity in gene expression parallels the decreased amplitude in RFI (genes rhythmically expressed in NR and LB but not in AR, as well as genes rhythmic in NR only; named respectively RRA and RAA), and genes that are rhythmic in AR-fed mice only (named AAR). Genes rhythmic under all three feeding paradigms (RRR genes) maintained a similar phase of expression and only exhibited a small decrease in amplitude, suggesting that RFI does not contribute much to their transcription (Figure II-3C, Figure II-4 A-C). In contrast, genes in the RRA and RAA categories, which represent a large fraction of the expressed mouse liver transcriptome (n = 2,229 genes), showed a significant dampening in gene expression with a robust decrease in amplitude that parallels the decrease in RFI oscillation (Figure II-3C, Figure II-4C). Phase analysis revealed that rhythmically expressed RRR, RRA, and RAA genes maintained a well-correlated phase of expression across the 24-hour day, yet most rhythmic genes were consistently phase-advanced by 1-2 hours in

NR- and AR-fed mice when compared to LB-fed mice (Figure II-4 A-B). Genes in the AAR category displayed an increased rhythmicity and amplitude in gene expression that is inversely correlated with the amplitude of RFI oscillation. Many of the AAR genes peak at a uniform phase at the end of the night / beginning of the day in AR-fed mice, potentially indicating activation by a single pathway and/or transcription factor (Figure II-4D). Pathway enrichment analysis indicate that AAR genes are involved in xenobiotic metabolism, response to infection, and protein processing in the endoplasmic reticulum (Figure II-4E). While the mechanisms underlying their rhythmic expression are unclear, we suspect that most AAR genes exhibit a peak of expression at the end of the night / beginning of the day in AR-fed mice in response to insufficient food intake at night compared to the physiological drive to feed. Taken together, these data indicate that the amplitude of RFI significantly contributes to the genome-wide oscillation in gene expression, and that more than 70% of the cycling hepatic transcriptome under *ad libitum* feeding lose rhythmicity in mice fed arrhythmically.

Rhythmic gene expression in the mouse liver is thought to be mostly driven by the hepatic circadian clock in a cell-autonomous manner[3,4,9]. To determine if the hepatic molecular clock is responsible for the RFI-dependent decrease in rhythmic gene expression, we examined the expression of core clock genes. We found that all core clock genes were in the RRR category (Figure II-5A, Figure II-6A), thus suggesting that the core molecular clock in the liver is not affected by RFI

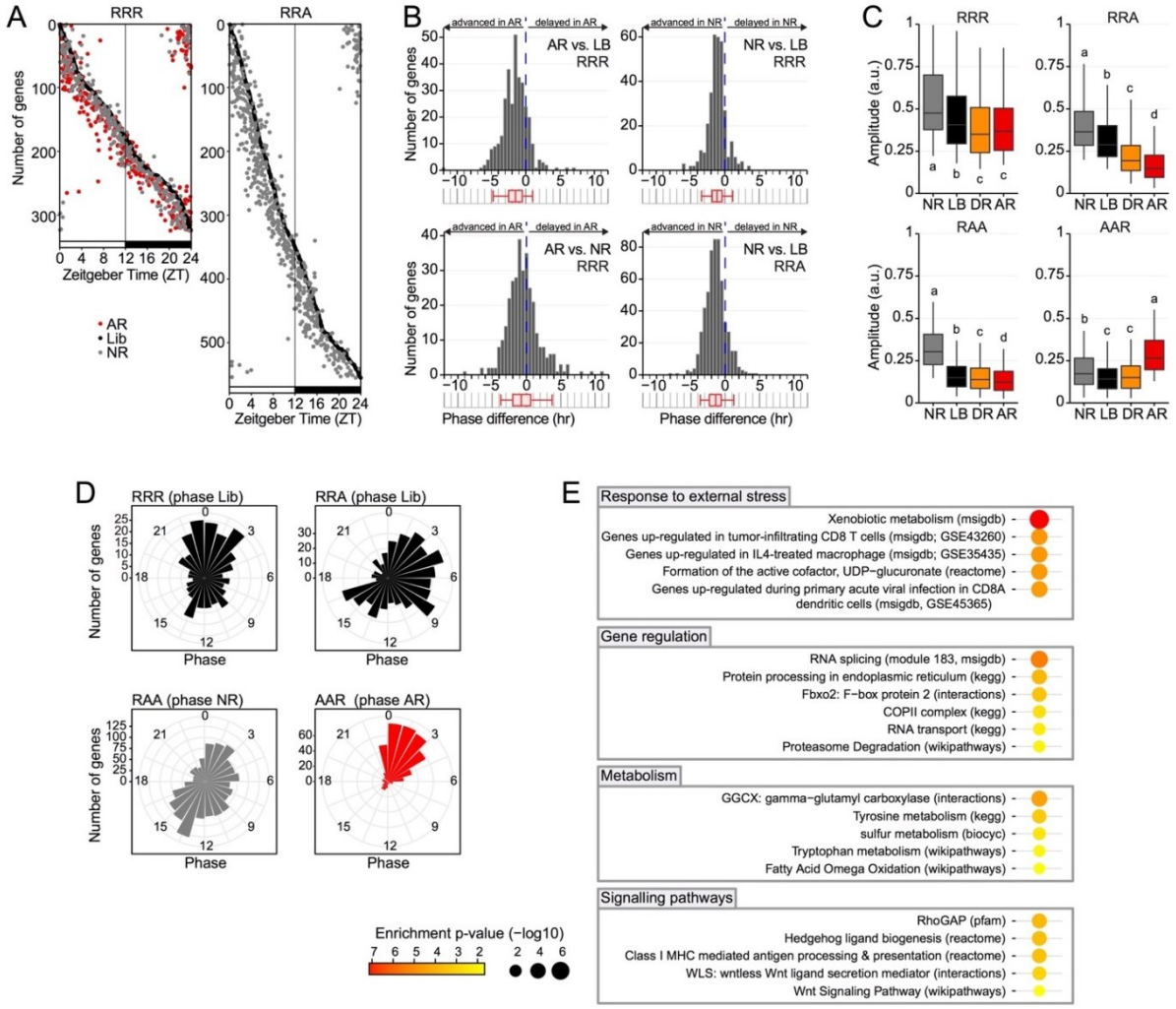

Figure II-4 (Supplement) Manipulation of rhythmic food intake impairs rhythmic gene expression in the mouse liver

(A) Phase of the 324 and 555 rhythmic genes within the RRR (genes rhythmic in NR-, LB-, and AR-fed mice) and RRA (genes rhythmic in NR- and LB-fed mice, and arrhythmic in AR-fed mice) categories for each feeding paradigm, respectively, organized by increasing phase in LB. **(B)** Phase difference of rhythmic expression between NR-, LB-, and AR-fed mice for the 324 and 555 rhythmic genes within the RRR and RRA categories, respectively. **(C)** Relative amplitude (rAMP) as reported by MetaCycle for all genes within each of the four categories, and sorted by feeding paradigm. **(D)** Rose plots of the phase of rhythmic genes within each category. Phases are taken such that LB is used whenever possible; otherwise, the phase from the most rhythmic feeding paradigm expressed in that category is used. Categories shown are RRR (genes rhythmic in NR-, LB-, and AR-fed mice; 324 genes), RRA (genes rhythmic in NR- and LB-fed mice, and arrhythmic in AR-fed mice; 555 genes), RAA (genes rhythmic in NR-fed mice only; 1674 genes), and AAR (genes rhythmic in AR-fed mice only; 495 genes). **(E)** Pathway enrichment for the AAR genes.

manipulation and that it does not significantly contribute to the RFI-mediated changes in RRA and RAA gene expression.

Based on these results, we hypothesized that the rhythmic expression of RRA/RAA genes is mostly driven by RFI, and that the 324 RRR genes are the only genes predominantly regulated by the hepatic clock. To test this hypothesis, we determined if clock-deficient mice fed only at night could maintain the rhythmic expression of RRA/RAA genes, but not RRR genes. To this end, we analyzed a public mouse liver RNA-Seq dataset in which wild-type and *Bmal1*$^{-/-}$ mice were fed only at night [15]. Visualization of gene expression revealed that most RRR genes lose rhythmicity in *Bmal1*$^{-/-}$ mice, confirming that RRR genes rely on a functional clock for rhythmic expression (Figure II-5B). Interestingly, RRR genes peaking at the end-of-night/beginning-of-day in wild-type mice become constitutively expressed at high levels in *Bmal1*$^{-/-}$ mice, whereas those peaking at the end-of-day/beginning-of-night become constitutively expressed at low levels (Figure II-5B, Figure II-6B). These results are consistent with direct transcriptional control by CLOCK:BMAL1, which binds DNA more potently in the middle of the day and whose direct target genes exhibit a similar pattern of expression in *Bmal1*$^{-/-}$ mice[22]. On the other hand, visualization of RRA/RAA genes in the liver of *Bmal1*$^{-/-}$ mice fed only at night revealed that most of these genes are expressed rhythmically, indicating that their rhythmic expression is driven by RFI and not by the hepatic clock (Figure II-5 B-C). Quantification of standardized gene expression binned by phase confirmed these results; RRR genes exhibit strong effects in median gene expression in response to

Bmal1 knockout but not to changes in RFI amplitude, whereas, conversely, RAA genes – and to a lesser extent RRA genes – show little to no response in *Bmal1*$^{-/-}$ mice but exhibit strong effects under manipulation of RFI (Figure II-5, Figure II-6B).

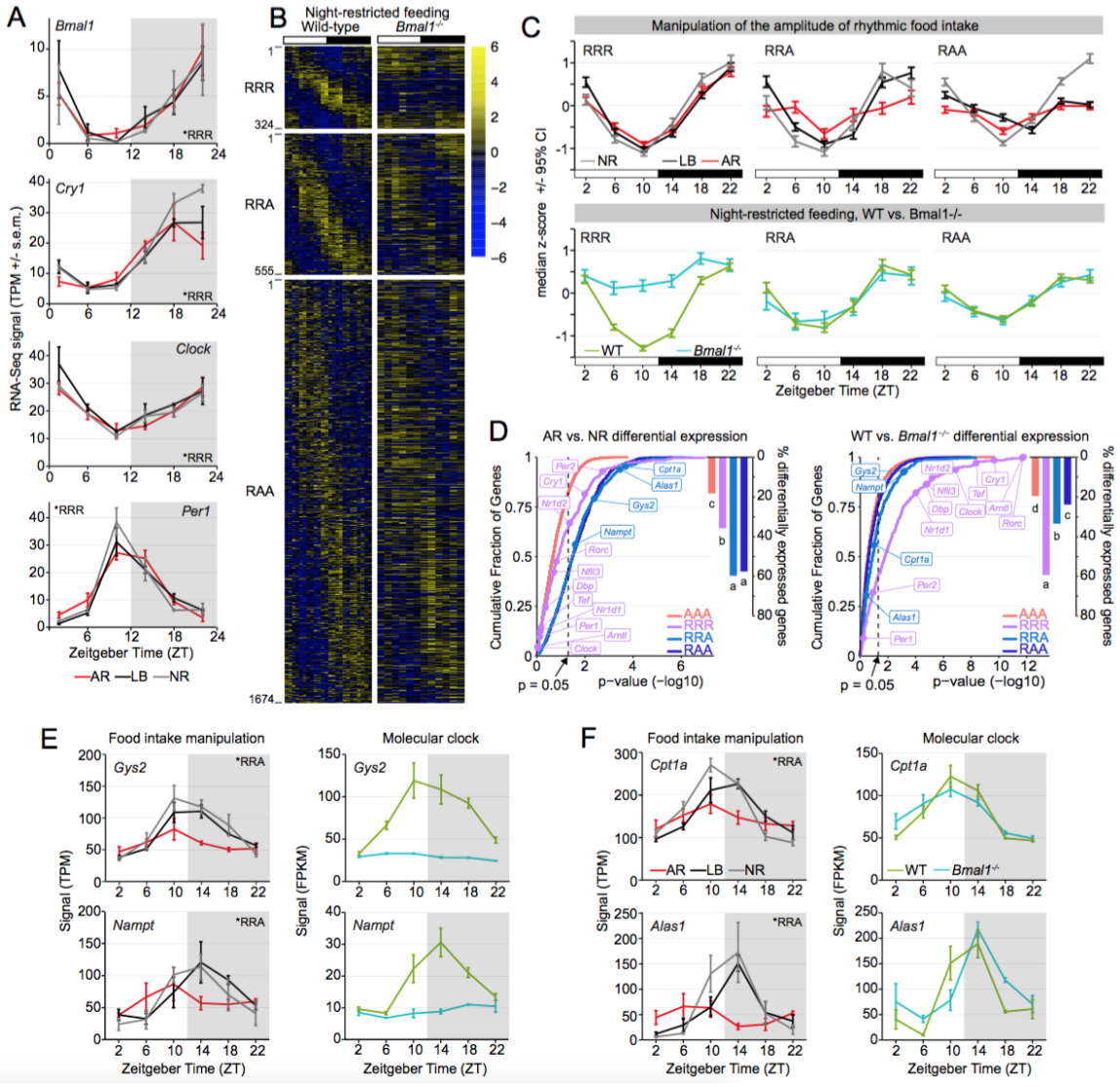

Figure II-5 Rhythmic food intake drives most hepatic rhythmic gene expression independently of the hepatic clock

(A) Expression of the core clock genes *Bmal1*, *Clock*, *Cry1* and *Per1* in the liver of mice fed arrhythmically (AR; red), *ad libitum* (LB; black), and only at night (NR; gray) (mean ± s.e.m; n = 3 per timepoint). **(B)** Standardized mouse liver gene expression for wild-type and *Bmal1*$^{-/-}$ mice fed only at night, derived from a public dataset[15].

Figure II-5 Continued

The RRR, RRA, and RRA categories are represented as in Figure II-3C with the same number and ordering of genes (n = 4 per timepoint for wild-type mice, n = 2 per timepoint for *Bmal1*$^{-/-}$ mice). **(C)** Quantification of the median standardized expression for the RRR, RRA, and RAA categories at each timepoint, binned by 4 h windows of phase and shown for the ZT20-24 bin. Error bars represent the 95% confidence interval. A figure including bins covering the 24 h day is provided as Figure II-6B. **(D)** Cumulative distributions of log-transformed p-values for differential rhythmicity in gene expression between AR and NR feeding paradigms (left) and between WT and *Bmal1*$^{-/-}$ backgrounds. p-values were obtained from the HANOVA metric of DODR analysis [23]. The category AAA (arrhythmic gene expression in NR-, LB-, and AR-fed mice) is shown as background. The percentage of differentially rhythmic genes is displayed for each category to the right. Groups with different letters are significantly different ($p<0.05$; Kolmogorov-Smirnov test). **(E-F)** Gene expression of four *Bmal1* target genes rhythmically expressed in *ad libitum* conditions, yet showing differences in rhythmic expression in response to both feeding and clock function (E) or changes in feeding only (F).

To further unveil the relative contribution of the hepatic clock *vs.* RFI in initiating rhythmic gene expression in the mouse liver, we performed a statistical analysis of differential rhythmicity using the program DODR [23]. Comparison of rhythmic expression between NR-fed and AR-fed mice revealed that most genes in the RRA and RAA categories are affected by RFI manipulation, and that the RRR genes were less affected (Figure II-5D). However, the effects of RFI on RRR genes were significantly higher than background (calculated using genes arrhythmically expressed in all three feeding paradigms, AAA), indicating that rhythmic expression of clock-controlled genes is still partially affected by RFI. Analysis of differential rhythmicity between wild-type and *Bmal1*$^{-/-}$ mice fed only at night also confirmed that genes in the RRR category were more affected by the disruption of molecular clock than genes in the RRA and RAA categories, which are at a level very close to that of the background and barely affected by *Bmal1* knockout. Interestingly, genes

whose rhythmic expression is preferentially regulated by RFI rather than by the clock (DODR NR- *vs.* AR-fed mice, $p \leq 0.05$; DODR wild-type *vs.* *Bmal1*$^{-/-}$ mice, $p > 0.05$) exhibit a phase distribution in NR-fed *Bmal1*$^{-/-}$ mice that is well correlated with the phase distribution in NR-fed wild-type mice, yet globally phase-advanced by 1-2 hours (Figure II-6C). This suggests that feeding time alone can set the phase distribution of a large fraction of the cycling transcriptome in a clock-deficient mouse similarly to the phase distribution observed in a wild-type mouse, and that the circadian clock delays the RFI-driven distribution of rhythmic gene expression by 1-2 hours. Taken together, our data demonstrate that the rhythmicity of most genes in the mouse liver is predominantly driven by the rhythm of food intake, and that the rhythmicity of only a few hundred genes is directly controlled by the cell-autonomous hepatic clock.

Results from the analyses of differential rhythmicity prompted us to identify rhythmic genes regulated by RFI, by the hepatic clock, or by both. As expected, almost all core clock genes were affected by *Bmal1* knockout but not by RFI manipulation, confirming that the molecular clock oscillations are resilient to changes in RFI amplitude (Figure II-5D, Figure II-6A). However, this was not the case for *Per1* and *Per2*, which are the entry point for the entrainment of mammalian circadian clocks[24-26]. *Per1* rhythmic expression was not affected by the molecular clock disruption or RFI manipulation, and continued to oscillate normally in both NR-fed *Bmal1*$^{-/-}$ mice and AR-fed wild-type mice (Figure II-5A, Figure II-6A). *Per2* rhythmic expression was only affected by RFI manipulation, with decreased

amplitude and a phase advance of 2.8 h from LB (Figure II-6A), which is consistent with reports showing that *Per2* expression is driven by systemic signals and not by the hepatic clock[4]. In addition, examination of four known BMAL1 target genes encoding rate-limiting enzymes showed that their rhythmic expression in the liver, which is assumed to driven by the hepatic clock, relies on RFI[27-30] (Figure II-5 E-F). Interestingly, the rhythmic expression of *Gys2* and *Nampt* is also impaired in NR-fed *Bmal1*$^{-/-}$ mice, indicating that their rhythmic expression is controlled by both the hepatic clock and RFI (Figure II-5E). However, *Cpt1a* and *Alas1* continue to cycle with similar phases and amplitudes in NR-fed *Bmal1*$^{-/-}$ mice, suggesting that their rhythmic expression is driven by RFI and not the hepatic clock despite being CLOCK:BMAL1 target genes (Figure II-5F).

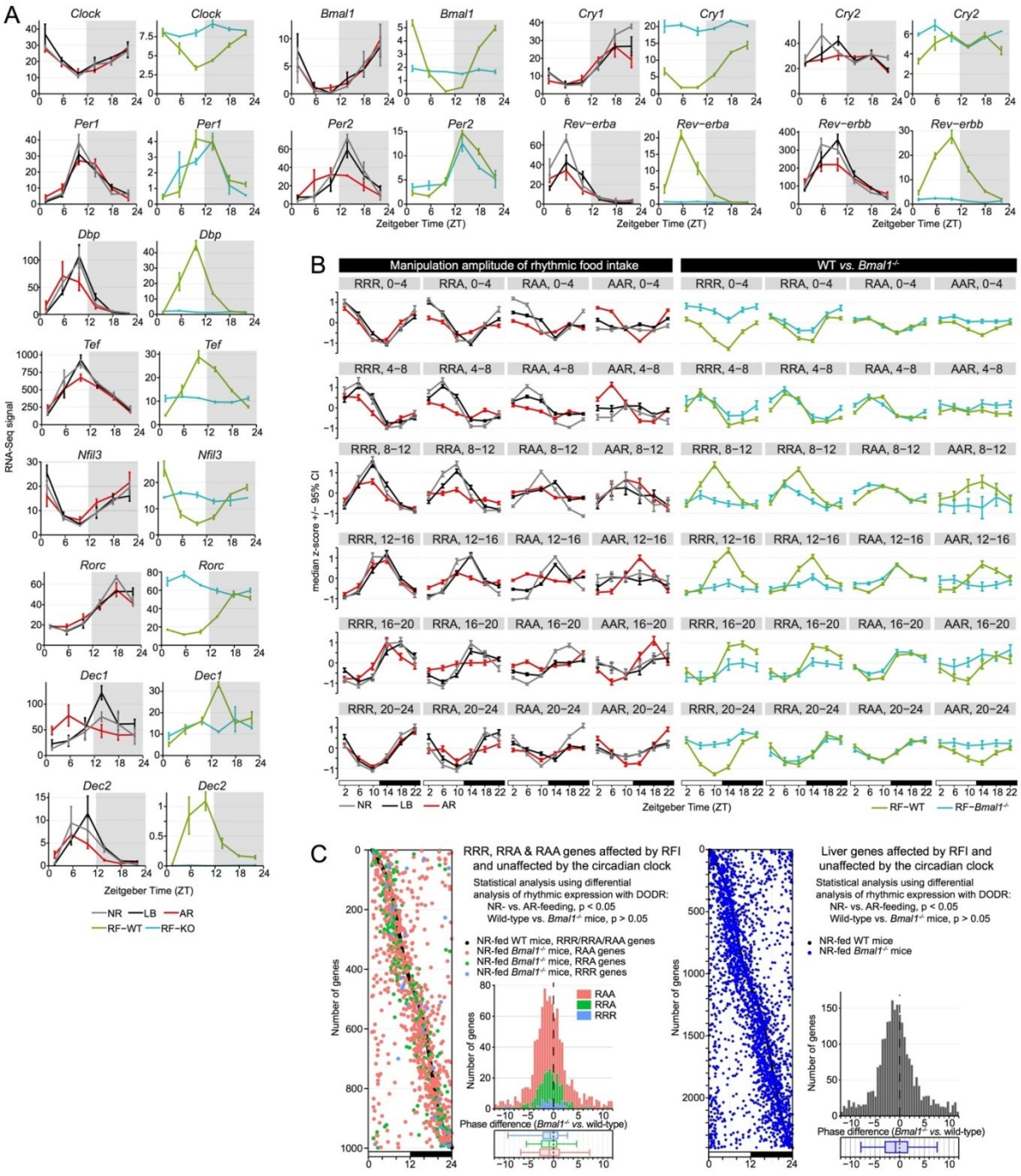

Figure II-6 (Supplement) Rhythmic food intake drives most hepatic rhythmic gene expression independently of the hepatic clock

(A) Clock gene mRNA expression in the mouse liver. Left: expression in the liver of mice fed arrhythmically (AR; red), *ad libitum* (LB; black), and only at night (NR; gray); datasets from this study. Right: expression in the liver of wild-type and *Bmal1*$^{-/-}$ mice fed only at night; dataset from Atger et al., 2015.

Figure II-6 Continued

(B) Genes within the RRR, RRA, RAA, and AAR categories (see description below) were binned by phase into 4-hour groups and their expression in both the RFI manipulation datasets and *Bmal1*$^{-/-}$ datasets (from [15]) were standardized separately. The standardized median value ± 95% confidence interval for each timepoint, binned group, and rhythmic category were then plotted for both datasets. RRR (genes rhythmic in NR-, LB-, and AR-fed mice; 324 genes), RRA (genes rhythmic in NR- and LB-fed mice, and arrhythmic in AR-fed mice; 555 genes), RAA (genes rhythmic in NR-fed mice only; 1674 genes), and AAR (genes rhythmic in AR-fed mice only; 495 genes). **(C)** Phase distribution and phase difference between wild-type vs. *Bmal1*$^{-/-}$ mice for genes significantly affected by rhythmic food intake (NR-fed *vs*. AR-fed gene expression; DODR p-value ≤ 0.05) but not affected by circadian clock disruption (wild-type *vs*. *Bmal1*$^{-/-}$ mice; DODR p-value > 0.05). Left: analysis on genes from the RRR, RRA, and RAA categories. Right: analysis on all genes expressed in the mouse liver. Mouse liver gene expression in wild-type and *Bmal1*$^{-/-}$ mice were retrieved from a public dataset[15].

The cell-autonomous hepatic clock is thought to drive rhythmic gene expression to temporally separate incompatible biochemical and metabolic processes[31]. Our findings that the rhythmic expression of several rate-limiting enzymes, which were thought to be directly regulated by the hepatic clock, actually relies on the rhythm of food intake prompted us to examine at the genome-wide level the genes and pathways that are regulated by RFI, the hepatic clock, or both. Remarkably, many metabolic pathways known to be rhythmic in the mouse liver were found to be regulated by RFI[32,33] (Figure II-7A). Many of them are involved in the regulation of carbohydrate and lipid metabolism, and include for example cholesterol and glycogen synthesis. This therefore suggests that RFI may contribute to the temporal coordination of metabolic pathways in the mouse liver without affecting the hepatic clock.

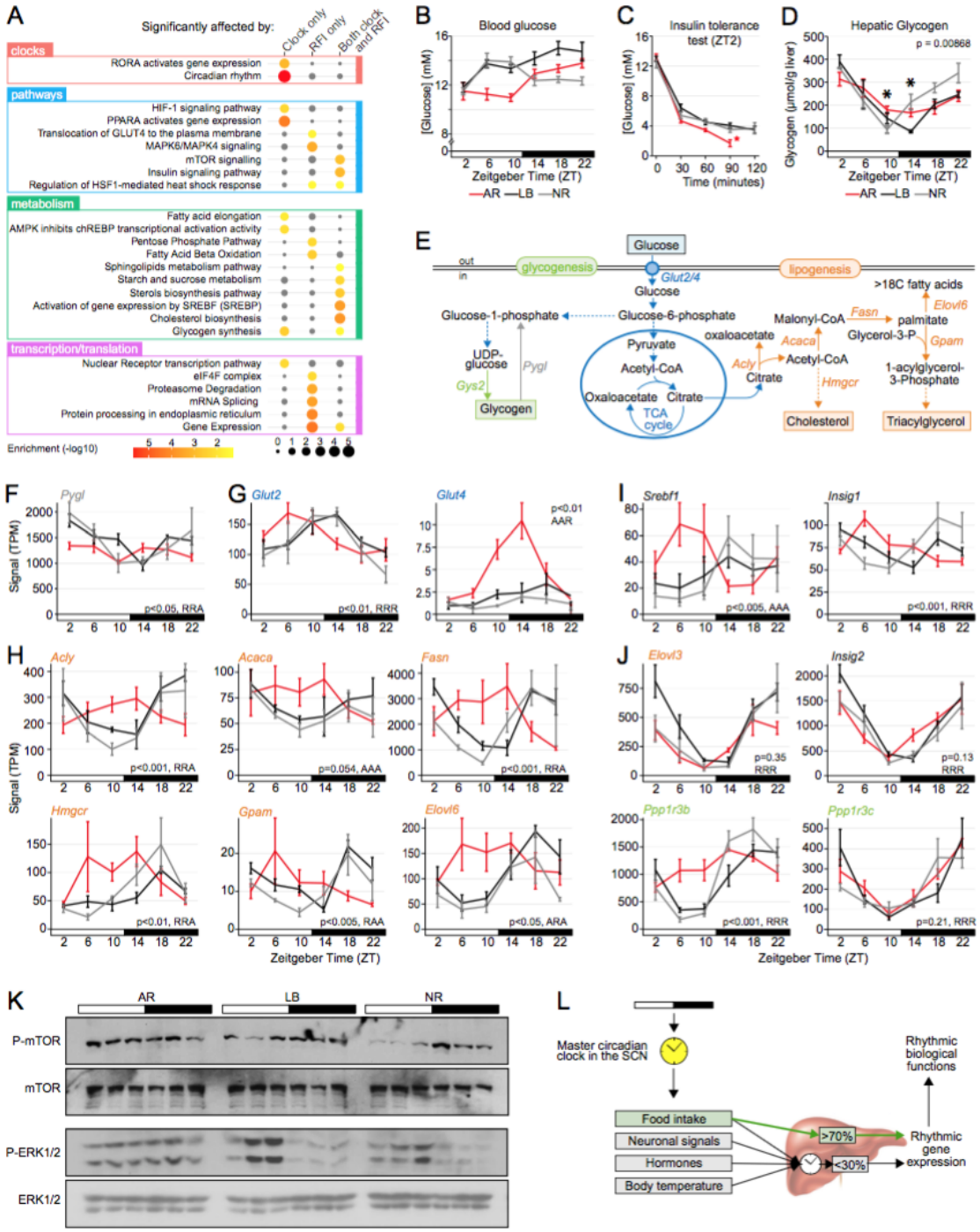

Figure II-7 Rhythmic food intake contributes to the timing of metabolic and signaling pathways independently of the hepatic clock

(A) Pathway enrichment for the RRR, RRA, and RAA genes (see Figure II-3C) based on whether their rhythmic expression is significantly affected by *Bmal1* knockout, RFI manipulation, or both.

Figure II-7 Continued

(B) Blood glucose levels at 6 timepoints (n = 12 mice per feeding paradigm; mean ± s.e.m). **(C)** Blood glucose levels before injection of insulin (t=0) and every 30 minutes after injection (n = 12 per feeding paradigm). Ten out of the twelve AR-fed mice were catatonic at t=90 and rescued with an injection of 20% glucose (see Methods). **(D)** Hepatic glycogen levels (n = 3 per feeding paradigm and timepoint; mean ± s.e.m; two-way ANOVA interaction p-value). The asterisks indicate $p < 0.05$ (one-way ANOVA) **(E)** Schematic of the glycogenesis and lipogenesis pathways in mammals. Rate-limiting enzymes and key genes are displayed. **(F-J)** Liver mRNA expression in AR-, LB-, and NR-fed mice for (F) glycogen phosphorylase *Pygl*, (G) the glucose transporters *Glut2* and *Glut4*, (H) the rate-limiting enzymes for lipogenesis and cholesterol biosynthesis, (I) the lipogenic transcription factor *Srebf1* and its co-regulator *Insig1*, and (J) paralog genes involved in glycogenesis and lipogenesis, but showing a different response to RFI manipulation. For each gene, the DODR p-value for AR *vs*. NR analysis and the rhythmic category are shown. **(K)** Protein expression in the liver of AR-, LB-, and NR-fed mice for phosphorylated and total mTOR and Erk1/2. **(L)** Graphical model showing that RFI drives upwards of 70% of the cycling transcripts in the mouse liver.

Based on the pathways influenced by RFI, we first examined if RFI manipulation impairs circulating blood glucose level, which is circadian and at trough levels at the dark:light transition in rodents[34]. We found that AR-fed mice exhibit an inverted rhythm in blood glucose levels when compared to NR-fed mice (Figure II-7B). Since responses to boluses of insulin or glucose are also clock-controlled[35], we performed an insulin tolerance test (ITT) at ZT2, *i.e.*, when blood glucose levels are similar between the 3 groups and differences between groups cannot be confounded by differences in blood glucose levels prior to insulin injection. Surprisingly, AR-fed mice were insulin hypersensitive. 83% (10 of 12) mice displayed hypoglycemia and catatonia 90 min post insulin injection and had to be rescued by a glucose injection, whereas LB- and NR-fed mice recovered with minimal problems (Figure II-7C). Because the rate-limiting enzyme for glycogen synthesis (glycogen synthase or *Gys2*) is arrhythmically expressed in AR-fed

mice[36,37] (Figure II-5F), we examined whether the hypersensitivity to insulin in AR-fed mice may be due to abnormal reserves of glycogen, resulting in impaired restoration of circulating blood glucose levels following insulin injection. Quantification of hepatic glycogen revealed that the rhythm of glycogen levels in the liver is dampened in AR-fed mice (Figure II-7D). Examination of the rhythmic expression of glycogen phosphorylase (*Pygl*), which codes for the enzyme responsible for glycogen breakdown in the liver, revealed that both catabolism and anabolism of glycogen is affected by RFI (Figure II-7 E-F). In addition, the expression of *Glut2* (aka *Slc2a2*), the main glucose transporter in hepatocytes, also shows a phase advance of 5.2 h from LB, suggesting that the availability of cellular glucose is shifted under AR-feeding[38] (Figure II-7G). The expression of the glucose transporter *Glut4* (aka *Slc2a4*), which is found primarily in adipose tissues and striated muscle, is also strongly up-regulated at the light:dark transition in the liver of AR-fed mice[39] (Figure II-7G).

Considering the possibility that intracellular glucose may be repurposed through lipogenesis instead of glycogenesis, we inspected the expression of rate-limiting lipogenic enzymes (Figure II-7E) and found that their expression was strongly impaired under arrhythmic feeding. While they are rhythmically transcribed with a peak at the end of the night in LB- and NR-fed mice, their expression is arrhythmic and almost out-of-phase in AR-fed mice (Figure II-7H). Importantly, expression of the master lipogenic transcription factor *Srebf1* and its co-regulator *Insig1* is also phase-advanced in AR-fed mice, suggesting that glycogenesis and

lipogenesis occur simultaneously in AR-fed mice whereas they normally occur sequentially in the liver of rhythmically-fed mice (Figure II-7I). Further investigation of the glycogenesis and lipogenesis pathways also revealed that paralog genes with similar rhythmic expression profiles in *ad libitum* fed mice exhibit a different response to RFI manipulation, *i.e*., only one of the two paralogs becomes arrhythmically expressed in AR-fed mice (Figure II-7J, Figure II-8A). This feature, which includes genes involved in fatty acid elongation (*Elovl3* and *Elovl6*), regulation of glycogen synthesis (*Ppp1r3b* and *Ppp1r3c*), and response to insulin (*Insig1* and *Insig2*), further suggests that disruption of rhythmic food intake can strongly impair the temporal organization of metabolic processes despite a functional cell-autonomous hepatic clock (Figure II-7J, Figure II-8A).

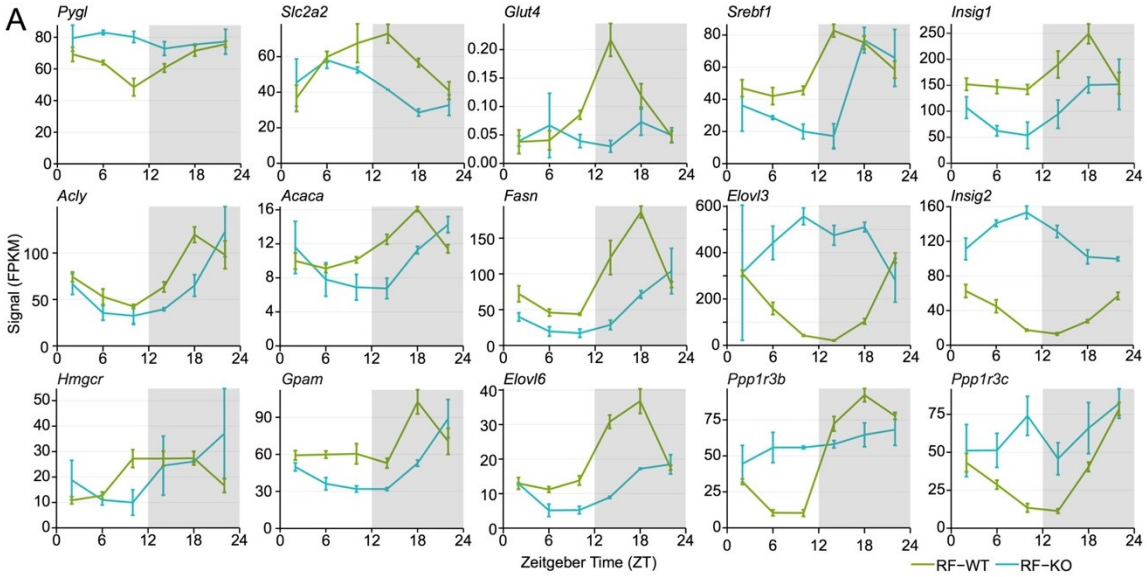

Figure II-8 (Supplement) Rhythmic food intake contributes to the timing of metabolic and signaling pathways independently of the hepatic clock

(A) Mouse liver mRNA expression for genes involved in glycogenesis, lipogenesis, and cholesterol biosynthesis in wild-type (green) and *Bmal1*$^{-/-}$ (cyan) mice fed only at night (datasets from Atger, et al. 2015). Effect of RFI manipulation on the expression of these 15 genes in the mouse liver is shown in Figure II-7 F-J.

To get insights into the mechanisms underlying the loss of rhythmic gene expression in AR-fed mice, we investigated the rhythmic activity of two major signaling pathways implicated in metabolism, mTOR and ERK1/2. While the total levels of mTOR and ERK1/2 proteins remains constant under all three feeding paradigms, we found that the rhythmic activation of these proteins via phosphorylation is strongly impaired in AR-fed mice (Figure II-7K). Specifically, while the activity of ERK1/2 and mTOR pathways are almost antiphasic in LB- and NR-fed mice, they occur coincidently throughout the 24-hr day in AR-fed mice (Figure II-7K). This suggests that the downstream targets of the mTOR and ERK1/2 pathways, which include several metabolic transcription factors, may contribute to

the loss of rhythmic gene expression in the livers of AR-fed mice. Taken together, our findings thus indicate that alteration in the rhythm of feeding can lead to observable changes in signaling and metabolic pathways without affecting the circadian clock (Figure II-7L).

We have demonstrated that RFI drives the majority of rhythms in hepatic gene expression, and contributes to the timing of signaling and metabolic pathways independently of the cell-autonomous molecular clock. It remains unknown, however, if the effects mediated by AR-feeding on glucose metabolism and lipogenesis are independent from the hepatic circadian clock, or if they originate from a desynchronization between clock-driven and RFI-driven rhythmic gene expression. Nevertheless, our findings that most of the rhythmic hepatic transcriptome is controlled by signals that originate from the SCN-driven rhythm of food intake rather than by the cell-autonomous hepatic clock raise the possibility that the contribution of RFI to rhythmic gene expression extends to other tissues, and that other SCN-driven cues may also participate in driving peripheral rhythmic gene expression. Our data also suggest that the master circadian clock in the SCN does not act solely to synchronize peripheral circadian clocks, but instead contributes more generally to circadian transcriptional programs body-wide. Reports that liver- and other tissue-specific clock-deficient mice exhibit substantial dysregulation of rhythmic gene expression and recapitulate some of the phenotypes observed in whole-body clock-deficient mice, yet continue to eat rhythmically, indicate that clock-driven and RFI-driven transcriptional programs are likely

intertwined. While the underlying mechanisms remain unclear, they may include the regulation of nutrients uptake from the portal vein or secretion of metabolites by the hepatic clock, further leading to an interdependent relationship between RFI and the cell-autonomous clock that helps maintain organismal health[3,4]. Finally, disruption of the clock has been shown to have far-reaching effects on aging and response to therapeutics, amongst others[40,41]. Our findings indicate that these effects could potentially be ameliorated through control of RFI, introducing an aspect of chronotherapy not yet explored.

Methods

Resources

Reagent or Resource	Source	Identifier
Antibodies		
Rabbit anti-mTOR	Cell Signaling Technology	Cat #2972
Rabbit anti-phospho-mTOR Ser2448	Cell Signaling Technology	Cat #2971
Rabbit anti-p44/42 MAPK	Cell Signaling Technology	Cat #9102
Rabbit anti-phospho-p44/42 MAPK Thr202/Tyr204	Cell Signaling Technology	Cat #4376
Donkey anti-Rabbit IgG secondary	GE Healthcare	NA934V
Chemicals, Peptides, and Recombinant Proteins		
TRIzol reagent	ThermoFisher	Cat #15596026
Isopropanol	EMD Millipore	Cat #PX1835-2
Ethanol	VWR	Cat #89125-176
Acid-Phenol/Chloroform, ph 4.5	ThermoFisher	Cat #AM9722
Chloroform	ThermoFisher	Cat #BP11451
Sodium acetate	ThermoFisher	Cat #AM9740
HEPES	Acros	Cat #75227-39-3
Glycerol	Sigma	Cat #G5516
EDTA	Sigma	Cat #E9884
Triton X-100	Sigma	Cat #T8787
NP-40	ThermoFisher	Cat #85124
DTT	Sigma	Cat #D9779
Phosphatase Inhibitor Cocktail	ThermoFisher	Cat #PI88266
Protease Inhibitor Cocktail	ThermoFisher	Cat #A32965
Sodium lauryl sulfate	ThermoFisher	Cat #S529-500
D-(+)-Glucose	Sigma	Cat #50-99-7

Table II-3 Reagents and resources

Reagent or Resource	Source	Identifier
Critical Commercial Assays		
Glucose Assay Reagent	Sigma	Cat #G3293
QuantSeq 3' mRNA-Seq Library Prep Kit	Lexogen	Cat #015.2X96
Glucometer	CVS	Cat #968574
BCA1 Kit	Sigma	Cat #B9643
QuantiFluor ssRNA	Promega	Cat #E3310
Deposited Data		
3' mRNA-Seq Feeding Data	This paper	GSE118967
RNA-Seq *Bmal1*$^{-/-}$ Data	[15]	GSE73554
Raw data: Mendeley	This paper	http://dx.doi.org/10.17632/t7gnz745kw.1
Experimental Models: Organisms/Strains		
Mouse: C57BL/6NCrl	Charles River Laboratories	Strain #027
Software and Algorithms		
ShortRead	[42]	NA
STAR	[43]	NA
GenomicRanges	[44]	NA
F24	[20]	NA
HarmonicRegression	[18]	NA
MetaCycle	[19]	NA
RAIN	[21]	NA
DODR	[23]	NA
HOMER	[45]	NA
Other		
Disposable Pellet Mixers and Cordless Motor	VWR	Cat #47747-358
Nitrocellulose Blotting Membrane	GE Healthcare	Cat #10600001
45mg dustless precision pellet	Bio-Serv	Cat F0165
G2 E-Mitter	Starr Life Sciences	NA
PVC Sheet	USPlastic	Cat #45095
Feeding Container	JewelrySupply	Cat #PB8301
24-hour Timer	General Electric	Cat #15119
4" PVC Tube	Home Depot	Cat #531103
4" Cap	Home Depot	Cat #39103/33463

Table II-3 Continued

Animals

C57BL/6 male mice were ordered from Charles River Laboratories (ages ranging from 43 to 49 days old), and maintained in individual cages on a 12 h light:12 h dark cycle (LD12:12) with a room temperature of 22 ± 1ºC. Animals were semi-randomly assigned to feeding groups such that starting body weight between all 3 feeding groups (NR, LB, AR) was not significantly different by one-way ANOVA (n = 20 mice per group). All animals were used in accordance with the guidelines set forth by the Institutional Animal Care and Use Committee of Texas A&M University (AUP #2016-0199).

Design of the feeding system

The feeding system we developed relies on an 8-compartment clear plastic round organizer (# PB8301, JewelrySupply) that is positioned on the top of a 24-hour timer (# 15119, General Electric), and stabilized by four screws drilled on the top of the timer and which get inserted between the organizer's compartments. The timer and food dispenser are inserted in a 4" PVC pipe (# 531103, Home Depot), and capped such that mice have access to one compartment every 3 hours (4" ABS Insert Test Cap with Knockout, #39103/33463, Home Depot). The whole system is then inserted in a standard mouse cage (N10 mouse cage, 7 1⁄2" x 11 1⁄2" x 5", #N10PLF, Ancare) drilled to accommodate the 4" PVC pipe. The entire cage/timer/pipe system is further stabilized by a custom-made support made of 1/4"

gray PVC (USPlastic), and connected to electric power using an extension cord (#145-017, Home Depot).

Pilot experiments were performed to ensure that the timer was effectively doing one full rotation every 24 hours, and that mice were not hoarding food pellets. We found that putting an excess of food in each compartment (typically at the beginning of the experiment to habituate mouse to the feeding system) was associated with increased hoarding behavior, and that decreasing amounts of food to 1.5 grams or less in each compartment extinguished the hoarding behavior.

Manipulation of the rhythm of food intake

Mice were fed with dustless precision pellets of 45 mg/pellet (# F0165 BioServ). Pellets are composed of 21.3% protein, 3.8% fat, 4% fibers, 8.1% ash, 54% carbohydrates, and <10% moisture. One gram of pellet is equivalent to 3.35 kcal.

Upon arrival, mice were randomly assigned to their final feeding paradigm (*ad libitum*, arrhythmic, or night-restricted feeding), and housed individually in their cages with *ad libitum* access to food and water for one week without using the feeding system to allow them to acclimate to their new surroundings (excess of food in one compartment, timer unplugged). After one week, all mice were fed with *ad libitum* access to food using the rotating food compartments (1.5 g per compartment) to allow them to acclimate to the feeding mechanism and to calculate the daily amount of food eaten for each mouse (baseline level of food intake). After

this full week under *ad libitum* feeding, we progressively adjusted the amounts of food available in each compartment for the arrhythmically-fed (AR) and night-restricted-fed (NR) mice every few days such that, after 2 weeks, all mice were on their final feeding paradigms (considered as week 0 for time of exposure to the feeding paradigm). We subjected mice to this progressive transition because sharp transition to AR- or NR-feeding paradigm results in a transitory decrease in the daily amount of food ingested per day, and a decrease in body weight. For all experiments, *ad libitum* (LB) mice had an excess of food placed within each compartment of the container (1.5 g), such that they never lacked for food. AR-fed mice had their daily food intake split evenly amongst the 8 compartments. NR-fed mice had their daily food split evenly amongst the 4 compartments corresponding to the night. All mice had *ad libitum* access to water.

Food containers were changed every day at ZT8 (3pm). Records of food placed within each compartment for each mouse, as well as food remaining after retrieving the container, were made to keep track of how much each mouse ate every 3 hours. Every 2 days, the total amount of food eaten by each mouse was assessed and potentially increased or decreased in order to maintain mice on their feeding paradigms. For example, mice on controlled feeding (AR or NR) that ate all pellets for two consecutive days were given two more pellets in opposite compartments such that their daily profile of food eaten did not change in rhythmicity. Similarly, mice that consistently had 4 or more pellets remaining for two consecutive days were given two less pellets, one each in opposite compartments.

The intent was to end each day with 1-2 pellets remaining in total for each mouse, indicating that the mouse was calorically satisfied without either a suspicion of starvation or an excess of food available.

Behavioral analysis

Mice aged 43-49 days were implanted with a sterile G2 E-Mitter (Starr Life Sciences) into the peritoneal cavity while anesthetized under a steady flow of 2% isoflurane in 100% O_2. Slow release Buprenorphine (1mg/kg) was injected intraperitoneally beforehand for pain relief. Mice were allowed to recover for 2 weeks before testing of the data collection system. Final data shown in the manuscript were collected for 1 week 3 months post-surgery. Data were collected over 10-minute intervals and binned into 30 minute intervals for analysis. Data collected during the first hour of recording was discarded to avoid bias from system initialization.

RNA extraction and processing

After 5 weeks of exposure to AR-, LB-, or NR-feeding paradigm, mice were anesthetized with isoflurane, decapitated, and the liver collected. The left lateral lobe was cut into three equivalent-sized pieces for RNA processing, with the remainder of the liver stored together. All collected tissues were flash-frozen in liquid nitrogen and stored at -80°C. One third of the left lateral lobe of the liver was used for RNA extraction and purification. RNA was extracted from frozen tissue

using TRIzol and following manufacturer's recommendations. Briefly, the frozen tissue was mixed with 300µL of TRIzol reagent, homogenized using a pellet mixer, and the volume brought to 1mL with 700µL of TRIzol reagent. 200µL chloroform was added, and the solution shaken and centrifuged at 12,000g for 15 minutes at 4°C. The aqueous phase was extracted and added to an equivalent amount of isopropanol. The resulting solution was that centrifuged at 12,000g for 10 minutes at 4°C, and the RNA pellet was washed with 1mL of 75% ethanol before being resuspended with 25µL RNase-free deionized water. Total RNA was then purified with an acid phenol/chloroform extraction, and precipitated by ethanol precipitation. The RNA pellet was then washed with 75% ethanol as described above, and finally resuspended in 25µL. Samples were quantified with a NanoDrop-1000 and with the Promega QuantiFluor ssRNA system, and quality / integrity of total RNA was assessed by gel electrophoresis.

Library generation and sequencing

RNA-Seq libraries were generated using the Lexogen QuantSeq 3'mRNA-Seq Library Prep Kit following manufacturer instructions, with 2µg of total RNA used as starting material. cDNA was PCR-amplified for 12 cycles following manufacturer recommendations for mouse liver tissue. Libraries were multiplexed in equimolar concentrations and sequenced in two runs using an Illumina NextSeq 500 platform (Brandeis University, USA).

Data processing

Sequenced reads were pre-processed with the R package ShortRead[42] to remove the first 12nt, remove low-quality bases at the 3' end, trim poly-A tails and embedded poly-A sequences, and remove all reads under 36nt in length. Reads were aligned to the mm10 transcriptome, assembly GRCm38.p4, with the STAR aligner[43] version 2.5.2b with options --outSAMstrandField intronMotif --quantMode GeneCounts --outFilterIntronMotifs RemoveNoncanonical. Secondary alignments were removed with samtools view -F 0x100. Read counts were summarized with the function countOverlaps from the R package GenomicRanges[44] and normalized by library size to a total of 1 million reads per library, resulting in a matrix of transcripts per million (TPM). Normalization to gene size was not performed, as we performed 3'-mRNA sequencing. Finally, only genes with greater than 1 TPM in at least 36 of the 72 samples were kept to form the final count matrix with 11536 genes.

To ensure quantification of the same transcriptome annotations between our dataset and that of [15], RNASeq data were downloaded from GEO, accession ID GSE73552. Reads were mapped to the mm10 transcriptome, assembly GRCm38.p4, using the STAR aligner version 2.5.2b with options --outSAMstrandField intronMotif --quantMode GeneCounts --outFilterIntronMotifs RemoveNoncanonical --outSAMtype BAM SortedByCoordinate --seedSearchStartLmax 15 --clip3pAdapterSeq TGGAATTCTCGGGTGCCAAGGAACTCCAGTCAC --outReadsUnmapped Fastx,

to replicate the original mapping procedure. Secondary alignments were removed with samtools view -F 0x100. Read counts were summarized with countOverlaps from GenomicRanges and normalized to FPKM values using DESeq2[46].

Rhythmicity analysis

Rhythmic analysis of the three major feeding paradigms (NR, LB, and AR) was performed with four programs: F24[20], MetaCycle[19], HarmonicRegression[18], and RAIN[21]. DR was not included in rhythmic analysis or used to form categories, but is shown in Figure II-3C to examine the effects of a feeding rhythm intermediate between LB and AR on rhythmic gene expression. Program-specific settings were as follows:

F24: iterations = 10000

MetaCycle: adjustPhase = 'predictedPer', combinePvalue = 'fisher'

HarmonicRegression: normalize = FALSE

RAIN: period = 24, deltat = 4, nr.series = 3, peak.border = c(0.2, 0.8), method = 'independent'

Resulting p-values were adjusted using the Benjamini-Hochberg method within the p.adjust function available in base R [47] to control for the false-discovery rate (FDR). Genes that were found to be rhythmic (BH-adjusted p-value –aka q-value– ≤ 0.05) in at least 2 of the 4 rhythmic programs per feeding paradigm were considered as rhythmic for that feeding paradigm (Table II-1). The rhythmic amplitude (rAMP) as

reported by MetaCycle was taken for all genes within each category and feeding paradigm.

Western blotting

Frozen liver tissue was incubated in 300µL of ice-cold RBS buffer (20mM HEPES, 50mM KCl, 10% glycerol, 2mM EDTA, 1% Triton X-100, 0.4% NP-40, 1X phosphatase inhibitor cocktail, 1mM DTT, and 1X Protease inhibitor cocktail) and homogenized on ice. Homogenate was centrifuged at 4°C for 10 minutes at high speed and the supernatant extracted. Protein abundance was quantified with the BCA1 kit (Sigma-Aldrich #B9643) following manufacturer instructions. Samples were run on SDS-PAGE gels and semi-dry transferred to a nitrocellulose membrane. Antibody information can be found in the Methods section. All of them were used at a concentration of 1:1000.

Glycogen assay

Hepatic glycogen was quantified through a method modified from Zhang, et al. 2017[48]. In brief, measured amounts of crushed frozen tissue were placed into tubes containing 500µL of 2M HCl (sample) or 2M NaOH (control) and heated on a hot block set to 95°C for one hour, shaken at 10 minute intervals. An equivalent amount of 2M NaOH (for samples) or 2M HCl (for controls) was added to neutralize the acidic and basic conditions, followed by centrifugation at 20000g for 10 minutes. 10µL was used for quantification with the Glucose Assay Reagent, following

manufacturer specifications, with a 0.5mM solution of dextrose used as a standard. Total glycogen was quantified by subtracting the signal of the undigested control from the digested sample and normalizing to the standard signal and tissue weight. Each batch (1 rhythm of each feeding paradigm) was normalized such that the total signal was equivalent between batches.

Blood glucose assay

0.5 mm of the tail tip of each mouse was removed and blood collected in a 25μL capillary until approximately half full. Each timepoint was spaced 3 days apart to avoid causing anemia from blood loss. Collected blood was expelled into a sodium heparin-coated microfuge tube, sealed, and shaken in order to coat the blood with heparin and prevent congealing. Samples were centrifuged at 10000g for 5 minutes and blood plasma collected from the upper layer. Plasma glucose was measured using a glucometer (CVS Health #968574). The lowest value that the glucometer could report was 20mg/dL (anything under this was reported as 'Low'), and so samples under this threshold were recorded as 20.

Insulin tolerance test

Food containers and water bottles were removed from each cage at ZT22, i.e., 4 hours before insulin injection. 5IU/kg body weight of insulin (Novalin R) was injected at ZT2. Blood was collected from the tail tip of each mouse just prior to injection and every 30 minutes afterwards for 2 hours, and glucose measured as

described for the blood glucose assay. 83% (10 of 12) of AR-fed mice were catatonic and unresponsive at the 90-minute collection time (responsiveness determined by testing the toe pinch reflex), and were rescued after blood collection at that time with an injection of 20% glucose at 0.1mL/10g mouse. As a result, blood from AR-fed mice was not collected at the 120-minute mark.

Quantification and statistical analysis

Statistical analyses were performed using core R functions. Plots of feeding-related gene expression (Figure II-5 A, E, F, Figure II-7 F-J, Figure II-6A) are displayed as mean TPM ± s.e.m, n = 3 mice per timepoint and feeding paradigm. Plots of gene expression profiles originating from the Atger, et al. 2015 datasets (Figure II-5 E-F, Figure II-6A, Figure II-8A) are displayed as mean FPKM ± s.e.m., n = 4 mice per timepoint for wild-type mice and n = 2 mice per timepoint for *Bmal1*$^{-/-}$ mice. Differences between groups (n = 18) in Figure II-C were analyzed with a repeated-measures two-way ANOVA, with Before|After and Day|Night as factors. # denotes a significant interaction (p-value < 0.05). Quantification of physiological data (Figure II- D-F) was performed by binning data into 30 minute intervals and is represented as mean ± s.e.m. with n = 7-8 mice per feeding paradigm. Differences between the three physiological measurements were determined through pairwise t-tests between matching measurements (*e.g.*, the body temperature during the day in NR-fed mice was only compared to the body temperature during the day in AR- and LB-fed mice).

Comparisons of rhythmic amplitude between feeding paradigms and rhythmic categories (Figure II-3B) was performed by taking the relative amplitude (rAMP) reported by MetaCycle[19] for all genes within each of the four categories, and sorted by feeding paradigm. Differences in rAMP between groups were determined by a Kruskal-Wallis test within each rhythmic category followed by a *post-hoc* Wilcoxon pairwise test with the Bonferroni correction.

Differential rhythmicity, detected as changes in peak phase and/or rhythmic amplitude, was tested within the two different datasets using DODR[23]. In the rhythmic feeding dataset, AR was tested against NR. In the dataset from Atger, et al. 2015, wild-type was tested against *Bmal1*$^{-/-}$. In both cases, genes were considered significant if the p-value of the resulting HANOVA test was less than or equal to 0.05. Differences in response to AR *vs.* NR and WT *vs*. *Bmal1*$^{-/-}$ between rhythmic categories were compared with pairwise Kolmogorov-Smirnov tests and are plotted as the $\log_{10}$-transformed p-values *vs.* the cumulative distribution of p-values within each group (Figure II-5D).

Comparisons between the feeding dataset and the dataset from Atger, et al. 2015 (Figure II-5C, Figure II-6B) were performed by binning all genes within the RRR, RRA, RAA, and AAR categories by phase into 6 bins representing 4 hours each, starting with ZT0. Expression data for each gene within both datasets was standardized to z-scores, and plotted as the median within each bin ± 95% confidence interval.

Blood glucose levels (Figure II-7B) and insulin tolerance blood glucose levels (Figure II-7C) were tested with a repeated measures two-way ANOVA (Paradigm × Timepoint). Individual timepoints are analyzed with a one-way ANOVA on paradigms. Hepatic glycogen (Figure II-7D) was analyzed for overall differences with a two-way ANOVA (Paradigm × Timepoint), with individual timepoints analyzed with a one-way ANOVA on paradigms.

Data and software availability

The RNA-Seq datasets generated in this paper are available at Gene Expression Omnibus (GEO) under the accession number GSE118967. Raw data for the western blot analysis are freely accessible via the following link: http://dx.doi.org/10.17632/t7gnz745kw.1 (Mendeley).

References

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

2 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).

3 Lamia, K. A., Storch, K. F. & Weitz, C. J. Physiological significance of a peripheral tissue circadian clock. *Proc Natl Acad Sci U S A* **105**, 15172-15177, doi:10.1073/pnas.0806717105 (2008).

4 Kornmann, B., Schaad, O., Bujard, H., Takahashi, J. S. & Schibler, U. System-driven and oscillator-dependent circadian transcription in mice with a conditionally active liver clock. *PLoS Biol* **5**, e34, doi:10.1371/journal.pbio.0050034 (2007).

5 Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111**, 16219-16224, doi:10.1073/pnas.1408886111 (2014).

6 Damiola, F. *et al.* Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes Dev* **14**, 2950-2961 (2000).

7 Stokkan, K. A., Yamazaki, S., Tei, H., Sakaki, Y. & Menaker, M. Entrainment of the circadian clock in the liver by feeding. *Science* **291**, 490-493, doi:10.1126/science.291.5503.490 (2001).

8 Kawamoto, T. *et al.* Effects of fasting and re-feeding on the expression of Dec1, Per1, and other clock-related genes. *J Biochem* **140**, 401-408, doi:10.1093/jb/mvj165 (2006).

9 Vollmers, C. *et al.* Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proc Natl Acad Sci U S A* **106**, 21453-21458, doi:10.1073/pnas.0909591106 (2009).

10 Chaix, A., Zarrinpar, A., Miu, P. & Panda, S. Time-restricted feeding is a preventative and therapeutic intervention against diverse nutritional challenges. *Cell metabolism* **20**, 991-1005, doi:10.1016/j.cmet.2014.11.001 (2014).

11 Hatori, M. *et al.* Time-restricted feeding without reducing caloric intake prevents metabolic diseases in mice fed a high-fat diet. *Cell Metab* **15**, 848-860, doi:10.1016/j.cmet.2012.04.019 (2012).

12 Saini, C. *et al.* Real-time recording of circadian liver gene expression in freely moving mice reveals the phase-setting behavior of hepatocyte clocks. *Genes Dev* **27**, 1526-1536, doi:10.1101/gad.221374.113 (2013).

13 Satoh, Y., Kawai, H., Kudo, N., Kawashima, Y. & Mitsumoto, A. Time-restricted feeding entrains daily rhythms of energy metabolism in mice. *Am J Physiol Regul Integr Comp Physiol* **290**, R1276-1283, doi:10.1152/ajpregu.00775.2005 (2006).

14 Izumo, M. *et al.* Differential effects of light and feeding on circadian organization of peripheral clocks in a forebrain Bmal1 mutant. *eLife* **3**, doi:10.7554/eLife.04617 (2014).

15 Atger, F. *et al.* Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc Natl Acad Sci U S A* **112**, E6579-6588, doi:10.1073/pnas.1515308112 (2015).

16 Mange, F. *et al.* Diurnal regulation of RNA polymerase III transcription is under the control of both the feeding-fasting response and the circadian clock. *Genome Res* **27**, 973-984, doi:10.1101/gr.217521.116 (2017).

17 van der Veen, D. R. *et al.* Impact of behavior on central and peripheral circadian clocks in the common vole Microtus arvalis, a mammal with ultradian rhythms. *Proc Natl Acad Sci U S A* **103**, 3393-3398, doi:10.1073/pnas.0507825103 (2006).

18 Luck, S., Thurley, K., Thaben, P. F. & Westermark, P. O. Rhythmic degradation explains and unifies circadian transcriptome and proteome data. *Cell Rep* **9**, 741-751, doi:10.1016/j.celrep.2014.09.021 (2014).

19 Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K. & Hogenesch, J. B. MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics* **32**, 3351-3353, doi:10.1093/bioinformatics/btw405 (2016).

20 Wijnen, H., Naef, F. & Young, M. W. Molecular and Statistical Tools for Circadian Transcript Profiling. **393**, 341-365, doi:10.1016/s0076-6879(05)93015-2 (2005).

21 Thaben, P. F. & Westermark, P. O. Detecting Rhythms in Time Series with RAIN. *Journal of Biological Rhythms* **29**, 391-400, doi:10.1177/0748730414553029 (2014).

22 Trott, A. J. & Menet, J. S. Regulation of circadian clock transcriptional output by CLOCK:BMAL1. *PLoS Genet* **14**, e1007156, doi:10.1371/journal.pgen.1007156 (2018).

23 Thaben, P. F. & Westermark, P. O. Differential rhythmicity: detecting altered rhythmicity in biological data. *Bioinformatics* **32**, 2800-2808, doi:10.1093/bioinformatics/btw309 (2016).

24 Albrecht, U., Zheng, B., Larkin, D., Sun, Z. S. & Lee, C. C. MPer1 and mper2 are essential for normal resetting of the circadian clock. *J Biol Rhythms* **16**, 100-104, doi:10.1177/074873001129001791 (2001).

25 Shearman, L. P., Zylka, M. J., Weaver, D. R., Kolakowski, L. F. & Reppert, S. M. Two period Homologs: Circadian Expression and Photic Regulation in the Suprachiasmatic Nuclei. *Neuron* **19**, 1261-1269, doi:10.1016/s0896-6273(00)80417-1 (1997).

26 Maywood, E. S., Mrosovsky, N., Field, M. D. & Hastings, M. H. Rapid down-regulation of mammalian Period genes during behavioral resetting of the circadian clock. *Proceedings of the National Academy of Sciences* **96**, 15211-15216, doi:10.1073/pnas.96.26.15211 (1999).

27 Hatanaka, F. *et al.* Genome-wide profiling of the core clock protein BMAL1 targets reveals a strict relationship with metabolism. *Mol Cell Biol* **30**, 5636-5648, doi:10.1128/MCB.00781-10 (2010).

28 Doi, R., Oishi, K. & Ishida, N. CLOCK regulates circadian rhythms of hepatic glycogen synthesis through transcriptional activation of Gys2. *J Biol Chem* **285**, 22114-22121, doi:10.1074/jbc.M110.110361 (2010).

29 Kaasik, K. & Lee, C. C. Reciprocal regulation of haem biosynthesis and the circadian clock in mammals. *Nature* **430**, 467-471, doi:10.1038/nature02724 (2004).

30 Ramsey, K. M. *et al.* Circadian clock feedback cycle through NAMPT-mediated NAD+ biosynthesis. *Science* **324**, 651-654, doi:10.1126/science.1171641 (2009).

31 Panda, S. Circadian physiology of metabolism. *Science* **354**, 1008-1015, doi:10.1126/science.aah4967 (2016).

32 Abbondante, S., Eckel-Mahan, K. L., Ceglia, N. J., Baldi, P. & Sassone-Corsi, P. Comparative Circadian Metabolomics Reveal Differential Effects of Nutritional Challenge in the Serum and Liver. *J Biol Chem* **291**, 2812-2828, doi:10.1074/jbc.M115.681130 (2016).

33 Eckel-Mahan, K. L. *et al.* Coordination of the transcriptome and metabolome by the circadian clock. *Proc Natl Acad Sci U S A* **109**, 5541-5546, doi:10.1073/pnas.1118726109 (2012).

34 La Fleur, S. E., Kalsbeek, A., Wortel, J. & Buijs, R. M. A suprachiasmatic nucleus generated rhythm in basal glucose concentrations. *Journal of neuroendocrinology* **11**, 643-652 (1999).

35 la Fleur, S. E., Kalsbeek, A., Wortel, J., Fekkes, M. L. & Buijs, R. M. A Daily Rhythm in Glucose Tolerance: A Role for the Suprachiasmatic Nucleus. *Diabetes* **50**, 1237-1243, doi:10.2337/diabetes.50.6.1237 (2001).

36 Ishikawa, K. & Shimazu, T. Daily rhythms of glycogen synthetase and phosphorylase activities in rat liver: Influences of food and light. *Life Sciences* **19**, 1873-1878, doi:10.1016/0024-3205(76)90119-3 (1976).

37 Zani, F. *et al.* PER2 promotes glucose storage to liver glycogen during feeding and acute fasting by inducing Gys2 PTG and G L expression. *Mol Metab* **2**, 292-305, doi:10.1016/j.molmet.2013.06.006 (2013).

38 Leturque, A., Brot-Laroche, E., Le Gall, M., Stolarczyk, E. & Tobin, V. The role of GLUT2 in dietary sugar handling. *Journal of Physiology and Biochemistry* **61**, 529-537, doi:10.1007/bf03168378 (2005).

39 Huang, S. & Czech, M. P. The GLUT4 glucose transporter. *Cell Metab* **5**, 237-252, doi:10.1016/j.cmet.2007.03.006 (2007).

40 Kondratov, R. V., Kondratova, A. A., Gorbacheva, V. Y., Vykhovanets, O. V. & Antoch, M. P. Early aging and age-related pathologies in mice deficient in BMAL1, the core componentof the circadian clock. *Genes Dev* **20**, 1868-1873, doi:10.1101/gad.1432206 (2006).

41 Gorbacheva, V. Y. *et al.* Circadian sensitivity to the chemotherapeutic agent cyclophosphamide depends on the functional status of the CLOCK/BMAL1 transactivation complex. *Proc Natl Acad Sci U S A* **102**, 3407-3412, doi:10.1073/pnas.0409897102 (2005).

42 Morgan, M. *et al.* ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* **25**, 2607-2608, doi:10.1093/bioinformatics/btp450 (2009).

43 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

44 Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput Biol* **9**, e1003118, doi:10.1371/journal.pcbi.1003118 (2013).

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

46 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

47 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).

48 Zhang, P. Analysis of Mouse Liver Glycogen Content. *Bio-protocol* **2**, e186, doi:10.21769/BioProtoc.186 (2012).

CHAPTER III

INVESTIGATION OF DIURNAL POLYADENYLATION SITE USAGE REVEALS DIFFERENTIAL REGULATION IN THE TRANSCRIPTION OF GENE ISOFORMS

Overview

Gene isoforms are mRNAs produced from the same locus, but that differ in their transcription start sites, protein coding DNA sequences, and/or untranslated regions. Consequently, different isoforms of the same gene can have altered gene function or even serve different biological functions. Conventional RNA-Seq strategies cannot accurately distinguish expression levels between distinct isoforms, and differences in gene expression studies almost always report total gene expression. However, increasing evidence indicates that differences in isoform usage may be important for the regulation of biological functions and for development of diseases such as cancer. Here, we aimed to define whether gene isoforms are subjected to differential regulation and expression by characterizing 24-hour rhythms in polyadenylation site (PAS) usage over the course of the day in the mouse liver. Conventional RNA-Seq experiments have shown that 15-30% of genes in the mouse liver are rhythmically expressed, and it is assumed that the comprising isoforms are expressed in a similar pattern. By performing 3'-end mRNA-Seq and using stringent criteria for defining differential rhythmic expression, we show that 15% of the genes with more than one PAS exhibit differential rhythmic expression. In particular, many genes known to be rhythmic in the mouse liver

harbor at least one constitutively expressed isoform, while several hundred genes characterized as arrhythmically expressed also exhibit a rhythmic isoform. Analysis of PAS usage in nuclear mRNA and single-cell data reveals that the vast majority of isoform-specific regulation does not involve post-transcriptional regulation (e.g., miRNA targeting, RNA half-life) or cell subtype-specific differences in expression. Finally, characterization of PAS usage in *Bmal1*$^{-/-}$ mice revealed that co-transcriptional regulation plays a large role in the expression of specific gene isoforms. Taken together, our results indicate that gene isoforms can be differentially regulated, and imply that gene isoforms can behave as distinct transcriptional units. Furthermore, our data suggest that conventional RNA-Seq strategies are not the most appropriate choice for detecting changes in gene isoform expression.

Introduction

The majority of genes in higher eukaryotes contain multiple sites at which RNA can be cleaved and polyadenylated, leading to the expression of distinct transcript isoforms[1,2]. The position of these alternative polyadenylation sites (APAS) can have large effects on 3' UTR length and/or protein sequence, and lead to transcripts being truncated or containing additional exonic or intronic sequences[3,4]. The length of the 3' UTR tail itself has large implications on RNA stability, mostly because it defines the probability for a 3'UTR to be targeted by RNA binding proteins (RBP), miRNA, or long non-coding RNA[5-7]. Importantly, it has recently

been shown that differences in the RNA transcript 3' UTR lengths can lead to profound differences in the resulting protein localization, suggesting crucial role for UTRs in the function of the protein they encode[8].

Investigation of gene expression at the genome-wide level is commonly achieved through the use of massively parallel RNA-Seq libraries, where analysis frequently quantifies expression across all potential isoforms and reports a single value for each gene. Thus, RNA-Seq analysis commonly relies on the assumption that all distinct isoforms of each gene are regulated in a similar manner. However, it remains unclear whether APAS transcript isoforms are for the most part regulated in a similar manner, or whether differential regulation of APAS isoform expression can occur, and how this may affect downstream biological functions. Recent evidence has shown that biological processes like pluripotency are regulated by APA[9,10], and that defects in APAS can lead to health defects including cancer[11-14]. The underlying mechanisms involve co-transcriptional loading of RBP, which promote cleavage and polyadenylation at proximal APAS to lead to increased expression of APAS isoforms with shorter 3' UTRs and decreased expression of isoforms with long 3'UTR. Thus, APA seem to mostly involve changes in the relative ratio of short *vs.* long APAS isoforms without affecting the overall level of transcription[13,15].

Regulation of rhythmic gene expression by the circadian clock has been described in numerous species and tissues, and about half of the transcriptome is rhythmically expressed in at least one tissue in mammals[16]. This widespread rhythmic expression is maintained by a transcriptional/translational negative

feedback loop, which in mammals is initiated by the heterodimeric transcription factor CLOCK:BMAL1[17]. In addition to transcription regulation, the steady-state rhythmic levels of transcripts are regulated post-transcriptionally[18,19]. Circadian rhythms are fundamental to the temporal organization of biological functions over the course of the day and night, and their dysregulation by genetic or environmental cues leads to a wide range of pathological disorders including metabolic syndromes, cancer, and cardiovascular disorders[20-23]. As with most RNA-Seq analysis, cycling transcriptomes have been reported at the gene level, and it remains unknown whether every APAS isoform of a rhythmically expressed gene is also rhythmically regulated. Moreover, the possibility remains that genes characterized as arrhythmically expressed may contain one or more cycling APAS isoform. In this study, we addressed whether APA can lead to widespread differential expression of distinct transcript isoforms by performing a comprehensive analysis of the diurnal mouse liver transcriptome using 3' mRNA-Seq and determining whether APAS isoforms from the same gene can exhibit differential rhythmic gene expression. Our results indicate that >10% of the expressed liver genes exhibit differential APAS transcript rhythmicity, and that the underlying mechanisms involve differential co-transcriptional APA, as well as to a lesser extent differential post-transcriptional regulation between APAS isoforms and cellular subtype expression of specific APAS isoforms.

Results

APAS isoforms exhibit differential rhythmicity in mouse liver

To determine if APA can lead to isoforms having different expression profiles, we investigated diurnal APAS usage using 3'mRNA-Seq in the liver of mice collected across the 24-hour day (Figure III-A). As previously reported[16], quantification of steady-state expression levels of unique reads across gene models revealed that almost 30% of the expressed mouse liver transcriptome is rhythmic (Figure III-2 A-B). To assay whether some genes exhibit rhythmic or arrhythmic expression of APAS isoforms, we first generated a database of PAS in mouse liver using 3'mRNA-Seq reads from over 100 biologically distinct libraries and compiling for more than one billion uniquely mapped reads. This process, which is detailed in the Methods section, identified 29,199 high-confidence PAS located in 10,160 genes (Figure III-B). Among these genes, 7693 (75.7%) had 2 or more PAS (Figure III-2C).

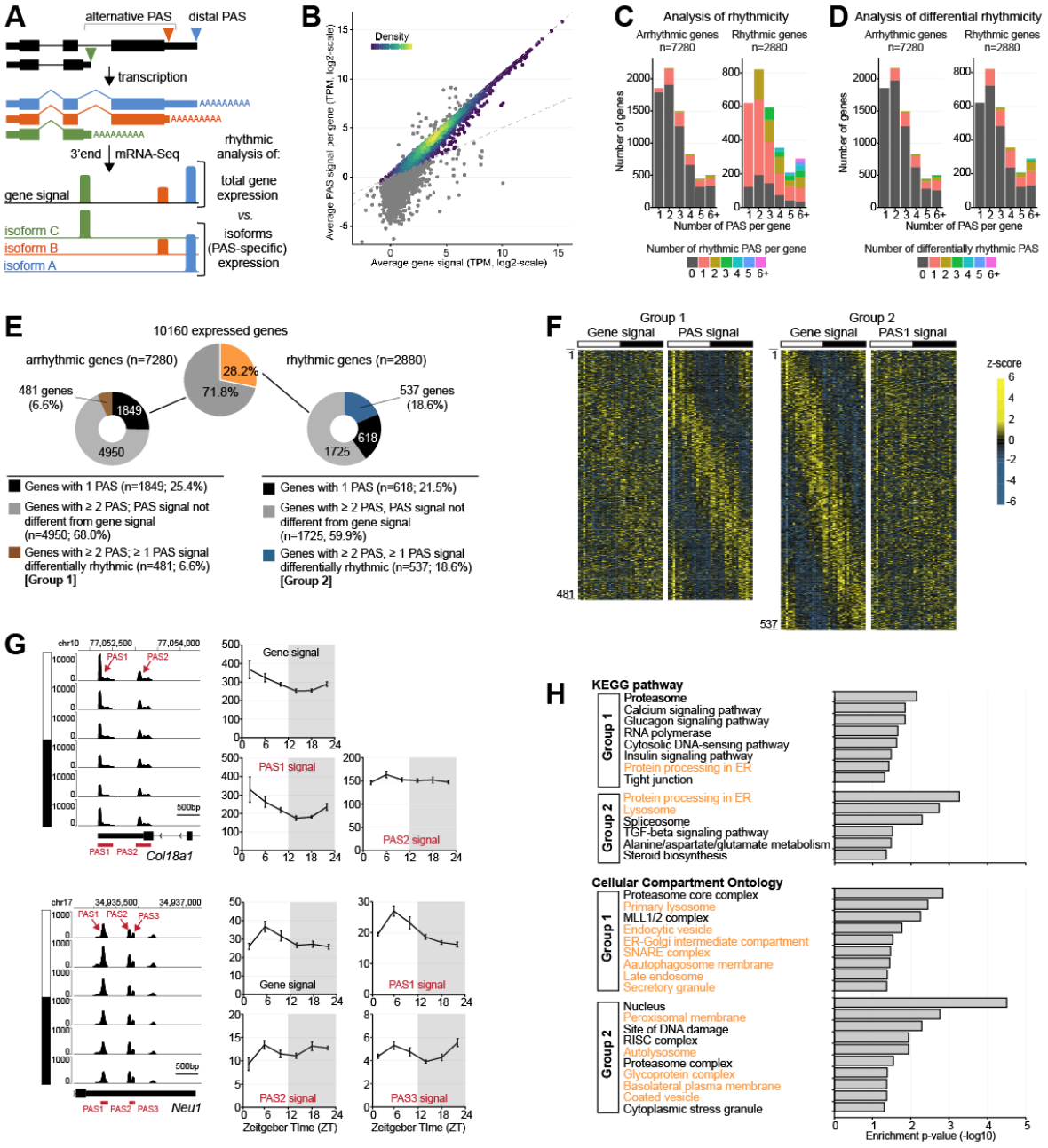

Figure III-1 APAS isoforms exhibit differential rhythmicity in mouse liver

(A) Diagram of how the expression of a gene may be decomposed into expression from multiple PAS. **(B)** Correlation between the expression of a gene and the sum of expression from its constituent PAS. PAS that were excluded (see Methods) are shaded gray. **(C-D)** Breakdown of number of PAS per gene by the rhythmicity of the gene and the number of rhythmic PAS per gene (C) or number of differentially rhythmic PAS per gene (D). **(E)** Breakdown of the number of PAS per gene, sorted into bins of 1 or 2+ PAS. Genes that will be considered for further analysis are labeled as Group 1 or Group 2.

Figure III-1 Continued

(F) Heatmaps of G1 and G2. Values are plotted in order of increasing phase down the y axis, and in order of increasing timepoint across the x axis, starting at ZT2. Each datapoint is colored according to standardized expression. **(G)** Gene model, PAS location, gene expression, and PAS expression of *Col18a1* and *Neu1* in total RNA. **(H)** KEGG pathway and Gene Ontology – Cellular Compartment analysis of G1 and G2.

We then used this PAS database and combined two stringent statistical analyses to determine if APAS isoforms from the same gene can exhibit differences in rhythmic expression. First, we compared the rhythmic expression, as assayed by combining four different statistical tools for rhythmicity (see Methods for details), between each gene and their respective APAS isoforms (Figure III-C). We found that PAS rhythmicity followed for the most part gene rhythmicity, with the majority of arrhythmic genes containing arrhythmic APAS isoforms and the majority of rhythmic genes containing rhythmic APAS transcripts (Figure III-C). However, PAS rhythmicity did not consistently match the rhythmicity of its corresponding gene, and many rhythmic genes harbored a combination of both rhythmic and arrhythmic PAS. Because this comparative analysis of rhythmicity only relies on thresholds and returns differences in rhythmicity between genes and APAS isoform profiles even for q-values being just below and above threshold, respectively, we performed a second analysis of differential rhythmicity between every APAS isoforms and their corresponding gene using the program DODR[24]. Using this second analysis, about 20% of the genes with 2 or more PAS displayed an APAS isoform being differentially rhythmic from gene signal (1,556 out of 7,693 genes), while no difference was found between PAS and gene rhythmicity for genes with one PAS

(Figure III-D). We combined the results of these two analyses to identify two groups of differentially expressed APAS isoforms (Figure III-E, Figure III-2D). The first group (Group 1) consists of 522 APAS isoforms representing 481 genes that are arrhythmic yet have a rhythmic PAS, while the second group (Group 2) consists of 699 PAS representing 537 genes that are rhythmic yet have an arrhythmic PAS (Figure III-E, Figure III-2C). Visualization of the differences between gene and PAS signals for these two groups with heatmaps (Figure III-F, Figure III-2D), along with IGV browser signals for two representative genes (*Col18a1* and *Neu1*; Figure III-G), illustrate our findings that APAS isoforms from a same gene can exhibit striking differences in their rhythmic expression. Based on our stringent analysis, these differences of rhythmic expression between APAS isoforms are widespread, and account for at least 10% of the genes expressed in mouse liver (1,018 out of 10,160), and 13.2% of the expressed genes with two or more PASs (1,018 out of 7,693).

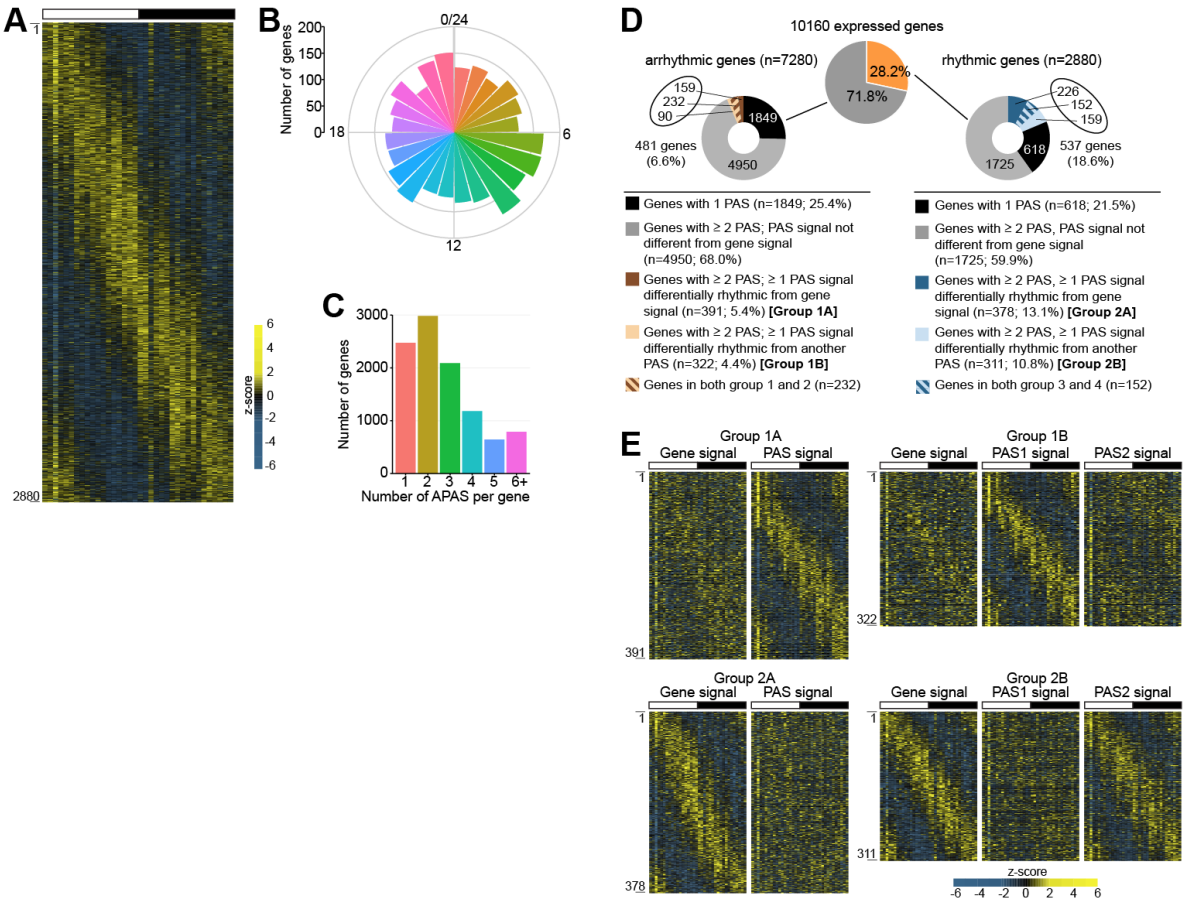

Figure III-2 (Supplement) APAS isoforms exhibit differential rhythmicity in mouse liver

(A) Heatmap of the 2880 genes found to be rhythmic in total RNA. Columns represent timepoints, starting at ZT2 and increasing left to right. Genes are sorted vertically by phase of expression. **(B)** Rose plot of the distribution of phases of the genes rhythmic in total RNA, in 1-hour bins. **(C)** Breakdown of the number of PAS per gene, sorted into bins of 1 or 2+ PAS. **(D)** Analysis of rhythmic expression identifies 4 groups of genes (G1a, G1b, G2a, G2b) that were combined into G1 and G2 (see Methods). **(E)** Heatmap of expression of G1a, G1b, G2a, G2b.

Previous studies have demonstrated that the products of APAS isoforms differing in their 3' UTR length but not in their coding sequence can be located in different cellular membrane compartment, *e.g.*, endoplasmic reticulum (ER) membrane *vs.* plasma membrane[8]. Interestingly, KEGG pathways and cellular

compartment ontology analyses revealed that genes in both groups 1 and 2 were significantly enriched for genes associated with membrane compartments, including lysosome, ER-Golgi, endocytic vesicle, autophagosome, endosome, and plasma membrane (Figure III-H). This suggests that differences in the rhythmicity between APAS isoforms may generate differences in the temporal control of intracellular functions.

Differentially expressed APAS isoforms are enriched for distal and proximal PASs, and are better associated with polysomes

Transcripts with longer 3'UTR have been shown to be under higher post-transcriptional regulation due to increased targeting by miRNA and RBP[25]. To get insights into the mechanisms that underlie differential APAS isoform expression, we mapped all PASs considered in our analyses and categorized them as distal 3'UTR, middle 3'UTR, and proximal 3'UTR based on their relative location across the 3'UTR (Figure III-3A). We also included a category labeled truncated for APAS located upstream the last exon and which generate a truncated protein upon translation. Finally, 1,051 PAS located up to 1 kb downstream the farthest annotated transcription termination site (TTS) were categorized as downstream (Figure III-3A).

Mouse liver genes exhibit a roughly equal distribution of proximal, middle, distal, and truncated PASs, with the remaining 3% being located downstream the annotated TTS (Figure III-3B). Genes with differential APA isoform expression

(groups 1 and 2) were enriched for PAS located in the middle 3'UTR, mostly because they have on average a higher number of PAS per gene (Figure III-3B). While differentially rhythmic APA isoforms were found for every category of PAS, those in group 1 (rhythmic APA isoform with arrhythmic gene signal) were enriched for proximal, distal and downstream PAS, thus suggesting that longer 3'UTR is not the sole factor contributing to rhythmic APA expression (Figure III-3C). On the other hand, differentially rhythmic APA isoforms in group 2 were enriched for distal and downstream PAS, suggesting that longer 3'UTR may contribute to isoform arrhythmic expression in rhythmically expressed genes (Figure III-3C). While not enriched compared to the global distribution, 253 genes exhibit a truncated APA isoform that is differentially rhythmic than the gene signal (109 in group 1, 144 in group 2). As many truncated transcripts result in incomplete proteins missing potential key regulatory domains, these truncated APAS can thus potentially represent dominant negative forms of proteins. We identified APAS for several genes in G1 and G2 that clearly demonstrate expression of an incomplete transcript, as illustrated with the gene *Stbd1* (Figure III-3 D-F).

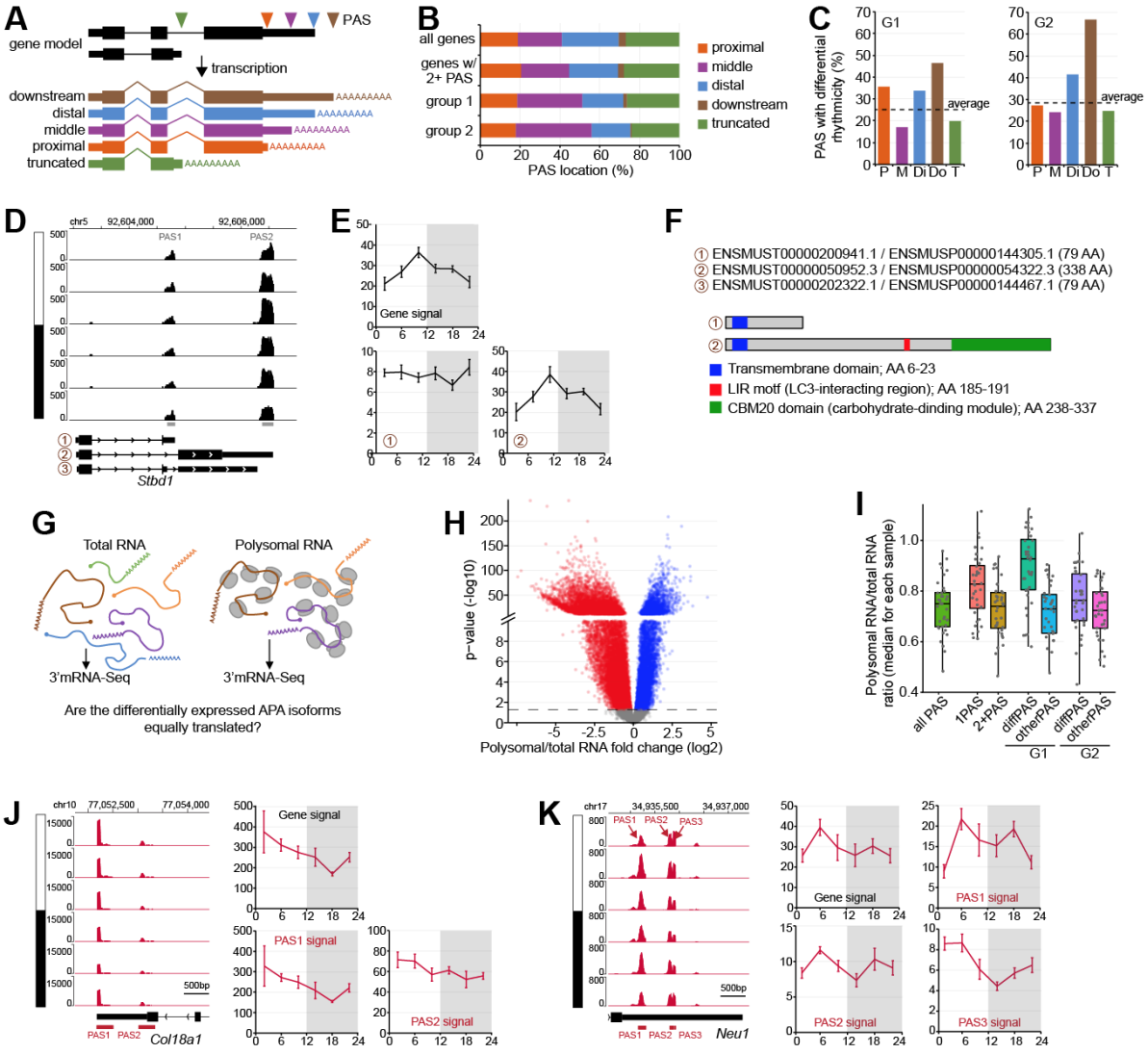

Figure III-3 Differentially expressed APA isoforms are enriched for distal and proximal PASs, and are better associated with polysomes

(A) Diagram of the different transcriptional outcomes of APAS. **(B)** Breakdown of the PAS location by for all genes, only those with 2 or more PAS, and G1 and G2. **(C)** Enrichment of each PAS location for differentially rhythmic PAS within G1 and G2, using the hypergeometric test. **(D-E)** Gene model, PAS location (D), gene expression, and PAS expression (E) of *Stbd1*. **(F)** The three known isoforms of *Stbd1*. **(G)** Overview of polysomal RNA. **(H)** Volcano plot for the log2 fold change of polysomal RNA over total RNA against the DESeq2-reported adjusted p-value (n=36). **(I)** Polysomal:total RNA ratios by sample (n=36) for all PAS, genes with only 1 PAS, genes with 2 or more PAS, differentially expressed G1 and G2, and the non-differentially rhythmic PAS for the same genes represented in G1 and G2. **(J-K)** Gene model, PAS location, gene expression, and PAS expression of *Col18a1* (J) and *Neu1* (K) in polysomal RNA.

To determine if the different APAS isoforms are actively translated, we performed 3' end RNA sequencing on polysomal RNA (Figure III-3G, Figure III-4A). We found that 13,687 (46.9%) of all PAS show a significant decrease in expression in polysomal RNA against total RNA (8,896 with log2 fold-change > 1), indicating that many APAS isoforms are translated less efficiently than suggested by their abundance in total RNA (Figure III-3H). On the other hand, 6,928 (23.7%) of all PAS are significantly increased in polysomal RNA (2,512 with log2 fold-change > 1) (Figure III-3H). As expected, non-coding RNA (ncRNA) such as *Malat1* and *Neat1* are among the least abundant RNA in polysomes (Figure III-4 B-D). Differentially expressed APAS isoforms in groups 1 ($p = 3.11E-12$) and 2 ($p = 6.09E-5$) are both significantly enriched in polysomes when compared to the other APAS isoforms of their parent genes, suggesting that their products are expressed and contribute to differential regulation in their rhythmic function (Figure III-3I). Interestingly, we observed differences in isoform translation for both *Col18a1* and *Neu1*, where one isoform is represented in polysomal RNA at a level similar to that of total RNA, whereas another isoform is far less abundant in polysomal RNA than total (Figure III-3J, Figure III-4E).

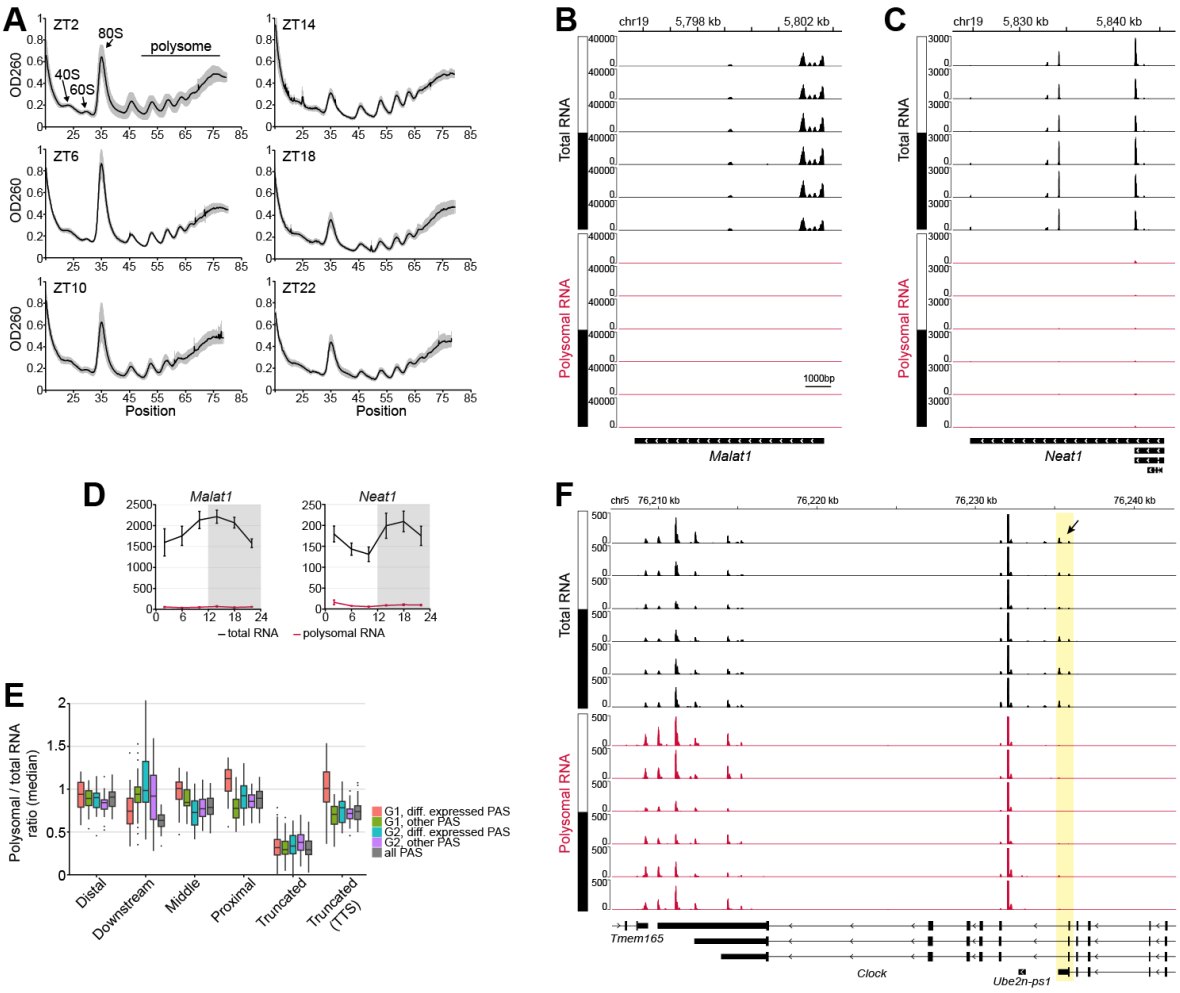

Figure III-4 (Supplement) Differentially expressed APA isoforms are enriched for distal and proximal PASs, and are better associated with polysomes

(A) Average profile of the polysomal RNA at each timepoint (n=6 by timepoint). Shaded area indicates s.e.m. **(B-C)** Gene model and PAS location and expression of two ncRNA, *Malat1* and *Neat1*, in total RNA and polysomal RNA. Very little expression is detected for these two genes in polysomal RNA. **(D)** Gene expression of *Malat1* and *Neat1*. **(E)** Polysomal/total RNA ratios for G1 and G2, the other non-differentially rhythmic PAS for the same genes represented in G1 and G2, and all PAS, separated by outcome of PAS location. **(F)** Gene model and PAS location and expression for *Clock*, which has an intronic PAS that could potentially lead to a dominant negative protein with no binding activity. However, expression of this PAS in polysomes is nonexistent, indicating the protein is not produced.

Post-transcriptional regulation significantly contributes to differential APAS isoform expression

While sequencing of total RNA is commonly used of as a snapshot of transcription, there are a large number of events that occur as an RNA transcript transitions between being transcribed in the nucleus to being available for translation in the cytoplasm. These post-transcriptional modifications represent potential events where differential regulation of transcripts can take place. To determine the contribution of post-transcriptional regulation to differential APAS isoform expression, we performed 3' end RNA-Seq of mouse liver nuclear RNA using the same mice as those used for total 3'end RNA-Seq, and compared nuclear RNA profiles and expression to those originating from total RNA sequencing (Figure III-5A). Comparison of the number of intronic reads between total RNA and nuclear RNA, which mostly originate from intronic polyA stretch being primed by the oligo-dT primer during library first strand synthesis, confirmed that 3'end RNA-Seq of nuclear RNA significantly increased the number of intronic reads (Figure III-6A). The rate at which transcripts are transcribed and degraded differs between transcripts, and differences in the relative amount of RNA transcript present in the nucleus *vs.* that of the cytoplasm can be used as a proxy for RNA half-life, with high nuclear/total RNA ratio indicating shorter half-life. Differential analysis of this ratio with DESeq2 for all mouse liver APAS isoforms revealed widespread differences, with 28.5% PAS (8,328 of 29,199) having a high nuclear/total RNA ratio and 49.4% PAS (14,428 of 29,199) with a low nuclear/total ratio (Figure III-5B).

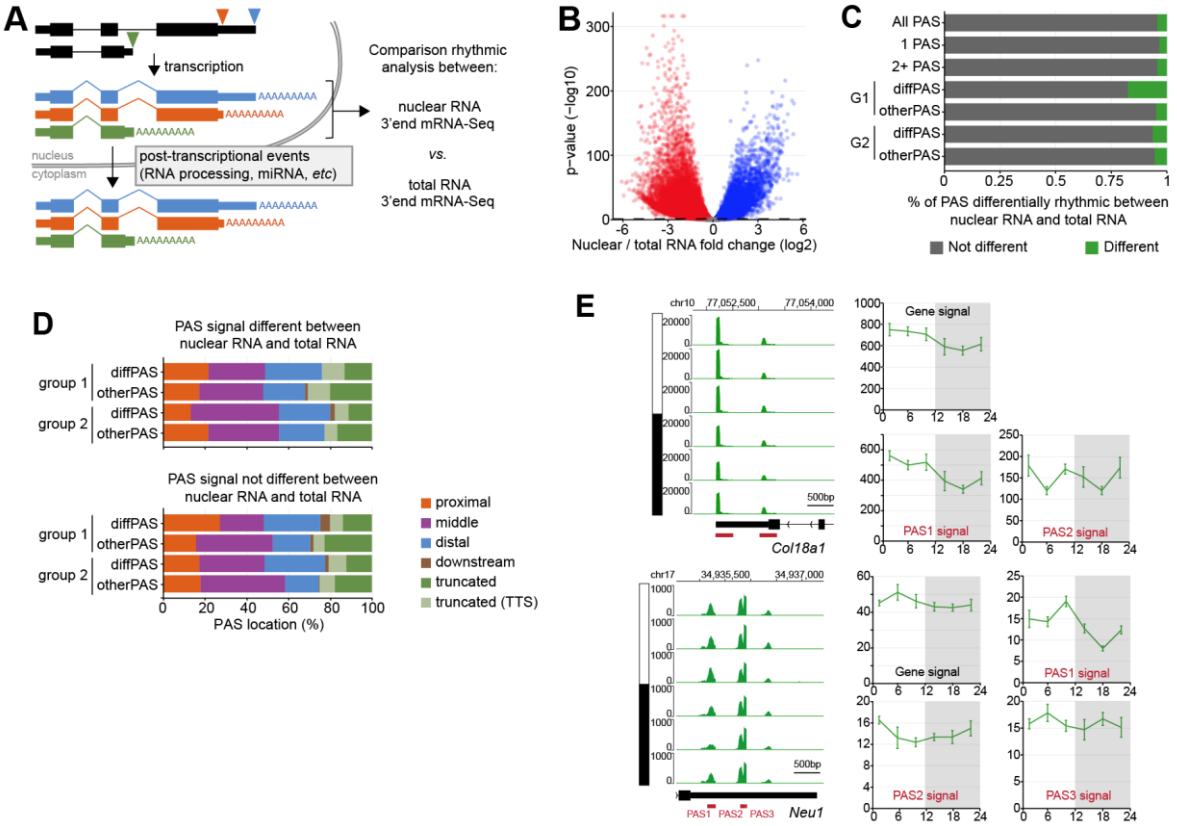

Figure III-5 Post-transcriptional regulation significantly contributes to differential APA isoform expression

(A) Diagram of nuclear RNA against total RNA. **(B)** Volcano plot for the log2 fold change of nuclear RNA over total RNA against the DESeq2-reported adjusted p-value (n=36). **(C)** Stacked bar graphs for the % PAS differentially rhythmic between nuclear RNA and total RNA for the 7 groups. **(D)** Breakdown of the PAS location for G1 and G2, separated by whether or not they are differentially rhythmic between total and nuclear RNA. **(E)** Gene model, PAS location, gene expression, and PAS expression of *Col18a1* and *Neu1* in nuclear RNA.

To determine if post-transcriptional regulation contributes to the differences in APAS isoform rhythmicity, we performed the same two-step rhythmicity analysis with nuclear RNA as that for total RNA (see Methods and above), and compared the nuclear RNA profiles with those of total RNA. We found that differentially rhythmic APAS isoforms in both groups 1 and 2 exhibit more differences in their

rhythmicity between nuclear and total RNA when compared to all APAS isoforms, indicating that post-transcriptional events likely contribute to the differential expression of these transcripts (Figure III-5C). For example, 17.5% of the rhythmic APAS isoforms in group 1 (rhythmic APAS isoform with arrhythmic gene signal) are arrhythmic at the nuclear RNA level, indicating that the rhythmicity of these APAS isoforms at the total RNA level is likely mediated post-transcriptionally. However, the vast majority of APAS isoforms did not display any significant differences in their rhythmicity between nuclear RNA and total RNA, pointing towards other mechanisms that might explain these differences (Figure III-5C). Rhythmic nuclear APAS isoforms exhibiting differences in their rhythmicity between nuclear and total RNA were enriched for distal PAS, whereas arrhythmic nuclear APAS isoforms being rhythmic in the total RNA fraction were enriched for middle PAS (Figure III-5D). Finally, no significant differences in rhythmic expression for any APAS isoforms were observed for the two genes *Col18a1* ($p = 0.23, 0.91$; DODR analysis) and *Neu1* ($p = 0.64, 0.045, 0.88$; DODR analysis) used as examples above, indicating that post-transcriptional modifications are not responsible for the differential expression of their APAS isoforms (Figure III-5E).

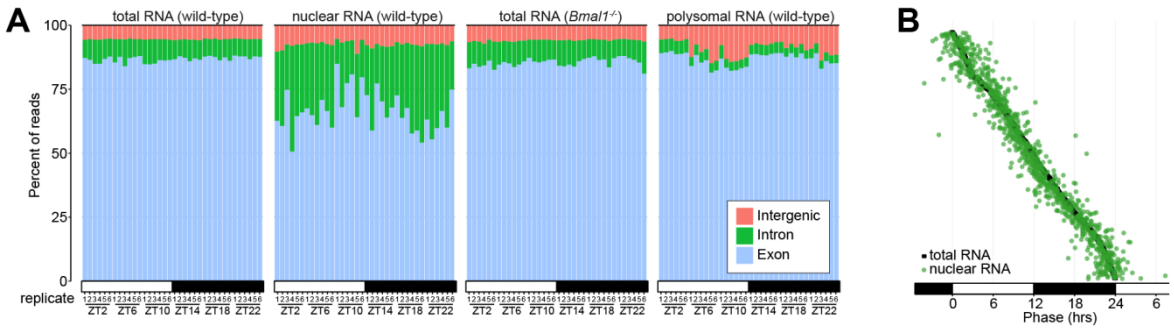

Figure III-6 (Supplement) Post-transcriptional regulation significantly contributes to differential APA isoform expression

(A) Percent of reads aligning to exons, introns, or intergenic for the four types of libraries sequenced in this study. **(B)** Correlation between the phases of rhythmically expressed PAS in total RNA and nuclear RNA.

Cellular subtype specificity contributes to differential APAS isoform expression

Another mechanism that may explain the differences in APAS isoform expression involves cell subtype specificity and cellular environment. While generally considered as a homogenous tissue, the liver is composed of epithelial cells and Kupffer cells in addition to the dominant population of hepatocytes. Moreover, the liver organization into lobules generates a gradient of high oxygen and nutrient conditions (pericentral) to low oxygen and nutrient conditions (periportal), which contributes to hepatocyte subtypes specialized in specific metabolic functions (Figure III-7A)[26,27]. To determine if the differentially expressed APAS isoforms are expressed in a cell subtype specific manner, we utilized a published mouse liver single-cell RNA-Seq dataset that explored differences in gene expression between cell subtypes[28]. This dataset was generated using MARS-Seq and thus sequenced 3' mRNA reads similar to our 3'mRNA-Seq,

thereby enabling the analysis of APAS isoform expression across the different liver cell subtypes.

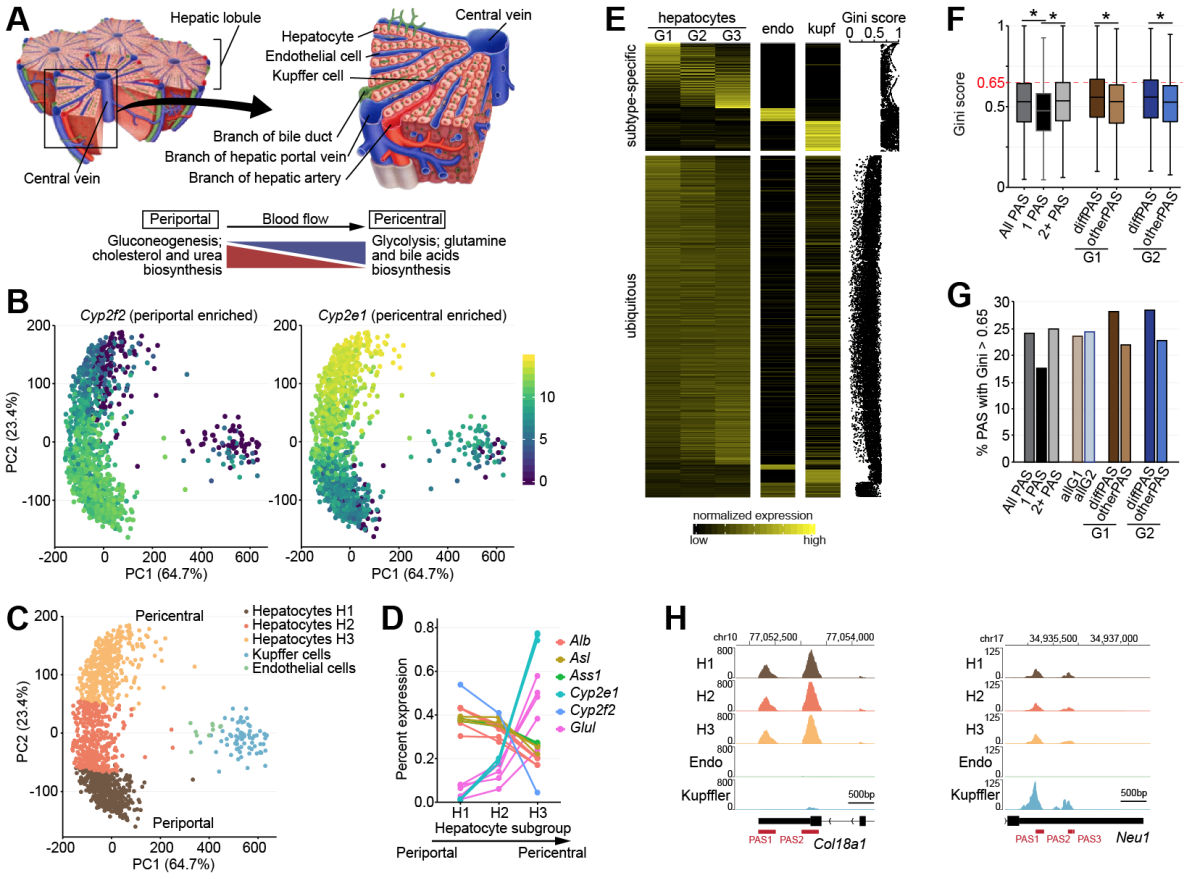

Figure III-7 Cellular subtype specificity also contributes to differential APA isoform expression

(A) Overview of the compartments of the liver, the liver lobules. High oxygen and nutrient blood flows from a triad of the bile duct, hepatic portal vein, and hepatic artery to the central vein leading to the heart. Hepatocytes take up oxygen and nutrients as the blood flows past, leading to a gradient in oxygenation and nutrient status in the blood and surrounding hepatocytes. **(B)** PCA plot of the distribution of single cells based on gene expression, using marker genes for hepatocytes, endothelial, and Kupffer cells. Cells are shaded according to their expression of *Cyp2f2* (left) and *Cyp2e1* (right). **(C)** The same PCA plot, shaded according to assigned group. Hepatocyte group H1 is enriched for periportal hepatocytes, while hepatocyte group H3 is enriched for pericentral hepatocytes. **(D)** PAS expression of the 6 genes used in the single-cell dataset study in hepatocyte subgroups H1-3, matching what was seen on a gene level. **(E)** Percent expression of each PAS in the 3 hepatocyte subgroups, endothelial, and Kupffer cells, such that each row adds up to 100% expression.

Figure III-7 Continued

The Gini coefficient for each PAS is shown on the right. PAS are sorted according to Gini coefficient and which of the 5 subgroups they have the highest expression in. **(F)** Boxplot of the Gini coefficient for all PAS, PAS for genes with only one PAS, PAS for genes with 2 or more PAS, and G1 and G2. **(G)** Percentage of the 7 groups with a Gini coefficient over 0.65. **(H)** Gene model and PAS position for *Col18a1* and *Neu1*, showing expression in the 5 subgroups.

Several genes known to be preferentially expressed at opposing ends of the liver lobules are commonly used to differentiate hepatocyte subtypes. Using these hepatocyte marker genes along with marker genes for Kupffer and endothelial cells, we spatially reconstructed by principal component analysis the liver subtypes to spread out the hepatocytes and identify cell subtypes. Expression analysis of two hepatocyte markers, the cytochrome P450 genes *Cyp2f2* and *Cyp2e1*, across more than 1,000 single cells showed biased expression towards the periportal and pericentral ends of liver lobules, respectively (Figure III-7B). Using this hepatocyte subtype zonation, we then split the hepatocyte population into 3 major groups, resulting in 5 total groups with the identified Kupffer and endothelial cells (Figure III-7C). Subsequent analysis of the APAS isoforms for six known hepatocyte marker genes (*Cyp2f2, Cyp2e1, Alb, Asl, Ass1, Glul*) confirmed that their expression was correctly partitioned across the liver lobule, indicating that our spatial reconstruction of the liver using single-cell data was successful (Figure III-7D).

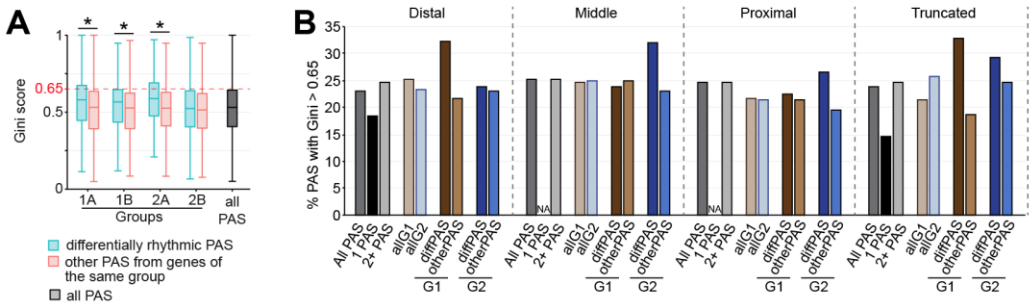

Figure III-8 (Supplement) Cellular subtype specificity also contributes to differential APA isoform expression

(A) Gini coefficient for the original groups 1A, 1B, 2A, and 2B, as well as all PAS. **(B)** Percentage of PAS with a Gini coefficient greater than 0.65, separated by the location of the APAS.

To determine whether APAS isoforms are expressed in specific cell subtypes or not, we sought to use a quantitative measurement that could be applied to the expression of each APAS isoform in all 5 liver subtypes. Several papers have extensively compared many different methods[29,30], yet no strong conclusions were reached on the most appropriate choice. Therefore, based on our own comparison of a few methods, we selected the Gini coefficient method as the most appropriate unbiased method for determining cell subtype specificity in mouse liver[30,31] Unsurprisingly, many APAS isoforms were expressed evenly with some small biases across the 5 groups, resulting in Gini coefficients around 0.5 (Figure III-7E). Quantification of the Gini coefficient for each APAS isoforms revealed that differentially rhythmic APAS isoforms in both group 1 and 2 exhibit significantly higher Gini coefficients than other transcripts, indicating that subtype specific expression contributes to differential rhythmic expression of APAS isoforms (Figure III-7F). To extend this finding, we set a Gini coefficient of 0.65 as the cutoff to

consider an APAS isoform as being expressed in specific cell subtypes or ubiquitously, based on the profile of Gini coefficients across all APAS isoforms, and the Gini coefficient for the different marker genes (Figure III-7E). Using this cut-off, 28% of the differentially rhythmic APAS isoforms in group 1 and 2 were expressed in a cell subtype-specific manner, whereas 22% of the other G1 and G2 APAS isoforms have a Gini coefficient higher than 0.65 (Figure III-7G). This suggests that specific subtype expression may contribute to about 6% of the differential rhythmic expression of APAS isoforms. Gini coefficient for *Col18a1* and *Neu1* APAS isoforms, along with heatmap visualization of their expression across mouse liver subtypes, revealed liver subtype specific expression was unlikely to be involved in the differential rhythmic expression of APAS isoforms (Figure III-7H).

Co-transcriptional regulation of differential APAS isoform expression in mouse liver

Because differential expression between APAS isoforms could not be solely explained by post-transcriptional regulation and cell subtype specific expression, we then examined if co-transcriptional events may contribute to differential APAS isoform expression. Increasing evidence indicates that many factors/proteins loaded onto the C-terminal domain (CTD) of RNA Polymerase II (Pol II) significantly influence the behavior of Pol II, *e.g.*, Pol II pausing and elongation rate [25,32]. These factors, which regulate 5' end RNA capping, RNA splicing, and 3' RNA cleavage and polyadenylation, have all been described to regulate mRNA synthesis and processing, and ultimately to contribute to RNA isoform diversity. For these

reasons, we hypothesized that the expression of APAS isoforms may be co-transcriptionally regulated, with the potential combination of initiating TFs, enhancer(s), and promoter usage influencing the loading of PAS-specific cleavage and polyadenylation factors, and eventually resulting in multiple APAS isoforms exhibiting differences in their relative expression (Figure III-10A). A corollary of this hypothesis implies that knocking out a TF can generate widespread changes in PAS usage by altering promoter events and co-transcriptional loading of RNA processing factors (Figure III-10A). To test this hypothesis, we sequenced the liver transcriptome over the 24-hour day using 3'mRNA-Seq of clock-deficient *Bmal1*$^{-/-}$ mice fed *ad libitum* (n = 36; n = 6 per timepoint), and quantified the expression of APAS isoforms.

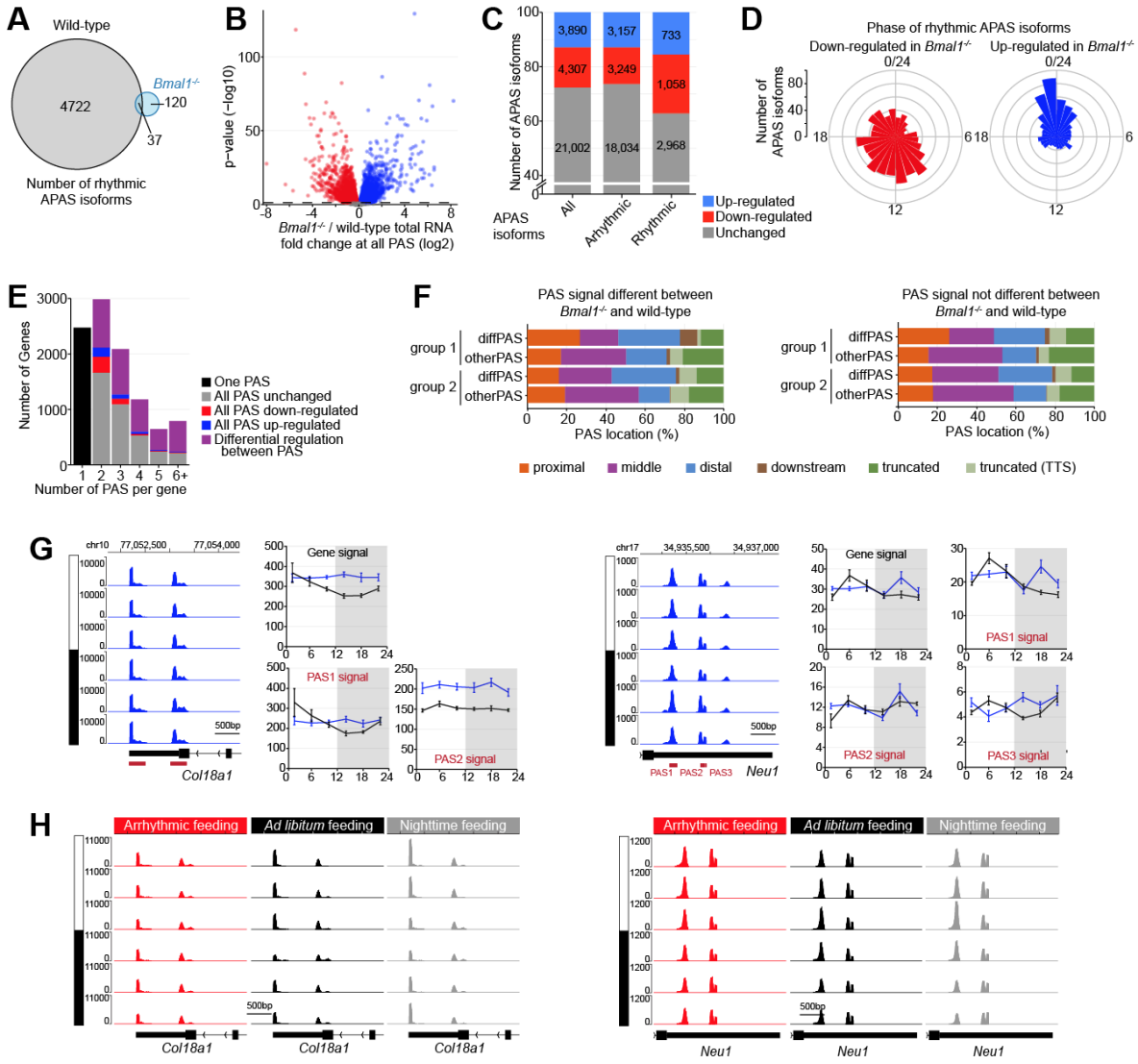

Figure III-9 Co-transcriptional regulation of differential APAS isoform expression in mouse liver

(A) Venn diagram of the PAS rhythmic only in WT total RNA (4722), only in *Bmal1*$^{-/-}$ total RNA (120), or both (37). **(B)** Volcano plot for the log2 fold change of *Bmal1*$^{-/-}$ total RNA over WT against the DESeq2-reported adjusted p-value (n=36). **(C)** The breakdown of genes up-regulated, down-regulated, or not changed from WT to *Bmal1*$^{-/-}$, separated by PAS rhythmicity. Those affected by the KO are enriched for rhythmic PAS (p = 1.38E-55, hypergeometric test). **(D)** Phases of rhythmic PAS significantly down- or up-regulated in *Bmal1*$^{-/-}$. **(E)** Breakdown of genes by how their PAS are affected by *Bmal1*$^{-/-}$ KO. Differential regulation indicates a mixture of effects, between not affected, down-regulated, or up-regulated. **(F)** Breakdown of PAS location for G1 and G2, separated by response to *Bmal1*$^{-/-}$. **(G)** Gene model, PAS location, gene expression, and PAS expression of *Col18a1* and *Neu1* in WT and *Bmal1*$^{-/-}$ total RNA. **(H)** Gene model and PAS signal for *Col18a1* and *Neu1* in arrhythmically fed, *ad libitum*, and night-restricted fed mice[33].

Bmal1 is a core component of the mammalian circadian clock, and its knockout in mouse abolishes molecular, physiological and behavioral rhythms. Not surprisingly, analysis of the rhythmic transcriptome in the liver of *ad libitum* fed *Bmal1*$^{-/-}$ mice revealed that almost all APAS isoforms are arrhythmically expressed (Figure III-9A). However, analysis of differential expression with DESeq2 revealed a profound effect of *Bmal1* knockout on the steady-state expression levels of APAS transcripts (Figure III-9B). About 28.1% of APAS isoforms (8,197 of 29,199) are significantly affected in *Bmal1*$^{-/-}$ mice, with a roughly event split between up- and down-regulation (47.5% vs. 52.5%, respectively; Figure III-9C). Interestingly, APAS isoforms misregulated in *Bmal1*$^{-/-}$ mice were enriched for rhythmic APAS isoforms (37.6% vs. 26.2% between rhythmic *vs.* arrhythmic APAS isoforms; $p = 6.09\times10^{-55}$, hypergeometric test; Figure III-9C). Consistent with the daytime binding of CLOCK:BMAL1 to DNA[34,35] and a direct effect of *Bmal1*$^{-/-}$ on APAS isoform differential expression, APAS isoforms down-regulated in *Bmal1*$^{-/-}$ mice were enriched in transcripts peaking around the day:night transition (ZT12), whereas APAS isoforms up-regulated in *Bmal1*$^{-/-}$ mice were enriched in transcripts peaking at the night:day transition (ZT0) (Figure III-9D).

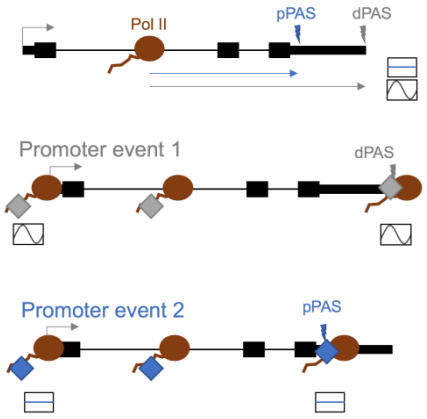

Figure III-10 (Supplement) Co-transcriptional regulation of differential APAS isoform expression in mouse liver

(A) Schematic of the process by which different promoter events may be co-transcriptionally loading elements onto the CTD arm of RNA Polymerase II, resulting in the independent production of different length transcripts and differential regulation of APAS isoforms.

To determine if *Bmal1* knockout consistently up- or down-regulate all APAS isoforms of a gene or if it can differentially affect the expression of APAS isoforms, we examined the 8,197 APAS isoforms misregulated in *Bmal1*$^{-/-}$ mice at a gene level. Of the genes affected in *Bmal1*$^{-/-}$ mice and containing 2 or more PAS, only a minority (10%) are regulated in a similar manner (Figure III-9F). Indeed, most of the genes containing an isoform affected by *Bmal1* knockout exhibit differential APAS isoform expression, *i.e.*, most genes harbored APAS isforms that are either up-, down-, or unchanged in *Bmal1*$^{-/-}$ mice. Importantly, this differential APAS isoform expression is predominant even for genes harboring only two PAS, indicating that this effect is not specific to genes containing a large number of PAS (Figure III-9E). This result thus indicates that knockout of a TF can affect the expression of specific APAS isoforms within a gene. This applies to the gene *Col18a1* (but not *Neu1*),

where one APAS isoform is expressed constitutively higher in *Bmal1*$^{-/-}$ mice than in wild-type mice, while the other isoform is expressed at a level similar between the two strains (Figure III-9G). Whether this effect of *Bmal1*$^{-/-}$ occurs directly, *i.e.*, co-transcriptionally in *cis*, or indirectly, *i.e.*, in *trans* with *Bmal1*$^{-/-}$ misregulating the expression of RBP or miRNA that regulate APA, remains to be determined. However, our findings that *Bmal1* knockout preferentially affects rhythmic APAS isoforms and that down-regulated isoforms peak a few hours after CLOCK:BMAL1 maximal DNA binding in wild-type mice, strongly suggest that some of the effects occur directly in cis and are co-transcriptional.

Taken together, our findings therefore suggest that changes of TF activity *in vivo* may differentially affect the expression of specific APAS isoforms. Moreover, they also suggest that conventional RNA-Seq analysis may underestimate how gene expression is affected by a particular treatment. To test this possibility, we used one of our publicly available 3' mRNA-Seq dataset that examined how the amplitude of feeding rhythms regulates rhythmic gene expression in mouse liver[33], and examined whether APAS isoforms may be differentially regulated by rhythmic food intake. *Col18a1*, and to a lesser extent *Neu1*, both show clear differential expression of their APAS isoforms depending on feeding rhythms, with one APAS isoform being expressed at a similar level across all feeding profiles, while the others show enhanced rhythmic expression in response to increasing amplitude of rhythmic food intake (Figure III-9H). Interestingly, these genes are not regulated in a similar manner. For example, the rhythmic APAS isoform for *Col18a1* is the distal

PAS, whereas it is the proximal PAS for *Neu1*. Thus, environmental factors may play a role in PAS usage, potentially by affecting how TFs are recruited to the chromatin and how they co-transcriptionally regulate PAS usage along with post-transcriptional regulation and cell subtype specific expression.

Discussion

Characterization of the rhythmic transcriptome in multiple tissues and species uncovered the pervasiveness of the circadian system on rhythmic gene expression. More than half of the genome in mammals is rhythmically expressed in at least one tissue, and virtually every biological function is under some sort of circadian regulation. As in other fields, analyses of RNA-Seq datasets have been predominantly performed by concatenating the different transcript variants in a single gene model, thereby assuming that the expression of all isoforms that form a gene is under similar transcriptional control. By analyzing the rhythmic 3'end transcriptome in mouse liver, we found that at least 10% of the genes have APAS isoforms exhibiting differential rhythmic expression. More than 500 rhythmically expressed genes harbor at least one arrhythmically expressed APAS isoform. Importantly, 481 genes analyzed as arrhythmically expressed exhibit a rhythmic APAS isoform, suggesting that these genes may contribute to circadian physiology despite overall gene signal suggesting otherwise. The majority of differential rhythmicity in APAS isoform expression appears to be regulated by co-transcriptional regulation of PAS usage, and to a lesser extent by post-

transcriptional events and cell subtype-specific expression. Importantly, our data also indicate that alteration of the environment, *i.e*., amplitude of feeding rhythms, can differentially impact APAS isoform expression, suggesting that conventional RNA-Seq analysis may not fully unravel the differences in gene expression between samples.

It remains unclear how PAS are selectively cleaved and polyadenylated. Our data indicate that knocking out a single transcription factor can misregulate the expression of over a quarter of all APAS isoforms. A comprehensive analysis focusing on the PAS-specific effects of TF knockouts would enlighten the molecular processes by which TFs regulate the transcription of specific APAS isoforms. As shown in Figure III-9, the APAS transcripts of some genes are wholly affected by *Bmal1* knockout and either up- or down-regulated, while the majority of genes are differentially affected. Therefore, it is conceivable that *Bmal1* is required for the transcription of all APAS transcripts in the former class, whereas it is required for only some of the APAS transcripts for the latter class. Interestingly, the recent findings that rhythmic transcription is associated with rhythmic interactions between CLOCK:BMAL1 enhancers and promoters[36,37] may provide a template for how specific cleavage and polyadenylation factors are loaded onto Pol II C-terminal domain and command termination of specific PAS at specific time of the day. The question also remains as to the combinations of factors required for proper transcription of each APAS isoforms. While our investigation has focused on TFs,

there are still many other factors that play a role in initiating and driving the expression of specific APAS isoforms such as RBPs, miRNAs, and lncRNAs[38-40].

An emerging aspect of APA resides in its drastic effects on protein localization, as exemplified by *Cd47* APAS isoforms where the protein encoded by the short 3' UTR isoform is located in the endoplasmic reticulum while long 3'UTR isoform encode for proteins located in the plasma membrane [8]. Our finding that differentially expressed APAS isoforms are enriched in membrane-associated proteins may indicate a more general role of 3'UTR length in protein subcellular localization. Future experiments investigating the role of our identified APAS isoforms (*e.g.*, *Col18a1* and *Neu1*) and the localization of their protein products could yield exciting prospects for not just regulation of APA, but also cell homeostasis and organismal health.

Finally, it is tempting to imply that our results also have deep implications on future studies looking into genome-wide regulation of gene expression. As mentioned above, APA can play a large role in understanding health risks and diseases [13,41]. However, the vast majority of RNA-Seq based gene expression studies have ignored isoform-specific information. As shown in the present study, a sizeable fraction of transcript isoforms can be differentially regulated, shedding light on the notion that not all gene isoforms behave in a similar manner. Since the vast majority of mammalian genes contain some form of APA [42], these studies may therefore miss crucial aspects in their analysis and results.

Methods

Animals

C57BL6/J and *Bmal1*$^{-/-}$ mice were raised in-house on a 12 hour light : 12 hour dark cycle (LD12:12), and maintained on *ad libitum* water and food. Mice were anesthetized with isoflurane, decapitated, and the liver collected. The left lateral lobe was cut into three equivalent-sized pieces for RNA processing, with the remainder stored together. All collected tissues were flash-frozen in liquid nitrogen and stored at -80°C. All animals were used in accordance with the guidelines set forth by the Institutional Animal Care and Use Committee (IACUC) of Texas A&M University (AUP #2019-0222).

Nuclear RNA isolation

50 to 100mg of previously frozen mouse liver tissue was suspended in 1X PBS and transferred to a 2mL glass homogenizer. Tissues were homogenized with pestle A 6 times and pestle B 4 times and then aliquoted into 2mL Eppendorf tubes containing 1mL homogenate solution (). Nuclei were washed twice with the hypotonic buffer, separated by centrifugation at 1500 g for 2 minutes at 4°C. Nuclei were then resuspended in hypotonic buffer, added to a sucrose cushion, and centrifuged at 20000 g for 10 minutes at 4°C. The supernatant was carefully removed and the nuclei washed twice with resuspension buffer, centrifuging at 1500 g for 2 minutes at 4°C, before being used in RNA extraction.

Polysomal RNA isolation

50 to 100 mg of previously frozen mouse liver (crudely ground) tissue was manually ground under LN2 with a mortar and pestle to a fine powder with 8 volumes of polysome buffer (20 mM HEPES pH 7.5, 250 mM NaCl, 10 mM MgCl2, 1% Triton X-100, 10 mM DTT, 10% glycerol, 100 µg/ml cycloheximide, 0.15 mg/ml PMSF, 0.2 mg/ml heparin, 1 µg/ml each of pepstatin, leupeptin and aprotinin). Ground tissue was thawed on ice and cellular debris was pelleted by centrifugation at 10,000 g for 10 m at 4 C. The cleared supernatant was measured for RNA content (Nanodrop) and 12-15 A260 units were loaded on a 13 ml sucrose density gradient (10 to 50% sucrose in 10 mM HEPES pH 7.5, 70 mM ammonium acetate, 5 mM magnesium acetate buffer prepared on a Biocomp Gradient Station) and spun for 2 hours in a pre-cooled SW-41 rotor in a Beckman Coulter Optima XE-90 ultracentrifuge at 41,000 rpm followed by approximately 1hr of a no-brake slow to stop. Gradient fractions were then collected in 14 tubes from the top (low density) to the bottom (high density) on a Biocomp fraction collector with continuous A260 monitoring for RNA content with a Triax flow cell (Biocomp).

RNA extraction and processing

RNA from all sources (total, nuclear, or polysomal) was extracted in the same manner. Total RNA for both WT and BMAL1$^{-/-}$ mice was generated from one-third of the left lateral lobe of the liver flash-frozen in liquid nitrogen and stored at -80°C and crushed using a mortar and pestle. Nuclear and polysomal RNA were

extracted from solution from their respective isolations. Briefly, either crushed frozen tissue or solution was mixed with 300μL of TRIzol reagent, homogenized with a pellet mixer, and the volume brought to 1mL with 700mL of TRIzol reagent. 200mL chloroform was added, and the solution shaken and centrifuged at 12,000 g for 15 minutes at 4C. The aqueous phase was extracted and added to an equivalent amount of isopropanol. The resulting solution was that centrifuged at 12,000 g for 10 minutes at 4C, and the RNA pellet was washed with 1mL of 75% ethanol before being resuspended with 25mL RNase-free deionized water. Total RNA was then purified with an acid phenol/chloroform extraction, and precipitated by ethanol precipitation. The RNA pellet was then washed with 75% ethanol as described above, and finally resuspended in 25mL. Samples were quantified with a NanoDrop-1000 and with the Promega QuantiFluor ssRNA system, and quality / integrity of total RNA was assessed by gel electrophoresis.

Library generation and sequencing

RNA-Seq libraries were generated using the Lexogen QuantSeq 3' mRNA-Seq Library Prep Kit following manufacturer instructions, beginning with 2μg of total RNA as starting material. cDNA was PCR-amplified for 12 cycles following manufacturer recommendations for mouse liver tissue. Libraries were multiplexed in equimolar concentrations and sequenced across multiple runs using an Illumina NextSeq 500 (Molecular Genomics Workspace, Texas A&M University, USA).

Data processing

Sequenced reads were pre-processed with the R package ShortRead [43] to remove the first 12nt, remove low quality bases at the 30 end, trim poly-A tails and embedded poly-A sequences, and remove all reads under 36nt in length. Reads were aligned to the mm10 transcriptome, assembly GRCm38.p4, with the STAR aligner [44] version 2.5.2b with options

--outSAMstrandField intronMotif–quantMode GeneCounts –outFilterIntronMotifs RemoveNoncanonical

Secondary alignments were removed with samtools view -F 0x100. Read counts were summarized with the function summarizeOverlaps from the R package GenomicRanges [45] using options

mode=IntersectionStrict inter.feature=FALSE

Libraries were filtered and normalized by library size using the Trimmed Mean of M-values (TMM) normalization [46] within edgeR [47] using default settings.

PAS mapping

Initial PAS definition was performed through the combination of two separate analyses. First, all total RNA reads as well as the reads from the 72 samples from [33] (=108 separate libraries) were trimmed to their 3' most mapped nucleotide, taking into account the CIGAR string. In the first analysis, the 3’ nucleotides were immediately put through peak calling by HOMER [48] to find a broad range of potential APAS using the following options:

makeTagDirectory -precision 3 -totalReads all -fragLength 1 -keepAll

findPeaks -strand separate -tbp 0 -fragLength 1 -size 10 -minDist 25 -

ntagThreshold 2 –region

resulting in 76018 prospective PAS. In the second analysis we attempted to define all APAS that represented the exact end of transcripts. Therefore, all 3' most nucleotide reads were filtered to only those containing at least 6 consecutive adenine residues at the 3' end in the original unmodified read, indicating that these reads are directly against the poly(A) tail. Next, we scanned the genome 20nt downstream of the mapped 3' end of these reads and removed those with 12 or more adenine residues in those 20nt, indicating that they existed due to internal priming events. Reads in both steps were filtered out using custom Perl scripts. Finally, PAS were defined through HOMER using the following options:

makeTagDirectory -precision 3 -totalReads all -fragLength 1 -keepAll

findPeaks -strand separate -tbp 0 -fragLength 1 -size 10 -minDist 25 -

ntagThreshold 5 –region

resulting in 31837 prospective APAS. The two sets of APAS were then concatenated and combined to yield a single list using the reduce function from GenomicRanges[45]. Finally, any APAS that were found to overlap with the mm10 blacklisted regions generated by ENCODE were removed[49,50]. A total of 86780 possible APAS resulted from these steps.

PAS filtering

PAS were annotated to genes using their overlaps with each gene part with summarizeOverlaps from GenomicRanges [45], with a final annotation performed using a priority list. TSS were defined as the region +0 to +100nt from the annotated TSS. TTS were defined as the region -20nt to +20nt from the annotated TTS. Finally, downstream regions were defined as up to 2kb from the end of the TTS above. Next, we removed all PAS that contained genomic poly(A) stretches that escaped detection in the previous steps. The region from -15 to +5 of the 3' end of every PAS was analyzed, and those with 12 or more adenine residues were removed. Since the downstream annotation can potentially result in ascribing APAS of a downstream gene to the upstream and unrelated gene, we removed all PAS that were annotated as downstream yet interior of another gene. Raw count values were then normalized by library size using the Trimmed Mean of M-values (TMM) normalization[46] within edgeR[47] using default settings.

We next looked at how well all of the PAS of each gene accounted for the gene expression seen in total RNA on a log2-scale. Any genes where the sum of expression of its PAS that overlap the CDS was less than half of its total expression or more than the total expression + 0.5 were removed (Figure III-B). Finally, we looked at the expression of each PAS and auto-included them if their contribution to any transcript was more than 10% of total, or if it had a mean TMM value over 5. Any PAS that had a maximum contribution to any transcript under 0.5% of total as well as a mean TMM value under 0.5 was automatically discarded. All PAS leftover

between these two ranges were tested for overlap with a publicly available database of APAS[51]. APAS from PolyA_DB were extended upstream to a total of 75nt, and any of our APAS that overlapped with those from PolyA_DB were kept. After filtering, a total of 29199 high-confidence APAS remained.

Rhythmicity and differential rhythmicity analysis

Rhythmicity analysis for every gene and PAS in the 4 paradigms was performed with four algorithms from three programs: F24[52,53], JTK_CYCLE and LS from MetaCycle[54], and HarmonicRegression[55]. The resulting p-values from all 4 algorithms were combined using Fisher's method into one p-value, all of which were then adjusted using the Benjamini-Hochberg method[56] within the p.adjust function available in base R to control for the false-discovery rate (FDR). Genes with a q-value under 0.05 were considered rhythmic for that paradigm.

Differential rhythmicity analysis was performed with the robustDODR algorithm within DODR[24]. Genes and PAS with a p-value less than 0.05 were considered as differentially rhythmic.

Differential expression analysis

All differential expression analysis was performed using DESeq2[57] using default settings.

Single cell reconstruction

All methods indicated in the original paper[28] were attempted to be followed as closely as possible. Single-cell data were downloaded from GEO and aligned to the mm10 transcriptome as above. Gene expression for all genes as well as the ERCC92 spike-in was performed using summarizeOverlaps from the GenomicRanges package[45] with the following options:

mode = "IntersectionStrict", singleEnd = TRUE, ignore.strand = FALSE,
inter.feature = FALSE

Cells were removed if the total of their reads mapping to ERCC92 was greater than 4% of the total of reads mapping to the genome[58]. Genes were removed from consideration if the total number of reads mapping to them was 0. All libraries were then normalized to library depth using the TMM normalization as above. Next, we summed the expression of markers for hepatocytes (*Apoa1, Apob, Pck1, G6pc, Ttr)*, endothelial cells (*Kdr, Egf17, Igfbp7, Aqp1*), and Kupffer cells (*Clec4f, Csf1r, C1qc, C1qa, C1qb*). Cells with a greater total for the endothelial or Kupffer marker genes were labeled as endothelial or Kupffer cells, respectively. Cells that had a higher total for both over the hepatocyte marker genes were discarded. All remaining cells were labeled as hepatocytes. Any hepatocyte cell with less than 1% of total expression coming from albumin (*Alb*) was discarded. In order to separate hepatocytes by their zonation profile, PCA analysis was performed on all hepatocytes using their expression for the 20 marker genes indicated in[27], and hepatocytes were split into three even groups (H1, H2, H3) based on their PC1

value. Finally, another PCA analysis was performed on all cells using a combination of the Braeuning hepatocyte, Kupffer, and endothelial marker genes (Figure III-7C).

For PAS expression in the single-cell data, the 3' end of all reads was taken and quantified across all accepted PAS (see PAS generation and filtering). The raw reads for all 5 groups (H1, H2, H3, Kupffer, and endothelial) were summed up. Due to the differing chemistries between 3' QuantSeq and MARS-Seq, some PAS did not match well, and so any PAS with fewer than 5 reads in total was removed. The remaining PAS were then TMM-normalized as above.

KEGG and GO analyses

All KEGG and GO analyses were performed using the kegga and goana functions available within limma[59]. Gene symbols were converted to Entrez IDs with AnnotationDbi[60].

Polysomal/total ratio

The polysomal/total RNA ratio was calculated by dividing expression in polysomal RNA by its expression in the equivalent sample from total RNA for every sample (n=36) and PAS. The resulting ratios were then grouped according to the 7 groups (Figure III-3I) and the median taken for every group by sample. The resulting 36 medians per group were then used for plotting and statistics.

References

1 Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470-476, doi:10.1038/nature07509 (2008).

2 Reyes, A. & Huber, W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Res* **46**, 582-592, doi:10.1093/nar/gkx1165 (2018).

3 Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* **18**, 18-30, doi:10.1038/nrm.2016.116 (2017).

4 Tian, B. & Manley, J. L. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* **38**, 312-320, doi:10.1016/j.tibs.2013.03.005 (2013).

5 Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A. & Burge, C. B. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**, 1643-1647, doi:10.1126/science.1155390 (2008).

6 Gong, C. & Maquat, L. E. lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* **470**, 284-288, doi:10.1038/nature09701 (2011).

7 Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464-469, doi:10.1038/nature07488 (2008).

8 Berkovits, B. D. & Mayr, C. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. *Nature* **522**, 363-367, doi:10.1038/nature14321 (2015).

9 Modic, M. *et al.* Cross-Regulation between TDP-43 and Paraspeckles Promotes Pluripotency-Differentiation Transition. *Mol Cell* **74**, 951-965 e913, doi:10.1016/j.molcel.2019.03.041 (2019).

10 Ye, J. & Blelloch, R. Regulation of pluripotency by RNA binding proteins. *Cell Stem Cell* **15**, 271-280, doi:10.1016/j.stem.2014.08.010 (2014).

11 Weng, T. *et al.* Cleavage factor 25 deregulation contributes to pulmonary fibrosis through alternative polyadenylation. *J Clin Invest* **129**, 1984-1999, doi:10.1172/JCI122106 (2019).

12 Rehfeld, A., Plass, M., Krogh, A. & Friis-Hansen, L. Alterations in polyadenylation and its implications for endocrine disease. *Front Endocrinol (Lausanne)* **4**, 53, doi:10.3389/fendo.2013.00053 (2013).

13 Gruber, A. J. & Zavolan, M. Alternative cleavage and polyadenylation in health and disease. *Nat Rev Genet* **20**, 599-614, doi:10.1038/s41576-019-0145-z (2019).

14 Mayr, C. & Bartel, D. P. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673-684, doi:10.1016/j.cell.2009.06.016 (2009).

15 Xu, C. & Zhang, J. Alternative Polyadenylation of Mammalian Transcripts Is Generally Deleterious, Not Adaptive. *Cell Syst* **6**, 734-742 e734, doi:10.1016/j.cels.2018.05.007 (2018).

16 Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111**, 16219-16224, doi:10.1073/pnas.1408886111 (2014).

17 Cox, K. H. & Takahashi, J. S. Circadian clock genes and the transcriptional architecture of the clock mechanism. *Journal of Molecular Endocrinology* **63**, R93-R102, doi:10.1530/jme-19-0153 (2019).

18 Preussner, M. & Heyd, F. Post-transcriptional control of the mammalian circadian clock: implications for health and disease. *Pflugers Arch* **468**, 983-991, doi:10.1007/s00424-016-1820-y (2016).

19 Kojima, S., Shingle, D. L. & Green, C. B. Post-transcriptional control of circadian rhythms. *J Cell Sci* **124**, 311-320, doi:10.1242/jcs.065771 (2011).

20 Rudic, R. D. *et al.* BMAL1 and CLOCK, two essential components of the circadian clock, are involved in glucose homeostasis. *PLoS Biol* **2**, e377, doi:10.1371/journal.pbio.0020377 (2004).

21 Shimba, S. *et al.* Brain and muscle Arnt-like protein-1 (BMAL1), a component of the molecular clock, regulates adipogenesis. *Proc Natl Acad Sci U S A* **102**, 12071-12076, doi:10.1073/pnas.0502383102 (2005).

22 Lamia, K. A., Storch, K. F. & Weitz, C. J. Physiological significance of a peripheral tissue circadian clock. *Proc Natl Acad Sci U S A* **105**, 15172-15177, doi:10.1073/pnas.0806717105 (2008).

23 Chaix, A., Lin, T., Le, H. D., Chang, M. W. & Panda, S. Time-Restricted Feeding Prevents Obesity and Metabolic Syndrome in Mice Lacking a Circadian Clock. *Cell Metab*, doi:10.1016/j.cmet.2018.08.004 (2018).

24 Thaben, P. F. & Westermark, P. O. Differential rhythmicity: detecting altered rhythmicity in biological data. *Bioinformatics* **32**, 2800-2808, doi:10.1093/bioinformatics/btw309 (2016).

25 Di Giammartino, D. C., Nishida, K. & Manley, J. L. Mechanisms and consequences of alternative polyadenylation. *Mol Cell* **43**, 853-866, doi:10.1016/j.molcel.2011.08.017 (2011).

26 Rappaport, A. M., Borowy, Z. J., Lougheed, W. M. & Lotto, W. N. Subdivision of hexagonal liver lobules into a structural and functional unit. Role in hepatic physiology and pathology. *The Anatomical Record* **119**, 11-33, doi:10.1002/ar.1091190103 (1954).

27 Braeuning, A. *et al.* Differential gene expression in periportal and perivenous mouse hepatocytes. *FEBS J* **273**, 5051-5061, doi:10.1111/j.1742-4658.2006.05503.x (2006).

28 Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* **542**, 352-356, doi:10.1038/nature21065 (2017).

29 Yanai, I. *et al.* Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650-659, doi:10.1093/bioinformatics/bti042 (2005).

30 Kryuchkova-Mostacci, N. & Robinson-Rechavi, M. A benchmark of gene expression tissue-specificity metrics. *Brief Bioinform* **18**, 205-214, doi:10.1093/bib/bbw008 (2017).

31 Wright Muelas, M., Mughal, F., O'Hagan, S., Day, P. J. & Kell, D. B. The role and robustness of the Gini coefficient as an unbiased tool for the selection of Gini genes for normalising expression profiling data. *Sci Rep* **9**, 17960, doi:10.1038/s41598-019-54288-7 (2019).

32 Proudfoot, N. J. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* **352**, aad9926, doi:10.1126/science.aad9926 (2016).

33 Greenwell, B. J. *et al.* Rhythmic Food Intake Drives Rhythmic Gene Expression More Potently than the Hepatic Circadian Clock in Mice. *Cell Rep* **27**, 649-657 e645, doi:10.1016/j.celrep.2019.03.064 (2019).

34 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).

35 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).

36 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).

37 Beytebiere, J. R. *et al.* Tissue-specific BMAL1 cistromes reveal that rhythmic transcription is associated with rhythmic enhancer-enhancer interactions. *Genes Dev* **33**, 294-309, doi:10.1101/gad.322198.118 (2019).

38 Liu, Y. *et al.* Cold-induced RNA-binding proteins regulate circadian gene expression by controlling alternative polyadenylation. *Sci Rep* **3**, 2054, doi:10.1038/srep02054 (2013).

39 Liaw, H. H., Lin, C. C., Juan, H. F. & Huang, H. C. Differential microRNA regulation correlates with alternative polyadenylation pattern between breast cancer and normal cells. *PLoS One* **8**, e56958, doi:10.1371/journal.pone.0056958 (2013).

40 Xiao, R. *et al.* Pervasive Chromatin-RNA Binding Protein Interactions Enable RNA-Based Regulation of Transcription. *Cell* **178**, 107-121 e118, doi:10.1016/j.cell.2019.06.001 (2019).

41 Creemers, E. E. *et al.* Genome-Wide Polyadenylation Maps Reveal Dynamic mRNA 3'-End Formation in the Failing Human Heart. *Circ Res* **118**, 433-438, doi:10.1161/CIRCRESAHA.115.307082 (2016).

42 Tian, B., Hu, J., Zhang, H. & Lutz, C. S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* **33**, 201-212, doi:10.1093/nar/gki158 (2005).

43 Morgan, M. *et al.* ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* **25**, 2607-2608, doi:10.1093/bioinformatics/btp450 (2009).

44 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

45 Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS Comput Biol* **9**, e1003118, doi:10.1371/journal.pcbi.1003118 (2013).

46 Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25, doi:10.1186/gb-2010-11-3-r25 (2010).

47 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

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

49 Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**, 355-364, doi:10.1038/nature13992 (2014).

50 Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep* **9**, 9354, doi:10.1038/s41598-019-45839-z (2019).

51 Wang, R., Nambiar, R., Zheng, D. & Tian, B. PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res* **46**, D315-D319, doi:10.1093/nar/gkx1000 (2018).

52 Wijnen, H., Naef, F. & Young, M. W. Molecular and Statistical Tools for Circadian Transcript Profiling. **393**, 341-365, doi:10.1016/s0076-6879(05)93015-2 (2005).

53 Hutchison, A. L. *et al.* Improved statistical methods enable greater sensitivity in rhythm detection for genome-wide data. *PLoS Comput Biol* **11**, e1004094, doi:10.1371/journal.pcbi.1004094 (2015).

54 Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K. & Hogenesch, J. B. MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics* **32**, 3351-3353, doi:10.1093/bioinformatics/btw405 (2016).

55 Hughes, M. E. *et al.* Harmonics of Circadian Gene Transcription in Mammals. *PLoS Genetics* **5**, e1000442, doi:10.1371/journal.pgen.1000442 (2009).

56 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).

57 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

58 Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol* **32**, 896-902, doi:10.1038/nbt.2931 (2014).

59 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47, doi:10.1093/nar/gkv007 (2015).

60 Pagès, H., Falcon, S., Carlson, M. & Li, N. AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor. *R package version 1.48.0*, doi:10.18129/B9.bioc.AnnotationDbi (2019).

CHAPTER IV

CONCLUSIONS, DISCUSSION, AND FUTURE DIRECTIONS

Rhythmic food intake and rhythmic gene expression

In Chapter II, I showed that mice fed on an arrhythmic schedule do not exhibit a shift in the phase of the circadian clock in the liver. The same also occurs under night-restricted feeding (NRF), where the phase of the clock does not shift. However, the amplitude of expression of many rhythmic genes increases, resulting in strongly rhythmic gene expression. Under day-restricted feeding (DRF), the phase of both the clock and many rhythmic genes shifts by up to 12 hours in the mouse liver[1,2]. Interestingly, the effects of NRF *vs*. DRF on rhythmic expression in clock deficient mice are not equal. NRF in clock-deficient mice results in many genes gaining or maintaining rhythmicity[3], while DRF results in far fewer rhythmically expressed genes than NRF[2]. If a clock-deficient mouse is completely devoid of rhythms, then there should be no functional difference between the two feeding paradigms. However, one additional factor is that the DRF clock-deficient mice were in constant darkness (DD), while the NRF clock-deficient mice were under 12 hours light:12 hours dark (LD). Recent studies have shown that light itself can have effects on tissues other than the SCN in systemic signals directing rhythmic gene expression[4]. Therefore, this discrepancy is potentially due to systemic signals from elsewhere in the body that maintain rhythmicity under clock

deficiency, such as body temperature or light, interacting with RFI-dependent gene expression.

It is unclear how the phase of core clock gene expression is shifted under DRF in response to rhythmic food intake (RFI). One possible avenue is that *Per2*, known for mediating some interactions between the circadian clock and nutrient status[5], may be initiating a reset of the circadian clock in response to feeding exclusively during the resting phase. A recent paper[6] explored this avenue and found that DRF increased levels of corticosterone at the end of the night, resulting in a cascade that ultimately delays the activation of *Rev-erbα*. Interestingly, this change happens quickly within the liver, but takes several days longer in the heart and muscle. Furthermore, this change has been shown to be dependent on glucocorticoids, where changing from NRF to DRF changes peripheral clock gene expression slowly, but the opposite happens rapidly[7].

The vast majority of studies into circadian rhythms and their relationship with feeding have been studied in the liver, which is extensively influenced by feeding, as nutrient-laden blood enters the liver via the portal vein. As such, the liver is in a unique position amongst tissues where much of the body's metabolism takes place. Previous papers have performed a cursory analysis into the rhythmicity of transcripts in other tissues[8,9]. However, it is relatively unknown how rhythmic food intake affects these other tissues, as well as what signal is thus synchronizing the response of these tissues to feeding. Some efforts have been made previously to identify metabolites that act upon the liver clock in response to food intake[10], but it is

unknown if these metabolites will have the same effect upon other tissues given the liver's unique position[11]. Therefore, I would like to investigate the response of an extensive list of tissues to rhythmic food intake. I would expect to see that tissues upstream as far as nutrient flow such as the liver and kidneys would contain more genes that depend on RFI for rhythmic gene expression, whereas tissues downstream would contain few, if any, genes that depend on RFI for expression. Instead, I would expect those tissues to contain primarily clock-controlled genes that will not deviate in rhythmic expression in response to RFI. Part of this work has already been completed in our lab in the heart, kidneys, and lungs, but an in-depth analysis of rhythmic gene expression has not yet been performed. Furthermore, similar to Chapter II, rhythmic gene expression in a NRF *Bmal1*$^{-/-}$ mouse would need to be assayed to identify those genes that are solely dependent on the clock for their expression. To add to this, I would like to investigate the synchronizing metabolite(s) that are responsible for driving rhythmic gene expression in other tissues in response to RFI. This would be performed by acute injection of potential metabolites using metabolite libraries or a vehicle, and assaying for changes in the expression of genes shown to be affected by RFI in the previous genome-wide analysis.

Pathways interacting with rhythmic food intake

As part of our conclusions for Chapter II (Figure II-7K), we noted that while total levels of mTOR, ERK1, and ERK2 are relatively consistent amongst all feeding

paradigms, levels of p-mTOR, p-ERK1, and p-ERK2 are rhythmic under NRF, and arrhythmic under arrhythmic feeding, suggesting that these pathways play a role in regulating the livers response to RFI. As discussed above, NRF (and DRF, to a lesser extent) have the ability to synchronize rhythmic gene expression, even in a clock-deficient mouse. However, NRF after an adrenalectomy does not provide the same synchronizing effect, indicating that systemic signals such as the hormone corticosterone must still be playing a role in the response to rhythmic food intake[12]. Therefore, one potential future direction would be to feed mice under the three feeding paradigms (AR, LB, NRF) while they are under the effects of antagonists for the mTOR (e.g., rapamycin) or ERK1/2 pathways[13]. If either of these pathways are directly or indirectly involved in mediating the response of rhythmic gene expression to RFI, then feeding mice under a NRF or *ad libitum* schedule will result in a phenotype and gene expression profile similar to that of arrhythmically-fed mice. Knowledge of the pathway involved in the body's response to RFI could pave the way to gaining the health benefits of time-restricted feeding, or "intermittent fasting" as the diet is publicly known[14], without requiring the use of the diet.

These experiments could also be repeated in clock-deficient mice. This should result in mice that contain very few rhythmic genes, as they cannot entrain gene expression to both clock-controlled signaling and RFI signaling, leaving only those genes that are sensitive to the remaining rhythmic systemic signals.

Directing polyadenylation events via co-transcriptional loading

In Chapter III, I found that many APAS transcripts are differentially expressed from their sibling transcripts, *i.e*., from APAS transcripts belonging to the same gene that are not differentially expressed. We then investigated the underlying mechanisms, and examined APA using nuclear RNA as a proxy for post-transcriptional modifications (Figure III-5), as well as across different tissue sub-types in a single-cell RNA-Seq dataset (Figure III-7). We did not find strong evidence indicating these two possible mechanisms are the main causes for the differential expression of APAS transcripts. Next, we examined expression of these APAS transcripts in a *Bmal1* KO strain (Figure III-9) to determine if the differences in transcript expression were due to co-transcriptional regulation, and noted over a quarter of our APAS were significantly up- or down-regulated upon *Bmal1* KO. The vast majority of these changes do not occur across all APAS transcripts of each gene, *i.e.* genes contain APAS that are differently regulated from each other in response to *Bmal1* KO. Thus, co-transcriptional factors influence the usage of APAS, resulting in differential expression between isoforms. These experiments could be repeated on a larger scale, involving many more transcription factor knockouts, to build a greater picture of the TF network that combine to activate transcription for each PAS.

Studies have shown that pausing of RNA Pol II is correlated with an increase in production of proximal PAS transcripts, likely due to the CPSF proteins on the CTD domain of Pol II being in close proximity to the PAS for longer periods of

time[15,16]. It has been hypothesized that certain transcription factors may be responsible for the usage of certain polyadenylation proteins during transcription[17]. Additionally, between initiation of transcription and elongation, Pol II pauses as it disengages from the promoter. During this time, the CTD domain of Pol II is available for CPSFs[18], as well as other factors such as RBPs, to be loaded onto the domain to aid in the process of transcription. In our model, we hypothesize that certain factors that are co-transcriptionally loaded onto the CTD arm are what actually direct Pol II cleavage and polyadenylation at certain APAS over other PAS (Figure IV-). In support of our hypothesis, a recent genome-wide survey of chromatin associated RBP found that many (but not all) RBPs are recruited at TSS and enhancers, that RBP binding to DNA is likely mediated by TFs, and thus occur at specific enhancers/promoters[19].

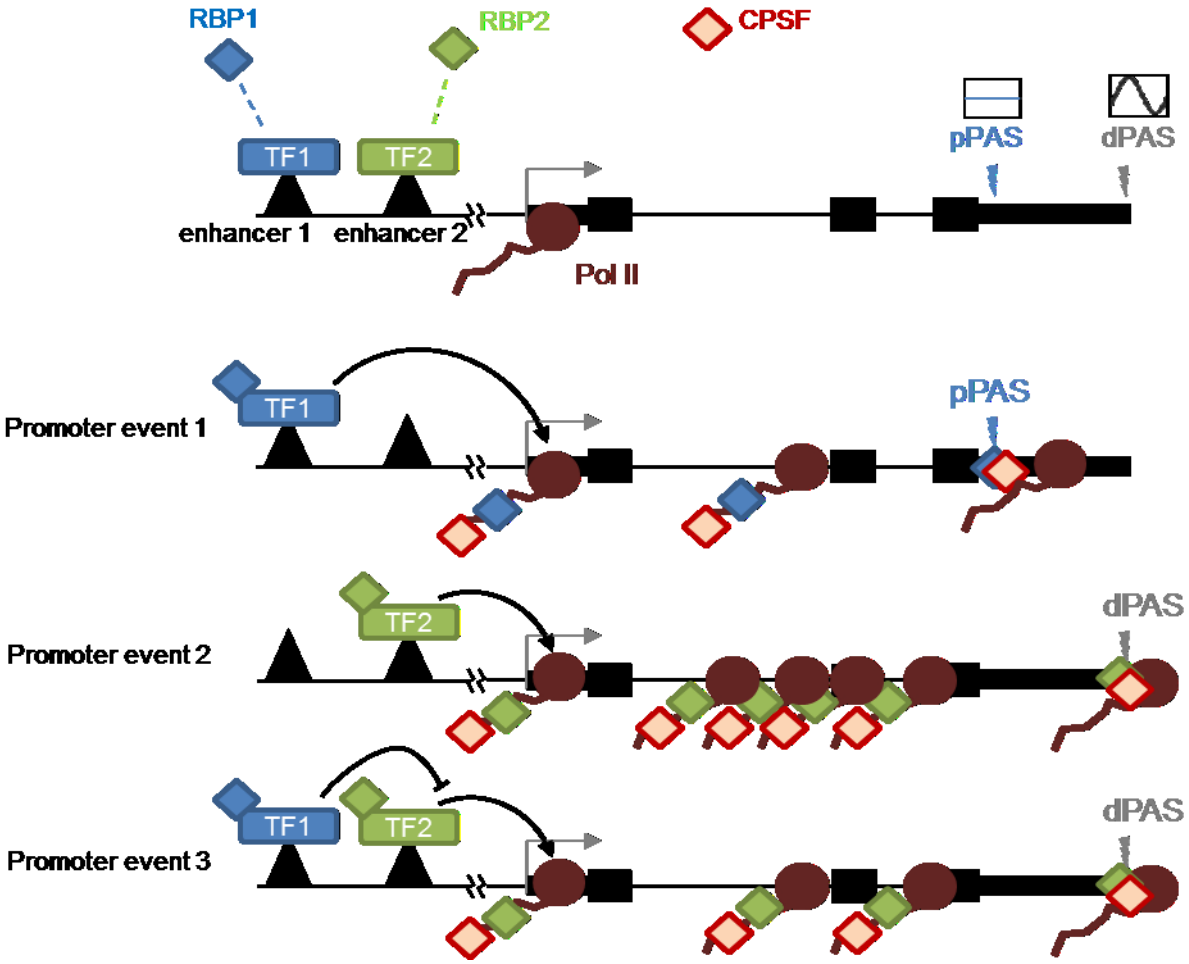

Figure IV-1 Postulated model for regulated of alternative PAS usage by co-transcriptional loading

Therefore, there are a number of directions that could be further investigated regarding Chapter III and our proposed model. The first involves exploring the *Bmal1* KO dataset in greater scope to ascertain whether these transcripts that gain or lose gene expression under BMAL1 KO have similarities in BMAL1 ChIP-Seq binding sites, protein domains, RBP binding sites, or MNase (chromatin) signal. The second involves analyzing the 3' UTR of differentially expressed APAS transcripts

and their sibling transcripts. If the 3' UTR length is playing a role in selection of these specific transcripts over others, I would predict to observe significant differences in the length of isoforms depending on the co-transcriptional proteins that are either used to initiate transcription or are loaded onto the CTD domain of Pol II. For example, *Bmal1* may be selecting for transcription of more proximal PAS over distal PAS, or vice versa, and this bias is then lost in *Bmal1* KO. In the third, I will investigate the 3' UTR RNA sequence to determine if there are any RBP motifs, or potentially use ChIP-Seq or PAR-CLIP evidence of RBP binding in order to determine if these transcripts are different due to binding of a rhythmically expressed RBP.

Finally, our model (Figure IV-) indicates there must be some form of combinatorial logic between TFs in determining PAS usage. This is further supported by the results from *Col18a1* expression in *Bmal1* KO (Figure III-9), where PAS 1 is expressed at a level identical to that in WT, whereas PAS2 is significantly upregulated in *Bmal1* KO. This therefore indicates that *Bmal1* is responsible for reducing the expression of PAS2, yet has no relation in determining the average expression of PAS1. The opposite case, where a TF acts concordantly to increase the usage of certain PAS, is not shown but is also a potential occurrence. Thus, we will perform a genome-wide analysis in order to test this model by examining the binding of many TFs to enhancers within each topologically associating domain (TAD) using ChIP-Seq. A database of published ChIP-Seq studies has been made available[20], and so can be used to perform a highly-dimensional analysis of which

TFs may be interacting with the promoters of differentially regulated genes. In the next step, we will examine publicly available datasets in the mouse liver of knockouts that use 3' end RNA sequencing to identify how PAS. If any strong candidates from the previous step do not already have usable datasets, we will generate our own knockouts and sequence their resulting PAS expression. If those TFs directly or indirectly act upon our genes found to be significantly differentially expressed under *Bmal1* KO, then they are co-transcriptional regulators with *Bmal1*. From here, we can use co-IP/MS on *Bmal1* and those TFs during the day timepoints to identify the recruited co-transcriptional elements that influence PAS usage.

References

1 Damiola, F. *et al.* Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes Dev* **14**, 2950-2961 (2000).

2 Vollmers, C. *et al.* Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proc Natl Acad Sci U S A* **106**, 21453-21458, doi:10.1073/pnas.0909591106 (2009).

3 Atger, F. *et al.* Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proc Natl Acad Sci U S A* **112**, E6579-6588, doi:10.1073/pnas.1515308112 (2015).

4 Koronowski, K. B. *et al.* Defining the Independence of the Liver Circadian Clock. *Cell* **177**, 1448-1462 e1414, doi:10.1016/j.cell.2019.04.025 (2019).

5 Schmutz, I., Ripperger, J. A., Baeriswyl-Aebischer, S. & Albrecht, U. The mammalian clock component PERIOD2 coordinates circadian output by interaction with nuclear receptors. *Genes Dev* **24**, 345-357, doi:10.1101/gad.564110 (2010).

6 Mukherji, A., Kobiita, A. & Chambon, P. Shifting the feeding of mice to the rest phase creates metabolic alterations, which, on their own, shift the peripheral circadian clocks by 12 hours. *Proc Natl Acad Sci U S A* **112**, E6683-6690, doi:10.1073/pnas.1519735112 (2015).

7 Le Minh, N., Damiola, F., Tronche, F., Schutz, G. & Schibler, U. Glucocorticoid hormones inhibit food-induced phase-shifting of peripheral circadian oscillators. *EMBO J* **20**, 7128-7136, doi:10.1093/emboj/20.24.7128 (2001).

8 Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111**, 16219-16224, doi:10.1073/pnas.1408886111 (2014).

9 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).

10 Landgraf, D. *et al.* Oxyntomodulin regulates resetting of the liver circadian clock by food. *Elife* **4**, e06253, doi:10.7554/eLife.06253 (2015).

11 Pocai, A. Action and therapeutic potential of oxyntomodulin. *Mol Metab* **3**, 241-251, doi:10.1016/j.molmet.2013.12.001 (2014).

12 Su, Y. *et al.* The role of feeding rhythm, adrenal hormones and neuronal inputs in synchronizing daily clock gene rhythms in the liver. *Mol Cell Endocrinol* **422**, 125-131, doi:10.1016/j.mce.2015.12.011 (2016).

13 Kidger, A. M., Sipthorp, J. & Cook, S. J. ERK1/2 inhibitors: New weapons to inhibit the RAS-regulated RAF-MEK1/2-ERK1/2 pathway. *Pharmacol Ther* **187**, 45-60, doi:10.1016/j.pharmthera.2018.02.007 (2018).

14 Longo, V. D. & Panda, S. Fasting, Circadian Rhythms, and Time-Restricted Feeding in Healthy Lifespan. *Cell Metab* **23**, 1048-1059, doi:10.1016/j.cmet.2016.06.001 (2016).

15 Nagaike, T. *et al.* Transcriptional activators enhance polyadenylation of mRNA precursors. *Mol Cell* **41**, 409-418, doi:10.1016/j.molcel.2011.01.022 (2011).

16 Neve, J. *et al.* Subcellular RNA profiling links splicing and nuclear DICER1 to alternative cleavage and polyadenylation. *Genome Res* **26**, 24-35, doi:10.1101/gr.193995.115 (2016).

17 Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat Rev Mol Cell Biol* **18**, 18-30, doi:10.1038/nrm.2016.116 (2017).

18 Hsin, J. P. & Manley, J. L. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**, 2119-2137, doi:10.1101/gad.200303.112 (2012).

19 Xiao, R. *et al.* Pervasive Chromatin-RNA Binding Protein Interactions Enable RNA-Based Regulation of Transcription. *Cell* **178**, 107-121 e118, doi:10.1016/j.cell.2019.06.001 (2019).

20 Mei, S. *et al.* Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. *Nucleic Acids Res* **45**, D658-D662, doi:10.1093/nar/gkw983 (2017).