Circadian Regulated Changes in Long non-coding RNA and Heterochromatin

By

JINHEE PARK

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the Degree of

Doctor of Philosophy

Graduate Program in Cell and Developmental Biology

Written under the direction of

William J. Belden, Ph.D.

And approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

(January 2019)

ABSTRACT OF THE DISSERTATION

Circadian Regulated Changes in Long non-coding RNA and Heterochromatin

By JINHEE PARK

Dissertation Director:

William J. Belden, PhD

The circadian clock governs gene expression for a large percentage of protein-coding genes in a tissue-specific manner. In this capacity, the clock maintains exquisite control of cell physiology and metabolism. The predominant regulatory mechanism of the clock is a transcriptional negative feedback loop that facilitates circadian-regulated facultative heterochromatin. The long-term consequence of disrupted diurnal rhythm, or mutations in core clock genes, is accelerated aging and an increased incidence of age-related diseases. However, the mechanisms underlying the precise pathways of the circadian clock and aging are not well understood. To understand the mechanisms of clock-regulated facultative heterochromatin in aging, I performed molecular experiments to examine the connections between BMAL1 and telomere homeostasis. I determined BMAL1 is associated with the telomeres and binding is conserved in zebrafish and mice. Expression of Telomere Repeat-containing RNA (TERRA), a long non-coding RNA (lncRNA) transcribed from the telomere has a diurnal rhythm in expression. In addition, there is a conserved rhythm in histone H3 lysine 9 tri-methylation (H3K9me3) at telomeres in zebrafish and mice. Given the rhythms in lncRNA and heterochromatin at the central clock gene(s) and telomeres, I set out to explore whether this was a genome-wide phenomenon, which may impact age-related redistribution of heterochromatin. I

performed RNA-Seq and H3K9me3 ChIP-Seq on zebrafish brain tissue at different times and different ages. The computational analysis of sequencing data followed by molecular confirmation revealed that the core clock genes maintain rhythmic expression regardless of age, but most diurnal genes change expression with age. Coincidently, there are diurnal and age-related changes in H3K9me3 that coincide with the changes in gene expression. Taken together, this study suggests a model where age-related redistribution of rhythmic facultative heterochromatin is potentially mediated by changes in diurnal lncRNA expression creating a circadian-chromatin regulatory network in aging.

Acknowledgments

I would like to thank my advisor, Prof. William J. Belden for the opportunity to join his lab and his huge patience to train me.

I am deeply grateful to the members of my thesis committee, Dr. Wendie Cohick, Dr. Karen Schindler, Dr. Zhiping Pang, and Dr. Elizabeth Snyder for their critical insights and review of my work. And I also thank Dr. Berry Jesse for his wonderful guidance as my TA advisor.

I would like to thank my lab member, Qiaoqiao Zhu, and former lab members, Hamidah Raduwan and Dr. Na Li as colleagues as well as good friends.

I thank Dr. Tracy Anthony and Emily Mirek for their enormous help with the mouse experiments, and I thank Dr. Elizabeth Snyder for helping me edit chapter 3 of my dissertation prior to it being submitted for publication.

I am also indebted to many individuals within the Cell and Developmental Biology program, especially Carolyn Ambrose, and the Department of Animal Sciences.

Finally, I dedicate this work to my dearest family.

TABLE OF CONTENTS

Page

ABSTRACT OF THE DISSERTATION ii

ACKNOWLEDGMENTS iv

TABLE OF CONTENTS v

LIST OF ILLUSTRATIONS viii

LIST OF TABLES x

GLOSSARY OF ABBREVIATIONS xi

CHAPTER 1: Review of literature 1

1.1 The mechanism of circadian clock 2

1.2 Chromatin regulation in the circadian clock 11

1.3 Long-non coding RNAs, heterochromatin and circadian clock 16

1.4 Circadian clock and aging 22

1.5 Circadian clock and telomere 24

CHAPTER 2: Circadian clock and telomere homeostasis 26

2.1 ABSTRACT 27

2.2 INTRODUCTION 28

2.3 RESULTS 32

WC-2 localizes to the telomere in *Neurospora* 32

BMAL1 association with the telomere 33

BMAL1 associates with the telomere in mice 35

Rhythmic expression of TERRA 36

Diurnal regulation of heterochromatin at telomere regions 38

Aging and stress affect the diurnal regulation of TERRA and heterochromatin 39

2.4 MATERIALS AND METHODS 57

2.5 DISCUSSION 63

CHAPTER 3: Genome-wide relationship among Circadian clock, heterochromatin and long-non coding RNAs with age 67

3.1 ABSTRACT 68

3.2 INTRODUCTION 69

3.3 RESULTS 72

Identification of long non-coding RNAs 72

Age-related changes to diurnal gene expression 74

Circadian Clock Remains Constant Regardless of Age transcripts 77

Age-specific Diurnal Genes with Overlapping Transcripts 78

Diurnal and Age-related Changes to Facultative Heterochromatin 80

3.4 MATERIALS AND METHODS 103

3.5 DISCUSSION 107

3.6 ACKNOWLEDGMENTS 112

CHAPTER 4: Conclusion 113

APPENDICES

A. *Bmal1* knock-out zebrafish by using homologous gene targeting 117

B. Circadian regulation of TIN2 between nucleus and cytoplasm 122

C. The effect of TIN2 knock-down in *bmal1*-luciferase rhythm 124

BIBLIOGRAPHY 126

LIST OF ILLUSTRATIONS

Figure Page

1 Schematic of the negative feedback regulation in circadian clock ---------- 9

2 Schematic model of RNAi-mediated heterochromatin assembly in *S. pombe*--21

3 WC-2 is associated with the telomeres in *Neurospora*----------40

4 BMAL1 at associates with the telomere in zebrafish and mice ----------41

5 A diurnal rhythm between BMAL1 and telomere in zebrafish and mice-------42

6 Zebrafish have a diurnal rhythm in TERRA expression ----------43

7 Diurnal Rhythm in TERRA requires BMAL1----------44

8 Rhythmic regulation of heterochromatin at telomere in zebrafish and mice --45

9 Aging alters the diurnal rhythms in TERRA expression and H3K9me3 at subtelomere in zebrafish----------46

10 Age-related changes in diurnal transcription of zebrafish brain tissue. --------83

11 The core clock remains rhythmic regardless of age ----------85

12 Representative examples of overlapping transcripts----------86

13 Diurnal rhythm in genome-wide H3K9me3 for each age group ----------88

14 Age-related changes in transcript abundance coincide with changes in heterochromatin----------89

15 Homologous gene targeting for *bmal1*$^{-/-}$ zebrafish ---------- 119

16 TIN2 does not show rhythmic translocation between the nucleus and cytosol ---------- 123

17 The knock-down of *Tin2* does not change the rhythm of *bmal1*-luciferase--- 125

Supplementary figure ---------- Page

1 Visualization of the binding of BMAL1 to telomere and TERRA expression in mice ---------- 47

2 The confirmation of the specificity of the BMAL1 antibody ---------- 49

3 qRT-PCR results in mice ---------- 50

4 Rhythmic TERRA expression in mouse RNA-Seq data ---------- 51

5 TERRA appears rhythmic in U2OS cells ---------- 52

6 Mouse TERRA northern blots ---------- 53

7 The effect of aging and stress on TERRA and H3K9me3 ---------- 54

8 Tophat2/Cufflinks pipeline ---------- 93

9 Hisat/Stringtie2 pipeline ---------- 94

10 Expression patterns of lncRNAs from early development through adult aged zebrafish ---------- 95

11 Age-dependent expression changes in diurnal genes ---------- 97

12 Disease Ontology enrichment analysis of diurnal transcripts at different ages ---------- 98

13 Expression of myl10, cishb, and myl1 sense and antisense ---------- 100

14 Expression and H3K9me3 for slc4a1a ---------- 101

15 Representative view of H3K9me3 broad peaks ---------- 102

LIST OF TABLE

Table		Page
1	**Circadian clock orthologues in various species**	**8**
2	**The sequence of qPCR oligos for zebrafish**	**56**
3	**The sequence of qPCR oligos for mouse**	**56**
4	**The sequence of oligos for TERRA probe**	**56**
5	**GO analysis of genes changed with age**	**91**
6	**The sequence of qPCR oligos in chapter 3**	**92**
7	**The sequence of qPCR oligos in appendix A**	**121**

GLOSSARY OF ABBREVIATIONS

ARD	Age-Related Disease
ACTH	Adrenal to adrenocorticotropic hormone
Bmal1$^{-/-}$	*Bmal1* Knock-out
CATP	CLOCK ATPase
CRH	Corticotropin releasing hormone
CCGs	Clock-Controlled Genes
CHD-1	chromodomain helicase DNA binding protein 1
ChIP	Chromatin immunoprecipitation
ChIP-seq	ChIP-sequencing
CK-1	Casein Kinase-1
CK1δ	CK-1 isoform delta
CRFH	Circadian Regulated Facultative Heterochromatin
CRY	CRYPTOCHROME
CSW-1	CLOCKSWITCH-1
CT	Circadian Time
CYC	CYCLE
Dbp	D site of albumin promoter binding protein
DBT	DOUBLETIME
DDX5	DEAD-box polypeptide 5
DDX9	DEAH-box protein 9
disiRNA	Dicer-independent small interfering RNA
DO	Disease Ontology
EM	Electron Microscopy
FFC	FRQ-FRH complex
FRH	FRQ-interacting RNA helicase
frq	*frequency*
GO	Gene Ontology
H/S	HISAT2/StringTie

H3K27me3	Histone H3 Lysine 27 (H3K9) tri-methylation
H3K4me3	Histone H3 Lysine 4 (H3K9) tri-methylation
H3K9ac	Histone H3 lysine 9 (H3K9) acetylation
H3K9me2	Histone H3 Lysine 9 (H3K9) di-methylation
H3K9me3	Histone H3 Lysine 9 (H3K9) tri-methylation
H3S10	Histone H3 Ser 10
HAT	Histone acetyltransferases
HDAC	Histone deacetylase
IGV	Integrative Genomics Viewer
KMT1	Histone H3 lysine 9 methyltransferase
lincRNA	long intergenic noncoding RNA
lncRNA	Long non-coding RNA
NAT	Natural Antisense Transcript
Oligo	Oligonucleotides
OT	Overlapping Transcripts
p15AS	p15 antisense
PER	PERIOD
PRC2	Polycomb Repressive Complex 2
PVN	Paraventricular nucleus
qPCR	Quantitative PCR
RIP-Seq	RNA immunoprecipitation-Sequencing
RNAi	RNA interference
RNA-Seq	RNA-Sequencing
RORα	RAR-related orphan receptor alpha
ROS	Reactive Oxygen Species
RT-qPCR	Reverse Transcription-quantitative PCR
SCN	Suprachiasmatic nucleus
SCG	Superior cervical ganglion
SETX	Senataxin
T/C	Tophat2/Cufflinks

TERRA	TElomere Repeat-containing RNA
TERT	Telomerase reverse transcriptase
TIM	Timeless
TSCC	Tongue Squamous Cell Carcinoma
TTFL	Transcription Translational Feedback Loops
WCC	WHITE COLLAR complex
WDR5	WD repeat-containing protein 5
WebGestalt	WEB-based GEne SeT AnaLysis Toolket
WT	Wild-Type
ZT	Zeitgeber Time

CHAPTER 1

Review of literature

Aging can be generally defined as a progressive decline or deterioration in physiological functions (1). Aging affects various aspects of our physiology and behaviors, including the circadian clock (2). Evidence indicates the circadian clock influences aging in important ways; however, the pathways underlying how the circadian clock and aging are intertwined are not well defined. Understanding the connection between the clock and aging is important because disrupted circadian rhythms are an early predictor and may contribute to neurodegenerative diseases including Alzheimer's, Parkinson's, and Huntington's (3, 4). Moreover, studying the relationship between the circadian clock and aging may provide valuable opportunities to promote healthy aging. In this chapter, I will provide an overview of molecular mechanisms that regulate the circadian clock. I will then relate some aspects of aging including telomere homeostasis to the circadian clock.

1.1 Mechanism of circadian clock

The circadian clock is an internal oscillator in organisms that synchronize internal timing to external environmental changes. *Drosophila*, *Neurospora*, and *Mus musculus* are three widely used model systems to study the genetics underlying circadian rhythms (*5-12*). Research in these models has revealed the circadian rhythm is functionally conserved across species and is predominantly controlled by negative feedback loops. The core clock proteins and the mechanisms were initially worked out in both *Drosophila* and *Neurospora* then homologs were found in mice.

Early genetic screens in *Drosophila* led to the identification of mutations that change circadian behavior, providing the framework to identify core clock genes. In

Drosophila, the key activator complex is composed of two transcription factors, CLOCK (CLK) and CYCLE (CYC). The CLK:CYC complex is a dimeric activator complex that controls the expression of *per* and *timeless* (*tim*) transcripts by binding E-box elements in their promoters (*13*). Following transcription and translation, PER and TIM accumulate in the cytoplasm during the night and form a heteromeric complex (*14*). The PER:TIM complex enters the nucleus, where it directly interacts with CLK:CYC. This interaction represses CLK:CYC transcriptional activity, resulting in suppression of transcription of *per* and *tim* (*15*). *Drosophila* DOUBLETIME (DBT) is a homolog to mammalian casein kinase-1 (CK-1) (*16*) and DBT promotes the progressive phosphorylation of PER in the cytoplasm (*17*). The hyperphosphorylated PERs trigger binding of the E3 ubiquitin ligase SLIMB (a member of the F-box/WD40 protein family) for degradation of PER through the ubiquitin-proteasome. As the sun rises, light triggers a conformational change in intracellular photoreceptor protein CRY to interact with TIM, resulting in rapid degradation of TIM (*18*). Because PER is unstable without TIM, PER is also degraded. The repression of CLK:CYC is then released, and transcription of *per* and *tim* resumes (*14*).

Similar to the circadian oscillators in *Drosophila*, the core clock of *Neurospora* is predominantly regulated by a negative feedback loop. In *Neurospora*, the WHITE COLLAR complex (WCC) is formed by WC-1 and WC-2, which bind to the c-box element at the *frequency* (*frq*) promoter to drive transcription of *frq* (*19*). After translation, FRQ protein associates with FRH (FRQ-interacting RNA helicase), and inhibits WCC activity leading to negative feedback inhibition (*20, 21*). FRH has been shown to stabilize FRQ, suggesting that the role of FRH may be to act as a partner protein to allow FRQ to

serve a multitude of functions (*22*). However, another model posits that it is the ATPase function of FRH that regulates the access of CK1a to phosphorylation sites in FRQ *in cis*, allowing for the proper phosphorylation of FRQ (*23*). In the course of a circadian period, FRQ is progressively phosphorylated by several kinases such as Casein Kinase-1 (CK-1), Casein Kinase-2 (CK-2), and Calcium/calmodulin-dependent protein kinase type 1 (CAMK1), and they control the stability of FRQ (*24*). Progressive phosphorylation facilitates the interaction of FRQ with the SCF-type ubiquitin ligase FWD-1, leading to degradation of FRQ through the ubiquitin-proteasome pathway (Fig. 1A) (*25-27*). In the *fwd-1* mutant, hyperphosphorylated FRQ accumulates because the normal ubiquitin-proteasome degradation pathway is blocked. This results in a defect in developmental rhythms such as daily spore production (*28*).

The framework of the circadian clock in mice is similar to both *Drosophila* and *Neurospora* (Fig.1 B) (*29-34*). One feedback loop is created by the transcriptional activators CLOCK:BMAL1 that drive expression of *Period (Per1/2/3)* and *Cryptochrome (Cry1/2)* genes by binding to E-box elements in their promoters. After translation, PER and CRY oligomerize with CK-1. Deletion of CK-1 causes PER to accumulate in the cytoplasm, suggesting CK1 is important for the regulation of PER phosphorylation and nuclear entry (*35*). The cytoplasmic PER complex is composed of PER, CRY and CK1δ, which then recruit GAPVD1, a cytoplasmic trafficking factor, creating an approximate 1.1 MDa complex that is subsequently transported into the nucleus. In the nucleus, PER and CRY become a large multisubunit complex (~1.9 MDa) containing RNA binding proteins, DNA helicases, and chromatin modifying enzymes (*36, 37*).

Affinity purification and electron microscopy (EM) indicate the murine PER complex has a dynamic composition that changes throughout the circadian cycle. In the repressive phase, the PER complex interacts with transcriptional co-repressors including chromatin-modifying enzymes. Thus, the PER complex contributes to repression of the CLOCK:BMAL1 target genes by changing chromatin states through deacetylation and methylation of Histone H3 lysine 9. First, SIN3-containing HDA

C1/2 in the PER complex removes acetyl groups from H3K9 at CLOCK:BMAL1 target sites eliminating the positive effects of acetylation. Then after approximately 4-hours, PER complexes recruit histone H3 lysine 9 methyltransferase (KMT1/Suv39H) and heterochromatin protein 1 (HP1γ) to establish non-permissive heterochromatin. HP1γ binds to H3K9me2 and H3K9me3 to link the Suv39H histone methyltransferase to the chromatin-bound PER complex. Therefore, the complex of HP1γ and Suv39H serves an important role for circadian H3K9me2 and H3K9me3 at the *Per1* and *Per2* promoters (*38, 39*). The PER complex also contains NONO (Non-POU domain-containing octamer-binding protein), a protein involved in RNA processing and WDR5 (WD repeat-containing protein 5), a component of a histone methyltransferase complex (*40*). In addition, RNA helicases such as DDX5 (DEAD-box polypeptide 5) and DDX9 (DEAH-box protein 9), a DNA/RNA helicase SETX (Senataxin) and the large subunit of RNA polymerase II are included in the PER complex (*36, 37, 40*).

In addition to a rhythm for *Per* and *Cry*, there is also a rhythm in *Bmal1* expression that is regulated by the second feedback loop mediated by RORα and REV-ERBα (*41, 42*). The binding of RORα at the *Bmal1* promoter activates *Bmal1* expression, whereas REV-ERBα inhibits *Bmal1* transcription. BMAL1 and REV-ERBα ChIP-Seq revealed thousands of binding sites throughout the genome (*43, 44*), suggesting an enormous number of transcripts in the cell are under direct control of the circadian clock.

Genes under control of the circadian clock are collectively called *clock-controlled genes (ccgs)* and some are regulated by the core clock proteins, which bind to enhancer elements in their promoter. The *ccgs* are expressed in a tissue- and phase-specific manner, so that clock outputs such as hormone release, metabolic control, and rhythmic

expression of proteins have timed oscillations (*10*). The importance of proper circadian gene expression is easily observed because misregulation of the circadian oscillator is implicated in numerous pathways and human diseases; including sleep disorders, metabolic disorders such as obesity and diabetes, age-related diseases like cancer (*45-47*). For example, *Per* mRNA is down-regulated in breast cancer cells and the overexpression of *Per1* results in growth inhibition of human cancer cell lines (*48, 49*) suggesting *Per* genes as tumor suppressors. It has also been reported that *Bmal1* suppresses the invasion of cancer cells by inhibiting the PI3K/Akt/MMP2 axis (*50*).

The mammalian circadian clock has a hierarchical organization consisting of a master clock located in the suprachiasmatic nucleus (SCN) of the hypothalamus and peripheral clocks found throughout the body. The peripheral clock and the master clock have identical molecular feedback regulation (e.g. the core clock), except the master clock in the SCN can be entrained by light:dark cycles. The master clock then synchronizes peripheral clocks in the different tissues (*51*). The communication between the master clock and clocks in peripheral tissues is achieved by a combination of neuronal and endocrine cues (*52-54*) Neuronal pathways such as the superior cervical ganglion (SCG) convey the entraining signal from the SCN to peripheral tissues to modulate hormone release. For example, light induces gene expression and glucocorticoid release in the adrenal gland via the SCN-SCG system. The peripheral clock in the adrenal gland also gates glucocorticoid production in response to adrenocorticotropin (ACTH) evoked by the SCN-pituitary-adrenal gland (HPA) axis.

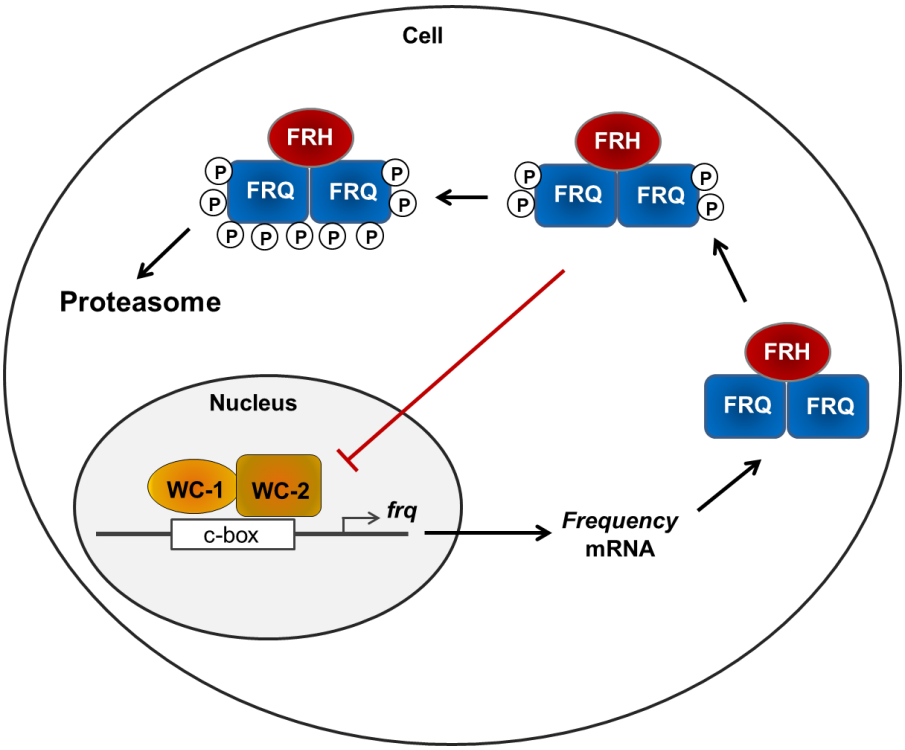
In addition to established models for circadian clock research, zebrafish is also a fabulous model offering many advantages and has been used as a system that is more

closely matched to humans. There is a high degree of conservation of circadian mechanisms across species, but important differences exist in the organization of nocturnal and diurnal circadian rhythms (55). Thus diurnal zebrafish is more applicable to human than the nocturnal mice. Zebrafish clocks also have high similarities to mammalian clocks and contain all the homologs of the mammalian core clock genes; many of which have been identified and cloned (*56*). The small size, low maintenance costs and high fecundity of zebrafish along with the availability of highly efficient genetic tools such as CRISPR/Cas9 make zebrafish an attractive model for genetic analysis of the circadian clock. Adding to the utility, cell lines derived from zebrafish embryos have a light-entrainable circadian clock, whereas mammalian cell lines require serum or dexamethasone shock to synchronize oscillations (*57*). The direct light sensitivity of cultured zebrafish cells serves as a great *in vitro* system to uncover the poorly understood processes of entrainment. Therefore, zebrafish is a powerful experimental system for circadian clock research and is one of the models used in my dissertation.

Table1. Circadian clock orthologues in various species

	Repressors		Activators	
Neurospora	**FRQ**	**FRH**	**WC-2**	**WC-1**
Fly	**PER**	**TIM**	**CLK**	**CYC**
Zebrafish	**PERs**	**CRYs**	**CLOCK**	**BMAL1**
Mice	**PERs**	**CRYs**	**CLOCK**	**BMAL1**

A


Cell
FRH
FRQ
FRQ
P
Proteasome
Nucleus
WC-1
WC-2
c-box
frq
Frequency
mRNA

B

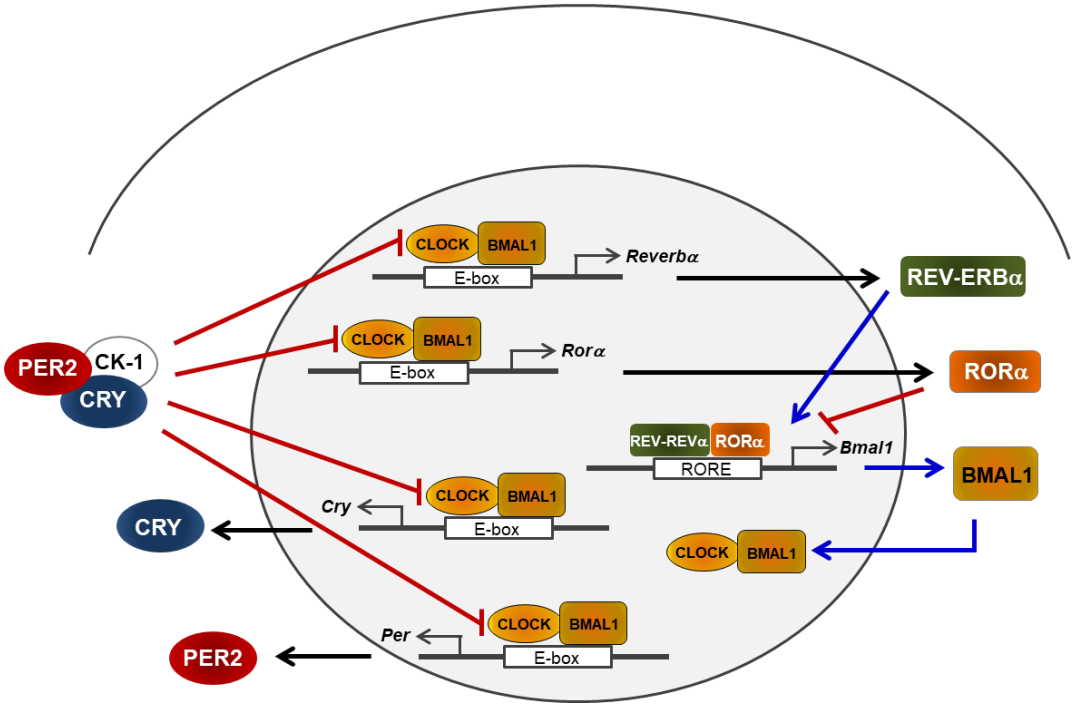

CLOCK
BMAL1
E-box
Reverbα
REV-ERBα
Rorα
RORα
PER2
CK-1
CRY
REV-REVα
RORE
Bmal1
BMAL1
Cry
Per

Figure 1. Schematic of the negative feedback regulation in circadian clock

(A) In *Neurospora*, transcription factors WC-1 and WC-2 form a heterodimer (orange ovals and squares). The WCC complex activates the transcription of *frq* transcripts. The FRQ and FRH (red oval and blue squares) complex repress WCC. The FRQ and FRH complex is progressively phosphorylated and degraded by the proteasome. (B) In vertebrates, CLOCK:BMAL1 heterodimer (orange ovals and squares) bind DNA of clock target genes such as *Per* and *Cry* at E-box elements and initiate the transcription of their RNA. The resulting PER and CRY proteins (red and blue ovals) form a dimer in the cytoplasm and translocate to the nucleus. The PER complex inhibits CLOCK:BMAL1 proteins from initiating further transcription. Next, RORα and REV-ERVβ induce a secondary loop to control *Bmal1* expression. RORα can bind to RORE element for *Bmal1* activation. REV-ERBβ works as a repressor in *Bmal1* transcription by competition at the same RORE sequence. For simple illustration, only single *Per* and *Cry* genes are shown. Red lines depict negative feedback repression, and the blue arrows signify positive regulation.

1.2 Chromatin regulation in the circadian clock

Because the circadian clock is predominantly controlled by transcriptional feedback loops, chromatin remodeling and chromatin modifications play an important role in clock regulation. The basic structure of chromatin is the nucleosome, which contains 147 base pairs of DNA wrapped around a histone octamer (two H2A, H2B, H3, and H4 histones) connected by linker DNA with histone H1 (*58*). The nucleosomes are packed into more compact structures by repeated folding of chromatin fibers into multiple loops and coils. The structure of chromatin undergoes various changes during different biological processes such as cell cycle or DNA repair (*59*). Euchromatin refers to chromatin that is less condensed and typically contains actively transcribed genes. Euchromatin is often characterized by high levels of the tri-methylation of Histone H3 lysine 4 (H3K4me3) and acetylation of Histone H3 lysine 27 (H3K27ac), which are both believed to be activating histone marks (*60*). In contrast, heterochromatin is more condensed and is largely believed to be transcriptionally silent (*61*). Facultative heterochromatin occurs within euchromatin when transcriptionally permissive chromatin is converted into heterochromatin at specific stages of development or the cell type. An example of constitutive heterochromatin is X-chromosome inactivation in female mammalian cells, which is heavily marked by tri-methylation of Histone H3 lysine 27 (H3K27me3) by Polycomb repressor complex 2 (PRC2) (*62*). Other regions that are believed to be constitutive heterochromatin include the telomere and centromere. In particular, telomeres and pericentric heterochromatin contain relatively high levels of tri-methylation of Histone H3 lysine 9 (H3K9me3) and HP1 (*63, 64*).

Dynamic alteration to nucleosome position within chromatin is achieved by a chromatin-remodeling enzymes and movement of the nucleosome is vital for many cellular events. For example, the rearrangement of chromatin from a closed state to a transcriptionally favorable state allows regulatory proteins to access specific genomic region(s). Conversely, opened chromatin can be returned to a compact state after transcription or when replication is complete (*65*). Regulation of chromatin and subsequent remodeling is governed by two major classes of protein: histone-modifying enzymes that add post-transcriptional modifications to the histones and ATP-dependent chromatin-remodeling enzymes that reorganize nucleosomes . ATP-dependent chromatin-remodeling complexes such as SWI/SNF (SWItching defective/ Sucrose NonFermenting) and CHD (Chromodomain Helicase DNA binding) use energy from the hydrolysis of ATP to repositioning nucleosomes along the DNA (*66*). In contrast, histone-modifying enzymes including histone lysine acetyltransferases (KAT), histone deacetylase (HDAC), and histone lysine methyltransferases (KMT) regulate this dynamic nature of chromatin by covalently modifying histone proteins (*67, 68*) which can then serve as a signaling platform for chromatin effectors. The activities of various chromatin-modifying enzymes can act in a combinatorial manner to elicit specific biological events in the regulation of DNA-based processes.

Since histone acetylation was first reported in 1964 (*69*), more than 100 different histone modifications have been identified. The modified histones function as a platform for the binding of specific effector proteins. Acetylated histones are bound by bromodomains chromatin-remodeling and effector proteins such as p300 (*70*). Similarly, chromodomain-containing proteins can recognize the methylated histones, leading to a

downstream cascade of events (*70*). In this way, the different histone modifications have distinct functions and regulatory mechanisms including transcription, repair, replication, and chromosome condensation (*71-74*). Aberrant regulation of chromatin structure is linked to some diseases such as cancer, neurological disorders and even aging-related defects (*75-77*).

A variety of covalent modifications to chromatin are also important for the circadian clock (*39, 57, 78-81*). The first experiments showing circadian changes to chromatin were light-inducible phosphorylation of histone H3 Serine 10 (H3S10P) in SCN (*58*). Later the histone acetyltransferase p300 was found to associate with CLOCK to generate rhythms in H3K9 and H3K14 acetylation on histones contained in the promoters of *ccgs* (*80*). Rhythms in histone acetylation at *Per* and *Cry* are synchronized with the cyclical expression of their mRNAs (*80, 82*). In contrast, during the repressive phase, H3K9 and H3K27 are methylated in the *Per* promoters (*39, 83*). Likewise, the *ccg Dbp* has rhythms in H3K9me2 and H3K9me3 that are added by histone H3 lysine 9 methyltransferases (KMT1/Suv39H). The daily rhythm in H3K9me2 and HP1 binding at *Dbp* are antiphasic to H3K9ac and H3K4me3 (*7, 39, 40*). Particularly, rhythms in H3K9me3 mediated by KMT1 at core clock genes appers to be a conserved mechanism in mice and *Neurospora*. In mice, KMT1/Suv39h as a part of PER complex regulates the rhythm in H3K9me3 and HP1γ binding at *Per2* promoter (*39*). Similarly, the central clock gene *frq* in *Neurospora* has a rhythm in H3K9me3 that is mediated by KMT1/DIM-5 and there is a rhythm in HP1 binding (*84*). Thus, the precise timing and combinatorial accumulation of histone modifications are conserved across species and essential for proper transcriptional regulation of circadian transcription.

The rhythm in histone modifications at clock genes and *ccgs* are tightly regulated by chromatin modifiers, which are recruited by circadian clock proteins such as CLOCK:BMAL1 and PER complex. Co-immunoprecipitation experiments have revealed Polycomb repressive complex 2 (PRC2) can associate with CLOCK:BMAL1. Polycomb repressive complex 2 (PRC2) contains a histone H3 lysine 27 methyltransferase (KMT6) that can add a mono- di- or tri-methyl group to histone H3 on lysine 27 to inhibit gene expression (*83*, *85*). KMT6/EZH2, the catalytic subunit of PRC2, is targeted to the *TERT* promoter in tongue squamous cell carcinoma (TSCC), and this targeting is believed to be mediated by BMAL1. The interaction between BMAL1 and the KMT6/EZH2 at the *TERT* promoter is thought to increase the susceptibility of cancer cells to the anti-cancer drug paclitaxel (*86-88*). In zebrafish, *kmt6/ezh2* appears to be regulated by the circadian clock and rhythmic expression of clock genes is disrupted in *ezh2* null mutant (*89* , *90*). Collectively, these results suggest that the clock is partially regulated by PRC2-dependent H3K27me3-mediated facultative heterochromatin; however, the pleiotropic phenotypes in *kmt6* knockdown or null strains thwart definitive conclusions. Furthermore, H3K27me3 does not appear to be required for the *Drosophila* or *Neurospora* clock.

In addition to histones modifications, chromatin remodeling is also a key facet of clock regulation. The first evidence of the importance of chromatin remodeling came from studies in *Neurospora*. CLOCKSWITCH-1 (CSW-1) and CATP (CLOCK ATPase) regulate the binding activity of WCC complex by changing chromatin structure around the C-box element during the repressive and activation periods making them key regulators of circadian transcription (*91*) (94). In addition, the chromodomain helicase DNA binding protein 1 (CHD-1) is needed to remodel chromatin structure at *frq* and loss

of *chd-1* cause constitutive DNA-methylation and heterochromatin at *frq* (*92*). *Neurospora* also requires SWI/SNF complex that interacts with WCC and is required both for circadian remodeling of the nucleosome at *frq* and rhythmic expression of *frq (93)*.

The closest mammalian homolog to *Neurospora* CHD-1 is Mi-2/CHD4, which is the core ATP-dependent chromatin-remodeling enzyme in the nucleosome remodeling and deacetylase (NuRD) corepressor. Biochemical isolation and mass-spectroscopy (MS) revealed the mouse PER complex contains subunits of NuRD (*94*). The NuRD complex is composed of HDAC1 and HDAC2, MTA2, and a methylated DNA binding protein MBD1 or MBD2. Proposed models of their function in the mammalian clock indicate that two NuRD subunits, CHD4 and MTA2, constitutively associate with CLOCK:BMAL1. Then, during negative feedback, the PER complex delivers the remaining NuRD subunits to DNA-bound CLOCK:BMAL1, which helps establish a non-permissive state for transcription. The existence of a CHD family member in both the Neurospora and mammalian clock indicate chromodomain helicase proteins are essential players in the circadian rhythm.

Despite the plethora of information on circadian chromatin, many of the precise mechanisms are still incomplete. However, it is clear that rhythm in facultative heterochromatin, specifically H3K9me3 is highly conserved across species as diverse as *Neurospora*, zebrafish, and mammals indicating that it is an essential component of the clock. Thus, studying circadian-regulated chromatin will provide useful insight into understanding the mechanisms of rhythmic biological pathways.

1.3 Long non-coding RNAs, heterochromatin, and circadian clock

The importance of facultative heterochromatin in the circadian clock warrants some background on how H3K9me3 is established in model systems and the best-studied system is pericentric heterochromatin in *S. pombe*. Pericentric heterochromatin formation in *S. pombe* is established by RNAi-mediated heterochromatin. RNAi is a gene-silencing mechanism that results from the introduction of homologous antisense RNA to neutralize targeted mRNA molecules. Earlier studies in *S. pombe* proposed a model for RNAi-mediated heterochromatin (Fig. 2) (*95, 96*). Briefly, nascent transcripts are generated by RNA polymerase II, and then RNA-directed RNA polymerase complex (RDRC) is recruited to synthesize a double-stranded RNA (dsRNA) from the transcripts. The dsRNA is then cleaved into siRNAs by Dicer1 and bound by Argonaute-associated proteins. The siRNAs are loaded into RNA-induced transcriptional silencing (RITS) complex. The active RITS recruits KMT1/CLR4 complex (CLRC) to establish H3K9me3, which is then bound by HP1(SWI6).

Circadian regulated facultative heterochromatin at *frq* in Neurospora is established in a somewhat annalogous mechanism to RNAi-mediated heterochromatin in *S. pombe*. Neurospora has a natural antisense transcript (NAT), *qrf* that is required to establish heterochromatin formation at central clock gene *frq* (*97*). The convergent expression *frq* and *qrf* leads to the production of small (approximately 22nt) RNA that are cleaved independent of Dicer to generate Dicer-independent siRNA (disiRNA) (*98, 99*). Normal *qrf* expression is needed for KMT1/DIM-5-mediated H3K9me3 heterochromatin HP1 binding and DNA methylation at *frq*. However, unlike pericentric heterochromatin in *S. pombe*, which requires RNA-dependent RNA polymerase and Dicer, transcription of

qrf is synthesized by RNA polymerase II and the disiRNA are cleaved independent of Dicer (*100*). The similarities and importance of these two RNA-guided heterochromatin systems serve an important basis for the research I performed in my thesis as my work hypothesized that similar a mechanism occurs at many other loci including telomeres, imprinted loci, and at *ccgs* where long non-coding RNAs (lncRNAs) help guide epigenetic modifications either in *cis* or *trans*.

Previous studies of genome-wide circadian transcription have mainly focused on the protein-coding genes (*101-103*). However, a recent systematic analysis reported that a significant number of long non-coding RNAs (lncRNAs) also show circadian expression in several species and these may help establish epigenetic modifications. For example, 112 lncRNAs showed differential expression between light and dark in the rat pineal gland (*104*). Mouse circadian transcriptome data suggests there are at least 1000 circadian lncRNAs from multiple mouse tissues (*105*). Further RNA-Seq studies in mouse liver found a class of circadian lncRNAs are associated with enhancer regions (eRNAs) that are bound by BMAL1 and REV-ERBα (*106*). The rhythmic expression of numerous lncRNAs suggests the circadian clock transcriptionally controls lncRNAs as well as coding genes but the role of the circadian lncRNAs, other than *qrf* in Neurospora, are unknown.

lncRNAs are a family of functionally active RNAs, defined as longer than 200 nucleotides that do not encode proteins. Most lncRNAs are transcribed from intergenic regions but also can be expressed from within introns and in the antisense direction (*107-110*). Large-scale sequencing projects in mammals indicate there are hundreds of thousands of lncRNAs; a far greater number than protein-coding genes (*111-113*). In

contrast to the protein-coding mRNAs, lncRNAs typically have a lower expression, stronger tissue specificity and lower sequence conservation (*114-118*). Alterations in the expression of lncRNAs coincide with diseases ranging from neurodegeneration to cancer, suggesting that lncRNAs may be vital modulators of physiologic state (*119, 120*). For instance, the lncRNA ANRIL regulates the transcription of three tumor suppressors, p16INK4a, p14ARF, and p15INK4b, which are important negative regulators of the cell cycle (*121*). Misregulated expression of ANRIL has been associated with the occurrence of several cancer types, including neuroblastoma (*122*) and acute lymphocytic leukemia (*123*).

lncRNAs can be found associated with chromatin-modifying enzymes. RNA immunoprecipitation followed by sequencing (RIP-seq) revealed large numbers of lncRNAs are associated with the repressive PRC2 complex. Emerging models suggest that in some instances, lncRNAs act as a linker by recruiting chromatin-modifying complexes to specific genomic loci to change the epigenetic state (*20, 124*). For instance, lncRNA *HOTAIR* is a 2.2 kb transcript originating from an intergenic region of the *HOXC* locus (*120*). The 5′ end of HOTAIR is associated with PRC2 that contains KMT6/EZH2 and mediates its recruitment to the 40kb *HOXD* cluster *in trans* (*125*). In a similar fashion, *lincRNA-p21* (long intergenic noncoding RNA-p21) is a transcriptional target of p53 transcribed from upstream of p21 gene on the opposite strand (*126*). The *lincRNA-p21* is associated with the H3K9 methyltransferase KMT1/SETDB1 and the DNA methyltransferase DNMT1 to repress target genes such as *Lin28a* and *Nanog* by inducing H3K9me3 and CpG methylation. The epigenetic silencing mediated by *lincRNA-p21* is necessary for blocking somatic cell reprogramming (*72*). In addition, lncRNA transcribed

from the telomere is involved in the epigenetic modulation of the telomere (*71*). Telomere repeat-containing RNA (TERRA) is a lncRNA containing UUAGGG repeats that vary in length from 100 bp to approximately 9 kb, depending on the location of the transcription start site (*75*). TERRA plays an important role in PRC2-dependent H3K27me3 formation and further establishment of H3K9me3, H4K20me3, and HP1 binding at telomeres (*77*). Despite the overwhelming evidence that lncRNAs contribute to establishing chromatin states, the precise mechanism of how lncRNAs control chromatin in a spatially and temporally coordinated manner is still not fully understood.

An example of spatial and temporal control of chromatin-mediated by a lncRNA occurs at *frq* in Neurospora. The *frq* locus has a large natural antisense transcript (NAT) that is required for circadian-regulated facultative heterochromatin. The *qrf* transcript oscillates antiphasic to *frq* and inhibits the expression of *frq* by establishing H3K9me3 and DNA methylation (*100*) (*119, 120*). In mice, there is also a *Per2* NAT (*Per2AS*) that oscillates antiphasic to *Per2* (*81*), but the function of *Per2AS* is still unknown. Based on the antiphasic expression between *Per2* and *Per2AS*, two models have been put forward. In the pre-transcriptional model, *Per2AS* prevents *Per2* mRNA transcription by interacting with the gene and *vice versa* as occurs in the classical transcriptional interference model. In the post-transcriptional model, a sense-antisense RNA duplex between *Per2* and *Per2AS* prevents translation and is rapidly degraded (*127*). However, these models are based solely on computational modeling and have not been empirically tested, so they remain entirely speculative. A separate possible mechanism that has not been modeled or tested is that *Per2AS* functions analogous to *qrf* in *Neurospora*. Thus, mammalian *Per2AS* could possibly guide H3K27me3 and/or H3K9me3. This idea is

rooted in analogous systems ranging from pericentric heterochromatin in *S. pombe* to imprinting and X-chromosome inactivation in mammals. Regardless, it is abundantly clear that the circadian clock functions to control the timed expression of lncRNAs that mediate chromatin modification at specific genome loci in a time-specific manner.

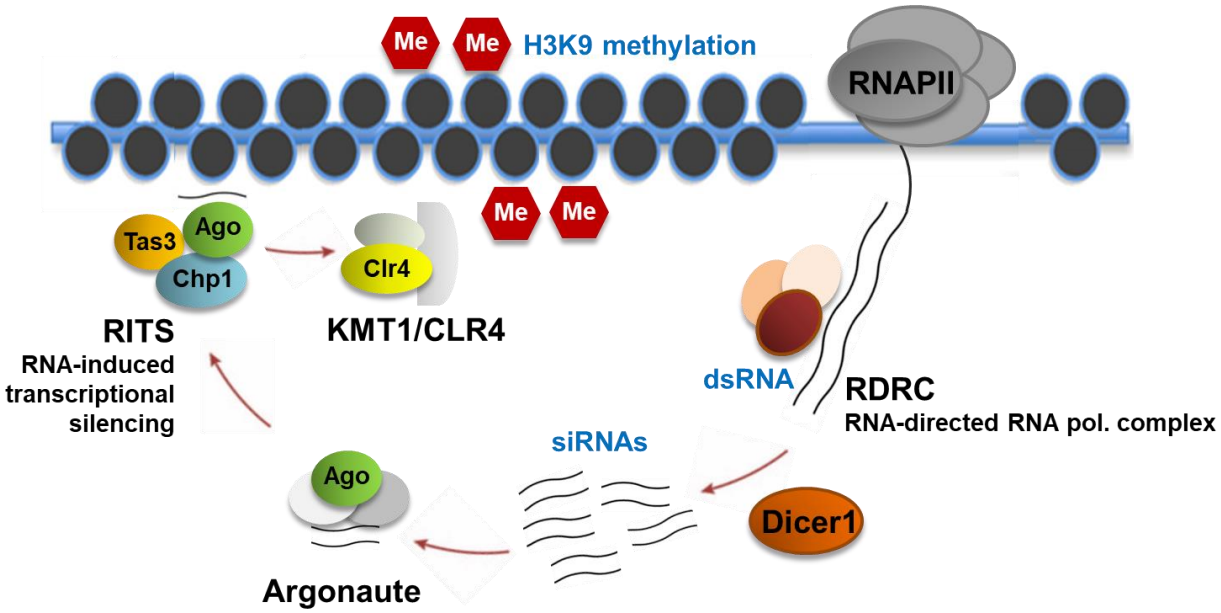


Figure 2. Schematic model of RNAi-mediated heterochromatin assembly in *S. pombe*. Nascent transcripts generated by RNA polymerase II. RNA-directed RNA polymerase complex (RDRC) is then recruited for the synthesis of dsRNA from the transcripts, which is cleaved into duplex siRNAs by Dicer1 and bound by Argonaute. These siRNAs are then processed into single-stranded siRNAs and loaded back into RITS completing the loop. Catalytically active RITS recruits the KMT1/CLR4 for spreading of epigenetic silencing into adjacent regions.

1.4 Circadian clock and aging

Aging is associated with a gradual loss of homeostatic regulation needed for maintenance and function of cells, tissues, and organs (*105, 106, 128*). For example, physiological changes in aging include a decline in fecundity and cardiac output along with increases in blood pressure and development of arteriosclerosis (*129, 130*). Aging affects all aspects of our physiology and behaviors, including circadian outputs. A well-known symptom of old age in mammals is the weakening of rhythmic behavior such as sleep/wake cycle and rhythms in the release of hormones including melatonin and cortisol, and fasting plasma glucose levels. Studies on *Drosophila* also revealed that a decline in sleep consolidation and weakened rest/activity rhythms occur in aging flies (*131*).

Physiological aging is accelerated by chronic disruption of clock function in mammals. For example, genetic disruption of clock genes manifests as a premature aging phenotype in mice. *Bmal1*$^{-/-}$ mice suffer from a series of conditions related to accelerated aging including sarcopenia, cataracts, cornea inflammation, osteoporosis, and premature hair loss (*132*). In addition, *Bmal1*$^{-/-}$ have significantly shorter lifespan compared to an isogenic background (*133*). *Clock*$^{-/-}$ mice also have a shorter average lifespan compared to wild-type mice and show advanced aging in response to low-dose irradiation (*133, 134*). *Bmal1*$^{-/-}$ mice have a more severe aging phenotype compared to *Clock*$^{-/-}$ because the CLOCK paralog NPAS2 can substitute for CLOCK in *Clock*$^{-/-}$ mice (*135, 136*). Therefore, it has been speculated that age-related changes in circadian clock may account for much of the physiological decline with age.

Recent genome-wide studies revealed a global reprogramming of the circadian transcriptome in a tissue-specific manner during aging (*137-140*). For instance, the age-induced changes in circadian gene expression in liver tissue are associated with biological pathways such as protein acetylation and NAD^+ metabolism. This finding expands on the NAD^+ deacetylation rhythm brought on by rhythm in the *Nampt* gene and suggests that reprogramming is due in part to age-related changes in acetylation brought on by slower liver metabolism. In addition, the extensive reprogramming of circadian transcriptome has recently been reported in aged stem cells (*137*). The subset of circadian controlled transcripts in aged stem cells changes from genes involved in homeostasis to those involved in tissue-specific stresses such as DNA damage or inefficient autophagy to cope with tissue-specific stress (*137*).

There are also numerous oscillating genes in young animals whose rhythms are lost with age, while other genes become rhythmic in aged animals (137, 138). However, the expression of the core clock maintains a rhythm regardless of age. Therefore, it appears the core clock is unaffected by age, but circadian output, including clock-associated physiological processes, changes dramatically. How age-related changes to circadian output occur without affecting the core circadian is currently unknown.

1.5 Circadian clock and telomeres

The long-term consequences of the disrupted circadian system are associated with accelerated telomere shortening and aging. For instance, rotating night shifts and short sleep duration are correlated with shortened telomere, a marker of aging. Moreover, deficiency of *TERT* gene in a genetic disorder dyskeratosis congenita leads to progressive shortening of telomeres and is associated with premature aging in adult (*141*). However, the molecular connection between the circadian clock and telomere homeostasis is still unknown.

Telomeres are a DNA and protein complex at the end of eukaryotic chromosomes (*142, 143*). The DNA sequence of the telomere is composed of TTAGGG repeats bound by shelterin, which serves to maintain genome integrity of linear chromosomes. The linearity of eukaryotic chromosomes causes two problems; an end-protection and an end-replication problem. The end-protection problem arises because the linear ends of the chromosome resemble a double strand break and need to be protected from inappropriate DNA repair. To deal with this problem, the telomere forms a T-loop structure at the end of chromosomes. The telomere terminates in a 3’ single-stranded overhang that invades the double-stranded telomeric repeats to form this T loop structure, which is essential for telomere protection (*144, 145*). The shelterin complex is also associated with the T-loop and protects the telomere ends from DNA damage response pathways (*146*).

In contrast, the end-replication problem occurs because DNA polymerase needs to travel past a section of DNA to synthesize the lagging strand and thus incomplete DNA replication occurs at chromosomes end (*147*) causing telomeres to become shorter after every cell division in somatic cells. To solve the end-replication problem, cells use

telomerase to add the TTAGGG repeat to the ends of chromosomes. Telomerase is an RNA-dependent DNA polymerase, composed of the catalytic subunit telomerase, which is a reverse transcriptase (TERT) and its RNA component (TERC) (*148*). Telomerase adds telomeric DNA to telomeres to maintain the telomere length using TERC as an RNA template (*144, 149*). In contrast to germline and stem cells, somatic cells have restricted telomerase expression leading to the erosion of telomeres, which limit cell division. Thus telomere shortening is observed with progressive cell division, indicating a pace of cellular aging (*150, 151*).

There is some evidence of a direct molecular connection between the clock and telomerase, supporting the idea that the circadian clock plays an important role in telomere maintenance. Human and mouse *TERT* mRNAs have a circadian rhythm and are proposed to be under direct control of CLOCK:BMAL1. *Clock*$^{-/-}$ mice do not appear to have a rhythm in *TERT* mRNA, and they have shortened telomeres relative to WT littermates (*152*). Also, hospital physicians with regular work schedules have a circadian oscillation of telomerase activity while emergency physicians working in night shifts lose the circadian rhythms of telomerase activity (152).

Taken together, these findings suggest there are connections between the circadian system and telomeres through control of TERT transcription and telomerase activity. Although more solid data are needed to support this theory, it can be speculated that the disruption of circadian clock causes misregulation of telomere homeostasis, contributing to accelerated aging and various human diseases. Thus, this thesis study to figure out that circadian regulated changes in lncRNAs and heterochromatin will provide a solid background to discover the precise mechanism underlying them.

CHAPTER 2

BMAL1 associates with telomeres to control rhythms in TERRA and heterochromatin

2.1 ABSTRACT

The circadian clock and aging are intertwined and impact cell physiology whereby disruption to normal diurnal rhythm causes accelerated aging, an increased incidence of age-related diseases and morbidity. At the cellular level, the aging is associated with telomere shortening. Telomere attrition also correlates with a disrupted circadian rhythm, and the clock regulates telomerase. However, the molecular pathway connecting telomere homeostasis and circadian clock is still unclear. In this chapter, I found that circadian transcription factors interact with telomere DNA suggesting a direct role of the circadian clock in telomere regulation. Through inspection of published ChIP-seq data, I found that White Collar 2 (WC-2) in *Neurospora* and BMAL1 in mice associate with the telomere. Subsequent ChIP analysis confirmed that the associations are rhythmic and I demonstrated that BMAL1 binding to the telomere is conserved in zebrafish and mice. I also found a circadian rhythm in Telomeric repeat-containing RNA (TERRA), a lncRNA transcribed from the telomere. The rhythm in TERRA was lost in *Bmal1*$^{-/-}$ mice proving a BMAL1-dependent rhythm in TERRA. Furthermore, H3K9me3 ChIP revealed a BMAL1-dependent rhythm in heterochromatin at the telomere, and this was lost in aged animals. Taken together, my results provide evidence that BMAL1 plays a direct role in telomere homeostasis by regulating rhythms in TERRA and heterochromatin which likely accounts for the increased telomere erosion that occurs coincident with circadian disruption.

2.2 INTRODUCTION

The circadian clock generates physiological and behavioral rhythms coinciding with 24-hour light and dark cycle. Other environmental changes such as food intake, seasonal variation, or redox status in the cells also affect the regulation of the circadian oscillator. Thus, the circadian clock plays a critical role in controlling biological events to anticipate external environmental changes (*5-7*). Studies dating back decades have explored the molecular mechanisms of circadian synchronization and entrainment in response to environmental alterations and revealed the regulation of the circadian clock is built upon a transcription feedback loop. In vertebrates, the positive arm of the clock is driven by the transcriptional activators CLOCK:BMAL1, which drive expression of the negative elements *Period* (*Per1, Per2,* and *Per3*) *and Cryptochrome* (*Cry1* and *Cry2*). As part of the negative feedback, HDAC1/2 in the per complex deacetylates the chromatin at *Per* loci and then establishes H3K9me3-dependent heterochromatin generating a repressive chromatin state (*36*). H3K9me3 is established by KMT1/Suv39H in the PER complex and H3K9me3 is this is bound by HP1, which is a transcriptional repressor that directly binds to the H3K9me (38, 39). A second feedback loop is generated by RORα and REV-ERVα, which activates and represses *Baml1* transcription, respectively (*41, 42*). The interlocked and cooperative feedback loops of the circadian clock generate a rhythm to help maintain accurate circadian timing. The clock also regulates the rhythms in *clock controlled genes (ccgs)* expression to generate phase-specific circadian outputs and optimal timing for biological processes (*34*).

Circadian disruption affects a multitude of physiological processes and is thus implicated in the development of many pathological states. For instance, shift workers at

night have an increased risk of metabolic syndromes, cardiovascular disease and cancer relative to the general population (*153*). Similarly, clock gene mutations in rodent models cause accelerated aging suggesting a connection between the circadian clock and aging. For example, mice lacking *Bmal1*, *Clock* or *Per1* and *Per2* have premature aging phenotypes and severely shortened lifespan (*132*). Moreover, the disrupted circadian clock is associated with telomere shortening and aberrant telomerase activity (*152*), suggesting circadian clock is likely involved in aging by regulating telomere homeostasis.

Telomeres are the specialized DNA-protein complex at the ends of each chromosome and are essential for maintaining genome integrity. Telomeres function as end-caps on the ends of chromosomes and are shortened during the cell cycle due to incomplete replication (*144, 154, 155*). In eukaryotes, incomplete replication on the lagging-strand causes telomeres to shorten each cell division [7,8]. This end-replication problem is solved by telomerase; an RNA-dependent DNA polymerase. Telomerase contains a reverse transcriptase (TERT) and an RNA component (TERC). TERT adds a six-nucleotide repeat TTAGGG to the 3' end of chromosomes using TERC as a template (*148*), which elongates the 3′ overhang to maintain the telomere length (*144, 149*). The linearity of eukaryotic chromosomes also results in an end-protection problem. When a chromosome breaks, the exposed DNA ends can activate the DNA damage response and repair pathways. These DNA damage response pathways require repression at telomeres to protect the telomere from end-to-end chromosome fusions or inappropriate recombination. In mammals, the end-protection problem is solved by a telomere-specific protein complex called shelterin, which helps forms the t-loop structure at telomere (*156, 157*). The shelterin complex protects telomere ends from DNA damage response as well

as non-homologous end joining (NHEJ) to block unnecessary recombination or translocations (*146*). Shelterin is composed of 6 subunits, TIN2, TRF1, TRF2, POT1, TPP1, and RAP1, that is required for the protection and replication of chromosome ends (*158, 159*). In addition to Shelterin, telomeres are packaged into heterochromatin containing H3K9me3, H4K20me3, and H3K27me3, which is essential for the protection of chromosome ends (*160*). Telomere H3K9me3 is added by histone lysine 9 methyltransferases (KMT1/Suv39), and heterochromatin protein 1 (HP1) binds to H3K9me3. Heterochromatin at the telomere is important for telomere capping and silencing (*161, 162*).

Due to the heterochromatin structure at telomeres, the telomeres were once believed to be transcriptionally silent. However, this changed with the discovery of a long non-coding RNA called Telomeric Repeat-containing RNA (TERRA) that is transcribed from the telomere. TERRA transcripts are heterogeneous with transcripts ranging in size between ~100 bp to 9 kb (*163*). TERRA has been implicated in telomere protection including heterochromatin formation (*164-166*), telomere elongation (*167*), and telomere replication (*168*). For instance, models have been proposed suggesting TERRA leads to recruitment of factors such as KMT1/Suv39h and HP1 to promote the heterochromatin around telomeres.

Disruptions to TERRA can directly or indirectly induce diseases such as astrocytoma, accelerated aging, and senescence (*169-172*). However, the mechanism of TERRA function is largely inferred because generating a loss-of-function TERRA model has proved difficult. As TERRA is synthesized from almost every chromosome, traditional genetic ablation approaches would be overly cumbersome and efforts to

knockdown TERRA using siRNA have been unsuccessful.

In this chapter, I found that White Collar-2 (WC-2, the ortholog of the transcription factor CLOCK) in *Neurospora crassa* (Neurospora) is enriched at the telomere suggesting a direct connection between the circadian clock and telomeres. I expanded this finding to animal models including zebrafish and mice and demonstrated a conserved system where clock transcription factors associate with the telomere. In addition to WC-2 in Neurospora, I found core clock protein BMAL1 in zebrafish and mice associate with a telomere, and its binding is rhythmic. I also determined there is a diurnal rhythm in TERRA expression in zebrafish and mice, and the diurnal rhythm is lost in *Bmal1*$^{-/-}$ mice indicating the rhythm in TERRA is under the control of the circadian clock. Lastly, I uncovered a diurnal rhythm in H3K9me3 at telomeres in zebrafish and mice, which is lost in *Bmal1*$^{-/-}$. These findings reveal a direct role for the circadian clock in telomere homeostasis by regulating rhythms in TERRA and H3K9me3. These findings may explain some of the advanced aging phenotypes that occur coincident with circadian disruption.

2.3 RESULTS

WC-2 localizes to the telomere in *Neurospora*

While examining published WC-2 chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) data, I observed that WC-2 appeared to localize to telomeres and localization was adjacent to Dicer-independent small interfering RNAs (disiRNA) (Fig. 3A) (*99, 173*). To confirm WC-2 localization at the telomere, I performed a standard ChIP followed by quantitative PCR (qPCR) using WC-2 antibodies and oligonucleotides adjacent to the telomere repeat. This WC-2 ChIP was performed on circadian entrained cultures with the *wc-2* deletion strain as a control for non-specific background. Results of the ChIP revealed a rhythm in WC-2 association with the telomeres (Fig. 3B). Moreover, the profile of WC-2 binding at the telomeres looked analogous to WC-2 association with the C-box in the *frequency* (*frq*) promoter (Fig. 3C). The discovery of WC-2 at telomeres suggested the circadian clock has a direct role in telomere regulation and implicates the clock in aging. However, studying telomere homeostasis and aging in *Neurospora* is hindered by numerous challenges including a lack of good aging assays, phenotypes, and reagents (*174, 175*). Moreover, observations in *Neurospora* may have little translational applications if the binding is not conserved in higher organisms; therefore, I proceeded to explore the regulation in both zebrafish and mice.

BMAL1 associates with the telomere

To determine whether circadian transcription factor binding was fungal-specific or a general clock mechanism, I obtained BMAL1 ChIP-seq data (GSM654882, GSE39977) from the GEO Database (www.ncbi.nlm.nih.gov/geo), mapped the reads to the Mus musculus 10 (mm10) genome, and examined BMAL1 association with the telomere (*81, 176*). Inspection of telomeres indicated BMAL1 binding was enriched in sequencing reads containing the telomere repeat suggesting BMAL1 localized to the end of chromosomes. Moreover, this binding appeared to have a diurnal rhythm (Fig. 4A, B). This finding was presumably missed in the original BMAL1 ChIP-seq analysis because the data was mapped to the mm9 assembly, which does not contain any telomere sequence (note, mm10 only has telomere sequence for chromosome 1, 2, 8, 14 and 18, Supplementary fig.1). To conclusively determine if BMAL1 localized to the telomeres and rule out potential ChIP-seq artifacts, I generated a BMAL1 antibody specific to the zebrafish isoform. Antibody specificity was tested via ChIP using oligonucleotides specific to the *Per2* E-box (Supplementary fig.2). Next, I examined the association of BMAL1 with the telomere by ChIP-slot blot on zebrafish brain tissue using the telomeric repeat (TTAGGG) as a probe at ZT2 and ZT10. The data indicates BMAL1 associates with telomere DNA and binding is higher at ZT10 relative to ZT2 (Fig. 4C). As further confirmation, I examined BMAL1 binding by ChIP-qPCR at the telomeres (*177*). The ChIP-qPCR further confirmed BMAL1 binding to the telomere at ZT10 and indicated an 18-fold increase relative to ZT2 (Fig. 4D). To test the rhythms in the BMAL1 binding to the telomere, I then proceeded to examine BMAL1 binding to the telomere by performing BMAL1 ChIP-qPCR from zebrafish brain tissue over a full 24-h diurnal cycle sampling

every 4 hours using oligonucleotides specific to a subtelomeric region on the left arm of chromosome 1. Figure 5A indicates there is a diurnal rhythm in BMAL1 associated with the subtelomeric region on chromosome 1 and binding peaked around ZT12.

BMAL1 associates with the telomere in mice.

Rhythmic binding of BMAL1 at the telomere may have major physiological consequences. Therefore, to conclusively determine if BMAL1 associates with the telomere, I obtained the *Bmal1*$^{-/-}$ mouse line (B6.129-Arntltm1Bra/J, Jackson lab) and used a commercially available BMAL1 antibody to establish BMAL1 telomere binding in mouse. The BMAL1-ChIP was performed in both WT and *Bmal1*$^{-/-}$ mice liver over a 24-hr diurnal cycle. To ensure proper entrainment, I measured the diurnal expression of *Per2* and *Bmal1* by RT-PCR. This analysis showed *Per2* and *Bmal1* transcripts had an anti-phasic pattern in WT and the diurnal rhythm was abolished in *Bmal1*$^{-/-}$ mice (Supplementary fig. 3A, B). Next, I examined whether BMAL1 localizes to the telomere by measuring the interaction between BMAL1 and the telomere at chromosome13. I determined there was a rhythm in BMAL1 binding that peaked between ZT0 - ZT4 in WT (Fig. 5B). Analysis of peak to trough levels (ZT0 vs. ZT12) indicated a significant change in BMAL1 binding, whereas the background amplicons in *Bmal1*$^{-/-}$ mice showed no significant difference among any time points or with the trough in WT (ZT12). As a further control for the BMAL1 ChIP, I also examine BMAL1 localization to the *Dbp* ,which is a *ccg* with a known E-box element in its promoter, and found a similar, but a slight phase delayed diurnal interaction that is consistent with previous reports (Fig. 5C) (*178*).

Rhythmic expression of TERRA

The association of clock transcription factors with the telomere in *Neurospora*, zebrafish, and mice led to the obvious question about the function of the BMAL1 binding to the telomere. Because TERRA is expressed from telomeres, I hypothesized that BMAL1 may regulate transcription of TERRA. As an initial test, I proceeded to examine if there was an oscillation in TERRA in zebrafish. Zebrafish were entrained under a 12:12-hour light-dark cycle for one week and sacrificed at ZT2 and ZT10, and both brain and liver tissue was collected. To observe the level of TERRA between the 2-time points, I performed a Northern blot on total RNA using a TERRA-specific oligonucleotide probe $(TTAGGG)_5$. The data indicate TERRA was higher at ZT2 relative to ZT10 in both brain and liver (Fig. 6A). To test rhythmic changes in TERRA transcripts in the full-time course, I proceeded to examine TERRA expression over a 24-h cycle sampling at 4-h intervals. A representative TERRA Northern blot from three independent biological replicates is shown in Figure 6B and suggests TERRA expression is rhythmic with a peak at ZT16-ZT0 and lowest between ZT4-ZT12. Quantification and statistical analysis of the 3 independent biological replicates confirmed the rhythm peaked during the night (Fig. 6C).

Next, I examined if the apparent rhythm in TERRA was circadian regulated and whether it was dependent on BMAL1. To do so, I performed TERRA Northern blots on mouse WT and *Bmal1*$^{-/-}$ liver over a 24-hr cycle. The TERRA Northern revealed a diurnal rhythm in wild-type liver that was absent in *Bmal1*$^{-/-}$ liver (Figure 7A). Quantification of the three independent biological replicates shows that TERRA peaked at the light to dark transition (Figure 7B). Zebrafish are diurnal while rodents are nocturnal and the rhythm in TERRA in both systems was consistent with the peak in expression occurring just

before their respective activity cycles. To further confirm the observed rhythm in TERRA expression, I inspected two previous circadian RNA-seq datasets performed on mouse liver (GSE39978, GSE36871) (*81, 179*) and found a rhythm in TERRA that peaked during subjective daytime (Supplementary fig. 4). I also examined TERRA in entrained human osteosarcoma U2OS cell line, which is one of well-known models for circadian research because they retain robust rhythms in most clock genes, and I found a dampened rhythm that peaked at ZT16 (Supplementary fig. 5).

Diurnal regulation of heterochromatin at telomere regions

It has been proposed that TERRA helps establish heterochromatin including H3K9me3, H4K20me3, and H3K27me3 at telomeres. The circadian clock also controls rhythms in heterochromatin. Accordingly, I hypothesized that the rhythmic TERRA expression mediated by clock proteins may affect heterochromatin structure at telomere. Therefore, I sought to determine if the rhythm in TERRA was accompanied by rhythms in heterochromatin. I began by performing H3K9me3 ChIP on cross-linked zebrafish brain tissue harvested at every 4-hour for 24-hours and detected a rhythm in H3K9me3 that peaked around ZT12 to ZT16 (Fig. 8A). In order to determine whether the H3K9me3 rhythm was dependent on a functional circadian oscillator and whether it was conserved in mice, I performed H3K9me3 ChIP in WT and *Bmal1*$^{-/-}$ mice. I observed a rhythm in H3K9me3 at the telomere on chromosome 13 in wild-type, but there was no apparent rhythm in H3K9me3 in *Bmal1*$^{-/-}$ (Fig.8B). As a control for the H3K9me3 rhythm, I measured H3K9me3 levels at *Dbp* in wild-type mice compared to in *Bmal1*$^{-/-}$ and confirmed a rhythm in WT, but not in *Bmal1*$^{-/-}$ consistent with a previous finding (Fig. 8C) (*178*). In WT mice, BMAL1 binding to *Dbp* was highest at ZT4 (Fig. 5B) as compared to H3K9me3, which peaked at ZT20 (Fig. 8B) representing an anti-phasic pattern between BMAL1 binding and the H3K9me3.

Aging and stress affect the diurnal regulation of TERRA and heterochromatin.

I next sought to gain insight into how and if the diurnal rhythm in TERRA changes with age and under stressed conditions with the idea that loss of normal TERRA expression may impact telomere integrity. Therefore, I tested TERRA expression at 3 different ages and under a stressed condition that included overcrowding and elevated ammonia nitrite levels (common stress for fish that occurs in the absence of regular water changes). In young and adult fish, TERRA displayed a rhythm in expression; however, the rhythm in TERRA appeared dampened in the brain of old and/or stressed zebrafish (Fig. 9A). Moreover, the diurnal rhythm of H3K9me3 in young fish disappeared in old animals (Fig. 9B). Consistent with findings in the brain, I also observed an age-related loss in TERRA and H3K9me3 in liver isolated from old and stressed animals (Supplementary fig. 7A and B) and further confirmed this in the skeletal muscle (Supplementary fig. 7C). These data suggest that the rhythm in TERRA and H3K9me3 is changed by aging and stress and supports the notion that circadian clock is related to telomere via rhythms in TERRA.

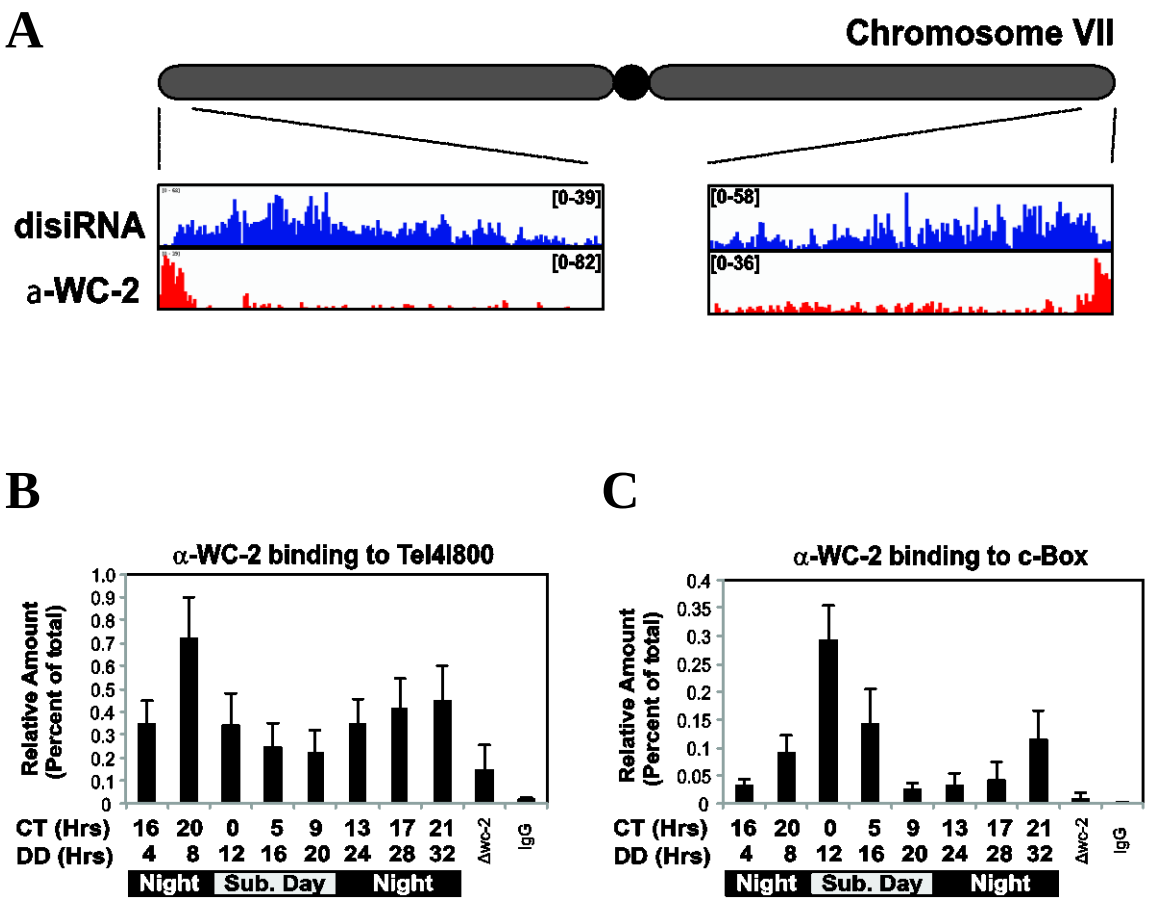


Figure 3. WC-2 is associated with the telomeres in *Neurospora*

(A) WC-2 ChIP-seq and Argonaute RIP-seq data in Neurospora showing WC-2 is localized to telomeres, and the binding is adjacent to Argonaute-associated Dicer-independent small interfering RNAs (disiRNA). Data were mapped to the *Neurospora* genome and visualized with Integrative Genomics Viewer (IGV). The interaction between WC-2 and telomere was confirmed by ChIP-qPCR under circadian entrainment (B) using the *frq* C-box as a control (C). The data were obtained from 4 biological replicates, and error bars represent the SEM.

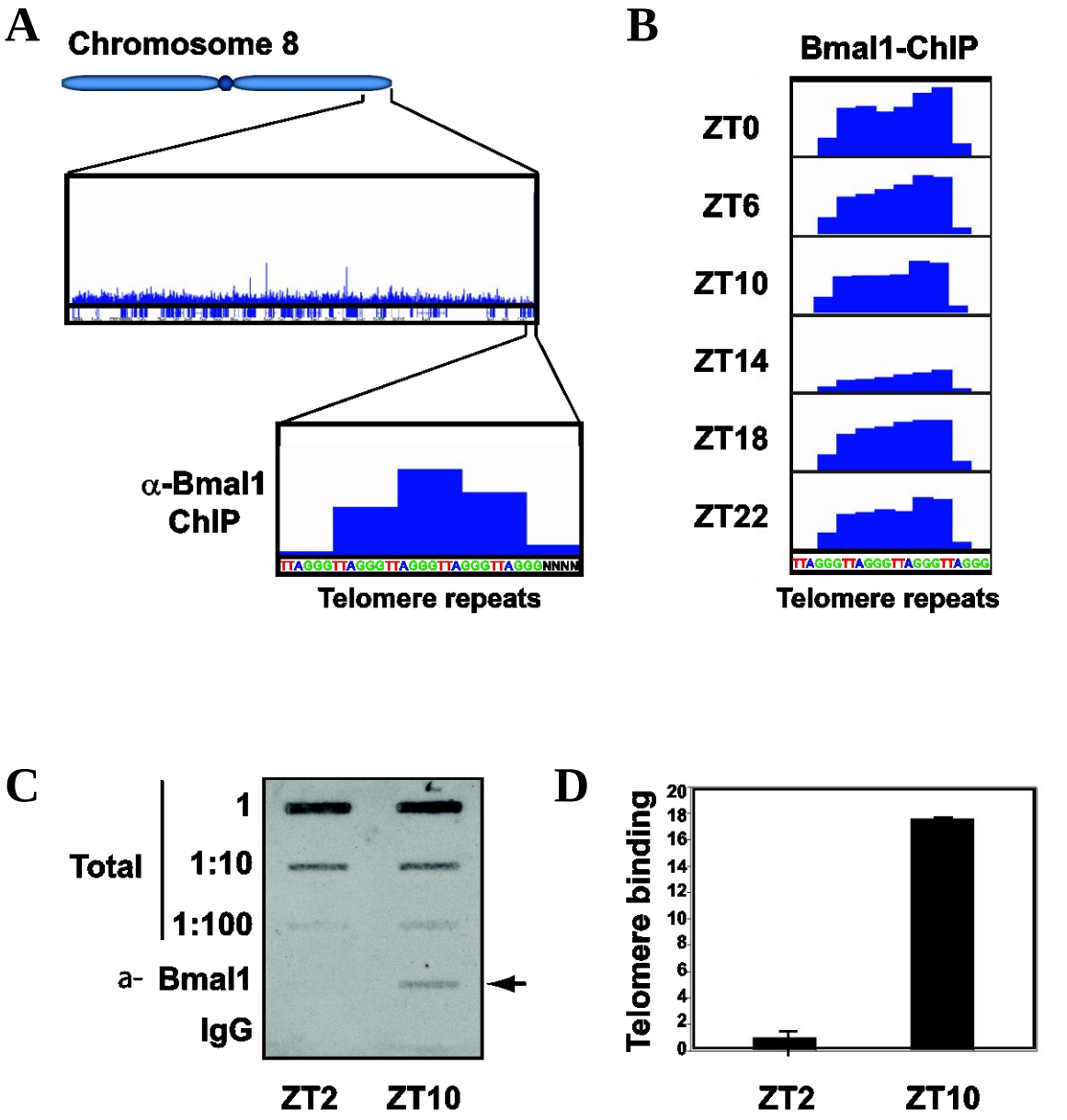


Figure4. BMAL1 at associates with the telomere in zebrafish and mice

(A) BMAL1-ChIP seq (GSM654882) indicates BMAL1 binding to the telomere repeat on mice chromosome 8. Zoomed image shows the BMAL1 peaks and TTAGGG repeats. (B) BMAL1 ChIP-seq over full circadian time course shows the association of BMAL1 to the telomere is rhythmic at chromosome 14. (C) BMAL1 ChIP-slot-blot from zebrafish skeletal muscle using the telomere repeat $(TTAGGG)_5$ probe. (D) The level of BMAL1 in zebrafish associated with the telomere determined by qPCR.

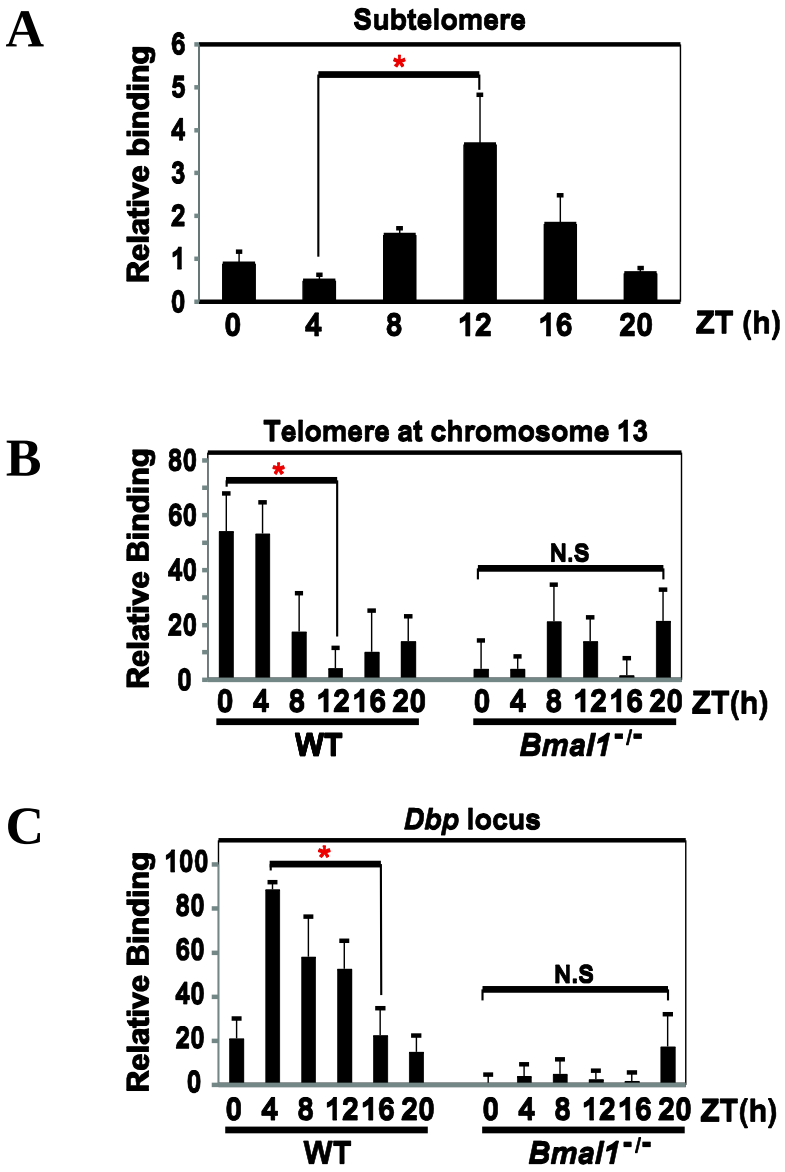


Figure 5. A diurnal rhythm between BMAL1 and telomere in zebrafish and mice. (A) BMAL1 ChIP from zebrafish brain tissue collected at designated time points showing BMAL1 associates with the subtelomere region on chromosome 1. The data are averages of 5 independent biological replicates. (B) BMAL1 ChIP revealed binding to the telomere on chromosome 13 in WT compared to *Bmal1*$^{-/-}$ mice. (C) Same as in B except binding was assayed for *Dbp* which is a known clock-controlled gene with the binding site of BMAL1. The data in B and C are from 4-6 independent biological replicates, and the error bars show SEM. Analysis was by one-way ANOVA followed by post hoc test (* ; $p \leq 0.05$).

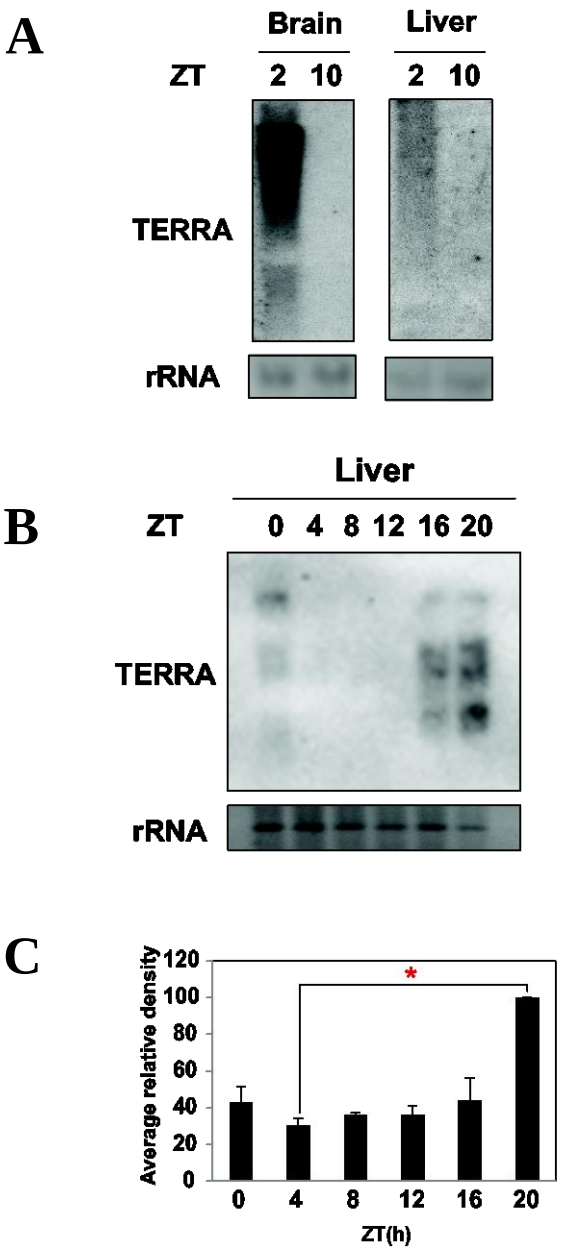


Figure 6. Zebrafish have a diurnal rhythm in TERRA expression.

(A) Northern blot of total RNA isolated from zebrafish brain and liver were probed for TERRA at ZT2 and ZT10. (B) A representative TERRA Northern blot from zebrafish liver over a full diurnal time course. (C) Data represents quantification of TERRA Northern blots from three independent biological replicates. The error bars show SEM. ANOVA was performed to determine the significant difference between samples. Analysis was by one-way ANOVA followed by post hoc test (* ; $p \leq 0.05$).

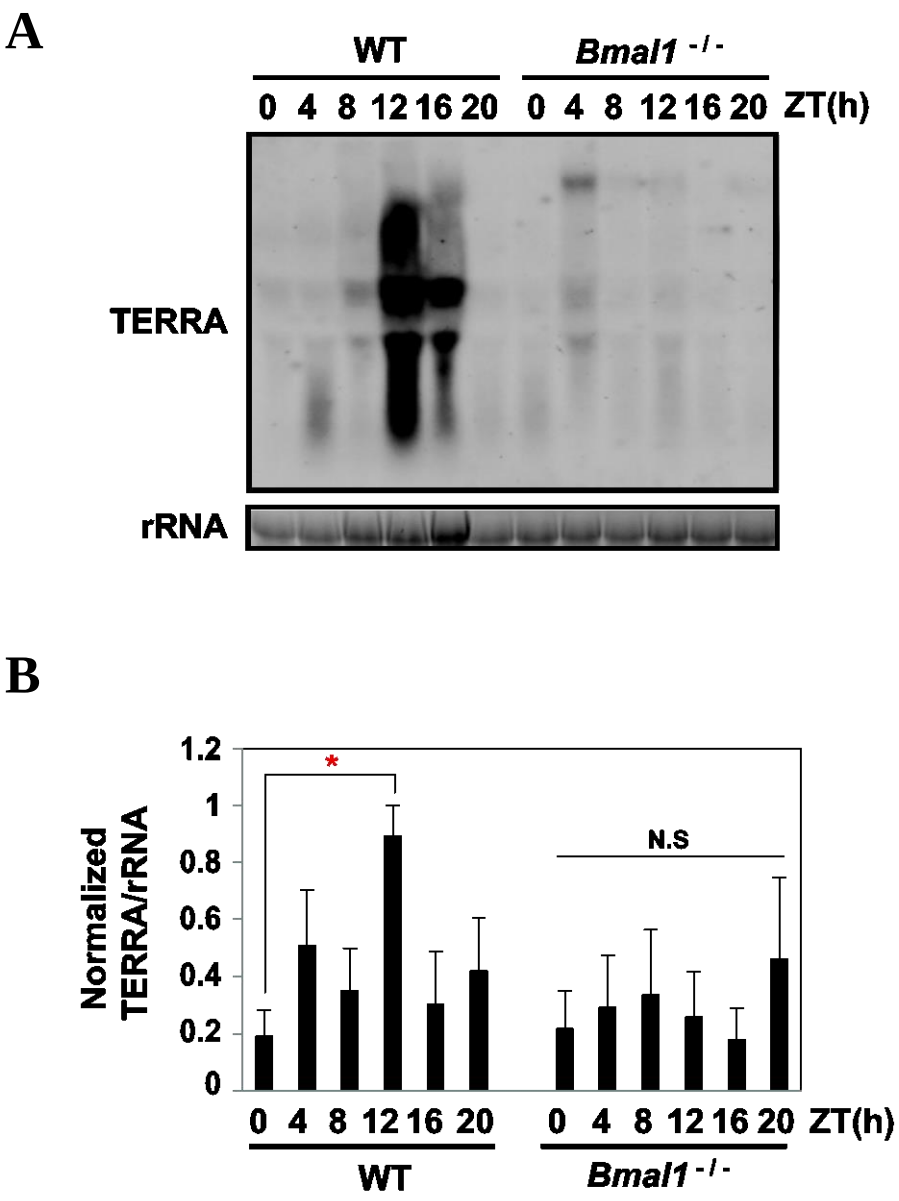


Figure 7. Diurnal Rhythm in TERRA requires BMAL1

(A) A representative TERRA Northern blot performed on total RNA isolated from WT and *Bmal1*$^{-/-}$ mice over a full-time course. (B) The quantification of 3 independent Northern blots was quantified by Image J and plotted as a bar graph. The error bars show SEM. The asterisks indicate $p \leq 0.05$. Analysis was by one-way ANOVA followed by post hoc test (* ; $p \leq 0.05$).

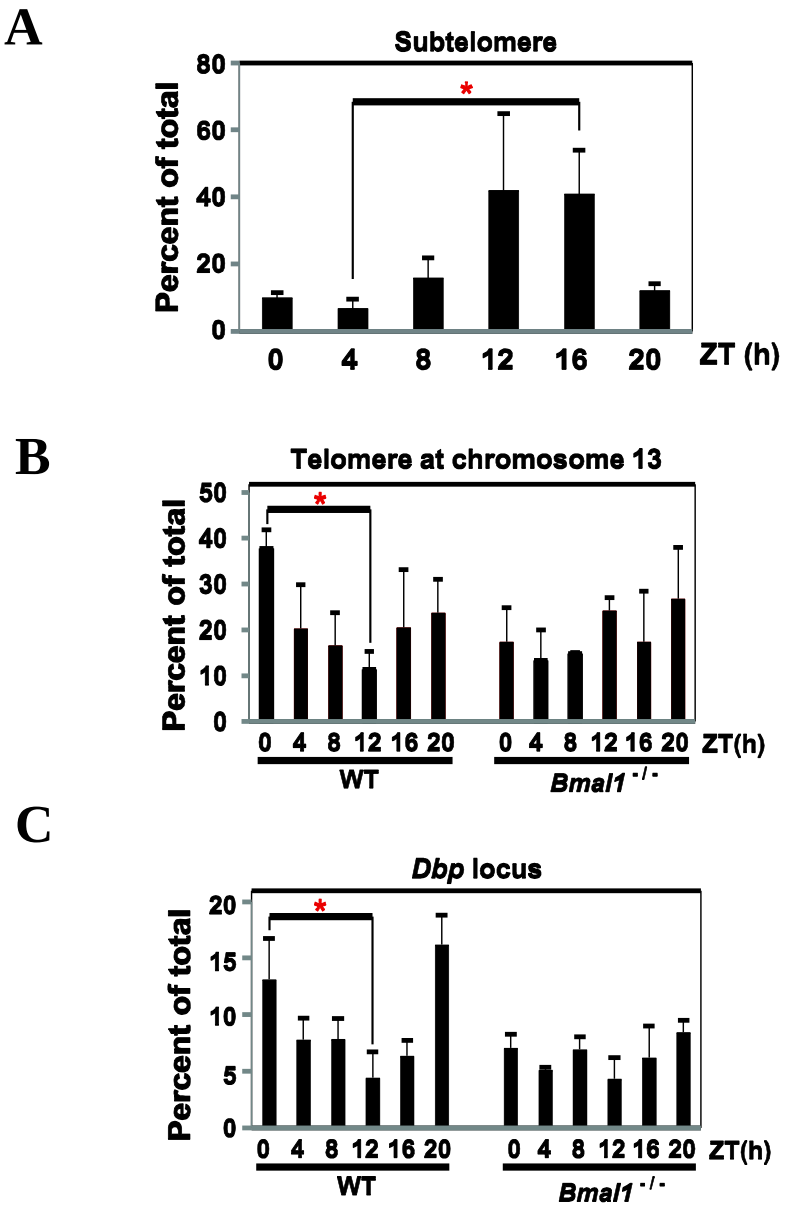


Figure 8. Rhythmic regulation of heterochromatin at telomere in zebrafish and mice

(A) H3K9me3 ChIP at subtelomere region of chromosome 1 in zebrafish brain tissues. Level of H3K9me3 was determined by qPCR using oligonucleotides in Table 2. (B) H3K9me3 levels at the telomere at chromosome 13 were determined by ChIP from mouse liver tissue in WT and *Bmal1*$^{-/-}$. (C) Same as in B except oligonucleotides were specific to *Dbp* locus. The data are averages from a minimum of 4 independent biological replicates. Error bars represent the SEM. Analysis was by one-way ANOVA followed by post hoc test (* ; $p \leq 0.05$).

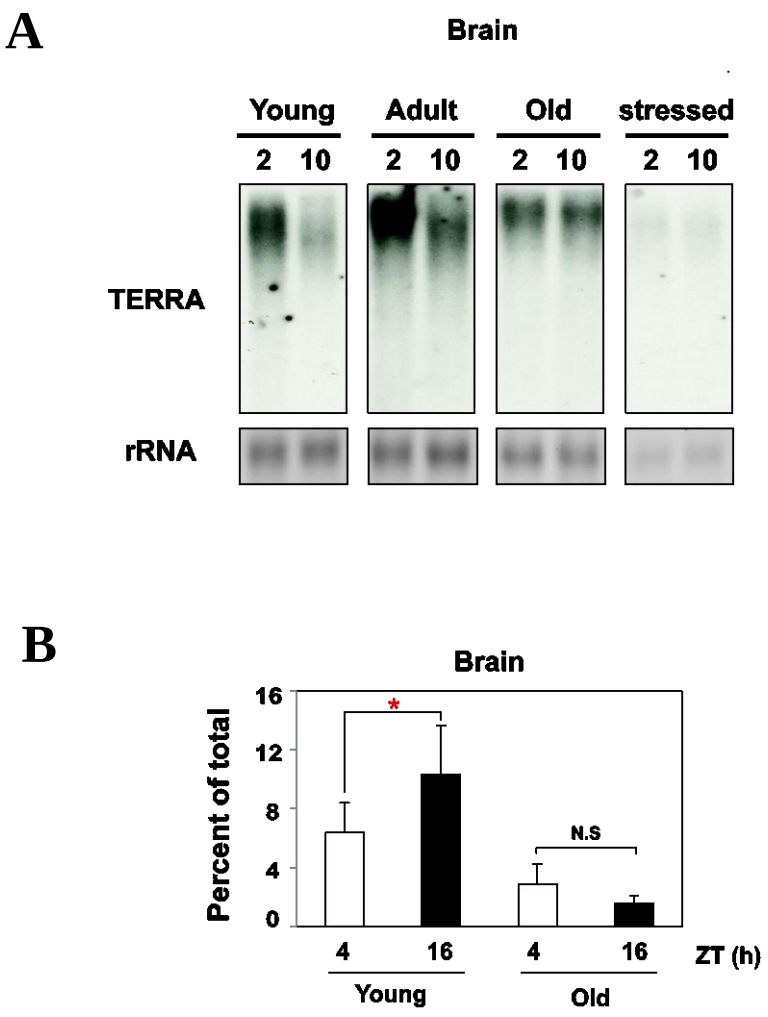
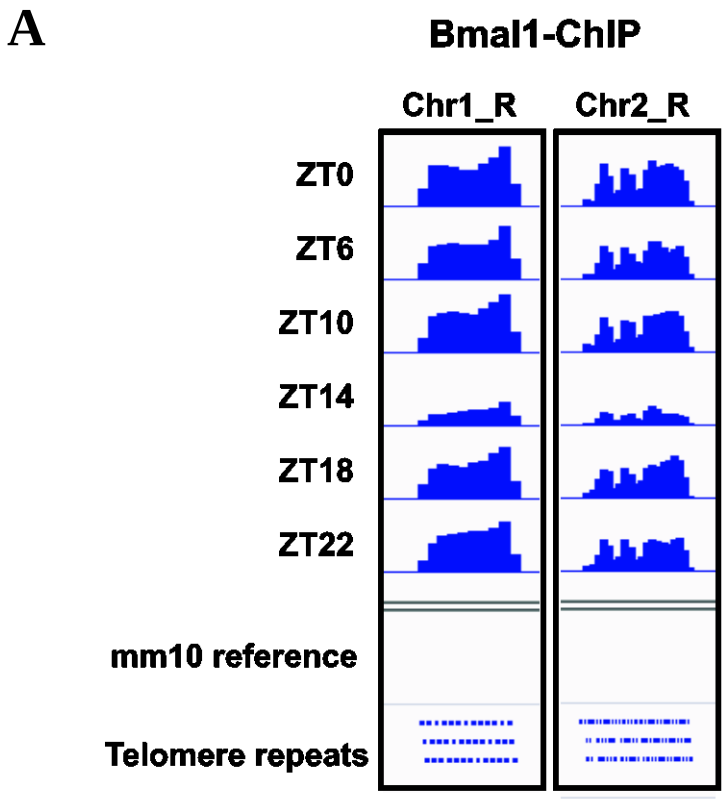
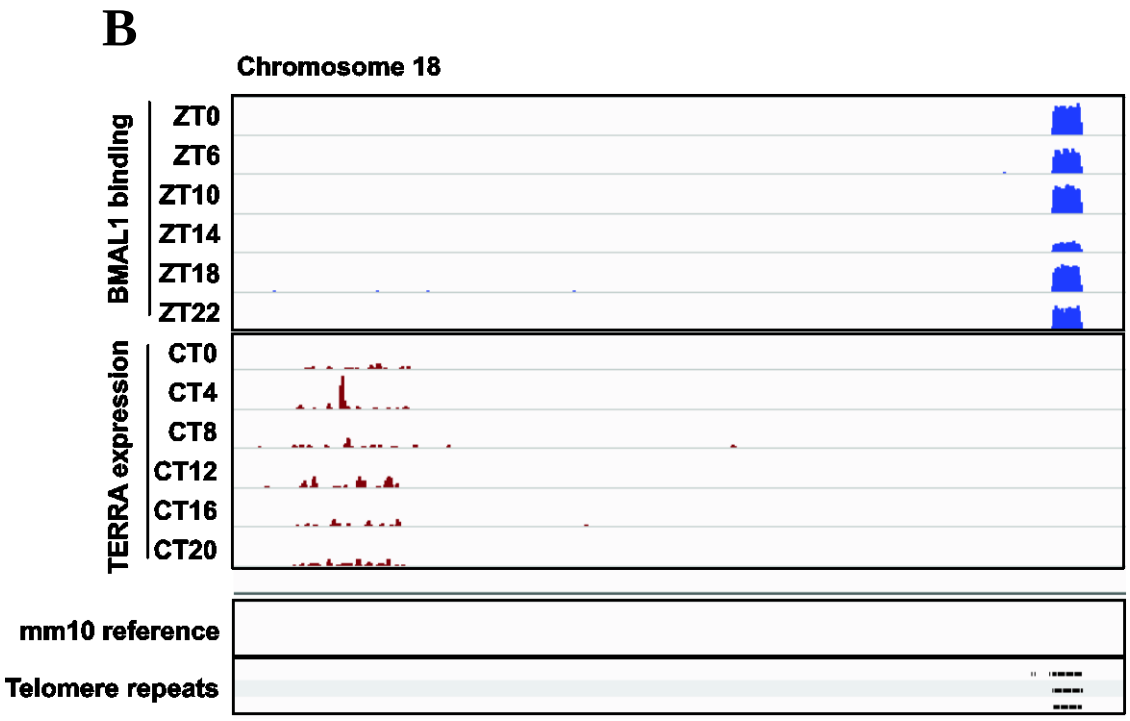


Figure 9. Aging alters the diurnal rhythms in TERRA expression and H3K9me3 at subtelomere in zebrafish

(A) A representative Northern blot examining diurnal TERRA expression between ZT2 and T10 at different ages and conditions (Young: 4M, Adult: 12M, Old: 20M, and Stressed). (B) Level of H3K9me3 at the subtelomere of chromosome 1 was measured by ChIP in zebrafish brain corresponding to different ages (Young; 4M, Old; 20M). The experiment is from 3 independent biological replicates. The error bars represent the SEM. Analysis was by t-test and the asterisk indicates $p \leq 0.05$.

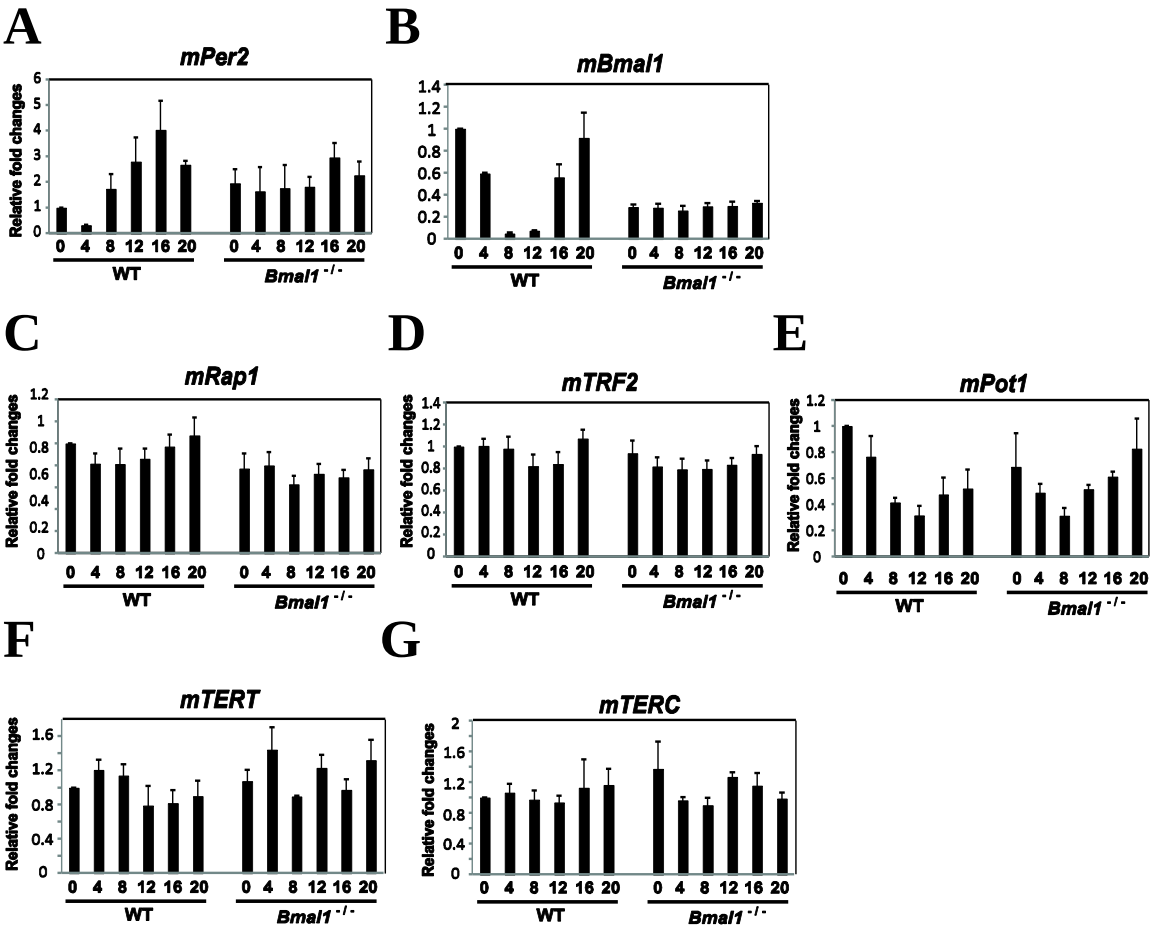

A
Bmal1-ChIP
Chr1_R
Chr2_R
ZT0
ZT6
ZT10
ZT14
ZT18
ZT22
mm10 reference
Telomere repeats


B
Chromosome 18
BMAL1 binding
ZT0
ZT6
ZT10
ZT14
ZT18
ZT22
TERRA expression
CT0
CT4
CT8
CT12
CT16
CT20
mm10 reference
Telomere repeats

Supplementary figure 1. Visualization of the binding of BMAL1 to telomere and TERRA expression in mice.

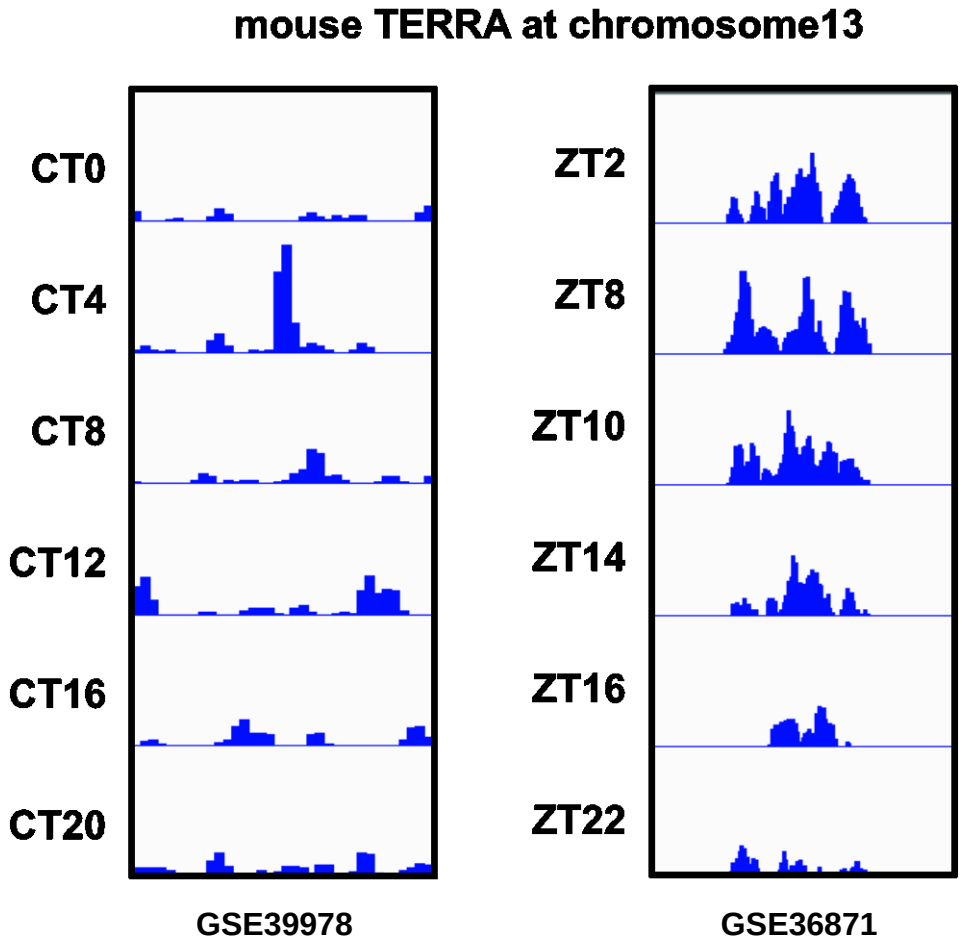
BMAL1 ChIP-Seq (GSM654882) indicates BMAL1 binds to the telomere on the right arm of chromosome 1 and 2. Data was mapped to the mm10 genome and visualized with IGV. (B) BMAL1 binding at the telomere of chromosome 18 and RNA-seq showing a probable TERRA transcript adjacent to the BMAL1 binding site.

Supplementary figure 2. Confirmation of the specificity of the BMAL1 antibody
The BMAL1 binding to at the *per2* E-box was examined by ChIP assay to test the specificity of lab-made antibody against BMAL1 peptides. After BMAL1 ChIP, qPCR was performed on zebrafish brain tissues at ZT2 and ZT10.



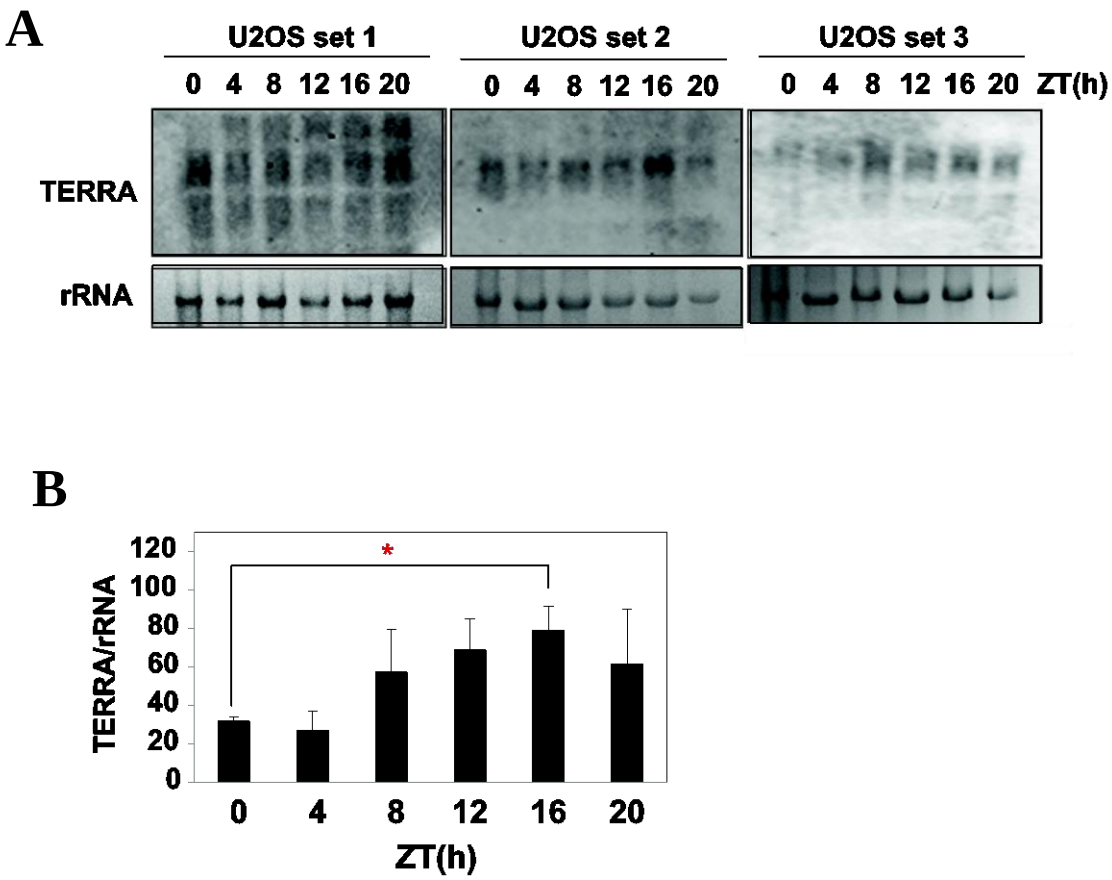
Supplementary figure 3. The level of transcripts of core clock genes and shelterin components in mice

RT-PCR of *Per2* (A) and *Baml1* (B) performed on liver tissue sampled every 4 h. (C-E) Expression of Shelterin components Rap1, TRF2, and Pot1 was measured by RT-PCR. (F, G) Expression of *TERT* and *TERC* transcripts was also determined by RT-PCR. The data are averages of 3 independent biological replicates. The error bars represent the SEM.



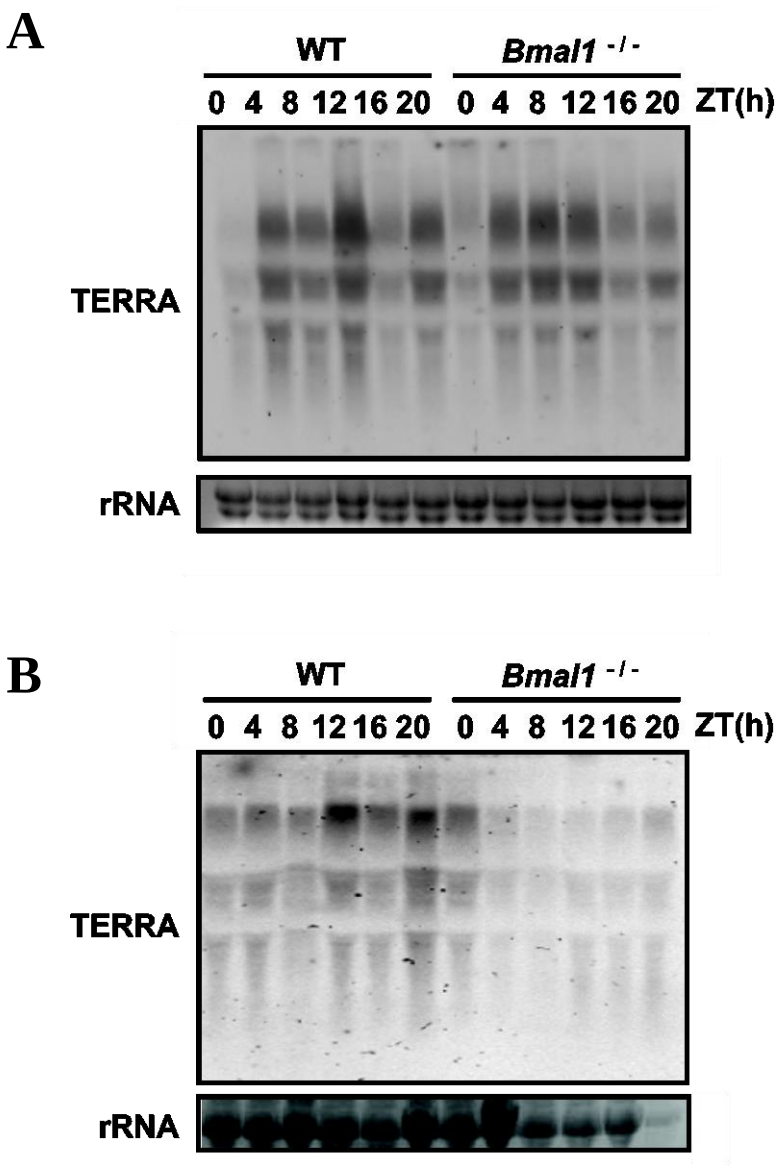
Supplementary figure 4. Rhythmic TERRA expression in mouse RNA-Seq data

TERRA expression in mouse liver tissue from two independent RNA-Seq experiments (GSE39978, GSE36871; replicate 2)



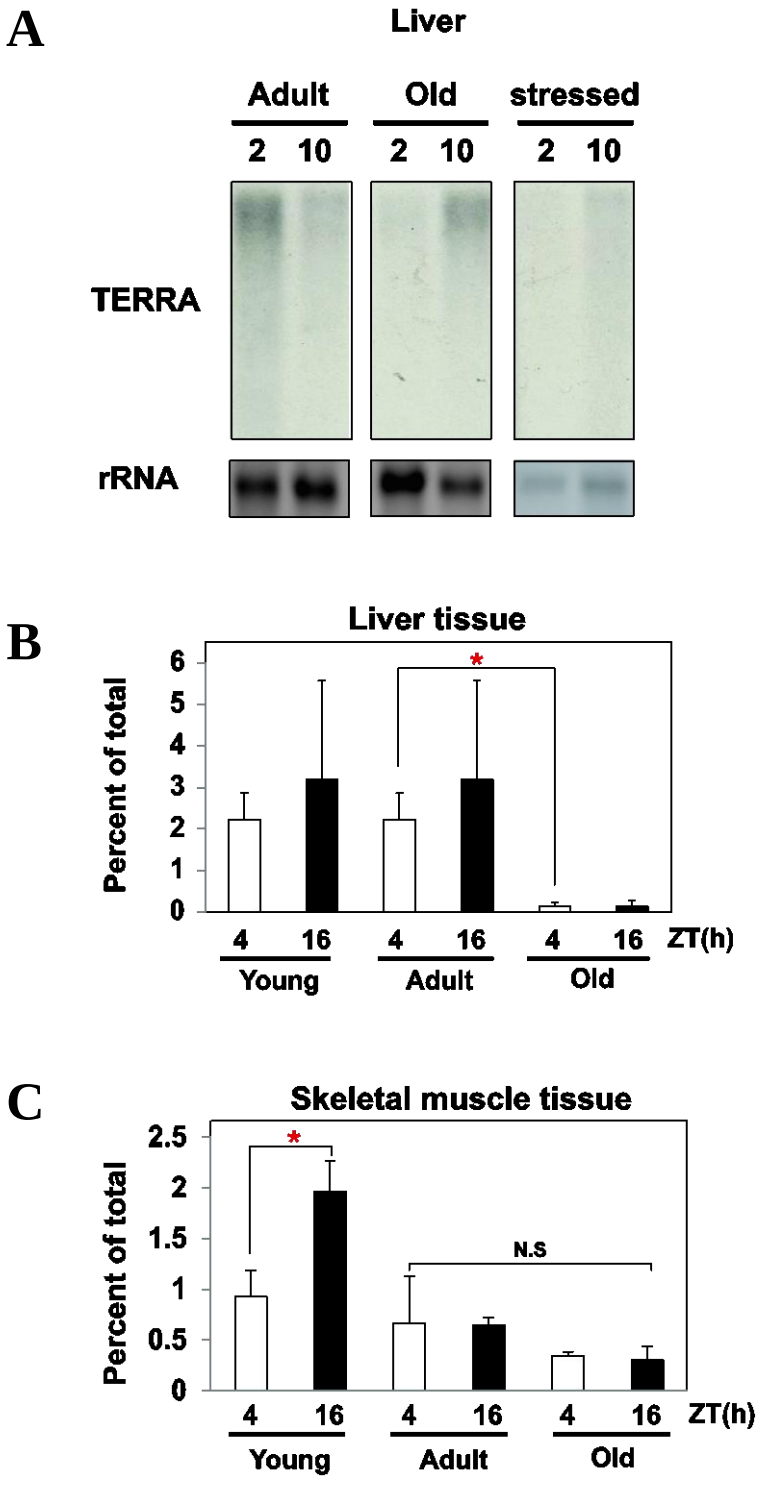
Supplementary figure 5. TERRA appears rhythmic in U2OS cells.

(A) Northern blots examining the level of TERRA transcript in human osteosarcoma cell line (U2OS) displayed for three independent biological replicates. (B) Quantification of the Northern blots from A was averaged and shown as a bar graph. Error bars show SEM and Analysis was by one-way ANOVA followed by post hoc test (* ; $p \leq 0.05$).



Supplementary figure 6. The diurnal rhythms in mouse TERRA

(A, B) Two additional TERRA Northern blots were done on RNA isolated from WT and *Bmal1*$^{-/-}$ mouse liver tissue. Data were combined with Figure 7A, and the quantification is shown in Figure 7B.


A
Liver
Adult
Old
stressed
2 10 2 10 2 10
TERRA
rRNA
B
Liver tissue
Percent of total
0 1 2 3 4 5 6
*
4 16 4 16 4 16 ZT(h)
Young
Adult
Old
C
Skeletal muscle tissue
Percent of total
0 0.5 1 1.5 2 2.5
*
N.S
4 16 4 16 4 16 ZT(h)
Young
Adult
Old

Supplementary figure 7. The effect of aging and stress on TERRA and H3K9me3
(A) TERRA Northern blots on RNA isolated from zebrafish liver at ZT2 and ZT10 under different conditions (Adult; 12M, Old; 20M, stressed n=2). (B) The level of H3K9me3 at the subtelomere of chromosome 1 from zebrafish liver was measured by ChIP for 3 different age groups (Young; 4M, Adult; 12M, Old; 20M). (C)Same as in B except the tissue was skeletal muscle. The data in B &C represent the average of 3 independent biological replicates. The error bars represent the SEM and analysis was by t-test (* ; $p \leq 0.05$).

Table 2. The sequence of qPCR oligos for zebrafish

Telomere_F	GGTTTTTGAGGGTGAGGGTGAGGGTGAGGGTGAGGGT
Telomere_R	TCCCGACTATCCCTATCCCTATCCCTATCCCTATCCCTA
Subtelo_F	TGTGTTCGTGGAGTGGATCTTG
Subtelo_R	GCGTGATCGTGGTTATGTTCA
Z_Ebox_R	CCATGCAAGCTATTGTAAAAGCA
Z_Ebox_R	TGTGGTCTCGCCCTGTTTG

Table 3. The sequence of qPCR oligos for mouse

mDbpI2_F	TGGGACGCCTGGGTACAC
mDbpI2_R	GGGAATGTGCAGCACTGGTT
mTelomere_chr13_F	CAAGTGTCCCCAGGAGATGT
mTelomere_chr13_R	CCACCATCACAGGTCACAAG
m_TERT_F	AGCAAAAACCTTCCTCAGCA
m_TERT_R	CCACAGGGAAGTTCACCACT
m_TERC_F	CATTAGCTGTGGGTTCTGGTCT
m_TERC_R	TCCTGCGCTGACGTTTGTTT
m_Trf2_F	TCAGCTGCTTCAAGTACAATGAG
m_Trf2_R	GGTTCTGAGGCTGTCTGCTT
m_Pot1_F	TGGCTTCGCCTCTTTGAC
m_Pot1_R	CCCGCAAAGCTTCTACCA
m_Rap1_F	GGATAGCGGAGAGCCACA
m_Rap1_R	GAGCGGCTTCCTCAAACA
m_Per2_F	CACGCTGGCAACCTTGAAGT
m_Per2_R	TGGTAGTACTCCTCATTAGCCTTCAC
m_Bmal1_F	GATCGAAAAAGCTTCTGCACAA
m_Bmal1_R	GGGTGGCCAGCTTTTCAA
m_eif2a_F	GCTGGGACGCCTAACCTACA
m_eif2a_R	GGATGAACGATTTCAAACATGCT

Table 4. The sequence of oligos for TERRA probe

pTelo250F	ATTCGCCCTTCCCTAACC
pTelo250R	TTCGCCCTTGTTAGGGTT

2.4 MATERIALS AND METHODS

Animal Care and Neurospora strains

Neurospora conidia were suspended in 2% liquid culture medium (LCM) (1x Vogel's salts, 2% glucose, 0.17% arginine) and grown in 100 mm Petri dishes overnight at 30 ℃ to generate mycelia mats. Plugs were cut and used to inoculate flasks containing 100 ml of 2% LCM and grown at 25 ℃ for 2 days. For circadian time course experiments, strains were entrained with a standard light to dark transfer and harvested after a timed incubation in the dark (4, 8, 12, 16, 20, 24, 28, and 32 h). Tissue was crosslinked with 1% formaldehyde, quenched with 100 mM glycine, harvested by filtration, frozen in liquid nitrogen, and then ground with a mortar and pestle in the presence of liquid nitrogen.

All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of Rutgers University. Wild-type zebrafish (SAT) were purchased from ZIRC (Zebrafish International Resource Center, Oregon) and housed according to procedures approved by the Policy on Housing of Vertebrate Animals Outside of Animal Facilities. Fish were fed twice daily and maintained under a 14-hour light: 10-hour dark cycle for breeding, or 12-hour:12-hour light:dark cycle for diurnal entrainment. Adult fish were kept in system water (conductivity 1500 ± 200 μs and pH 7.0-7.5). Embryos and young larvae were maintained in egg water (30 mg/l Instant ocean in deionized water). Fish were sacrificed by emersion in cold MS-222 (300 mg/l, Sigma) and dissected under PBS.

To perform experiments in BMAL1 deficient condition, WT and *Bmal1* $^{-/-}$ mice were used(*180*). Male and female C57BL/6J mice (8–12 wk old) carrying a homozygous

deletion of *Bmal1* were used alongside wild-type controls (B6.129-*Arntl*^{tm1Bra}/J, Jackson Laboratories). Male and female heterozygous *Bmal1* $^{-/-}$ mice were bred. For the genotyping, ear tissues from 8 week-old mice were collected, and genomic DNA was extracted. PCR-based genotyping was performed according to the protocol provided by Jackson laboratory, and male and female WT and homozygous *Bmal1*$^{-/-}$ mice were sorted and used for experiments. For time-course sampling, mice were maintained under a 12-hour:12-hour light-dark cycle and sacrificed at 4-hour intervals.

Antibodies

I used a collection of custom-made and commercial rabbit polyclonal antibodies. BMAL1 antibodies that recognize the mouse isoform were purchased from Abcam (ab3350). I also generated a BMAL1 antibody that recognizes the zebrafish isoform. Briefly, three custom peptides corresponding to the proposed surface-exposed regions of BMAL1 (Peptide #1 CSPGGKKIQNGGTPD, #2 CSSSDTAPRERLIDA, #3 CSTNCYKFKIKDGSF) was used as the immunogen, and resulting antibodies were affinity purified using Sulfolink immobilization kit (Thermo scientific, 44995) following manufacturer's guidelines. Briefly, the 3 peptides were resuspended in 2 mercaptoethylamine-HCl (2-MEA) solution at 37℃ for 1.5 hours. The reduced peptides were mixed with the SulfoLink resin and coupled to the beads by rocking for 15 min and then allowed to settle for 30 min at room temperature. The protein concentration of the flow-through and the non-bound fraction was compared to determine the coupling efficiency. After coupling, the crude sera were loaded into the SulfoLink Column for affinity purification. Antibody bound to resin was washed three times with Tris-buffered

saline and eluted in 0.1M glycine-HCl (pH 2.5). Elutes were neutralized by adding 1M Tris-HCl (pH8.5). Antibodies specific to H3K9me3 were purchased from Abcam (Abcam, ab8898).

Chromatin immunoprecipitation

The ChIP experiments followed the general procedure described previously (*181*) but modified for zebrafish and mouse tissue. Isolated zebrafish tissue was cross-linked with 1% formaldehyde for 10 min at room temperature then quenched with 0.1M Glycine for an additional 10 min. The cross-linked tissues were snap-frozen in liquid nitrogen and stored at -80 ℃. The tissue was homogenized with a micropestle in the presence of 100 μl FA lysis buffer [0.05 M Hepes (pH 7.4), 0.15 M NaCl, 0.001 μM EDTA, 1% Triton X-100, 0.5% Deoxycholic acid] containing protease inhibitors (2.0 μg/ml leupeptin, 2.0 μg/ml pepstatin A, 1.0 mM PMSF). Additional cell disruption and crude chromatin shearing were achieved by sonication at low power (2 × 20 sec at 10% power using a cup sonicator). Lysates were transferred into polystyrene sonication tubes and sonicated again (6 × 20 sec at 20% power). The resulting lysates were cleared of cellular debris by centrifugation at 277 × rpm for 10 min. The sonication regime consistently yielded chromatin sheared to an average size of 500 bp. In a typical ChIP, I used 2.0 mg of sheared chromatin for BMAL1 and 200 μg of chromatin with H3K9me3 with either BMAL1 or H3K9me3 antibody prebound to protein A-conjugated magnetic beads (Dynabeads). The ChIP was washed five times with RIPA buffer and then eluted twice with 0.1 M sodium bicarbonate, 1.0% SDS for 15 min at 42 ℃. The cross-links were reversed by adding 2 μl of 5M NaCl and incubated for a minimum of 4-hours at 65 ℃.

Protein was removed by the addition of 4 μl of proteinase K (10 mg/ml), 4 μl of 1.0M Tris-HCl (pH 6.5), 2 μl of 0.5M EDTA(pH 8.0), and incubated at 42 ℃ for 1-hour. DNA was purified by a phenol/chloroform extraction. The relative levels of BMAL1 or H3K9me3 at *per2* E-box and telomere were determined by qPCR using oligonucleotides contained in Table 2. PCR conditions to detect *per2* E-box were 95℃ for 15 s, 60 ℃ for 1 min and conditions for telomere were 40 cycles of 95℃ for 15 s, 54℃ for 2 min (*182*).

ChIP was performed on mouse liver using isolated nuclei. Tissue was harvested and chopped in small pieces on the ice and cross-linked with 1% formaldehyde for 10 min, then quenched with 0.1M glycine for 10 min at room temperature. Nuclei were prepared from the cross-linked liver tissue as follows. Tissue was homogenized with a micropestle in 1 ml of ice-cold buffer A [250 mM sucrose, 5 mM $MgCl_2$, and 10 mM Tris–HCl (pH 7.4)]. The crude lysate was centrifuged at 600 x g for 10 min at 4 ℃ to pellet the nuclei. The supernatant was discarded, and the nuclei were gently resuspended in 1ml of ice-cold buffer A and centrifuged again at 600 x g for 10 min at 4 ℃. The crude nuclei pellet was resuspended in 1 ml of ice-cold buffer B [2.0 M sucrose, 1 mM $MgCl_2$, and 10 mM Tris– HCl (pH 7.4)], mixed and centrifuged at 16,000 x g at 4 ℃ for 30 min. The tube was inverted and pushed gently against a paper towel, removing most of the upper layer. The nuclei was resuspended in 30 μl FA lysis buffer [0.05 M Hepes (pH 7.4), 0.15 M NaCl, 0.001 μM EDTA, 1% Triton X-100, 0.5% deoxycholate] containing protease inhibitors (2.0 μg/mL leupeptin, 2.0 μg/mL pepstatin A, 1.0 mM PMSF) and transferred to a polystyrene tube and sonicated 8 × 30 sec at 20% power. The resulting lysates were cleared of cellular debris by centrifugation at 13000 rpm. 200 μg of sheared extract was used for BMAL1 ChIP and 100 μg of for H3K9me3 ChIP. For each sample, 3

μl of antibody was prebound to 30 μl of magnetic beads overnight. qPCR was performed using oligonucleotides contained in Table 3. Statistical significance was determined using a P value was calculated by one-way analysis of variance (ANOVA) followed by a Bonferroni test: $*P < 0.05$

Northern and slot blot

Total RNA from zebrafish tissues was isolated by Trizol (Invitrogen) following the manufacturer's protocol. 3-5 μg of total RNA were incubated for 15 min at 65 ℃ in RNA loading buffer (1X MOPS, 56.8% formamide, 20.4% formaldehyde, 11% RNA loading dye [1mM EDTA pH8, 0.23% bromophenol blue, 50% glycerol]) and resolved by electrophoresis on 1.2% agarose gel containing 5% formaldehyde for 3 hours at 70 V in 1X MOPS buffer (2 mM EDTA, 20 mM MOPS 5 mM Sodium acetate). Gels were rinsed two times with distilled water then soaked in 10 X SSC for 30 min then transferred to a hybond N^+ membrane by capillary transfer. RNA was UV-crosslinked to the membrane and hybridized with DIG-labeled TERRA specific probe at 65℃ overnight. Membranes were washed with 2X SSC/0.1% SDS at 42℃ two times and 0.1% SSC/0.1% SDS at 65 ℃ three times. The membrane was then incubated in DIG blocking buffer for 2-hours followed by incubation with anti-digoxigenin Fab fragments for 30 min. The membranes were washed with 1X maleic acid/0.3% tween six times 10 min and visualized by CDP-star (Roche).

For the telomere slot blot, ChIP DNA was prepared in 300 μl of denaturation solution (0.4 M NaOH, 10 mM EDTA) then boiled at 95 ℃ for 10 min, and spotted on the hybond N^+ membrane under a vacuum. Membranes were pre-hybridized in DIG easy

hyb (Roche) for 2-hours and then hybridized overnight with a DIG-labeled oligonucleotide probe specific to the telomere $(TTAGGG)_5$. The signal detection was performed by following Roche guidelines.

In addition to the DIG-labeled oligonucleotide probe, I also subcloned a portion of the telomere repeat to make PCR generated probes. Briefly, pSXneo279 (T2AG3) was obtained from Addgene (plasmid #12403) and used as a template to amplify a fragment containing TTAGGG repeats. The PCR amplified TTAGGG product was cloned into pCR4-TOPO vectors (Invitrogen). After sequencing, one clone contained 35 TTAGGG repeats and was used with DIG Probe Synthesis Kit (Roche Diagnostics, see table 4 for the oligonucleotide sequence).

Analysis of ChIP-Seq and RNA-Seq

Published WC-2 ChIP-seq (*173*) and BMAL1-ChIP seq data (*81, 176*) were downloaded and mapped to the corresponding reference genome, NC10 or mm10. Visualization of binding to the telomere was done using the Integrative Genomics Viewer (IGV) (*183*). To observe TERRA expression, two independent published RNA-Seq data (*81, 179*) from murine liver tissue were downloaded and analyzed by mapping with the mm10 reference genome. The rhythmic TERRA was illustrated in the IGV viewer.

2.5 DISCUSSION

In chapter 2, I found that circadian clock transcription factors are associated with telomere DNA using multitude analyses across a range of species. Re-examination of two independent ChIP-seq experiments (*81, 173*) indicates that BMAL1 interacts with telomeres. Subsequent confirmation in different systems ranging from fungi to animals demonstrates that WC-2 in *Neurospora* and BMAL1 in zebrafish and mice are localized to telomeres suggesting a conserved evolutionary process across multiple species. In addition, I determined the binding of clock proteins to the telomere is diurnally regulated. The binding of BMAL1to the telomeres is required for diurnal TERRA expression and H3K9me3 at the telomere as the rhythms in TERRA and H3K9me3 at telomere are disrupted in the absence of BMAL1. Further experimentation in old animal indicates rhythmic TERRA expression and heterochromatin at telomere are lost with age. Taken together, these results suggest that the circadian clock has an essential role in telomere homeostasis by regulating telomeric lncRNA and heterochromatin at the telomere.

Research indicates there are many age-related phenotypes in clock mutant animal models, suggesting that the circadian oscillator plays an important role in counteracting aging (*132, 184-187*). Previous efforts to uncover the mechanisms has revealed some direct connections between circadian clock and telomerase (*152, 188*). The enzyme TERT and its activity were found to be under circadian control in mice (*152*). It has been reported that reconstitution of TERT in senescent fibroblasts is necessary for circadian entrainment (*188*). However, my analysis indicated there were no statistically significant rhythm changes in TERT under diurnal condition (Supplemental Figure 3F). It is unclear why I was unable to detect a rhythm in TERT but presumaly occurred because I was

unable to obtain age-matched mice same as the previous study. I also examined changes in mRNA expression of the telomerase RNA component, m*TERC*, and shelterin components including *mRap1*, *mTRF2* and *mPot1* over time to test the possibility that circadian clock controls telomere homeostasis by regulating transcriptional controls of *TERC* or shelterin complex. However, as with *TERT* I was unable to detect a rhythm. Therefore, it suggests that the rhythmic regulations at telomere result from rhythmic binding of clock protein BMAL1 to telomere, rather than rhythmic transcriptional control of telomerase.

In this study, it is proposed the interaction of circadian clock on telomere to regulate chromatin structure at telomere. The disappearance of diurnal TERRA regulation and the lower TERRA expression in *Bmal* $^{-/-}$ mice (Fig.7) suggests that BMAL1 is a transcriptional regulator of TERRA and presumably binds via a noncanonical E-box in the subtelomeric region. The telomere repeat (GGGTTA) is different from the canonical E-box sequence (CACGTG), but the bHLH domains of CLOCK and BMAL1 prefer a non-canonical 7 bp E-box sequence (AACGTGA and CATGTGA). When viewed in this contect there is a significant amount of identity between the non-canonical E-box relative to the telomere repeat sequence (AACGTGA vs. AGGGTTA). Given those observations, and my finding that the circadian transcription factors bind in multiple different organisms, it seems highly plausible that BMAL1 binds the telomere repeat.

In addition, I also found a BMAL1-dependent rhythm in TERRA and H3K9me3. The rhythms in H3K9me3 and TERRA suggest telomere homeostasis is rhtyhmic and controlled by a complex interplay between facultative heterochromatin, lncRNA, and the circadian clock although the exact mechanism will require futher study. An interesting

facet of TERRA maturation is a proposed mechanism whereby it is cleaved into Dicer-independent telomere-specific small RNAs (tel-sRNAs). This tel-sRNAs share the same strand specificity as TERRA and tel-sRNAs are regulated by epigenetic mechanisms such as KMT12/MLL1 and KMT1/Suv39h1/2, suggesting the tel-sRNAs are subject to epigenetic regulation (*189*). Because KMT1/Suv39h is a component of the PER complex (*38, 39*), it is interesting to speculate that circadian regulation of TERRA may serve to mediate circadian heterochromatin at the telomere via rhythms in tel-sRNAs. Support for this idea can be found in *Neurospora* where it is clear that disiRNA co-localize to regions around the telomere, adjacent to WC-2 binding. However, efforts to characterize TERRA in *Neurospora* proved challenging, and I was unable to detect it by Northern suggesting it is a highly unstable transcript that is rapidly converted to disiRNAs. Based on this and the known potential for lncRNAs to directly regulate hetertochromatin as in X-chromosome inactivation, and RNAi-mediated heterochromatin; is seems reasonable that TERRA may be directly involved in mediating rhythms in telomere heterochromatin (*63, 97, 124*).

Circadian regulation of telomeres, including rhythm in TERRA and heterochromatin formation, has important clinical implications as shortened telomeres correlate with many human diseases. For example, mutations in the WRN gene (Werner syndrome), BLM gene (Bloom syndrome,) FANCB gene (Fanconi anemia) and DNA repair genes such as Ataxia telangiectasia (ATM) cause accelerated telomere shortening and premature aging. Moreover, shorter telomere length is associated with many types of cancers, and constitutive telomerase activity is one of the transformative events in cancer cells to overcome replicative senescence. Therefore, protection of the telomere from environmental stress and the maintenance of telomere structure could be important for

healthy aging. The data presented in this study suggest that alterations to the normal diurnal regulation of telomere homeostasis likely lead to increased senescence when normal circadian rhythms are disrupted and may impact genome integrity. Whether this mechanism is the driving the connect between circadian dysregulation and cancer, as is seen with the higher incidence of cancer in shift-workers, remains an open question.

In conclusion, it is clear that BMAL1 is localized to the telomere and there is a BMAL1-dependent diurnal rhythm in TERRA and heterochromatin at the telomere. These findings and accumulating evidence linking the circadian clock and accelerated aging provide a more comprehensive understanding of possible mechanisms and may lead to therapeutic interventions to enhance circadian regulated telomere homeostasis in aging and aging-related diseases.

CHAPTER 3

Age-related changes in diurnal non-coding RNA coincides with changes in genome-wide facultative heterochromatin

(Published in BMC Genomics)

3.1 ABSTRACT

Background: Disrupted diurnal rhythms cause accelerated aging and an increased incidence of age-related disease and morbidity. The circadian clock governs cell physiology and metabolism by controlling transcription and chromatin. The goal of this study is to further understand the mechanism of age-related changes to circadian chromatin with a focus on facultative heterochromatin and diurnal non-coding RNAs.

Results: I performed a combined RNA-Seq and ChIP-seq at two diurnal time-points for three different age groups to examine the connection between age-related changes to circadian transcription and heterochromatin in neuronal tissue. My analysis focused on uncovering the relationships between long non-coding RNA (lncRNA) and age-related changes to histone H3 lysine 9 tri-methylation (H3K9me3), in part because the Period (Per) complex can direct facultative heterochromatin and models of aging suggest age-related changes to heterochromatin and DNA methylation. My results reveal that lncRNAs and circadian output change dramatically with age, but the core clock genes remain rhythmic. Age-related changes in clock-controlled gene (ccg) expression indicate there is age-dependent circadian output that changes from anabolic to catabolic processes during aging. Also, there are diurnal and age-related changes in H3K9me3 that coincide with changes in transcription.

Conclusions: The data suggest a model where some age-related changes in diurnal expression are partially attributed to age-related alterations to rhythmic facultative heterochromatin. The changes in heterochromatin are potentially mediated by changes in diurnal lncRNA creating an interlocked circadian-chromatin regulatory network that undergoes age-dependent metamorphosis.

3.2 INTRODUCTION

Early circadian research showed non-circadian day lengths negatively impact longevity (*190*), even though the free-running rhythm only changes slightly with age (*191*). More recent reports indicate circadian disruption contributes to increased morbidity (*192*) and diseases including; metabolic disorders (*193-195*), depressive disorders (*196*), cardiovascular disease (*197*), cancer (*198*), advanced aging (*132*). In addition, age-related changes to the circadian clock result in a dampened endocrine and neuroendocrine hormone rhythm (*199*) concurrent with a decrease in Clock and Bmal1 expression, while the rhythm in Per genes is largely unaffected (*200*). There are also extensive age-related changes to clock-controlled genes (ccg) in the liver that correlate with age-related changes to metabolism which occurs with a change in acetylation (*137, 138*). As a result of these reports and others, there is a growing appreciation of the connection between circadian disruption, advanced aging and an increased incidence of age-related maladies, yet the mechanism underlying these remains cryptic.

The core circadian rhythm is maintained by a transcriptional negative feedback loop where the dimeric transcriptional activators CLOCK and BMAL1 (*201-204*) drive expression of the Period (Per1, Per2, Per3) (*205-207*) and Cryptochrome (Cry1, Cry2) genes (*208, 209*). Once expressed, the PERs associate with the CRYs and form a multisubunit complex (approximately 2 MDa) that contains repressive chromatin modifiers including a histone H3 lysine 9 methyltransferase (KMT1/Suv39h), RNA binding proteins, and the nucleosome remodeling and deacetylase (NuRD) corepressor (*38-40, 94*). In a second feedback loop, RORα and REV-ERBα have a positive and negative role respectively in regulating rhythms in Bmal1 expression (*41, 42*). The clock

regulates, either directly or indirectly, a large percentage of protein-coding genes (*78, 81, 101, 210-213*) and controls genome-wide chromatin states that oscillate between permissive and non-permissive states (*78, 81, 214*). Facultative heterochromatin at the central clock gene(s) in Neurospora, Drosophila and mammals is observed during the repressive phase and is characterized by histone H3 lysine 9 di- and tri-methylation (H3K9me2 & H3K9me3) and HP1 binding (*39, 84, 97, 178, 215, 216*). The extent of circadian-regulated facultative heterochromatin (CRFH) throughout the genome is unknown, but it does occur at the D-element binding protein (*Dbp*) suggesting, at the very least, it is present at a ccg (*178*).

Long non-coding RNAs (lncRNAs) and natural antisense transcripts (NATs) are known to aid in establishing repressive chromatin states and can function in both cis and trans (*217, 218*). As such, they are a potential mechanism by which the circadian clock could establish permissive or non-permissive chromatin states to regulate gene expression. Support for this idea comes from numerous reports documenting a wide array of rhythmic ncRNA (*78*) including a NAT that originates from Per2 (*78, 81*). Based on studies on the Neurospora frequency NAT, it is possible that Per2AS may be necessary to help heterochromatin and assist in feedback repression (*97*).

Current theories on aging suggest that changes in heterochromatin may be responsible for age-related changes in gene expression and alterations in H3K9me3 may function as a driver of aging (*219, 220*). For example, in models of Werner syndrome, a premature aging disorder in humans, there are genome-wide changes to H3K9me3. The changes in H3K9me3 in Werner syndrome are likely due to WRN-guided heterochromatin because of WRN associates with KMT1/SUV39H and HP1 (*220, 221*).

The connections among the circadian clock, heterochromatin, lncRNA, and aging led us to examine the role of lncRNA in age-related changes to the clock-chromatin circuit and ascertain whether this may be contributing to age-related diseases. To do so, I performed a comprehensive RNA-sequencing (RNA-seq) and H3K9me3 ChIP-seq on zebrafish brain tissue from animals that were 4, 12, and 20 months (M) old at zeitgeber time (ZT) 4 and ZT16. I observed age-related changes to diurnal H3K9me3 and RNA expression, with a subset of diurnal genes showing changes in H3K9me3. The results suggest that age-related changes to circadian regulated transcription occur in part due to changes in heterochromatin.

3.3 RESULTS

Identification of long non-coding RNAs

There are age-related changes to circadian transcription that coincide with changes in physiology, but the connection to changes in circadian noncoding RNA and chromatin are lacking. To further define age-related changes to circadian transcription, including changes in lncRNA expression, I performed a multi-dimensional RNA-sequencing (RNA-seq) experiment with a focus on lncRNA transcript discovery. I isolated total RNA from zebrafish brain tissue at two diurnal time points (ZT4 and ZT16) for 3 different ages (4M, 12M, 20M). I processed the ribo-depleted, stranded RNA-seq data using Tophat2/Cufflinks (T/C) (*222*) and/or HISAT2/StringTie (H/S) (*223*) with the zebrafish genome GRCz10. In My analyses, I also included stranded RNA-seq data from developing zebrafish embryos (*224*) to enhance lncRNA identification. In order to detect the known per2AS transcript with H/S pipeline, I had to use an FPKM cut-off of 0.1 in StringTie. Stipulating per2AS identification as criteria in StringTie generated an excessive number of single-exon transcripts (>150,000), so I proceeded with the analysis using T/C results (Supplementary Fig. 8and 9). The Cufflinks transcripts were further subdivided by class code and processed with bedtools intersect to identify potential overlapping lncRNA. This analysis revealed there were potentially 14,773 overlapping transcripts (OT). Next, I processed the OT with PLEK (predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme) to predict potential NATs (*225*) with the notion these may play a role in chromatin regulation. Overall, this analysis revealed 5,299 lncRNAs overlap with another gene, either protein-coding gene or overlapping lncRNAs. The inclusion of RNA-seq data from developing zebrafish

embryos allowed us to perform a comprehensive analysis of lncRNA from early development through aged adult animals. To further parse the data, I took the 146,910 transcripts from Cufflinks and found that 87,327 transcripts are not contained in the zebrafish GRCz10 genome assembly. These 87,327 transcripts were processed with PLEK to find lncRNA (*225*) yielding 50,524 potential lncRNAs. Next, I filtered the 50,524 potential lncRNAs based on expression stipulating that a transcript must change in at least one sample set to separate regulated transcription versus nonspecific transcriptional noise. In total, 17,702 potential lncRNAs displayed differential expression from embryogenesis to aged adult.

Further clustering of 17,702 transcripts into 8 groups based on expression revealed that expressions of some lncRNA are restricted to specific times during embryo development, while others are restricted to adult fish (Supplementary Fig. 10). In general, I found specific lncRNAs are largely confined to brain tissue, and they were either not present during development or were low abundance. One group (Group 6, Supplementary Fig. 10G) displayed a noticeable diurnal oscillation at 12M where transcripts were highly expressed at ZT4, but largely absent at ZT16. These same lncRNAs appeared constant regardless of the time-of-day at other ages.

Age-related changes to diurnal gene expression

Next, I proceeded to examine diurnal expression across all three age groups. Transcripts were split into protein-coding genes and transcripts not currently annotated in the GRCz10 reference assembly (unannotated transcripts). Among the 87,327 unannotated transcripts, PLEK predicted roughly 58% were lncRNA (Supplementary Fig. 10). Analysis of diurnal expression comparing ZT4 to ZT16 ($q \leq 0.05$) revealed 273, 604, and 193 diurnal transcripts in 4, 12, and 20 M age groups, respectively (Fig. 10A). Unannotated transcripts represented 48.7% and 39.4% in 4 and 20 M groups, while they represented 83.4% of diurnally expressed transcripts in the 12 M group. Annotated diurnally expressed genes for each age group can be found in online. At each age, a portion of the diurnally expressed transcripts overlapped with another transcript on the opposite DNA strand. In total, 9.5% (26 of 273), 6.6% (40 of 604), and 12.4% (24 of 193) of diurnally expressed transcripts at 4, 12, and 20 M contained an overlapping transcript.

To examine age-related changes in diurnal gene expression, diurnally expressed genes from the three different ages were split into two groups. Those that peaked at ZT4 were classified as morning-specific, and those that peaked at ZT16 were classified as evening-specific. Comparison across ages revealed the vast majority of diurnal genes change with age and only a small subset (32 morning-specific and 7 evening-specific) of genes maintained diurnal expression regardless of age (Fig. 10B,C). The morning-specific genes included per2 and cry1a, while evening-specific genes included bmal1/arntl1a and clock (Fig. 11C). I further examined the age-related changes to diurnal gene expression by clustering transcripts elevated at ZT4 relative to ZT16 for each age (Supplementary Fig. 11). In general, morning and evening specific genes changed

expression as the animals aged.

To further define the physiological consequence of age-related changes in diurnal gene expression, I calculated gene ontology (GO) enrichment for diurnally expressed genes as a function of age using WEB-based GEne SeT AnaLysis (WebGestalt) Toolkit (*226*, *227*) (Fig. 10D). This analysis showed diurnally expressed genes differed in their ontology as a function of age. For example, in 4 M-old fish diurnally expressed genes were mainly associated with nucleosome/chromatin assembly, protein-DNA complex assembly and response to stimuli such as light, abiotic factors, and stress. In 12 M fish, the top enriched GO terms included circadian rhythm along with biosynthetic and metabolic processes. In 20 M fish, the most significant GO terms included glucose metabolism and carbohydrate or monosaccharide catabolism. The results indicate that age-related changes in diurnal genes are chronologically clustered into different biological processes that include cellular growth and anabolic processes in young animals, changing to genes involved in cellular maintenance in adults, and switching to catabolism in older animals.

Next, I performed disease enrichment on diurnal-regulated genes at each age to compare age-related changes with known disease markers. Disease Ontology (DO) analysis was conducted using DOSE (*228*) and revealed a correlation between gene expression and age-related diseases (Supplementary Fig. 12). In addition to mood and bipolar disorder that appeared to manifest in older animals, genes implicated in diseases originating in other tissues were also present, suggesting that separate alterations in clock output in peripheral tissue may be a causative event in disease manifestation. Furthermore, I used the Age-Related Disease (ARD) database (*229*) and found an

association between gene expression changes and age-related cognitive decline. For example, I found that *c3a.1* and *anxa3b*, genes implicated in neurodegenerative diseases such as Alzheimer's disease and Parkinson's (*230*), lost their diurnal rhythm at 20M .

Circadian Clock Remains Constant Regardless of Age

I next focused on the 39 genes (32 morning-specific and 7 evening-specific) that maintained diurnal expression regardless of age (Fig. 11A). The GO analysis revealed the vast majority of genes are involved in circadian clock regulation (Fig. 11B). This subset of genes includes *arntl2*, *clocka*, *cry1a*, *cry2*, *per1b*, and *per2* as well as *hsf2*, *bhlhe40*, *fech*, *nr1d1* and *aldh4a1*, which all show distinctive diurnal expression (Fig. 11C). Consistent with other reports (*137*, *138*), these data indicate that core clock genes are largely unaffected by age, suggesting there is an undefined age-related network that alters circadian output beyond the normal plasticity found in circadian oscillators.

Age-specific Diurnal Genes with Overlapping Transcripts

A fraction of diurnally expressed genes had a corresponding overlapping transcript so I examined how the expression of these pairs changed with age with the idea that they may engage in transcriptional interference of representing chromatin regulating lncRNA and its cognate *cis* target. This concept is rooted in the observation that lncRNAs are known regulators of chromatin and can influence gene expression in both *cis* and *trans* (*218*). In support of this notion, I found that 9.5, 6.6 and 12.4% of diurnally expressed genes for the 3 age groups had an OT. Expression of representative diurnal genes with corresponding OT is in shown in Fig. 12A-D. As previously noted per2, which maintained a diurnal rhythm regardless of age, has a corresponding antisense transcript (per2AS), although in the RNA-seq *per2AS* lacks a readily apparent rhythm for the 2 diurnal time points tested here (Fig. 12A). However, when I examine a 24-hr time course using strand-specific RT-PCR, an antiphasic rhythm is apparent (data not are shown). Diurnal expression of other genes with OT was restricted to specific ages. For example, myl10 has a diurnal expression at 4 M but not 12 and 20 M. Like *per2AS, myl10AS* does not have a detectable rhythm at any age (Fig. 12B, Supplementary Fig.13A).

In contrast, *cishb* and *cishbAS* both have a diurnal expression with a similar phase (both transcripts are elevated at ZT4 and reduced at Z16) (Fig. 12C, Supplementary Fig.13B). Finally, *myl1* exhibited diurnal expression at 20 M, but not 4M and 12 M and its OT, *myl1AS* did not show a detectable rhythm at any age examined (Fig. 12D, Supplementary Fig. 13C). RT-PCR results for these OT pairs are consistent with the RNA-seq data, confirming OT are present at a subset of diurnal genes but do not

necessarily display antiphasic rhythms; a finding consistent with results from liver (*78*). However, my data is limited to 2 diurnal time-points, so I cannot rule out the possibility that these OTs have an oscillation with a different diurnal phase. Also, the antisense transcripts were all expressed at much lower levels, so detecting differences were inherently more difficult.

Diurnal and Age-related Changes to Facultative Heterochromatin

Given my observation of many diurnally expressed lncRNAs and their potential role as regulators of chromatin, I proceeded to examine if there were age-related changes in facultative heterochromatin, specifically H3K9me3, which might account for the age-related changes in gene expression. I performed H3K9me3 chromatin immunoprecipitation (ChIP) from crosslinked zebrafish brain tissue at ZT4 and ZT16 at 4, 12 and 20 months of age. Chromatin enriched with H3K9me3 was identified using MACS2 ($q \leq 0.01$) (*231*) and differential diurnal distribution within and between ages defined by DiffBind (*232*). Diurnal heterochromatin between ZT4 and ZT16 was observed at all ages ($p \leq 0.05$) and H3K9me3 levels changed throughout the genome concurrent with a change in age. For example, there were 3,424 rhythmic H3K9me3 peaks at 4 M, 3,417 at 12 M and 9,717 at 20 M (Fig. 13A a-c). To understand the distribution of H3K9me3 between coding and non-coding loci and the extent to which H3K9me3 oscillated at OT, the H3K9me3 peaks were broken down into those occurring in or near protein-coding genes (Fig. 13A d-f) or at loci containing predicted lncRNA (Fig. 13A g-i). The peaks were further classified into ones that occurred near OTs, and these are highlighted in green. At 4M, 14.2% of the diurnal H3K9me3 peaks occurred at annotated genes with an OT, while 8.3% occurred at a lncRNA with an OT; at 12M 22% occurred at annotated genes with an OT, and 13.6% occurred at a lncRNA with an OT; and at 20 M, 11.2% of peaks occurred at annotated genes with an OT, and 22.2% occurred at a lncRNA with an OT. I next examined the differences in the genome location of H3K9me3 that occurred in protein-coding genes relative to regions containing lncRNA (Fig. 13B). At protein-coding genes, H3K9me3 tended to cluster in promoter regions 1-3

kb upstream of transcription start site, in introns and intergenic regions, whereas at lncRNAs, H3K9me3 mostly occurs in intergenic regions. This phenomenon was likely observed because lncRNAs lack a 5' or 3' UTR and tend to be smaller with less splicing than their protein-coding counterparts. Collectively, the data indicate there are genome-wide diurnal oscillations in H3K9me3 that is changed in an age-dependent manner.

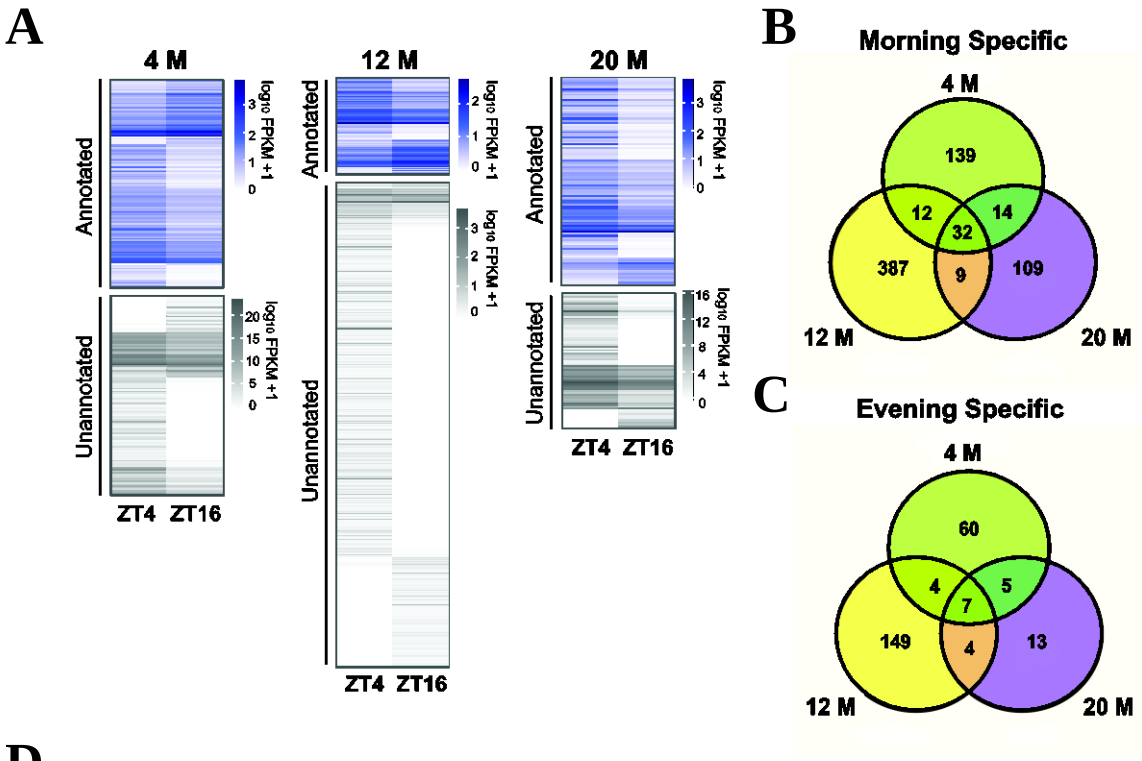
To further define age-related changes in facultative heterochromatin and examine the impact on age-related changes in gene expression, I began by selecting genes that displayed differential expression between 4 M and 20 M at the same time point ($q \leq 0.05$). I found 266 transcripts had an age-related change at ZT4 and 462 at ZT16 (Fig. 14A and B). At ZT4, 76 annotated and 190 unannotated transcripts changed expression between 4 and 20 M and at ZT16, 167 annotated and 295 unannotated transcripts displayed age-related changes (online). Next, I sorted the genes with age-related changes into four groups: Group 1 was morning-specific genes that decrease with age, Group 2 was morning-specific genes that increase with age, Group 3 was evening-specific genes that decrease with age and Group 4 was evening-specific genes that increase with age. Consistent with previous findings (*138*), GO analysis revealed age-related changes to metabolic function were prevalent (Table 5). Other notable findings in My analysis indicate that cardiovascular fitness and potential regenerative decline with age (Group 1) whereas responses to hypoxia and low oxygen are endemic to aged animals (Group 4).

I next examined age-related changes to H3K9me3. Analysis of H3K9me3 between 4 and 20 M suggested H3K9me3 levels change for large portions of the genome as organisms age (Fig. 14C). This analysis shows an overall decrease in significant H3K9me3 peaks in 4 M versus 20 M at ZT4 and the opposite at ZT16. Changes occurred

at protein-coding genes and lncRNA and were particularly pronounced for lncRNAs, some of which contained sense-antisense pairs. When compared with gene expression, H3K9me3 enrichment generally showed the expected and opposite pattern. For example, H3K9me3 was elevated near *slc4a1a* at 4 M relative to 20 M, and *slc4a1a* showed a corresponding decrease in expression at 20 M compared to 4 M (Fig. 14D, additional fig.12 a-c). In this case, loss of H3K9me3 at ZT4 near slc4a1a with increasing age resulted in a gain of morning-specific diurnal expression that was exclusive to 20 M brain.

Conversely, I also observed the opposite and found an age-related loss in diurnal expression and a corresponding increase in heterochromatin. This phenomenon is represented by fbxo32 (Fig. 14E), which displayed a diurnal oscillation at 4M in the RNA-seq that was lost at 12 and 20 M. In this instance, H3K9me3 levels increased with age at ZT16, which corresponded with the loss of expression. Despite the age-related changes in diurnal expression of the protein-coding genes and heterochromatin, the corresponding overlapping lncRNA did not always show an overt rhythm or expression differences among ages, indicating there is far more complexity to the system than just changes to lncRNA expression.

Of note, the size of H3K9me3 peaks around slc4a1a and fbxo32 were approximately the size of a single nucleosome (150-200 bp), as were many of the peaks in and around protein-coding genes. There were also larger heterochromatin domains containing broader peaks, and these typically occurred around unannotated transcripts (Supplementary Fig. 15). In certain instances, the broad peaks had diurnal changes, but there were no apparent oscillations in the corresponding transcript.


A
4 M
12 M
20 M
Annotated
Unannotated
log10 FPKM +1
ZT4 ZT16
B
Morning Specific
4 M
139
12
14
32
387
9
109
12 M
20 M
C
Evening Specific
4 M
60
4
5
7
149
4
13
12 M
20 M

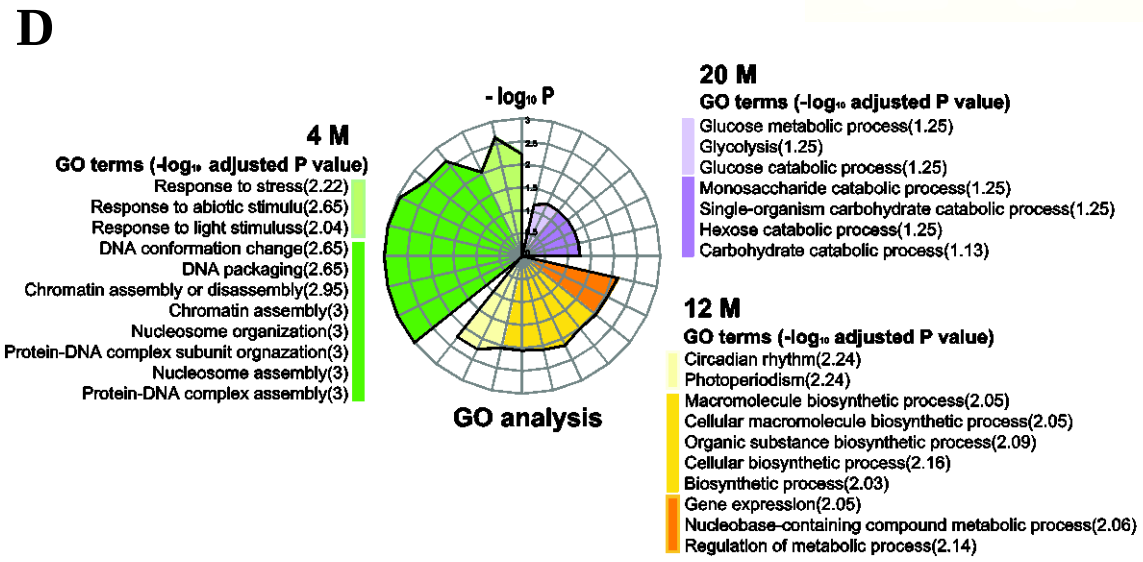

D
4 M
GO terms (-log10 adjusted P value)
Response to stress(2.22)
Response to abiotic stimulu(2.65)
Response to light stimuluss(2.04)
DNA conformation change(2.65)
DNA packaging(2.65)
Chromatin assembly or disassembly(2.95)
Chromatin assembly(3)
Nucleosome organization(3)
Protein-DNA complex subunit orgnazation(3)
Nucleosome assembly(3)
Protein-DNA complex assembly(3)
- log10 P
GO analysis
20 M
GO terms (-log10 adjusted P value)
Glucose metabolic process(1.25)
Glycolysis(1.25)
Glucose catabolic process(1.25)
Monosaccharide catabolic process(1.25)
Single-organism carbohydrate catabolic process(1.25)
Hexose catabolic process(1.25)
Carbohydrate catabolic process(1.13)
12 M
GO terms (-log10 adjusted P value)
Circadian rhythm(2.24)
Photoperiodism(2.24)
Macromolecule biosynthetic process(2.05)
Cellular macromolecule biosynthetic process(2.05)
Organic substance biosynthetic process(2.09)
Cellular biosynthetic process(2.16)
Biosynthetic process(2.03)
Gene expression(2.05)
Nucleobase-containing compound metabolic process(2.06)
Regulation of metabolic process(2.14)

Figure 10. Age-related changes in diurnal transcription of zebrafish brain tissue
(A) Heatmaps illustrate diurnally regulated genes from zebrafish brain tissue at 4 M, 12 M and 20 M (FDR ≤0.05). Transcripts were sorted into annotated (blue color scale) and unannotated genes (gray color scale). The list of genes and locus information is contained in supplementary tables 1a-1c. The overlap of age-related changes in diurnal gene expression was determined for (B) morning-specific genes (elevated at ZT4) and (C) evening-specific genes (elevated at ZT16). (D) Radial graph depicting the GO analysis from age-specific diurnal genes (P ≤0.01). Functionally related GO terms are clustered based on age (4 M; green, 12 M; yellow and 20 M; purple).

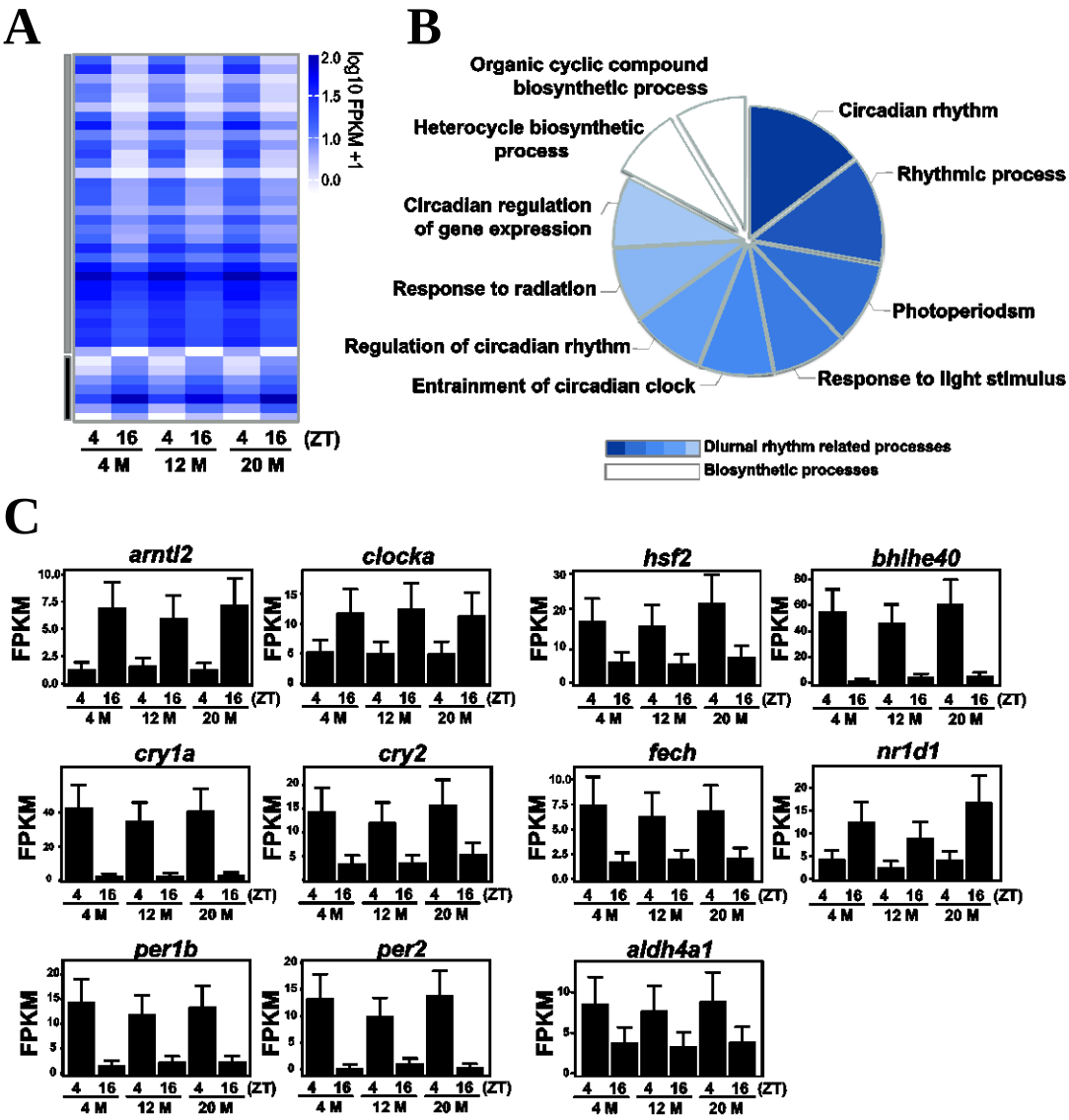


Figure 11. The core clock remains rhythmic regardless of age

(A) Expression of the 39 genes that maintained oscillations across all age groups. (B) Functional categorization of the GO terms for the 39 genes whose diurnal expression did not change with age. Genes related to diurnal rhythm-related function (shaded blue) or other biosynthetic processes (shaded white). (C) Relative expression at both times across all age groups for a subset of genes that did not change with age (FDR < 0.05).

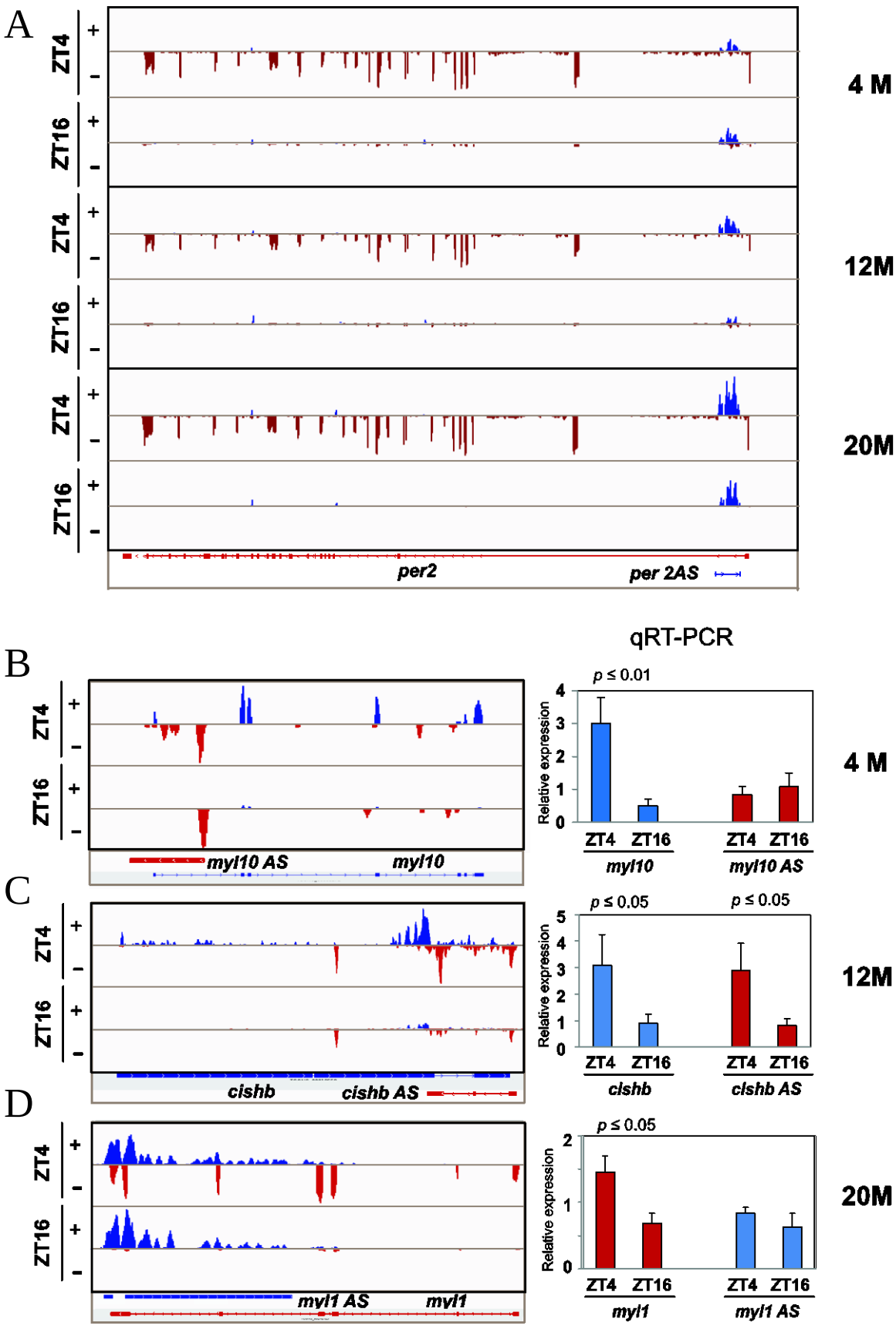

A
ZT4
ZT16
ZT4
ZT16
ZT4
ZT16
4 M
12M
20M
per2
per 2AS
B
qRT-PCR
ZT4
ZT16
myl10 AS
myl10
p ≤ 0.01
Relative expression
ZT4 ZT16 ZT4 ZT16
myl10
myl10 AS
4 M
C
cishb
cishb AS
p ≤ 0.05
p ≤ 0.05
cishb
cishb AS
12M
D
myl1 AS
myl1
p ≤ 0.05
myl1
myl1 AS
20M

Figure 12. Representative examples of overlapping transcripts

IGV traces of the (A) per2 locus showing relative expression of transcripts that originate from the plus strand (blue) and the minus strand (red) across all three age-groups. Representative IGV traces and validation of sense and antisense transcripts by quantitative RT-PCR of diurnal transcripts with corresponding antisense transcripts for each representative age-group (B) myl10 at 4 M, (C) cishb at 12M and (D) myl1 at 20M. Error bars present SEM and the p-values are shown on the graphs.

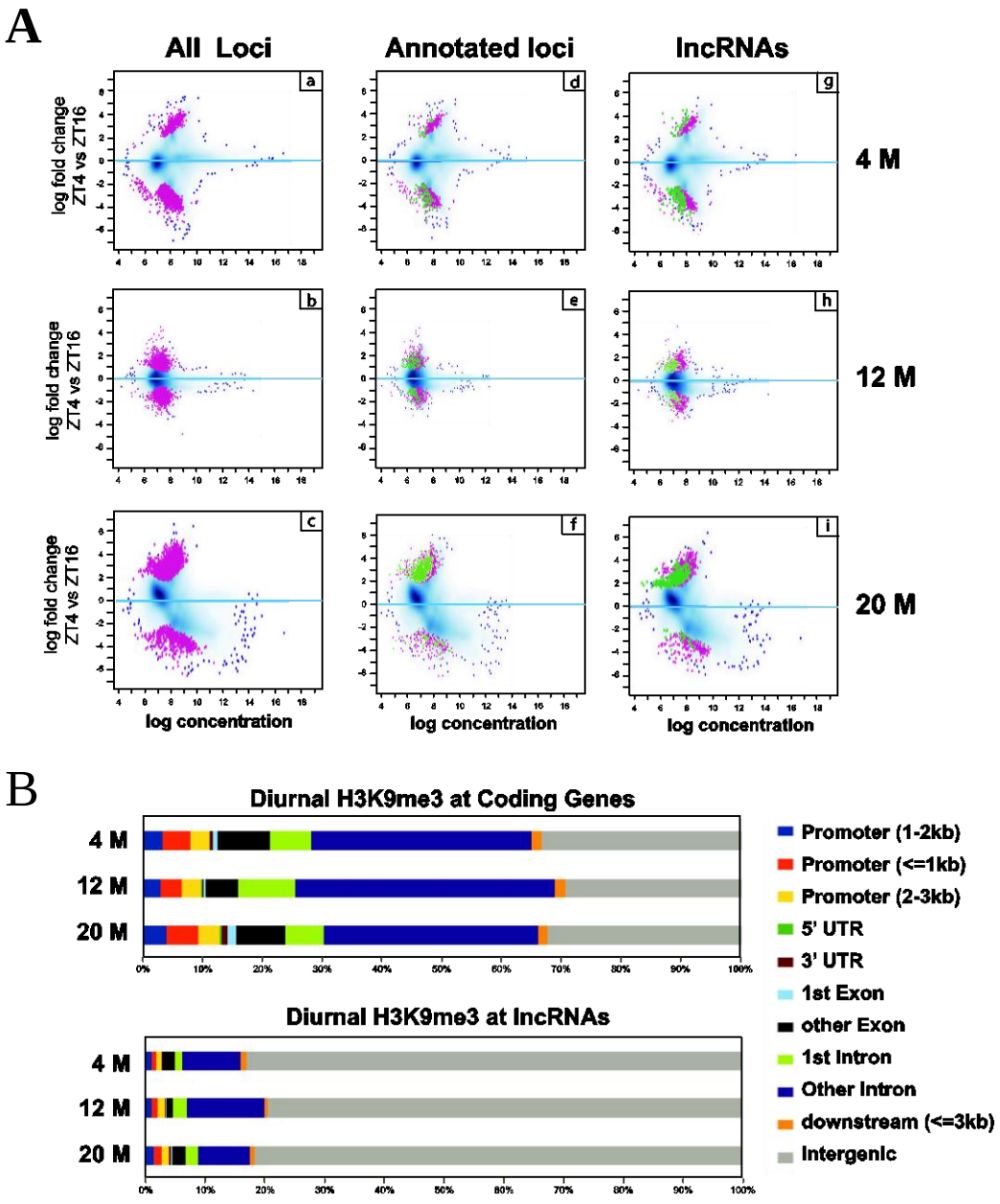


Figure 13. Diurnal rhythms in genome-wide H3K9me3 for each age group

(A) Diurnal changes in H3K9me3 levels between ZT4 and ZT16 for all 3 age groups. Loci with significant diurnal changes in H3K9me3 are shown in red. Positive values indicate H3K9me3 peaking at ZT4 and negative values peak at ZT16. The plots are further subdivided into annotated loci or predicted lncRNAs. The green spots indicate loci with OT. (B) Relative genome feature location of diurnal H3K9me3 for coding genes and lncRNAs.

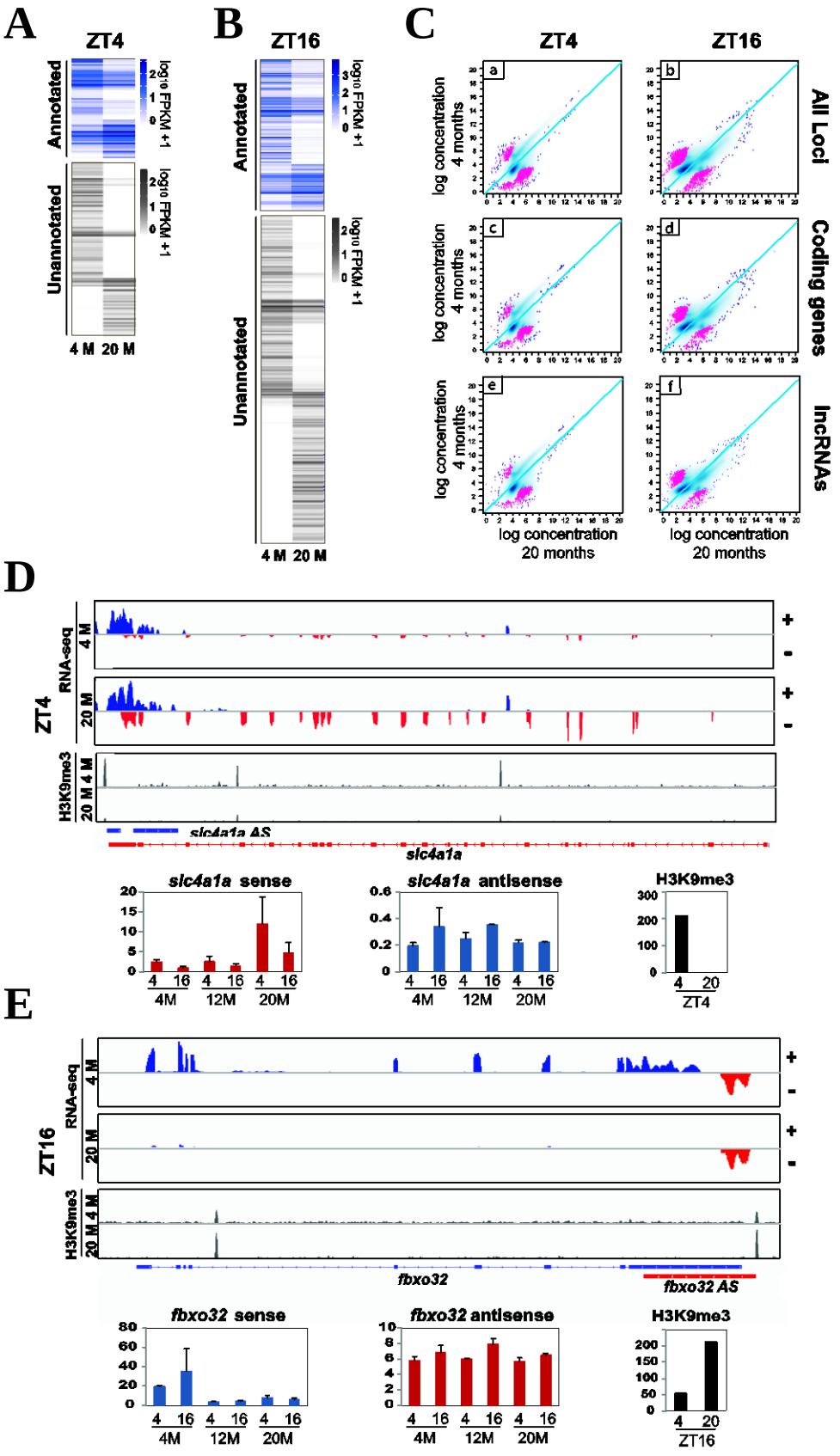

A
ZT4
Annotated
Unannotated
log10 FPKM +1
4 M 20 M
B
ZT16
Annotated
Unannotated
log10 FPKM +1
4 M 20 M
C
ZT4
ZT16
All Loci
Coding genes
lncRNAs
log concentration 4 months
log concentration 20 months
D
ZT4
RNA-seq
4 M
20 M
H3K9me3
20 M 4 M
slc4a1a AS
slc4a1a
slc4a1a sense
slc4a1a antisense
H3K9me3
4M 12M 20M
ZT4
E
ZT16
RNA-seq
4 M
20 M
H3K9me3
20 M 4 M
fbxo32
fbxo32 AS
fbxo32 sense
fbxo32 antisense
H3K9me3
4M 12M 20M
ZT16

Figure 14. Age-related changes in transcript abundance coincide with changes in heterochromatin.

(A) Genes that display age-related changes in expression between 4 M and 20 M at ZT4 (FDR ≤0.02). The genes are further subdivided into 50 annotated (blue scale) and 136 unannotated genes (Grayscale). (B) Same as in (A) except these loci have age-related changes at ZT16. At ZT16 there were 178 annotated genes (top) and 296 novel genes (bottom) that changed between 4M and 20M. Genes represented in the heatmaps are contained in supplementary tables 2a and 2b. (C) Scatter plots showing age-related changes in H3K9me3 between 4M and 20M at ZT4 and ZT16 for all loci described in (a) and (b). Loci with significant diurnal changes in H3K9me3 are shown in red. The data are further subdivided into protein-coding genes and lncRNAs. Representative examples of age-related changes in expression and H3K9me3 at genes with OT are visualized with IGV. (D) At ZT4, slc4a4 has an age-related increase in expression and an age-related decrease in H3K9me3. (E) At ZT16 there is an age-related loss in fbxo32 and a corresponding increase in H3K9me3.

Table 5. GO analysis of genes changed with age

The most enriched GO clusters are selected from 4 groups of genes and ranked by their -log10 p-value ($p \leq 0.01$).

Morning-Specific

Group 1: Decrease with Age		Group 2: Increase with Age	
GO annotation	adj P value	GO annotation	adj P value
heart contraction	3.70E-03	short-chain fatty acid metabolic process	6.24E-05
heart process	3.70E-03	regulation of biological quality	2.00E-04
fin regeneration	4.70E-03	homeostatic process	6.00E-04
circulatory system process	5.50E-03	fatty acid metabolic process	1.40E-03
blood circulation	5.50E-03	ion homeostasis	1.70E-03
tissue regeneration	7.80E-03	chemical homeostasis	2.20E-03
fin morphogenesis	1.48E-02	monocarboxylic acid metabolic process	3.10E-03
purine-containing compound biosynthetic process	1.88E-02	cellular lipid metabolic process	1.14E-02
response to wounding	1.93E-02	carboxylic acid metabolic process	2.08E-02
appendage development	1.93E-02	oxoacid metabolic process	2.22E-02

Evening-Specific

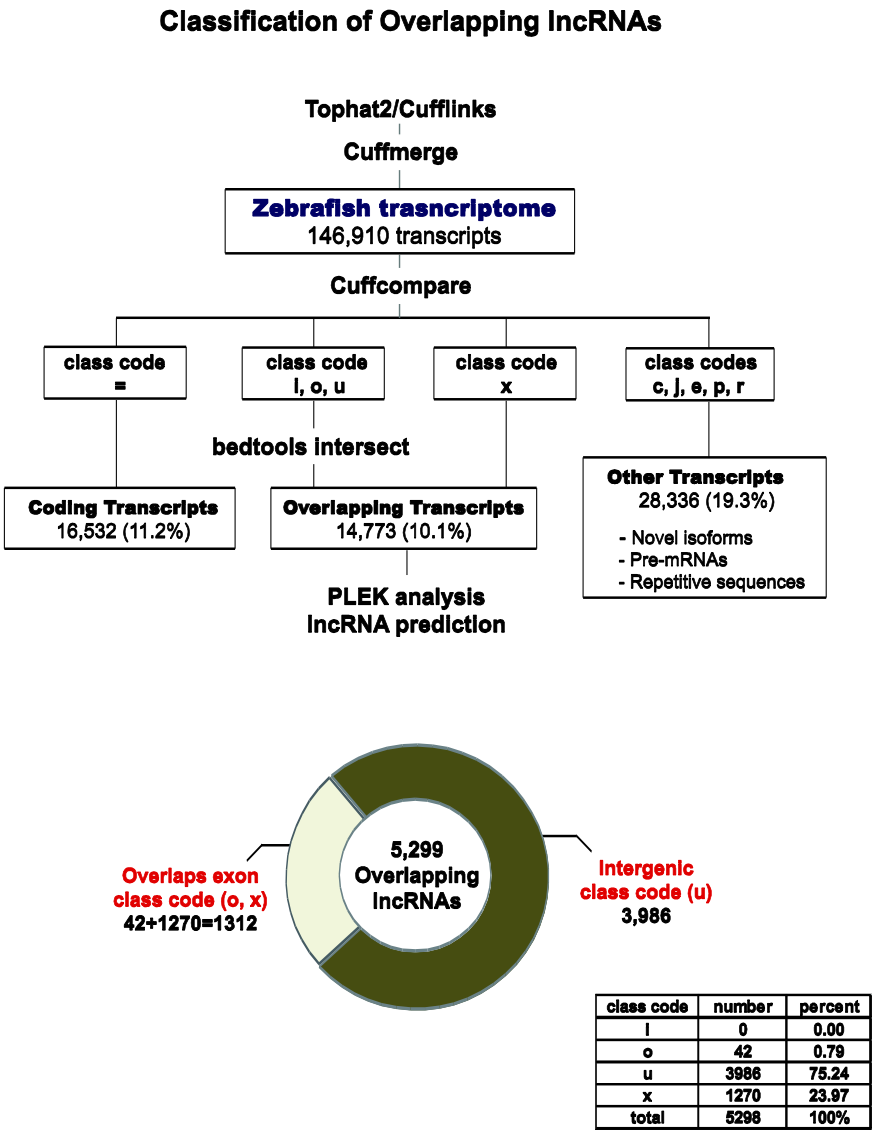
Group 3: Decrease with Age		Group 4: Increase with Age	
GO annotation	adj P value	GO annotation	adj P value
glycolysis	1.92E-06	endothelial cell migration	2.00E-04
monosaccharide metabolic process	3.32E-06	immune system process	1.00E-03
hexose metabolic process	3.32E-06	response to decreased oxygen levels	1.10E-03
glucose catabolic process	4.92E-06	response to hypoxia	1.10E-03
monosaccharide catabolic process	5.66E-06	response to oxygen levels	1.10E-03
single-organism carbohydrate catabolic process	5.66E-06	sprouting angiogenesis	1.50E-03
hexose catabolic process	5.66E-06	erythrocyte homeostasis	2.70E-03
carbohydrate catabolic process	6.48E-06	erythrocyte differentiation	2.70E-03
glucose metabolic process	1.52E-05	embryonic hemopoiesis	2.90E-03
generation of precursor metabolites and energy	2.00E-04	homeostasis of number of cells	3.10E-03

Table 6. The sequence of qPCR oligos in chapter 3

myl_10_AS_F	GTAGGCCCAGTAAAGCAGC
myl_10_AS_R	ATAGCCCATATTTCTAAGGTAATC
myl_10_S_F	CCCTTGGATGTGGCTGGTAA
myl_10_S_R	TCTCCGTGGGTGATGACGTA
cishb_AS_F	GGACTTGTATTCGCGTTGATTTT
cishb_AS_R	CAGCAGAATGTGAGGTGACAGTTA
cishb_S_F	AGCGCATTAAAAAGCATTGGA
cishb_S_R	GGAGAGAGGTGTCTGGATGCA
myl1_AS_F	GCAGTGGCGCAGTAGGTAGTG
myl1_AS_R	GCCGAGCCTCAAACCAAAG
myl1_S_F2	CCCACCGCTGATGATATGGT
myl1_S_R2	TGCAGCATAGGCAGGAAACC

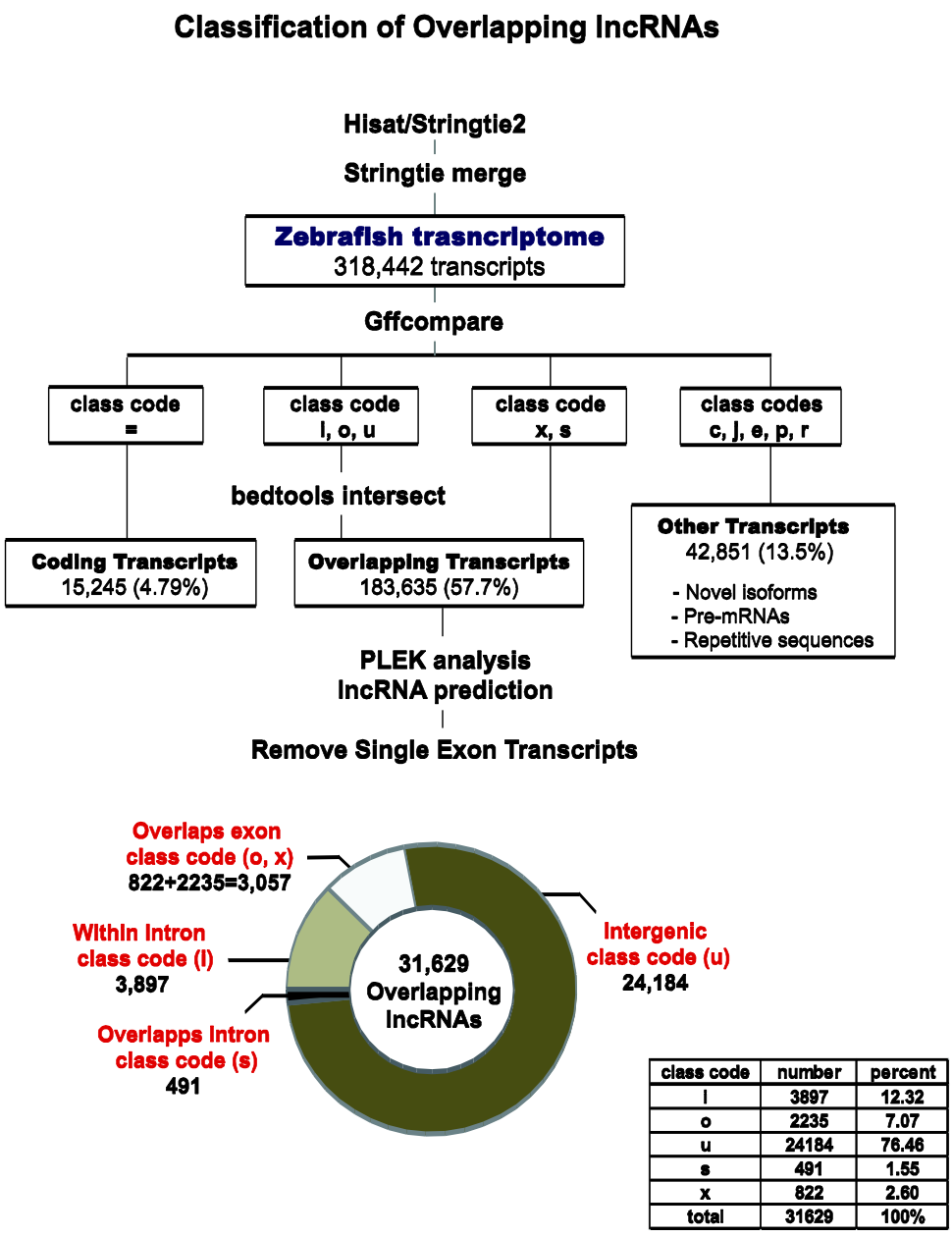
Slc4a1a_F	AAGAGATCCGGCCCTCTGA
Slc4a1a_R	CGCTTGCGTTGCAGTATGG
Slc4a1a_AS_F	GCCTTCAGAACTGCCTTAATCC
Slc4a1a_AS_R	CTGAAGAATATTTCCAGCACCTTGT
Fbxo32_F	TGTCCGGCCGATTAAAGAAC
Fbxo32_R	CCTGAACCAGACTGCACAGAGA
Fbxo32_AS_F	CGGTTAAAGGTTTGGGATCA
Fbxo32_AS_R	ATGGCGCATTGATGAACATA

ChIP_slc4a1a_peak1_F	CATCTGAGCCGAGGTTCGA
ChIP_slc4a1a_peak1_R	CAGTAGGTAGTGCTGTCGCTTCA
ChIP_Fbxo32 peak1_F	TTAATCGTCTTGCCAATAGGG
ChIP_Fbxo32 peak1_R	TTTCTAGCAAACAGCTTTTAGGA



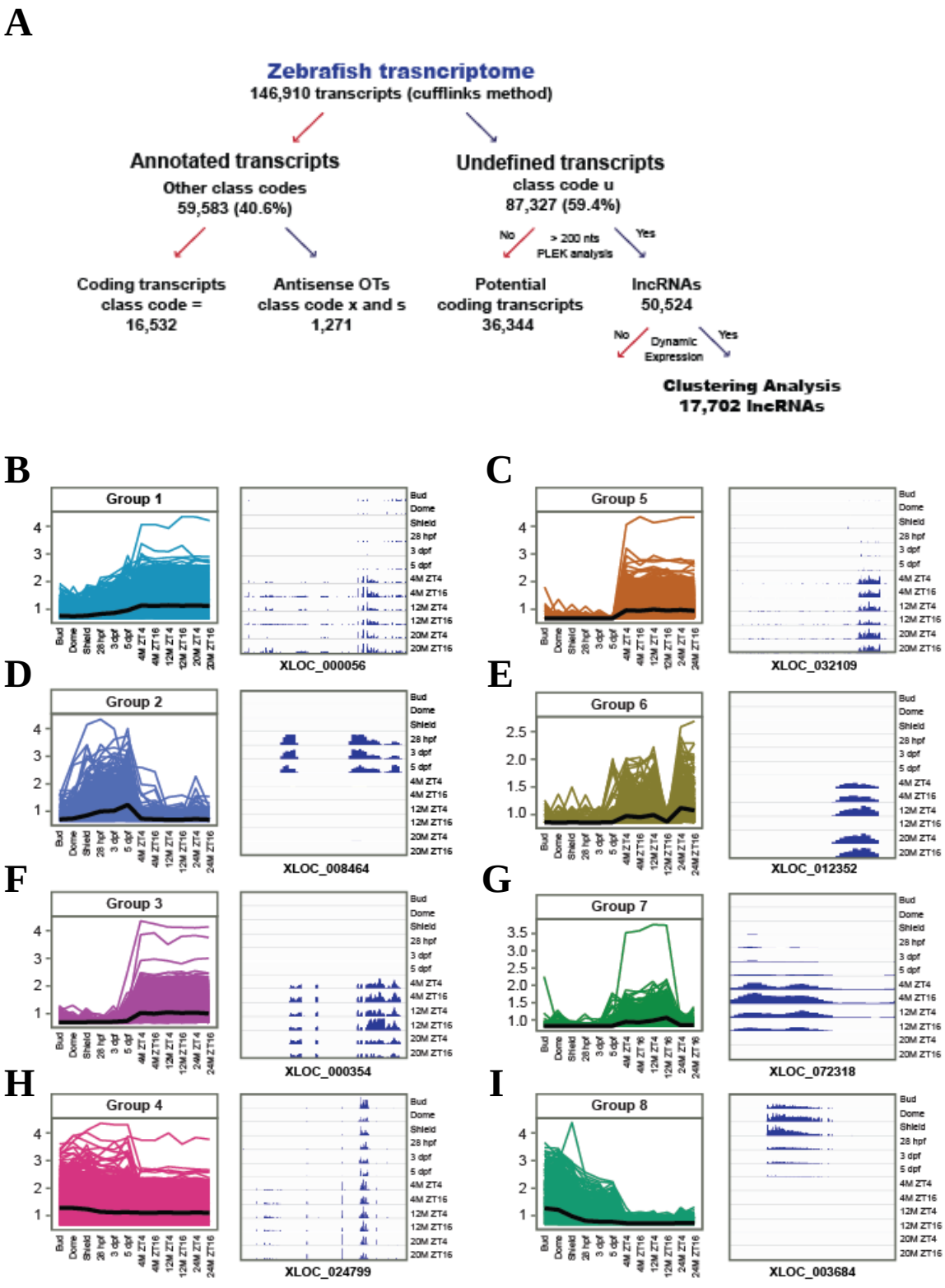
Supplementary Figure 8. Tophat2/Cufflinks pipeline

The T/C pipeline generated 146,910 transcripts (isoforms). The cutoff of FPKM in cufflinks was 0.05 (default setting). Among them, 16,532 was protein-coding transcripts (class code =). In order to discover OTs, class code x and s was extracted from the merged GTF file which is created after cuffmerge. The lists of class code i, o and u were also extracted and processed by using bedtools intersect (option -S) to sort only antisense OTs. In total, 14,774 OTs were identified through the tophat2-cufflinks pipeline. Then, PLEK analysis was done with the 14,774 OTs to predict lncRNAs from them. As a result of PLEK, I found 5,299 OT lncRNAs. Over 75 percent of the OT lncRNAs are unannotated intergenic from class code u.



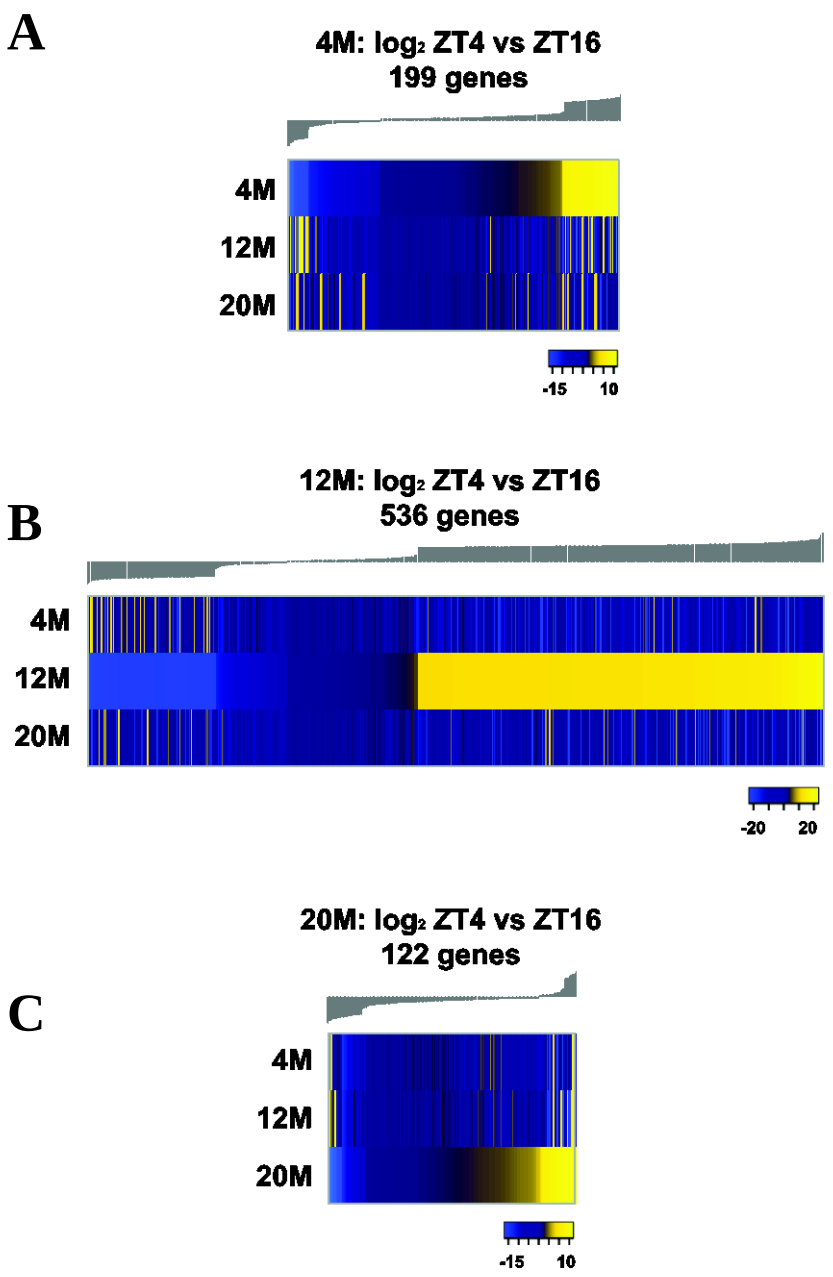
Supplementary figure 9. Hisat/Stringtie2 pipeline

The H/S pipeline discovered 318,442 transcripts at isoforms level. The FPKM cutoff of stringtie merge step was 0.1. As a result, I found 15,245 protein-coding transcripts (class code =) and 183,635 OTs. Among 183,635 OTs, PLEK analysis predicted 31,629 lncRNAs.


A
Zebrafish trasncriptome
146,910 transcripts (cufflinks method)
Annotated transcripts
Other class codes
59,583 (40.6%)
Undefined transcripts
class code u
87,327 (59.4%)
No
> 200 nts
PLEK analysis
Yes
Coding transcripts
class code =
16,532
Antisense OTs
class code x and s
1,271
Potential
coding transcripts
36,344
lncRNAs
50,524
No
Dynamic
Expression
Yes
Clustering Analysis
17,702 lncRNAs
B
Group 1
XLOC_000056
C
Group 5
XLOC_032109
D
Group 2
XLOC_008464
E
Group 6
XLOC_012352
F
Group 3
XLOC_000354
G
Group 7
XLOC_072318
H
Group 4
XLOC_024799
I
Group 8
XLOC_003684
Bud
Dome
Shield
28 hpf
3 dpf
5 dpf
4M ZT4
4M ZT16
12M ZT4
12M ZT16
20M ZT4
20M ZT16
24M ZT4
24M ZT16

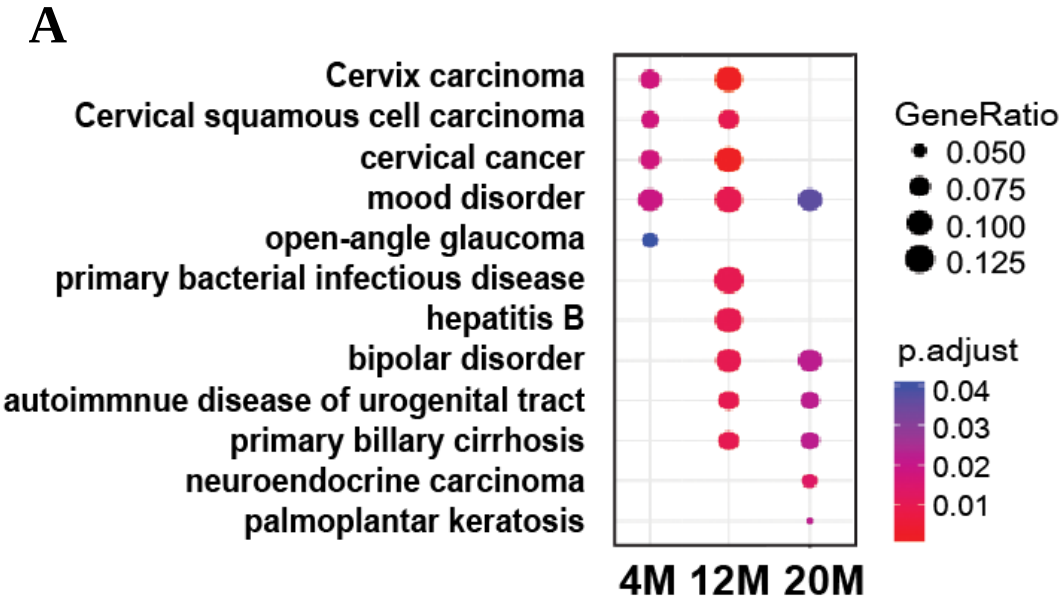
Supplementary Figure 10. Expressional patterns of novel lncRNA during developmental and aging processes in zebrafish

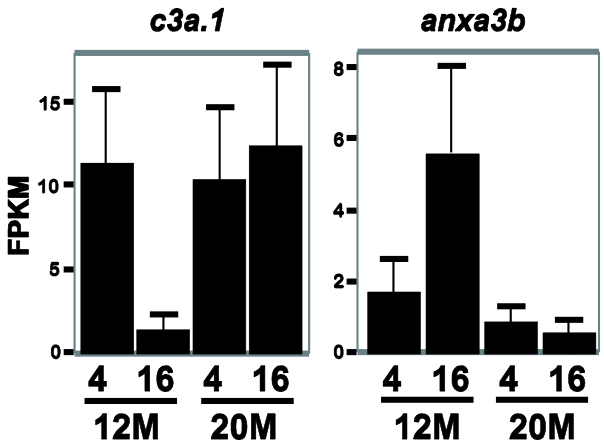
Overview of the RNA seq filtering pipeline to identify novel lncRNAs showing dynamic expression in developmental process and aging. Among 86,327 undefined transcripts, 50,524 lncRNA are sorted by PLEK analysis based on k-mer values. (a-h) The dynamic 17,702 undefined lncRNA were classified under eight groups based on similarity of expression patterns. (Black thick lines are an average pattern of each group). Representative expression of transcripts from each group was visualized by the IGV browser (See left panels).

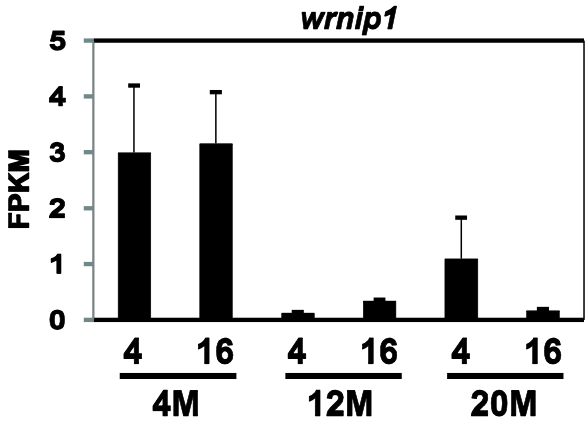


Supplementary figure 11. Age-dependent expression of diurnal genes

Heatmaps represent log_2 fold changes (FC) of FKPM values of diurnal genes between ZT4 and ZT16. The lists of diurnally regulated genes in exclusively 4 M (A), 12M (B) and 20M (C) are depicted in heatmaps, and the log2 FC are plotted as a bar chart on the top of each heatmap.

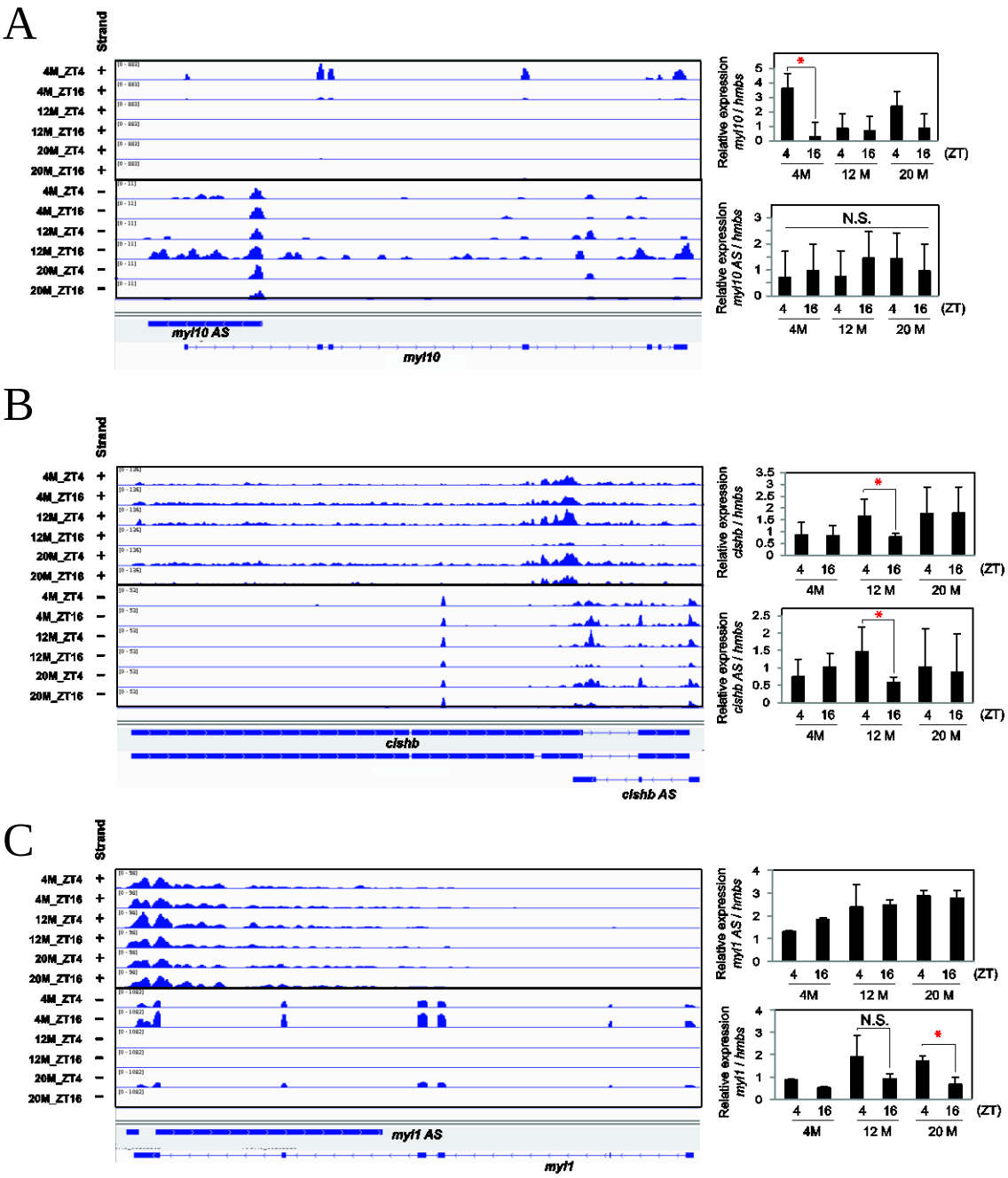

A
Cervix carcinoma
Cervical squamous cell carcinoma
cervical cancer
mood disorder
open-angle glaucoma
primary bacterial infectious disease
hepatitis B
bipolar disorder
autoimmnue disease of urogenital tract
primary billary cirrhosis
neuroendocrine carcinoma
palmoplantar keratosis
4M 12M 20M
GeneRatio
0.050
0.075
0.100
0.125
p.adjust
0.04
0.03
0.02
0.01


B
c3a.1
anxa3b
FPKM
15
10
5
0
8
6
4
2
0
4 16 4 16
12M 20M
4 16 4 16
12M 20M


C
wrnip1
FPKM
5
4
3
2
1
0
4 16 4 16 4 16
4M 12M 20M

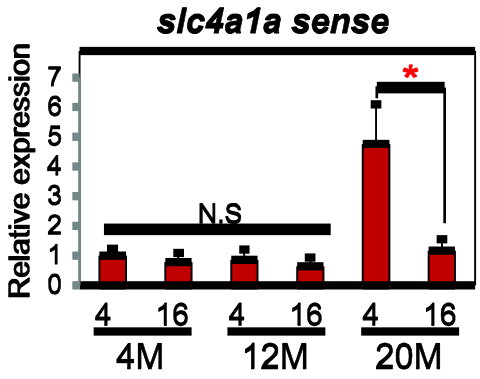
Supplementary Figure 12. Disease Ontology Enrichment Analysis of diurnal transcripts at different ages

(A) The most top-ranked enriched DO terms of diurnal genes in the specific age are illustrated in a dot graph ($p \leq 0.05$). (B) Age-specific diurnal genes which correlate with age-related diseases are visualized in bar charts. (C) Aging-related changes in *wrnip1* transcripts.

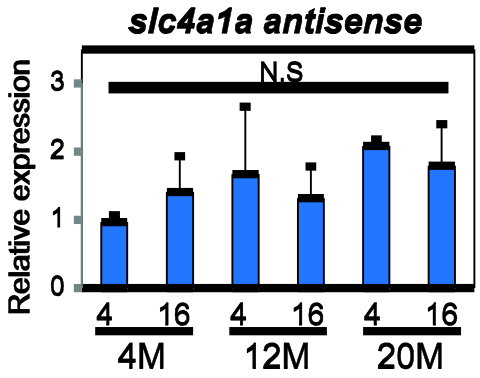


Supplementary Figure 13. Expression of myl10, cishb, and myl1 sense and antisense transcripts. IGV illustration showing relative expression of sense and antisense transcripts originated from (A) myl10, (B) cishb and (C) myl1 for both time points for all three age-groups. The panel on the right shows relative expression validated by RT-PCR (n=3). Error bars represent the SEM with asterisks indicating p-value < 0.05.

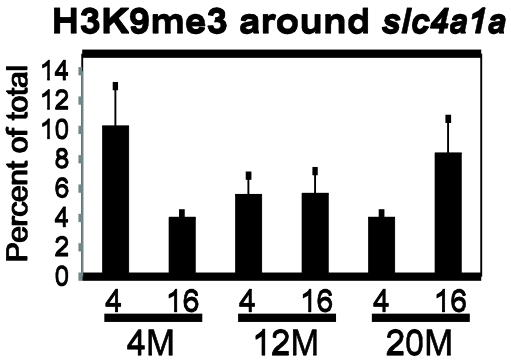
A



B

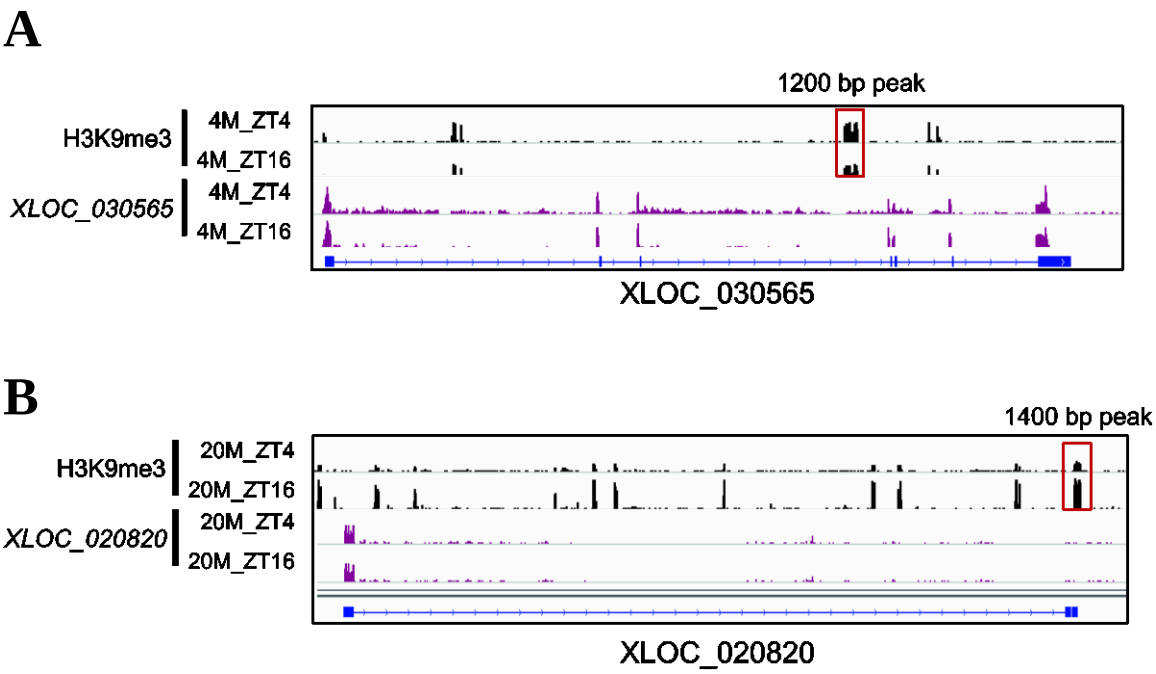


C



Supplementary Figure 14: Expression and H3K9me3 for slc4a1a

(A-B) The expression of slc4a1a sense and antisense transcripts was confirmed by RT-PCR (n=3). (C) H3K9me3 ChIP-qPCR at slc4a1a. Error bars represent the SEM and asterisks indicates $p<0.05$.



Supplementary Figure 15. Representative view of H3K9me3 broad peaks

Broad peaks over 1kb around unannotated transcripts were observed and illustrated in the IGV viewer. The H3K9me3 broad peaks were found in an unannotated transcript XLOC_030565 (A) and downstream of an unannotated transcript XLOC_020820 (B). The broader peaks spanning multiple nucleosomes showed a diurnal rhythm between ZT4 and ZT16.

3.4 MATERIALS AND METHODS

Animal Care

Wild-type zebrafish (Danio rerio) (SAT) were obtained from ZIRC (Zebrafish International Resource Center, Oregon), housed in system water (conductivity 600 ppm, pH 7.4) at 27 C, and fed twice daily. For diurnal entrainment, the fish were placed under a 12:12 light:dark cycle, but otherwise maintained at 14:10 light:dark cycle for breeding. Embryos and young larvae were raised in egg water (30 mg/L instant ocean in deionized water). For tissue extraction, fish were sacrificed by emersion in cold MS-222 (300 mg/l, Sigma), decapitated and dissected under Phosphate Buffered Saline (PBS) using microdissection tools.

RNA-Seq

4 Month (M), 12 M and 20 M zebrafish were maintained under a 12:12 light:dark cycle for one week prior to tissue removal. Zebrafish were sacrificed at ZT4 and ZT16 and brain tissue was removed as described (*233*). Tissue was immediately snap frozen in liquid nitrogen and RNA was extracted from all samples simultaneously using TRIzol (Invitrogen). RNA-Seq libraries (n=2) were processed at the Columbia Genome Center (New York, NY). Briefly, ribosomal RNA was depleted with RiboZero Gold (Illumina) and converted to cDNA with TruSeq Stranded Library Kit (Illumina). The libraries were sequenced on an Illumina 2500 to a depth of 90 million 100 bp paired-end reads per library. The data are deposited in Gene Expression Omnibus (GEO) under the accession number GSE109856.

ChIP-Seq

Zebrafish were maintained under identical conditions described for the RNA-seq. Brain tissue was removed and immediately cross-linked with 1% formaldehyde for 10 min then quenched with 0.1 M glycine for 10 min prior to being snap frozen in liquid nitrogen. Frozen tissue was lysed by mechanical disruption in the presence of ChIP lysis buffer [50.0 mM Hepes (pH 7.4), 150 mM NaCl, 1.0 mM EDTA, 1% Triton X-100, 0.1% deoxycholate] containing protease inhibitors (2.0 μg/mL leupeptin, 2.0 μg/mL pepstatin A, 1.0 mM PMSF). Additional cell disruption and chromatin shearing were done in a Misonix cup sonicator 5 times for 30 s at 20% power. The resulting lysates were cleared of cellular debris by centrifugation at 16,000 x g for 1 min. The average size of sheared chromatin was 500 bp. For the ChIP, 200 g of lysate was mixed with 3 μg H3K9me3 antibody (Abcam ab8898) prebound to Protein A magnetic beads (Dynal) then incubated overnight at 4oC. The ChIP samples were washed 5x with RIPA buffer (10 mM Tris pH 8.0, 1mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% sodium deoxycholate, 140 mM NaCl) and eluted with 0.1M sodium bicarbonate, 1% SDS. Isolated samples were heated at 65 oC to reverse crosslinks and protein was removed by the addition of proteinase K. DNA was further purified by phenol-chloroform extraction then precipitated with 1/10 volume 3M NaOAc pH 5.4 and 2.5 volumes 95% ethanol. Purified DNA from each H3K9me3 ChIP done in triplicate was used for library preparation. DNA was sequenced 50 million 50-bp single-end reads on the Illumina HiSeq 2500 at Duke University Genome Center. The data are deposited in Gene Expression Omnibus (GEO) under the accession number GSE109856. H3K9me3 from the ChIP-seq was validated by qPCR using oligonucleotides contained in Table 6.

RNA-seq Analysis

The RNA-seq data were processed using two different analysis pipelines roughly outlined in Supplementary figures 1 and 2 for transcript discovery. The first method used Tophat2-Cufflinks-Cuffdiff and second used Hisat2-Stringtie-Cuffdiff (*222, 223, 234*). To enhance transcript discovery, I also included previously published RNA-seq dataset from developing zebrafish embryos (GSE32989) (*224*). The combined analysis yielded a single GTF file for differential expression. Paired-end reads were aligned to zebrafish reference assembly GRCz10 (http://www.ensembl.org/Danio_rerio/Info/Index) with TopHat2 (version 2.0.9) using the GRCz10 GTF file as a guide. The mapped reads were assembled with Cufflinks (version 2.2.1). All transcripts from Cufflinks were combined to create new GTF file using Cuffmerge with the default FPKM cutoff of 0.05. The differential gene expression was performed using Cuffdiff (*222*). Data was also mapped with HISAT2, and a separate transcript file was created using StringTie (*223*). The StringTie minimum FPKM was set 0.1 because that was the maximum level capable of detecting the known Per2AS transcript. All subsequent statistical analyses were done in R Studio using the bioconductor package CummeRbund. Visualization of transcripts was done using the Integrative Genomics Viewer (IGV) (*183*).

ChIP-seq Analysis

The reads of H3K9me3 ChIP-seq were mapped to the zebrafish GRCz10 genome using Bowtie2 (*235*). Bam files of each triplicate were further processed using MACS2 for peak calling (FDR $\leq$ 0.01) (*231*). Differential H3K9me3 enrichment between different conditions was analyzed by Diffbind and visualized with MAplot (*232*). ChIPseeker was

used to analyze genomic features of the H3K9me3 enrichment (*236*). The H3K9me3 peaks were visualized with IGV (*183*).

Quantitative Reverse Transcription PCR

Zebrafish (n=3) were maintained and entrained under identical conditions described for RNA-seq. Brain tissue was removed and immediately snap frozen in liquid nitrogen. TRIzol reagent was used for the extraction of total RNA from the brain tissue. 500 g of total RNA was then used as a template with the SuperScript III First-Strand Synthesis kit (Invitrogen) to produce cDNA. The specific targets were amplified by RT-PCR using oligonucleotides in table6 using *hmbs* as the internal control. T-tests were performed to calculated p-value to determine significant difference between samples.

Gene and Disease Ontology

WebGestalt (WEB-based GEne SeT AnaLysis Toolkit) was used for Gene Ontology (GO) enrichment using an adjusted p-value ≤ 0.01 as the cutoff, with multiple-hypothesis correction to calculate the enrichment p-value. The GO clusters with significant p-value were taken for further analysis. The top GO terms were converted to log10 adjusted p-value and visualized with a radial graph or in tabular format (*226*, *227*). In order to use human Disease Ontology database, the zebrafish genes were first matched to their human orthologues by using Ensembl biomart (www.ensembl.org/biomart) and I used Bioconductor DOSE packages with p-value ≤ 0.05 cut to identify corresponding genes.

3.5 DISCUSSION

Using multiple genomic approaches, the work reported here was designed as an initial test for the hypothesis that diurnal gene expression and age-related dysregulation is driven, in part, by lncRNA-guided heterochromatin changes. This hypothesis arose from a number of previous observations. First, H3K9me2 and H3K9me3-dependent facultative heterochromatin occurs at the clock genes during the repressive phase of the circadian cycle (*39, 84, 216*). Second, the Per-complex contains KMT1/SUV39 along with a number of noncoding RNA binding proteins including NONO (*39, 40*). Third, the Per2 gene has an antisense transcript and in *Neurospora*, the *frq* NAT natural antisense transcript (NAT) *qrf* is needed for H3K9me-dependent facultative heterochromatin (*78, 81, 97*). Forth, there are multiple examples, from single-cell eukaryotes to mammals that reveal H3K9me3 can be directed by lncRNA ranging from centromeres in S. pombe (*237*) to H19-guided imprinting in mammals (*238*). And finally, emerging models on aging suggest there is a redistribution of H3K9me3 and DNA methylation, which is either a direct cause or a consequence of aging. Taken together, these findings suggest age-related changes to the circadian transcriptome may occur in part due to age-related changes to circadian lncRNA, which could then alter the underlying heterochromatin. The data presented here is by no means definitive proof this as an all-inclusive model, but does demonstrate a likelihood this occurs at some loci. The data also reveal other mechanisms must exist to account for changes in gene expression independent of H3K9me3.

In good agreement with previous observations, this report shows global reprogramming of the circadian transcriptome during aging even though the underlying circadian oscillator remains largely unaffected. For example, in the liver, age-induced

changes in circadian expression are accompanied by changes in protein acetylation and NAD+ metabolism (*138*). Similarly, extensive reprogramming of circadian transcriptome in aged stem cells causes a switch in expression of genes involved in homeostasis, to those involved in tissue-specific stress such as DNA damage or inefficient autophagy (*137*). My study contributes new insights into mechanisms underlying age-related changes to diurnal expression and further implicates epigenetic changes that include alterations to circadian lncRNA and genome-wide heterochromatin as potential regulators. An additional objective of this study was to further identify lncRNAs in zebrafish brain. Considering the majority of expressed transcripts in mammals do not appear to code for proteins (*112*), these data should serve as a valuable resource to the zebrafish community. In particular, lncRNAs have diverse roles in regulation ranging from transcriptional interference in yeast, to X chromosome inactivation, imprinting, and epigenetic silencing in mammals, including directing H3K9me3 and H3K27me3 (*239*) so a more complete zebrafish noncode is warranted. Due to the growing evidence lncRNAs are targets of circadian regulation in a variety of tissues and organisms, it is important to get a base-line measure of lncRNAs to begin to understand their role in the clockworks. For example there are 112 lncRNAs showing differential expression in the light and dark in rat pineal gland (*104*) and mouse circadian transcriptome data suggests there are at least 1000 circadian regulated lncRNAs observed in multiple tissues (*210*). Here, I expanded the zebrafish noncode by adding brain to a developmental dataset and explored whether any of these lncRNAs coincided with diurnal changes in H3K9me3. Further, I examined how lncRNA expression changed with age and time-of-day. Among the over 140,000 transcripts isoforms identified via Cufflinks, more than 50 percent were previously

unannotated in the GRCz10 and 50,524 are predicted to be lncRNAs by PLEK (both long intergenic non-coding RNA and OT). It remains to be determined how many of these newly identified transcripts are under direct control of the circadian clock and whether they function as heterochromatin regulators in either *cis or trans*.

My findings of age-related changes to diurnal transcription are consistent with other reports showing the core circadian oscillator is relatively stable regardless of age, but ccg expression undergoes chronological changes over an animals lifetime. Specifically in the study not only are protein-coding *ccgs* changing with age, but also a significant number of rhythmic lncRNA also change as organism's age. From gene ontology analysis of protein-coding genes, it is clear the circadian transcriptome changes from anabolic to catabolic processes as one ages. The dynamic nature of age-related changes to clock output between 4M and 12M suggests the existence of a complex regulatory mechanism that partially resides at the level of chromatin and becomes unstable with age (20M). This is quite consistent with the age-related changes in acetylation given deacetylation precedes heterochromatin.

I also examined age-related changes in diurnal genes by disease ontology analysis in hopes of identifying key factors that could provide insight into age-related disease. I specifically examined changes in the expression of Werner syndrome helicase (WRN) because mutations in WRN protein cause an advanced aging disorder known as adult progeria. WRN, like PER2, is associated with KMT1 (SUV39h) and HP1 and there is a reduction in heterochromatin in WRN null cells (*221*). I did not observe any age-related changes to WRN (data not shown), nor did it have a diurnal expression (at least at the phase I examined). However, WRN interacting protein 1 (wrnip1) was lost with age,

so it is interesting to speculate that an interaction between Wrn and KMT1 might occur through Wrnip1 and this could potentially account for some of the age-related changes in heterochromatin. Alternatively, other alterations in heterochromatin may be mediated by changes in lncRNAs that potentially interact with the PER2 complex and direct Per-associated KMT1 to different loci; a speculative notion at this stage, but one that needs further study. Regardless, it is clear there is a complex regulatory system below the level of the master oscillator controlling circadian output that changes with age and occurs concurrent with changes in lncRNA, H3K9me3, and must include other chromatin changes.

As a final note, it is important to emphasize that the goal of this study was not to identify circadian protein-coding genes per se, this has been done exhaustively by others. As such, I intentionally ignored established guidelines for circadian rhythm research (*240*) and instead set out to examine a potential mechanism for age-related changes to gene expression, with the main focus being lncRNA discovery and heterochromatin. To that end, I limited my analysis to two diurnal time points, sequenced to extremely high depth (90 million paired-end reads), and preserved strandedness to augment transcript discovery. This focus on diurnal regulation proved fruitful as it allowed us to identify a much broader range of lncRNAs and analyze H3K9me3 in greater detail. It is now clear that a more comprehensive analysis of age-related changes to lncRNAs and corresponding chromatin structural changes on the circadian time scale is warranted to uncover more detailed mechanism(s) underlying aging.

Conclusions: There are age-related changes to lncRNA and H3K9me3 that coincide with changes in *ccg* expression. Based on the known function of lncRNA guiding epigenetic modifications in cellular events such as imprinting, X-chromosome inactivation, and NAT-mediated heterochromatin, it is possible that a subset of diurnal lncRNA may guide age-related epigenetic changes that affect *ccg* expression without altering the master circadian oscillator. This theory is supported by my observations showing diurnal and age-related changes in gene expression, including lncRNAs, coincide with diurnal and age-related changes in H3K9me3. This notion provides one potential solution to the perplexing observation that clock output is dynamic as organism's age, but the core clock does not change appreciably (*137, 138*). My findings reveal that chromatin structure, in the form of circadian-regulated facultative heterochromatin, is one facet that determines whether a gene is rhythmic in young or older adults or when it is expressed in given block of time over one's lifespan. Based on these and other findings, this may be one of many underlying causes of age-related disease manifestation observed with circadian disruption.

3.6 ACKNOWLEDGMENTS

I thank Dr. Elizabeth Snyder and members of the Belden Lab for critical comments and discussions.

CHAPTER 4

Conclusions

The circadian clock is a biological oscillator that functions to anticipate external environmental changes. The molecular mechanism of the circadian clock is evolutionarily conserved from fungi to mammals and it is predicted that roughly 40 percent of all protein-coding genes in mammals are regulated by the circadian clock. Accordingly, the clock is a key regulator in various physiologic and metabolic pathways including the aging process. Mutations in core clock genes or disruption in diurnal rhythms cause premature aging and an increased incidence of age-related disease (*132, 184-187*). Thus, it has been proposed that age-dependent alteration in the circadian clock is involved in age-related physiological changes. Indeed, telomere shortening as a hallmark of aging correlates with disrupted circadian rhythms and the clock regulates telomerase, suggesting the molecular connection between the circadian clock and aging is through telomere regulation. Supporting this notion are recent bioinformatics studies demonstrating genome-wide reprogramming in circadian transcripts during aging (*137, 138*). Although accumulated evidence indicates an essential role of circadian clock in aging, the precise molecular mechanisms underlying circadian clock and aging remain unclear.

In chapter2, I found circadian clock proteins are associated with the telomere, suggesting a direct role for the clock in telomere homeostasis. Examination of published ChIP-seq data first hinted that WC-2 in *Neurospora* and BMAL1 in mice interacted with telomeres. I confirmed the association between BMAL1 and telomere and found it is conserved in zebrafish and mice. I also determined BMAL1-dependent rhythms in TERRA and heterochromatin at the telomere. Taken together, my results suggest that BMAL1 is directly involved in telomere homeostasis by controlling rhythms in TERRA

and heterochromatin at the telomere. Even though I determined BMAL1 binds to telomere DNA in zebrafish and mice, it is still unclear how the association is achieved. Thus, EMSA assay should be performed to study further how BMAL1 interact with telomere DNA. In addition, to understand how the BMAL1 binding to telomere is related to telomere structure, it could be included to observe the interaction between BMAL1 to shelterin complex components such as RAP1, TRF,1and POT1.

To further describe the mechanism of age-related changes to circadian chromatin, in chapter 3 I performed a genome-wide study with a focus on facultative heterochromatin and diurnal non-coding RNAs. The combined RNA-Seq and ChIP-Seq at two diurnal time-points for three different age groups revealed a connection between age-related changes to circadian transcription and heterochromatin in the zebrafish brain. My results revealed that lncRNAs and *ccgs* change dramatically with age, but the core clock genes remain rhythmic. The age-related changes in *ccg* expression showed that age-dependent circadian outputs change from anabolic to catabolic processes during aging. Coincidently, I observed diurnal and age-related changes in H3K9me3 suggesting age-related changes in diurnal transcripts may be due to age-related alterations in facultative heterochromatin. Future research could include other modification such as H3K27me3, which is associated with gene activation, providing more comprehensive understanding among circadian clock, lncRNAs, and chromatin modifications in gene repression and activation processes.

The direct role of the circadian clock in telomere homeostasis may provide insight to understand the connection of circadian rhythms in aging. Moreover, targeting the circadian clocks to regulate telomere homeostasis opens a novel direction in anti-

aging therapy. This, in conjunction with the genome-wide study connecting diurnal heterochromatin and lncRNAs with age provides a solid background for further research on the interplay between the clocks, chromatin structure, and lncRNAs in the aging process.

APPENDIX

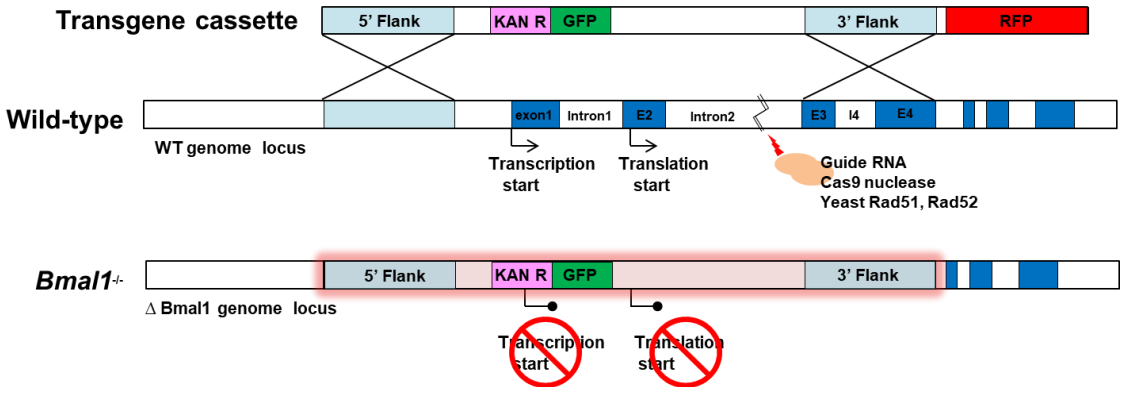
A. Construction of *bmal1* knock-out zebrafish using homologous gene targeting

One of the goals of in chapter 2 was to examine BMAL1, TERRA and H3K9me3 rhythms in a *bmal1* knockout. To accomplish this, I attempted but failed to make a homozyogous *bmal1*$^{-/-}$ zebrafish line. Below is a discussion of the work I did trying to obtain this fish line.

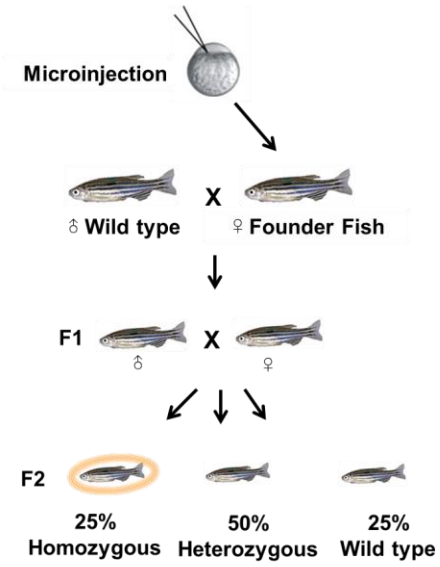
To establish the *bmal1*$^{-/-}$ zebrafish, I attempted to knock out the endogenous *bmal1* gene by targeted gene replacement using CRISPR/Cas9. I micro-injected a homologous gene targeting cassette shown below (Fig.14A). First, I created a targeting construct using a yeast homologous recombination (*241*). The 5' and 3' flanks for *bmal1* knockout contruct were amplified by PCR (see table 7 for oligonucleotide sequences) and inserted into pZHGT3 linearized with the meganuclease *I-Ceu*I by homologous recombination in yeast FY2. The 5' flank fragment contained approximately 3 kb upstream of transcriptional start site and the 3' flank fragment was approximately 3 kb immediately downstream of translational start site. To make transgene cassette for *bmal1* locus, The pZHGT3-*bmal1*KO was linearized with *I-SceI* and purified by gel extraction. Next, the gRNA was designed to target *bmal1* genomic locus between exon 1 and 2, and Cas9 mRNA with a poly A tail was synthesized using by vitro transcription (Ambion). I also added purified yeast Rad51 and Rad52 proteins into the micro-injection mixture in order to improve the efficiency of homologous recombination. The rad51 and 52 CDS were cloned into pET24a vectors respectively. The resulting vectors were transformed into BL21(DE3)-RIL and induced with 0.5 mM IPTG for 4 h. The purification of bacterial lysates containing Rad51 or Rad52 proteins was conducted by fast protein liquid

chromatography (FPLC) dialyzed with 1X Danieau's solution with 20% glycerol. The concentration of Rad51 and Rad52 proteins was adjusted 2.5 μM, Lastly, three morpholinos (MO) targeting Ku70, Lig4, and XRCC4 mRNA were included into the micro-injection mixture in order to inhibit non-homologous end joining (NHEJ). The concentration of morpholinos was 50 μM and the morpholinos were boiled at 95°C for 5 min and cooled on ice right before adding to the micro-injection mixture. Only one-cell stage eggs were used for microinjection. The 25 μl of micro-injection mixture including 1 μl of Rad51/52 proteins, 2.5 μl of cas9 mRNA, 2.5 μl of *bmal1* gRNA, 2.5 μl of each morpholino (XRCC4, Lig4, Ku70), 2 μl of transgene cassette, 2 μl of phenol red dye, and 9.5 μl of water. 2 nl of the mixture was injected into fertalized eggs at the one-cell stage. After injection, the eggs were separated into 48-well plates. 2 days later, healthy eggs were transferred into a 10 mm Petri dish (~25 eggs per one Petri dish). 3 month after the micro-injection, I genotyped the fish by Southern blot using genomic DNA from fin-clip using DIG-labeled GFP DNA as a probe. The GFP-positive founder fish were bred with WT zebrafish to validate germ-line transmission of *bmal1* knock-out . Southern blot was repeated on F1 fish male and female fish were selected for further breeding to generate a homozygous *bmal1* knock-out zebrafish. I conducted PCR based genotyping to find *bmal1*$^{-/-}$ zebrafish among F2 generation (see table 6). Despite significant attempts, I was unable to obtain a *bmal1*$^{-/-}$ homozygous zebrafish. It is possible that the transgene inserted into the non-specific genomic locus or the *bmal1* homozygous fish could not survive past 30 days post-fertilization.

A


Transgene cassette
5' Flank
KAN R
GFP
3' Flank
RFP
Wild-type
exon1
Intron1
E2
Intron2
E3
I4
E4
WT genome locus
Transcription start
Translation start
Guide RNA
Cas9 nuclease
Yeast Rad51, Rad52
Bmal1-/-
5' Flank
KAN R
GFP
3' Flank
Δ Bmal1 genome locus
Transcription start
Translation start

B


Microinjection
♂ Wild type
X
♀ Founder Fish
F1
X
F2
25%
Homozygous
50%
Heterozygous
25%
Wild type

C

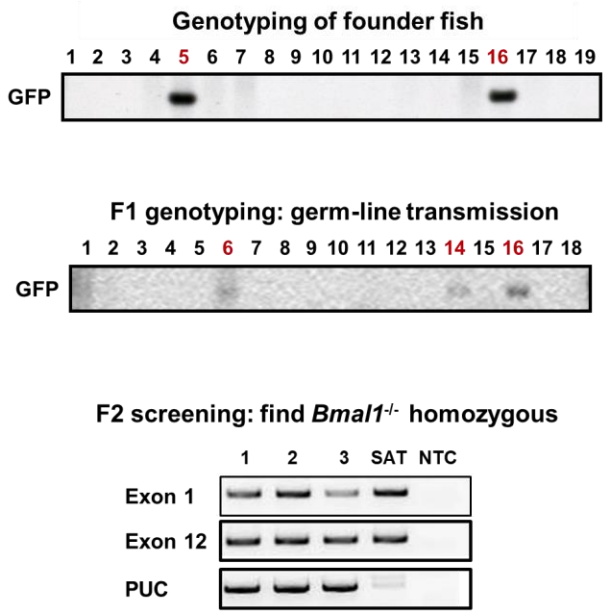

Genotyping of founder fish
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
GFP
F1 genotyping: germ-line transmission
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
GFP
F2 screening: find Bmal1-/- homozygous
1 2 3 SAT NTC
Exon 1
Exon 12
PUC

Figure 15. Homologous gene targeting for *bmal1*$^{-/-}$ zebrafish

(A) A schematic representation gene targeting cassette to produce *bmal1*$^{-/-}$ zebrafish. (B) The breeding scheme to obtain *bmal1*$^{-/-}$ homozygous zebrafish (C) Genotyping based on Southern blot using GFP as a probe. A PCR-based genotyping on F2 embryos by using oligonucleotides specific to the *bmal1* exon1, exon12, and PUC region in the transgenic cassette.

Table 7. The oligo sequence of appendix A

5F_F	ACTAAACTCACTAGGGATAACAGGGTAATACACAGCAGT CACAATGTAC
5F_R	GTAATACGACTCACTATAGGGCGAATTGGGTACCGTAAG CTAAATGGTTCGTAG
3F_F	CGATATCAAGCTTATCGATACCGTCGACCTCGAGGGACTA CAACCGCAAGAGG
3F_R	AATGACCCCGTAATTGATTACTACCTCAGCGGGCGGTCAA TTAAGCAAC
Exon 12 F	CAGACCGTAAGAGCTTTTGCACC
Exon 12 R	CTCTGGTCTACGAACACAAAC
Exon 1 F	CACAGTTCCTCACGGACCG
Exon 1 R	GGTTTGTGGCGAAGCTCTG
pUC_F	CTAAATACATTCAAATATGTATCCG
pUC_R	GCATGGCGGTAATACGGTTATC

B. Circadian regulation on shuttling of TIN2 between the nucleus and cytoplasm

During the course of my thesis it was increasingly clear circadian clock regulated telomere homeostatis. These finding made examine circadian regulation of Shelterin activity and how this may communicate with cellular metabolism; another cellular process controlled by the circadian clock. TIN2 is a component of the Shelterin complex it appears TIN2 shuttles between the nucleus and mitochondria to regulate metabolism and ROS production (*242*). Also, over a third of the mitochondrial proteins accumulate in a circadian manner to serve catabolic and oxidative functions of mitochondria (*243*). Based on these results coupled with my findings at the telomere, I hypothesized there may be circadian translocation of TIN2 from nucleus to the mitochondria. To test this idea, I examine trafficking of TIN2 between the nucleus and cytoplasm by subcellular fractionation into a nuclear fraction and a cytoplasmic fraction in two different cell lines (U2OS and NIH3T3) (*244*). Cells are grown to 100 % confluency, entrained with 0.1 mM dexamethasone for 2 hours and harevested at designated time. Nuclear and Cytoplasmic were was carried out using the Rapid Efficient and Practical (REAP) protocol (*244*). 20 μg of proteins from each time point were subjected to SDS-PAGE and western blotting using specific antibodies to the TIN2, Cytochrome C as a cytoplasmic/mitochondrial marker, and Histone H3 as a nucleus marker. For quantification of the western blot, the ImageJ program was used. The data indicate no circadian trafficking of TIN2 translocation between the nucleus and cytoplasm in either U2OS and NIH3T3 (Figure. 15).

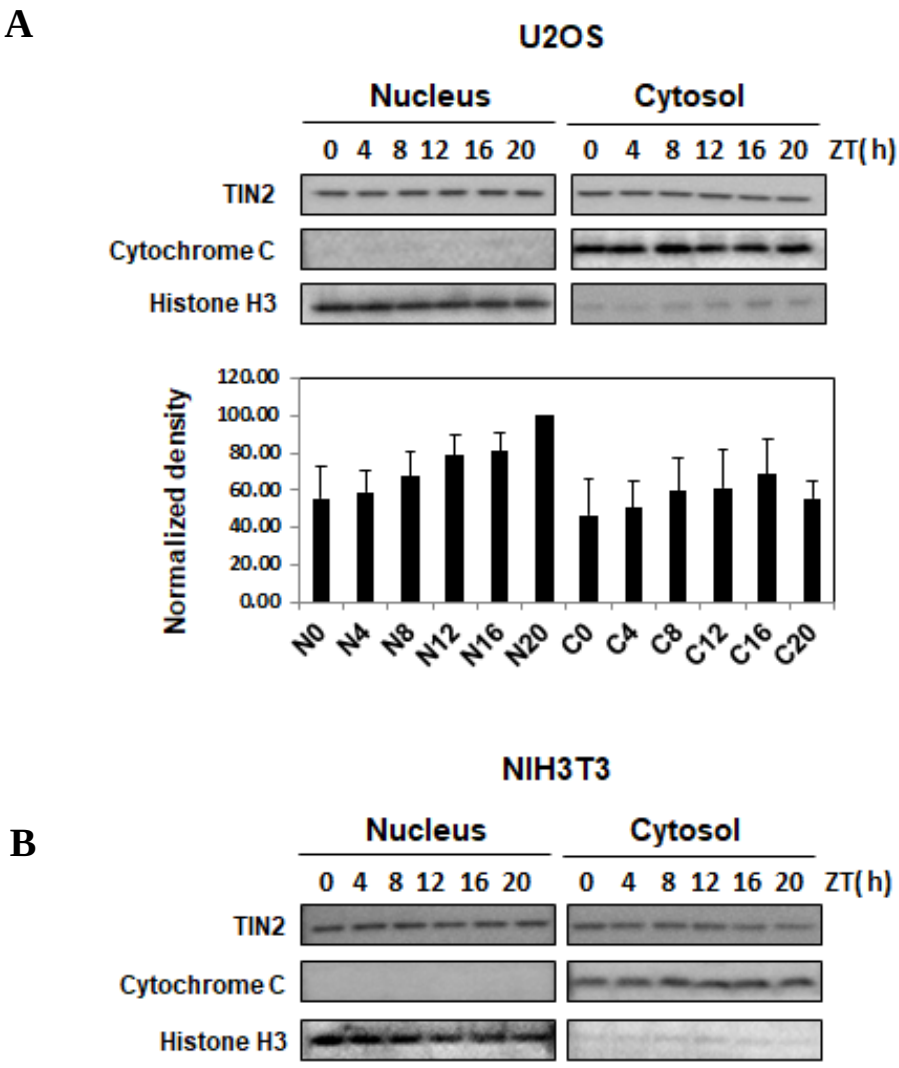


Figure 16. TIN2 does not show rhythmic translocation between the nucleus and cytosol. A representative western blot of nuclear and cytosolic fractions of the U2OS (A) and NIH3T3 (B). Cytochrome C was used as a cytoplasmic and mitochondria marker, histone H3 as a nuclear marker to exclude contamination during cell fraction isolation. The experiments in U2OS were independently performed on 3 biological replicates. Quantification of western blots in U2OS from three replicates was depicted as a bar graph. The error bars show SEM. NIH3T3 experiment was performed with one set of sample.

C. The effect of TIN2 knock-down in *bmal1*-luciferase rhythms

I also examine if there was feedback regulation from the telomere to the circadian clock. In order to test whether disrupted telomere homeostasis is associated with the circadian disruption, I performed a TIN2 knockdownand examine a rhythms in *bmal*1 luciferase. I transfected U2OS cells containing bmal1-luciferase with a *Tin2* siRNA that contained three different siRNA seqeuces targeting *Tin2*. The *Tin2* knock-down efficiency of the siRNAs was validated by real-time PCR (Figure 16A). Based on the result, I selected two siRNAs for further analysis. The *Tin2* siRNAs were transfected using Lipofectamine 3000 at 80 percent cell confluency. The cells were then entrained with 0.1mM of dexamethasone for 2 hours then media containing D-luciferin was added to the transfected cells. Bioluminescence of *bmal1*-luciferase was detected every hour for a 72 hour period. The data indicate that knockdown of *Tin2* mRNA does not affect the period of *bmal1* rhythms (Figure 16B). Although this produced negative data there may still be an effect of other shelterin components and I did not test TRF1, TRF2, POT1, or RAP1..

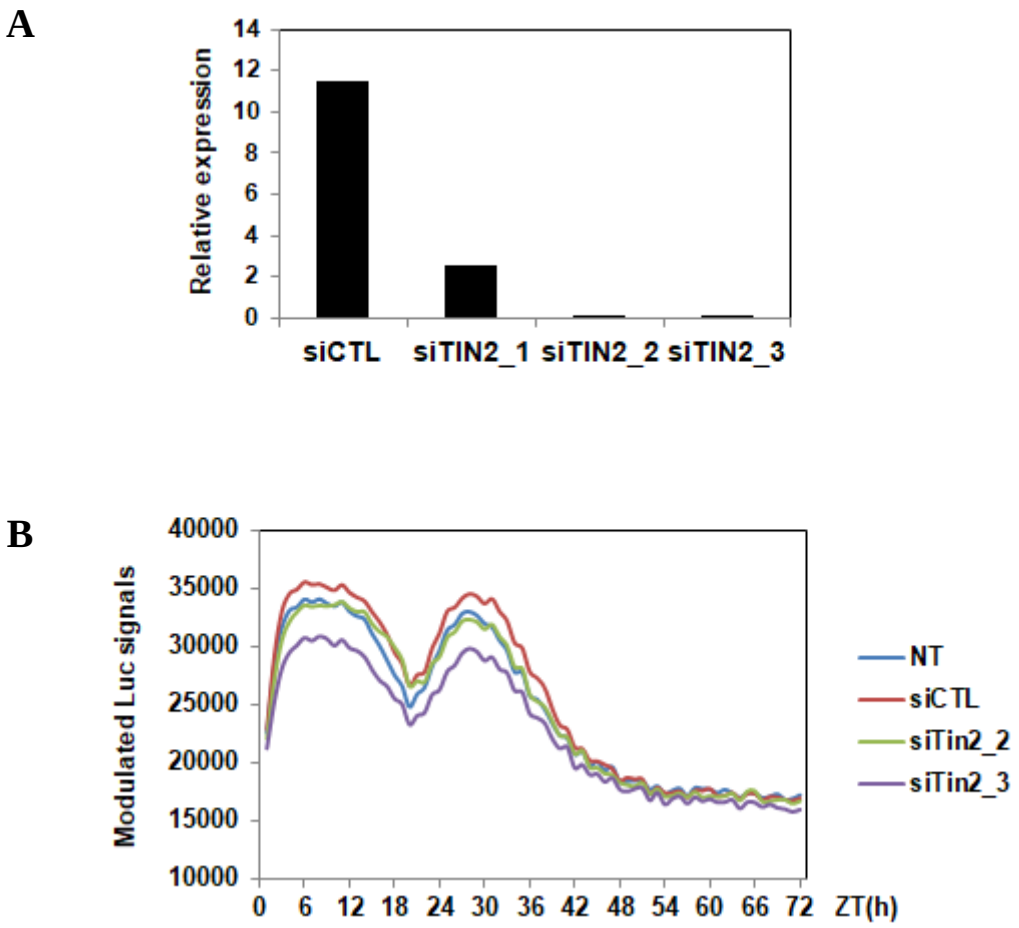


Figure 16. The knock-down of *Tin2* does not change the rhythm of *bmal1*-luciferase

(A) Down-regulation of *Tin2* mRNA in U2OS cells determined by q RT-PCR. *eif2a* was used as an internal control (siCTL: scrambled sequence, siTin2_1: 5'-GAATCCTCCTCAGCAACAA-3', siTin2_2: 5'-AGCAACAAGAATCCTCCTC-3', siTin3: 5'-CAGCAACAAGAATCCTCCT-3'). (B) Bioluminescence traces of *bmal1*-luciferase from U2OS cells treated with siCTL as a control or siRNAs targeting Tin2 (NT : non-treated sample).

BIBLIOGRAPHY

1. C. Lopez-Otin, M. A. Blasco, L. Partridge, M. Serrano, G. Kroemer, The hallmarks of aging. *Cell* **153**, 1194-1217 (2013).
2. A. A. Kondratova, R. V. Kondratov, The circadian clock and pathology of the ageing brain. *Nat Rev Neurosci* **13**, 325-335 (2012).
3. J. Mattis, A. Sehgal, Circadian Rhythms, Sleep, and Disorders of Aging. *Trends Endocrinol Metab* **27**, 192-203 (2016).
4. S. M. Abbott, A. Videnovic, Chronic sleep disturbance and neural injury: links to neurodegenerative disease. *Nat Sci Sleep* **8**, 55-61 (2016).
5. D. Bell-Pedersen *et al.*, Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat Rev Genet* **6**, 544-556 (2005).
6. P. E. Hardin, S. Panda, Circadian timekeeping and output mechanisms in animals. *Curr Opin Neurobiol* **23**, 724-731 (2013).
7. U. Schibler, P. Sassone-Corsi, A web of circadian pacemakers. *Cell* **111**, 919-922 (2002).
8. J. C. Dunlap, Molecular bases for circadian clocks. *Cell* **96**, 271-290 (1999).
9. C. H. Ko, J. S. Takahashi, Molecular components of the mammalian circadian clock. *Hum Mol Genet* **15 Spec No 2**, R271-277 (2006).
10. S. M. Reppert, D. R. Weaver, Coordination of circadian timing in mammals. *Nature* **418**, 935-941 (2002).
11. C. Heintzen, Y. Liu, The Neurospora crassa circadian clock. *Adv Genet* **58**, 25-66 (2007).
12. M. Brunner, T. Schafmeier, Transcriptional and post-transcriptional regulation of the circadian clock of cyanobacteria and Neurospora. *Genes Dev* **20**, 1061-1074 (2006).
13. R. Allada, N. E. White, W. V. So, J. C. Hall, M. Rosbash, A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless. *Cell* **93**, 791-804 (1998).
14. H. Zeng, Z. Qian, M. P. Myers, M. Rosbash, A light-entrainment mechanism for the Drosophila circadian clock. *Nature* **380**, 129-135 (1996).
15. P. E. Hardin, Molecular genetic analysis of circadian timekeeping in Drosophila. *Adv Genet* **74**, 141-173 (2011).
16. B. Kloss *et al.*, The Drosophila clock gene double-time encodes a protein closely related to human casein kinase Iepsilon. *Cell* **94**, 97-107 (1998).
17. J. L. Price *et al.*, double-time is a novel Drosophila clock gene that regulates PERIOD protein accumulation. *Cell* **94**, 83-95 (1998).
18. P. Emery *et al.*, Drosophila CRY is a deep brain circadian photoreceptor. *Neuron* **26**, 493-504 (2000).
19. S. K. Crosthwaite, J. C. Dunlap, J. J. Loros, Neurospora wc-1 and wc-2: transcription,

photoresponses, and the origins of circadian rhythmicity. *Science* **276**, 763-769 (1997).

20. P. Cheng, Q. He, Q. He, L. Wang, Y. Liu, Regulation of the Neurospora circadian clock by an RNA helicase. *Genes Dev* **19**, 234-241 (2005).
21. D. L. Denault, J. J. Loros, J. C. Dunlap, WC-2 mediates WC-1-FRQ interaction within the PAS protein-linked circadian feedback loop of Neurospora. *EMBO J* **20**, 109-117 (2001).
22. J. M. Hurley, L. F. Larrondo, J. J. Loros, J. C. Dunlap, Conserved RNA helicase FRH acts nonenzymatically to support the intrinsically disordered neurospora clock protein FRQ. *Mol Cell* **52**, 832-843 (2013).
23. L. Lauinger, A. Diernfellner, S. Falk, M. Brunner, The RNA helicase FRH is an ATP-dependent regulator of CK1a in the circadian clock of Neurospora crassa. *Nat Commun* **5**, 3598 (2014).
24. Y. Liu, J. Loros, J. C. Dunlap, Phosphorylation of the Neurospora clock protein FREQUENCY determines its degradation rate and strongly influences the period length of the circadian clock. *Proc Natl Acad Sci U S A* **97**, 234-239 (2000).
25. N. Y. Garceau, Y. Liu, J. J. Loros, J. C. Dunlap, Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. *Cell* **89**, 469-476 (1997).
26. C. T. Tang *et al.*, Setting the pace of the Neurospora circadian clock by multiple independent FRQ phosphorylation events. *Proc Natl Acad Sci U S A* **106**, 10722-10727 (2009).
27. B. D. Aronson, K. A. Johnson, J. J. Loros, J. C. Dunlap, Negative feedback defining a circadian clock: autoregulation of the clock gene frequency. *Science* **263**, 1578-1584 (1994).
28. Q. He *et al.*, FWD1-mediated degradation of FREQUENCY in Neurospora establishes a conserved mechanism for circadian clock regulation. *EMBO J* **22**, 4421-4430 (2003).
29. T. K. Darlington *et al.*, Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. *Science* **280**, 1599-1603 (1998).
30. K. Kume *et al.*, mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell* **98**, 193-205 (1999).
31. T. K. Sato *et al.*, Feedback repression is required for mammalian circadian clock function. *Nat Genet* **38**, 312-319 (2006).
32. H. Tei *et al.*, Circadian oscillation of a mammalian homologue of the Drosophila period gene. *Nature* **389**, 512-516 (1997).
33. M. J. Zylka, L. P. Shearman, D. R. Weaver, S. M. Reppert, Three period homologs in mammals: differential light responses in the suprachiasmatic circadian clock and oscillating transcripts outside of brain. *Neuron* **20**, 1103-1110 (1998).
34. C. L. Partch, C. B. Green, J. S. Takahashi, Molecular architecture of the mammalian

circadian clock. *Trends Cell Biol* **24**, 90-99 (2014).

35. J. P. Etchegaray *et al.*, Casein kinase 1 delta regulates the pace of the mammalian circadian clock. *Mol Cell Biol* **29**, 3853-3866 (2009).
36. R. P. Aryal *et al.*, Macromolecular Assemblies of the Mammalian Circadian Clock. *Mol Cell* **67**, 770-782 e776 (2017).
37. K. Padmanabhan, M. S. Robles, T. Westerling, C. J. Weitz, Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex. *Science* **337**, 599-602 (2012).
38. H. A. Duong, M. S. Robles, D. Knutti, C. J. Weitz, A molecular mechanism for circadian clock negative feedback. *Science* **332**, 1436-1439 (2011).
39. H. A. Duong, C. J. Weitz, Temporal orchestration of repressive chromatin modifiers by circadian clock Period complexes. *Nature structural & molecular biology* **21**, 126-132 (2014).
40. S. A. Brown *et al.*, PERIOD1-associated proteins modulate the negative limb of the mammalian circadian oscillator. *Science* **308**, 693-696 (2005).
41. T. K. Sato *et al.*, A functional genomics strategy reveals Rora as a component of the mammalian circadian clock. *Neuron* **43**, 527-537 (2004).
42. N. Preitner *et al.*, The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251-260 (2002).
43. Y. Oishi *et al.*, Bmal1 regulates inflammatory responses in macrophages by modulating enhancer RNA transcription. *Sci Rep* **7**, 7086 (2017).
44. A. Bugge *et al.*, Rev-erbalpha and Rev-erbbeta coordinately protect the circadian clock and normal metabolic function. *Genes Dev* **26**, 657-667 (2012).
45. C. G. Parks *et al.*, Employment and work schedule are related to telomere length in women. *Occup Environ Med* **68**, 582-589 (2011).
46. A. A. Prather *et al.*, Shorter leukocyte telomere length in midlife women with poor sleep quality. *J Aging Res* **2011**, 721390 (2011).
47. G. Liang *et al.*, Associations between rotating night shifts, sleep duration, and telomere length in women. *PLoS One* **6**, e23462 (2011).
48. S. T. Chen *et al.*, Deregulated expression of the PER1, PER2 and PER3 genes in breast cancers. *Carcinogenesis* **26**, 1241-1246 (2005).
49. S. Gery *et al.*, The circadian gene per1 plays an important role in cell growth and DNA damage control in human cancer cells. *Mol Cell* **22**, 375-382 (2006).
50. C. H. Jung *et al.*, Bmal1 suppresses cancer cell invasion by blocking the phosphoinositide 3-kinase-Akt-MMP-2 signaling pathway. *Oncol Rep* **29**, 2109-2113

(2013).

51. F. K. Stephan, I. Zucker, Circadian rhythms in drinking behavior and locomotor activity of rats are eliminated by hypothalamic lesions. *Proc Natl Acad Sci U S A* **69**, 1583-1586 (1972).
52. E. D. Buhr, J. S. Takahashi, Molecular components of the Mammalian circadian clock. *Handb Exp Pharmacol*, 3-27 (2013).
53. F. Damiola *et al.*, Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes Dev* **14**, 2950-2961 (2000).
54. K. A. Stokkan, S. Yamazaki, H. Tei, Y. Sakaki, M. Menaker, Entrainment of the circadian clock in the liver by feeding. *Science* **291**, 490-493 (2001).
55. E. Challet, Minireview: Entrainment of the suprachiasmatic clockwork in diurnal and nocturnal mammals. *Endocrinology* **148**, 5648-5655 (2007).
56. D. Whitmore, N. S. Foulkes, U. Strahle, P. Sassone-Corsi, Zebrafish Clock rhythmic expression reveals independent peripheral circadian oscillators. *Nat Neurosci* **1**, 701-707 (1998).
57. C. Crosio, N. Cermakian, C. D. Allis, P. Sassone-Corsi, Light induces chromatin modification in cells of the mammalian circadian clock. *Nat Neurosci* **3**, 1241-1247 (2000).
58. R. D. Kornberg, Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-871 (1974).
59. A. Babu, R. S. Verma, Chromosome structure: euchromatin and heterochromatin. *Int Rev Cytol* **108**, 1-60 (1987).
60. T. Zhang, S. Cooper, N. Brockdorff, The interplay of histone modifications - writers that read. *EMBO Rep* **16**, 1467-1481 (2015).
61. S. I. Grewal, S. Jia, Heterochromatin revisited. *Nat Rev Genet* **8**, 35-46 (2007).
62. P. Trojer, D. Reinberg, Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* **28**, 1-13 (2007).
63. P. Bernard, R. Allshire, Centromeres become unstuck without heterochromatin. *Trends Cell Biol* **12**, 419-424 (2002).
64. S. Schoeftner, M. A. Blasco, A 'higher order' of telomere regulation: telomere heterochromatin and telomeric RNAs. *EMBO J* **28**, 2323-2336 (2009).
65. V. Dror, F. Winston, The Swi/Snf chromatin remodeling complex is required for ribosomal DNA and telomeric silencing in Saccharomyces cerevisiae. *Mol Cell Biol* **24**, 8227-8235 (2004).
66. M. Vignali, A. H. Hassan, K. E. Neely, J. L. Workman, ATP-dependent chromatin-remodeling complexes. *Mol Cell Biol* **20**, 1899-1910 (2000).

67. F. Ma, C. Y. Zhang, Histone modifying enzymes: novel disease biomarkers and assay development. *Expert Rev Mol Diagn* **16**, 297-306 (2016).
68. R. Marmorstein, R. C. Trievel, Histone modifying enzymes: structures, mechanisms, and specificities. *Biochim Biophys Acta* **1789**, 58-68 (2009).
69. V. G. Allfrey, R. Faulkner, A. E. Mirsky, Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* **51**, 786-794 (1964).
70. M. J. Bottomley, Structures of protein domains that create or recognize histone modifications. *EMBO Rep* **5**, 464-469 (2004).
71. M. A. Dawson, T. Kouzarides, Cancer epigenetics: from mechanism to therapy. *Cell* **150**, 12-27 (2012).
72. P. Chi, C. D. Allis, G. G. Wang, Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nat Rev Cancer* **10**, 457-469 (2010).
73. P. Tessarz, T. Kouzarides, Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol* **15**, 703-708 (2014).
74. A. J. Ruthenburg, H. Li, D. J. Patel, C. D. Allis, Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8**, 983-994 (2007).
75. J. L. Ronan, W. Wu, G. R. Crabtree, From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet* **14**, 347-359 (2013).
76. C. Lu *et al.*, Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* **352**, 844-849 (2016).
77. P. W. Lewis *et al.*, Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* **340**, 857-861 (2013).
78. C. Vollmers *et al.*, Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell Metab* **16**, 833-845 (2012).
79. L. Aguilar-Arnal *et al.*, Cycles in spatial and temporal chromosomal organization driven by the circadian clock. *Nat Struct Mol Biol* **20**, 1206-1213 (2013).
80. J. P. Etchegaray, C. Lee, P. A. Wade, S. M. Reppert, Rhythmic histone acetylation underlies transcription in the mammalian circadian clock. *Nature* **421**, 177-182 (2003).
81. N. Koike *et al.*, Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338**, 349-354 (2012).
82. A. M. Curtis *et al.*, Histone acetyltransferase-dependent chromatin remodeling and the vascular clock. *J Biol Chem* **279**, 7091-7097 (2004).
83. J. P. Etchegaray *et al.*, The polycomb group protein EZH2 is required for mammalian circadian clock function. *J Biol Chem* **281**, 21209-21215 (2006).
84. C. E. Ruesch *et al.*, The histone H3 lysine 9 methyltransferase DIM-5 modifies chromatin at frequency and represses light-activated gene expression. *G3 (Bethesda)* **5**,

93-101 (2014).

85. Y. Qiu *et al.*, Eosinophils and type 2 cytokine signaling in macrophages orchestrate development of functional beige fat. *Cell* **157**, 1292-1308 (2014).
86. Q. Tang *et al.*, Circadian Clock Gene Bmal1 Inhibits Tumorigenesis and Increases Paclitaxel Sensitivity in Tongue Squamous Cell Carcinoma. *Cancer Res* **77**, 532-544 (2017).
87. Z. Li *et al.*, The polycomb group protein EZH2 is a novel therapeutic target in tongue cancer. *Oncotarget* **4**, 2532-2549 (2013).
88. C. Wang *et al.*, EZH2 Mediates epigenetic silencing of neuroblastoma suppressor genes CASZ1, CLU, RUNX3, and NGFR. *Cancer Res* **72**, 315-324 (2012).
89. Y. Zhong, Q. Ye, C. Chen, M. Wang, H. Wang, Ezh2 promotes clock function and hematopoiesis independent of histone methyltransferase activity in zebrafish. *Nucleic Acids Res* **46**, 3382-3399 (2018).
90. G. Caretti, M. Di Padova, B. Micales, G. E. Lyons, V. Sartorelli, The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev* **18**, 2627-2638 (2004).
91. W. J. Belden, J. J. Loros, J. C. Dunlap, Execution of the circadian negative feedback loop in Neurospora requires the ATP-dependent chromatin-remodeling enzyme CLOCKSWITCH. *Mol Cell* **25**, 587-600 (2007).
92. W. J. Belden, Z. A. Lewis, E. U. Selker, J. J. Loros, J. C. Dunlap, CHD1 remodels chromatin and influences transient DNA methylation at the clock gene frequency. *PLoS Genet* **7**, e1002166 (2011).
93. B. Wang, A. N. Kettenbach, S. A. Gerber, J. J. Loros, J. C. Dunlap, Neurospora WC-1 recruits SWI/SNF to remodel frequency and initiate a circadian cycle. *PLoS Genet* **10**, e1004599 (2014).
94. J. Y. Kim, P. B. Kwak, C. J. Weitz, Specificity in circadian clock feedback from targeted reconstitution of the NuRD corepressor. *Mol Cell* **56**, 738-748 (2014).
95. A. Verdel *et al.*, RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672-676 (2004).
96. D. V. Irvine *et al.*, Argonaute slicing is required for heterochromatic silencing and spreading. *Science* **313**, 1134-1137 (2006).
97. N. Li, T. M. Joska, C. E. Ruesch, S. J. Coster, W. J. Belden, The frequency natural antisense transcript first promotes, then represses, frequency gene expression via facultative heterochromatin. *Proc Natl Acad Sci U S A* **112**, 4357-4362 (2015).
98. Y. Dang, L. Li, W. Guo, Z. Xue, Y. Liu, Convergent transcription induces dynamic DNA methylation at disiRNA loci. *PLoS Genet* **9**, e1003761 (2013).
99. H. C. Lee *et al.*, Diverse pathways generate microRNA-like RNAs and Dicer-

independent small interfering RNAs in fungi. *Mol Cell* **38**, 803-814 (2010).

100. Z. Xue *et al.*, Transcriptional interference by antisense RNA is required for circadian clock function. *Nature* **514**, 650-653 (2014).
101. S. Panda *et al.*, Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* **109**, 307-320 (2002).
102. A. I. Su *et al.*, Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A* **99**, 4465-4470 (2002).
103. H. R. Ueda *et al.*, System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nat Genet* **37**, 187-192 (2005).
104. S. L. Coon *et al.*, Circadian changes in long noncoding RNAs in the pineal gland. *Proc Natl Acad Sci U S A* **109**, 13319-13324 (2012).
105. T. Niccoli, L. Partridge, Ageing as a risk factor for disease. *Curr Biol* **22**, R741-752 (2012).
106. S. I. Rattan, Theories of biological aging: genes, proteins, and free radicals. *Free Radic Res* **40**, 1230-1238 (2006).
107. P. Carninci *et al.*, The transcriptional landscape of the mammalian genome. *Science* **309**, 1559-1563 (2005).
108. M. E. Dinger, K. C. Pang, T. R. Mercer, J. S. Mattick, Differentiating protein-coding and noncoding RNA: challenges and ambiguities. *PLoS Comput Biol* **4**, e1000176 (2008).
109. M. C. Frith *et al.*, Discrimination of non-protein-coding transcripts from protein-coding mRNA. *RNA Biol* **3**, 40-48 (2006).
110. J. M. Perkel, Visiting "noncodarnia". *Biotechniques* **54**, 301, 303-304 (2013).
111. P. Carninci, Y. Hayashizaki, Noncoding RNA transcription beyond annotated genes. *Curr Opin Genet Dev* **17**, 139-144 (2007).
112. E. P. Consortium *et al.*, Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799-816 (2007).
113. P. Kapranov *et al.*, RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**, 1484-1488 (2007).
114. M. N. Cabili *et al.*, Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* **25**, 1915-1927 (2011).
115. M. Guttman *et al.*, Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223-227 (2009).
116. C. Kutter *et al.*, Rapid turnover of long noncoding RNAs and the evolution of gene expression. *PLoS Genet* **8**, e1002841 (2012).
117. N. Mukherjee *et al.*, Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat Struct Mol Biol* **24**, 86-96 (2017).
118. I. Ulitsky, A. Shkumatava, C. H. Jan, H. Sive, D. P. Bartel, Conserved function of

lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537-1550 (2011).

119. T. Gutschner *et al.*, The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* **73**, 1180-1189 (2013).
120. O. Wapinski, H. Y. Chang, Long noncoding RNAs and human disease. *Trends Cell Biol* **21**, 354-361 (2011).
121. E. Pasmant *et al.*, Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res* **67**, 3963-3969 (2007).
122. E. Pasmant *et al.*, Role of noncoding RNA ANRIL in genesis of plexiform neurofibromas in neurofibromatosis type 1. *J Natl Cancer Inst* **103**, 1713-1722 (2011).
123. I. Iacobucci *et al.*, A polymorphism in the chromosome 9p21 ANRIL locus is associated to Philadelphia positive acute lymphoblastic leukemia. *Leuk Res* **35**, 1052-1059 (2011).
124. S. Tu, G. C. Yuan, Z. Shao, The PRC2-binding long non-coding RNAs in human and mouse genomes are associated with predictive sequence features. *Sci Rep* **7**, 41669 (2017).
125. C. P. Ponting, P. L. Oliver, W. Reik, Evolution and functions of long noncoding RNAs. *Cell* **136**, 629-641 (2009).
126. G. Storz, An expanding universe of noncoding RNAs. *Science* **296**, 1260-1263 (2002).
127. D. Battogtokh, S. Kojima, J. J. Tyson, Modeling the interactions of sense and antisense Period transcripts in the mammalian circadian clock network. *PLoS Comput Biol* **14**, e1005957 (2018).
128. N. Barzilai, D. M. Huffman, R. H. Muzumdar, A. Bartke, The critical role of metabolic pathways in aging. *Diabetes* **61**, 1315-1322 (2012).
129. L. Hayflick, Biological aging is no longer an unsolved problem. *Ann N Y Acad Sci* **1100**, 1-13 (2007).
130. T. B. Kirkwood, D. P. Shanley, The connections between general and reproductive senescence and the evolutionary basis of menopause. *Ann N Y Acad Sci* **1204**, 21-29 (2010).
131. K. Koh, J. M. Evans, J. C. Hendricks, A. Sehgal, A Drosophila model for age-associated changes in sleep:wake cycles. *Proc Natl Acad Sci U S A* **103**, 13843-13847 (2006).
132. R. V. Kondratov, A. A. Kondratova, V. Y. Gorbacheva, O. V. Vykhovanets, M. P. Antoch, Early aging and age-related pathologies in mice deficient in BMAL1, the core componentof the circadian clock. *Genes Dev* **20**, 1868-1873 (2006).
133. N. L. Nadon, Exploiting the rodent model for studies on the pharmacology of lifespan extension. *Aging Cell* **5**, 9-15 (2006).

134. Y. V. Dubrovsky, W. E. Samsa, R. V. Kondratov, Deficiency of circadian protein CLOCK reduces lifespan and increases age-related cataract development in mice. *Aging (Albany NY)* **2**, 936-944 (2010).
135. J. P. Debruyne *et al.*, A clock shock: mouse CLOCK is not required for circadian oscillator function. *Neuron* **50**, 465-477 (2006).
136. J. P. DeBruyne, D. R. Weaver, S. M. Reppert, CLOCK and NPAS2 have overlapping roles in the suprachiasmatic circadian clock. *Nat Neurosci* **10**, 543-545 (2007).
137. G. Solanas *et al.*, Aged Stem Cells Reprogram Their Daily Rhythmic Functions to Adapt to Stress. *Cell* **170**, 678-692 e620 (2017).
138. S. Sato *et al.*, Circadian Reprogramming in the Liver Identifies Metabolic Pathways of Aging. *Cell* **170**, 664-677 e611 (2017).
139. R. C. Kuintzle *et al.*, Circadian deep sequencing reveals stress-response genes that adopt robust rhythmic expression during aging. *Nat Commun* **8**, 14529 (2017).
140. K. L. Eckel-Mahan *et al.*, Reprogramming of the circadian clock by nutritional challenge. *Cell* **155**, 1464-1478 (2013).
141. T. Vulliamy *et al.*, The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* **413**, 432-435 (2001).
142. S. B. Cohen *et al.*, Protein composition of catalytically active human telomerase from immortal cells. *Science* **315**, 1850-1853 (2007).
143. C. B. Harley, B. Villeponteau, Telomeres and telomerase in aging and cancer. *Curr Opin Genet Dev* **5**, 249-255 (1995).
144. C. W. Greider, E. H. Blackburn, A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature* **337**, 331-337 (1989).
145. C. W. Greider, Telomeres do D-loop-T-loop. *Cell* **97**, 419-422 (1999).
146. N. Arnoult, J. Karlseder, Complex interactions between the DNA-damage response and mammalian telomeres. *Nat Struct Mol Biol* **22**, 859-866 (2015).
147. C. J. Webb, Y. Wu, V. A. Zakian, DNA repair at telomeres: keeping the ends intact. *Cold Spring Harb Perspect Biol* **5**, (2013).
148. D. Fu, K. Collins, Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. *Mol Cell* **28**, 773-785 (2007).
149. C. W. Greider, E. H. Blackburn, Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* **43**, 405-413 (1985).
150. N. R. Forsyth, W. E. Wright, J. W. Shay, Telomerase and differentiation in multicellular organisms: turn it off, turn it on, and turn it off again. *Differentiation* **69**, 188-197 (2002).
151. L. Hayflick, The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp Cell Res* **37**,

614-636 (1965).

152. W. D. Chen *et al.*, The circadian rhythm controls telomeres and telomerase activity. *Biochem Biophys Res Commun* **451**, 408-414 (2014).
153. X. S. Wang, M. E. Armstrong, B. J. Cairns, T. J. Key, R. C. Travis, Shift work and chronic disease: the epidemiological evidence. *Occup Med (Lond)* **61**, 78-89 (2011).
154. M. Z. Levy, R. C. Allsopp, A. B. Futcher, C. W. Greider, C. B. Harley, Telomere end-replication problem and cell aging. *J Mol Biol* **225**, 951-960 (1992).
155. C. W. Greider, Telomeres and senescence: the history, the experiment, the future. *Curr Biol* **8**, R178-181 (1998).
156. T. de Lange, How telomeres solve the end-protection problem. *Science* **326**, 948-952 (2009).
157. T. de Lange, How shelterin solves the telomere end-protection problem. *Cold Spring Harb Symp Quant Biol* **75**, 167-177 (2010).
158. T. de Lange, Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* **19**, 2100-2110 (2005).
159. W. Palm, T. de Lange, How shelterin protects mammalian telomeres. *Annu Rev Genet* **42**, 301-334 (2008).
160. A. Galati, E. Micheli, S. Cacchione, Chromatin structure in telomere dynamics. *Front Oncol* **3**, 46 (2013).
161. I. Roig *et al.*, Female-specific features of recombinational double-stranded DNA repair in relation to synapsis and telomere dynamics in human oocytes. *Chromosoma* **113**, 22-33 (2004).
162. B. Perrini *et al.*, HP1 controls telomere capping, telomere elongation, and telomere silencing by two different mechanisms in Drosophila. *Mol Cell* **15**, 467-476 (2004).
163. C. M. Azzalin, P. Reichenbach, L. Khoriauli, E. Giulotto, J. Lingner, Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798-801 (2007).
164. H. P. Chu *et al.*, TERRA RNA Antagonizes ATRX and Protects Telomeres. *Cell* **170**, 86-101 e116 (2017).
165. I. Lopez de Silanes *et al.*, Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. *Nat Commun* **5**, 4723 (2014).
166. Z. Deng, J. Norseen, A. Wiedmer, H. Riethman, P. M. Lieberman, TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres. *Mol Cell* **35**, 403-413 (2009).
167. M. Graf *et al.*, Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* **170**, 72-85 e14 (2017).
168. R. L. Flynn *et al.*, Alternative lengthening of telomeres renders cancer cells

hypersensitive to ATR inhibitors. *Science* **347**, 273-277 (2015).

169. S. Sampl *et al.*, Expression of telomeres in astrocytoma WHO grade 2 to 4: TERRA level correlates with telomere length, telomerase activity, and advanced clinical grade. *Transl Oncol* **5**, 56-65 (2012).
170. A. Porro, S. Feuerhahn, P. Reichenbach, J. Lingner, Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. *Mol Cell Biol* **30**, 4808-4817 (2010).
171. E. Cusanelli, P. Chartrand, Telomeric repeat-containing RNA TERRA: a noncoding RNA connecting telomere biology to genome integrity. *Front Genet* **6**, 143 (2015).
172. C. Wang, L. Zhao, S. Lu, Role of TERRA in the regulation of telomere length. *Int J Biol Sci* **11**, 316-323 (2015).
173. K. M. Smith *et al.*, Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for neurospora white collar complex. *Eukaryot Cell* **9**, 1549-1556 (2010).
174. C. Wu *et al.*, Characterization of chromosome ends in the filamentous fungus Neurospora crassa. *Genetics* **181**, 1129-1145 (2009).
175. C. B. Harley, A. B. Futcher, C. W. Greider, Telomeres shorten during ageing of human fibroblasts. *Nature* **345**, 458-460 (1990).
176. G. Rey *et al.*, Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. *PLoS Biol* **9**, e1000595 (2011).
177. C. Wright, G. Herbert, R. Pilkington, M. Callaghan, S. McClean, Real-time PCR method for the quantification of Burkholderia cepacia complex attached to lung epithelial cells and inhibition of that attachment. *Lett Appl Microbiol* **50**, 500-506 (2010).
178. J. A. Ripperger, U. Schibler, Rhythmic CLOCK-BMAL1 binding to multiple E-box motifs drives circadian Dbp transcription and chromatin transitions. *Nat Genet* **38**, 369-374 (2006).
179. J. S. Menet, J. Rodriguez, K. C. Abruzzi, M. Rosbash, Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *Elife* **1**, e00011 (2012).
180. M. K. Bunger *et al.*, Mop3 is an essential component of the master circadian pacemaker in mammals. *Cell* **103**, 1009-1017 (2000).
181. H. Raduwan, A. L. Isola, W. J. Belden, Methylation of histone H3 on lysine 4 by the lysine methyltransferase SET1 protein is needed for normal clock gene expression. *J Biol Chem* **288**, 8380-8390 (2013).
182. R. M. Cawthon, Telomere measurement by quantitative PCR. *Nucleic Acids Res* **30**, e47 (2002).
183. J. T. Robinson *et al.*, Integrative genomics viewer. *Nat Biotechnol* **29**, 24-

26 (2011).

184. M. Ruger, F. A. Scheer, Effects of circadian disruption on the cardiometabolic system. *Rev Endocr Metab Disord* **10**, 245-260 (2009).

185. M. K. Bunger *et al.*, Progressive arthropathy in mice with a targeted disruption of the Mop3/Bmal-1 locus. *Genesis* **41**, 122-132 (2005).

186. E. L. McDearmon *et al.*, Dissecting the functions of the mammalian clock protein BMAL1 by tissue-specific rescue in mice. *Science* **314**, 1304-1308 (2006).

187. A. Jenwitheesuk, C. Nopparat, S. Mukda, P. Wongchitrat, P. Govitrapong, Melatonin regulates aging and neurodegeneration through energy metabolism, epigenetics, autophagy and circadian rhythm pathways. *Int J Mol Sci* **15**, 16848-16884 (2014).

188. Y. Qu *et al.*, Telomerase reconstitution contributes to resetting of circadian rhythm in fibroblasts. *Mol Cell Biochem* **313**, 11-18 (2008).

189. F. Cao *et al.*, Dicer independent small RNAs associate with telomeric heterochromatin. *RNA* **15**, 1274-1281 (2009).

190. C. S. Pittendrigh, D. H. Minis, Circadian systems: longevity as a function of circadian resonance in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **69**, 1537-1539 (1972).

191. C. S. Pittendrigh, S. Daan, Circadian oscillations in rodents: a systematic increase of their frequency with age. *Science* **186**, 548-550 (1974).

192. A. J. Davidson *et al.*, Chronic jet-lag increases mortality in aged mice. *Curr Biol* **16**, R914-916 (2006).

193. H. Reinke, G. Asher, Circadian Clock Control of Liver Metabolic Functions. *Gastroenterology* **150**, 574-580 (2016).

194. J. Bass, J. S. Takahashi, Circadian integration of metabolism and energetics. *Science* **330**, 1349-1354 (2010).

195. C. B. Green, J. S. Takahashi, J. Bass, The meter of metabolism. *Cell* **134**, 728-742 (2008).

196. D. Landgraf *et al.*, Genetic Disruption of Circadian Rhythms in the Suprachiasmatic Nucleus Causes Helplessness, Behavioral Despair, and Anxiety-like Behavior in Mice. *Biol Psychiatry* **80**, 827-835 (2016).

197. M. E. Young, M. S. Bray, Potential role for peripheral circadian clock dyssynchrony in the pathogenesis of cardiovascular dysfunction. *Sleep Med* **8**, 656-667 (2007).

198. S. Masri, K. Kinouchi, P. Sassone-Corsi, Circadian clocks, epigenetics, and cancer. *Curr Opin Oncol* **27**, 50-56 (2015).

199. M. A. Hofman, D. F. Swaab, Living by the clock: the circadian pacemaker in older people. *Ageing Res Rev* **5**, 33-51 (2006).

200. D. E. Kolker *et al.*, Aging alters circadian and light-induced expression of clock genes in golden hamsters. *J Biol Rhythms* **18**, 159-169 (2003).

201. M. P. Antoch *et al.*, Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. *Cell* **89**, 655-667 (1997).
202. D. P. King *et al.*, Positional cloning of the mouse circadian clock gene. *Cell* **89**, 641-653 (1997).
203. M. H. Vitaterna *et al.*, Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science* **264**, 719-725 (1994).
204. J. B. Hogenesch, Y. Z. Gu, S. Jain, C. A. Bradfield, The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors. *Proc Natl Acad Sci U S A* **95**, 5474-5479 (1998).
205. B. Zheng *et al.*, The mPer2 gene encodes a functional component of the mammalian circadian clock. *Nature* **400**, 169-173 (1999).
206. N. Cermakian, L. Monaco, M. P. Pando, A. Dierich, P. Sassone-Corsi, Altered behavioral rhythms and clock gene expression in mice with a targeted mutation in the Period1 gene. *EMBO J* **20**, 3967-3974 (2001).
207. L. P. Shearman, X. Jin, C. Lee, S. M. Reppert, D. R. Weaver, Targeted disruption of the mPer3 gene: subtle effects on circadian clock function. *Mol Cell Biol* **20**, 6269-6275 (2000).
208. R. J. Thresher *et al.*, Role of mouse cryptochrome blue-light photoreceptor in circadian photoresponses. *Science* **282**, 1490-1494 (1998).
209. G. T. van der Horst *et al.*, Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature* **398**, 627-630 (1999).
210. R. Zhang, N. F. Lahens, H. I. Ballance, M. E. Hughes, J. B. Hogenesch, A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111**, 16219-16224 (2014).
211. C. Kramer, J. J. Loros, J. C. Dunlap, S. K. Crosthwaite, Role for antisense RNA in regulating circadian clock function in Neurospora crassa. *Nature* **421**, 948-952 (2003).
212. J. M. Hurley *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 (2014).
213. K. F. Storch *et al.*, Extensive and divergent circadian gene expression in liver and heart. *Nature* **417**, 78-83 (2002).
214. J. S. Takahashi, Transcriptional architecture of the mammalian circadian clock. *Nat Rev Genet* **18**, 164-179 (2017).
215. W. J. Belden, Z. A. Lewis, E. U. Selker, J. J. Loros, J. C. Dunlap, CHD1 Remodels Chromatin and Influences Transient DNA Methylation at the Clock Gene *frequency*. *PLoS genetics* **7**, e1002166 (2011).
216. P. Taylor, P. E. Hardin, Rhythmic E-box binding by CLK-CYC controls daily cycles in per

and tim transcription and chromatin modifications. *Mol Cell Biol* **28**, 4642-4652 (2008).

217. M. Magistri, M. A. Faghihi, G. St Laurent, 3rd, C. Wahlestedt, Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends Genet* **28**, 389-396 (2012).
218. M. Guttman *et al.*, lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**, 295-300 (2011).
219. C. Sidler, O. Kovalchuk, I. Kovalchuk, Epigenetic Regulation of Cellular Senescence and Aging. *Front Genet* **8**, 138 (2017).
220. A. Tsurumi, W. X. Li, Global heterochromatin loss: a unifying theory of aging? *Epigenetics* **7**, 680-688 (2012).
221. W. Zhang *et al.*, Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science* **348**, 1160-1163 (2015).
222. C. Trapnell *et al.*, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562-578 (2012).
223. M. Pertea, D. Kim, G. M. Pertea, J. T. Leek, S. L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* **11**, 1650-1667 (2016).
224. A. Pauli *et al.*, Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res* **22**, 577-591 (2012).
225. A. Li, J. Zhang, Z. Zhou, PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinformatics* **15**, 311 (2014).
226. B. Zhang, S. Kirov, J. Snoddy, WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research* **33**, W741-748 (2005).
227. J. Wang, D. Duncan, Z. Shi, B. Zhang, WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic acids research* **41**, W77-83 (2013).
228. G. Yu, L. G. Wang, G. R. Yan, Q. Y. He, DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **31**, 608-609 (2015).
229. R. Tacutu *et al.*, Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. *Nucleic acids research* **41**, D1027-1033 (2013).
230. D. H. Cribbs *et al.*, Extensive innate immune gene activation accompanies brain aging, increasing vulnerability to cognitive decline and neurodegeneration: a microarray study. *J Neuroinflammation* **9**, 179 (2012).
231. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

232. R. Stark, G. Brown. (2011).
233. T. Gupta, M. C. Mullins, Dissection of organs from the adult zebrafish. *J Vis Exp*, (2010).
234. C. Trapnell *et al.*, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511-515 (2010).
235. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
236. G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).
237. R. Martienssen, D. Moazed, RNAi and heterochromatin assembly. *Cold Spring Harb Perspect Biol* **7**, a019323 (2015).
238. S. V. Dindot, R. Person, M. Strivens, R. Garcia, A. L. Beaudet, Epigenetic profiling at mouse imprinted gene clusters reveals novel epigenetic and genetic features at differentially methylated regions. *Genome Res* **19**, 1374-1383 (2009).
239. R. Bonasio, R. Shiekhattar, Regulation of transcription by long noncoding RNAs. *Annu Rev Genet* **48**, 433-455 (2014).
240. M. E. Hughes *et al.*, Guidelines for Genome-Scale Analysis of Biological Rhythms. *J Biol Rhythms* **32**, 380-393 (2017).
241. T. M. Joska, A. Mashruwala, J. M. Boyd, W. J. Belden, A universal cloning method based on yeast homologous recombination that is simple, efficient, and versatile. *J Microbiol Methods* **100**, 46-51 (2014).
242. L. B. Sullivan, J. H. Santos, N. S. Chandel, Mitochondria and telomeres: the promiscuous roles of TIN2. *Mol Cell* **47**, 823-824 (2012).
243. A. Neufeld-Cohen *et al.*, Circadian control of oscillations in mitochondrial rate-limiting enzymes and nutrient utilization by PERIOD proteins. *Proc Natl Acad Sci U S A* **113**, E1673-1682 (2016).
244. K. Suzuki, P. Bose, R. Y. Leong-Quong, D. J. Fujita, K. Riabowol, REAP: A two minute cell fractionation method. *BMC Res Notes* **3**, 294 (2010).