CROSSTALK BETWEEN LONG NON-CODING RNAS AND CIRCADIAN CHROMATIN

By

QIAOQIAO ZHU

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 Endocrinology and Animal Biosciences

Written under the direction of

William J. Belden

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2020

ABSTRACT OF THE DISSERTATION

Crosstalk between long non-coding RNAs and circadian chromatin

By QIAOQIAO ZHU

Dissertation Director:

William J. Belden, Ph.D.

The circadian rhythm is governed by transcriptional negative feedback facilitated by oscillating histone modifications and chromatin remodeling. The circadian rhythm is entrained by external zeitgebers (light and temperature) and are conserved in *Neurospora*, *Drosophila*, *zebrafish*, and mammals. The clock gene *frequency* in *Neurospora* and *Period2* in vertebrates have natural antisense transcripts (NATs) whose function is not understood. In this dissertation I examined the connection among the circadian clock, non-coding RNAs and heterochromatin formation on both on the genome-wide and locus-specific level using a multi-organism approach. I performed a genome-wide study, using RNA-seq and ChIP-seq to understand the role of H3 lysine 4 methyltransferase (KMT2/SET-1) and histone H3 lysine 9 methyltransferase (KMT1/DIM-5) in *Neurospora* to understand the role of 2 seemingly opposing modifications. Integrated analysis of RNA-seq and ChIP-seq showed crosstalk and redistributions between histone H3 lysine 4 tri-methylation (H3K4me3) and histone H3 lysine 9 tri-methylation (H3K9me3). I also examined how perturbing the expression of a diurnal lncRNA affected

downstream heterochromatin formation at the telomeres. The core circadian clock controls rhythms in *TERRA*, a long noncoding RNA that originates from telomeres and my research shows alcohol disrupts the diurnal rhythm in TERRA and heterochromatin at the telomere, which in theory makes telomeres more susceptible to DNA damage. I also examined the *Per2* NAT, *Per2AS* and found the diurnal rhythm in *Per2AS* is dependent on BMAL1. Using the ChIRP-MS, I identified *Per2AS*-interacting proteins. Specifically, I found hnRNP M interacts with *Per2A*S and hnRNP M is required to maintain the normal amplitude and period of *Per2.* Furthermore, I demonstrate that hnRNP M is necessary for H3K9me3 and H3K27me3 at *Per2*. These findings support a model where *Per2AS* may serve as scaffold for hnRNP M and other associated proteins that assist in heterochromatin formation at *Per2*. Collectively, this dissertation furthers our understanding of the circadian clock, non-coding RNAs, and circadian regulated facultative heterochromatin formation.

ACKNOWLEDGMENTS

First, I would like to extend my most sincere appreciation to my advisor, Dr. Belden, for his guidance and support through my graduate education. His passionate attitude towards science inspired me to keep motivated and overcome the difficulties with my research, as he always told me about science "If something was easy, someone already done it". Without his encouragement, I would not be brave enough to try new solutions and make my projects productive. I appreciate all his contributions of time, help, and ideas to make my journey of Ph.D. life memorable.

Besides my advisor, I would like to thank my committee members, Dr. Dipak K. Sarkar, Dr. Tracy G. Anthony and Dr. KiBum Lee. Thank you all for your kind support and great mentorship. You have all been extremely helpful and I appreciate the great advice you all provided during graduate school.

I would like to thank to all the collaborators of the dissertation projects. Thanks to previous lab members, Mukund Ramakrishnan and Jinhee Park for the contributions on the *Neurospora* project. Thanks to Dr. Joshua Gamsby' lab from USF help with the luciferase assay. Thanks to Dr. Tracy Anthony and Emily Mirek's help on the mouse experiments for our previous BMAL1 project. Thanks for Dr. Bagnell's lab and Dr. Amanda Jetzt for help with my cell culture experiments. Thanks for Nadia Rachdaoui's thoughtful comments on alcohol project. Thanks to Dr. Elizabeth Snyder's discussion with *Per2AS* project and kindly provided me all the resources whenever I needed from her lab. Also, I would like to thank China Scholarship Council (CSC#201406270116) for their support.

I would like to thank all the members in Belden lab. Specially, I would like to thank Jinhee Park for teaching me a lot on the research and teaching assistance experience. It has been so great to overlap with you for four years and have you as friend in graduate school to share all the pressure and joy in our research and everyday life.

I would like to thank all the faculty staff and graduate students of Animal Science department for my support. I would like to thank Dr. Cohick and Dr. Bagnell, for their kind support in the Endocrinology and Animal Science Biosciences program and meet this wonderful "family". Thank you Dr. Bagnell, for your kindness and constant support through my graduate studies. You are an excellent graduate student director and you have always been extremely helpful since the first day I joined the program till I finished my dissertation. Thanks to the TA and fellowship opportunity from animal science department to support me finishing my work. Thanks to Kathy Manger who always helped me with all sorts of academic questions. Thanks to Tory Gombosi, Damien Carroll-DeBose and Jeanette Mullins. For their kindly help and support.

Finally, I dedicate this work to my dearest family, especially my parents and my husband Xiaobing Zhang. Thanks for encourage me study abroad and for their endless love and all the emotional support to overcome all the difficulties in my graduate study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Figures

Supplemental Figures

GLOSSARY OF ABBREVIATIONS (A to Z)

adh	alcohol dehydrogenase
BMAL1	Brain and Muscle ARNT-Like 1
CATP	Clock ATPase
ccg	*clock-controlled genes*
CHD1	Chromodomain Helicase DNA-binding 1
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation with DNA sequencing
ChIRP	chromatin isolation by RNA purification
ChIRP-MS	Comprehensive identification of RNA-binding proteins by mass spectrometry
CLOCK	Circadian Locomotor Output Cycles Kaput
CRFH	circadian-regulated facultative heterochromatin
CRY	CRYPYOCHROME
CSW-1	Clockswitch
CTCF	CCCTC-binding factor
Dbp	*albumin D element-binding protein*
disiRNAs	Dicer-independent small interfering RNAs
dsRNA	double-stranded RNA
ERI-1	enhanced RNAi-1
eRNA	enhancer RNA

FAD	cysteinyl-flavin adenine dinucleotide
FRH	FRQ-interacting RNA Helicase
frq	*frequency*
H3K27me3	histone H3 lysine 27 trimethylation
H3K4me	histone H3 lysine 4 methylation
H3K9ac	Histone H3 lysine 9 acetylation
H3K9me	histone H3 lysine 9 methylation
H4K20me3	trimethylated histone H4 at lysine 20
HDAC3	histone deacetylase 3
HNF6	hepatocyte nuclear factor 6
hnRNP	Heterogenous nuclear ribonucleoprotein
HP1	Heterochromatin Protein 1
JARID1a	JumonjiC and ARID domain-containing histone lysine demethylase 1a
KAT	lysine acetyltransferase
KMT1/DIM-5	Histone H3 lysine 9 methyltransferase
KMT2/SET-1	Histone H3 lysine 4 methyltransferase
LADs	lamina-associated domains
lncRNA	long non-coding RNA
MEFs	mouse embryonic fibroblasts
miRNA	micro RNA

Nampt	*nicotinamide phosphoribosyl-transferase*
NAT	natural antisense transcript
NCoR	nuclear receptor co-repressor
NuRD	nucleosome remodeling and deacetylase complex
PARP1	PAR polymerase 1
PER	PERIOD
Per2AS	*Per2* antisense transcript
PolII	RNA polymerase II
PSF	polypyrimidine tract-binding protein-associated splicing factor
qrf	*frequency antisense transcript*
RdRp	RNA-dependent RNA polymerase
RITS	RNA-induced transcriptional silencing
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
siRNAs	small interfering RNAs
SIRT	Sirtuin
SWI/SNF	SWItch/Sucrose NonFermentable
TERC	*telomerase RNA component*
TERRA	*telomeric repeat-containing RNA*
TERT	telomerase reverse transcriptase

TES	transcriptional end site
tim	Timeless
TRF2	Telomeric Repeat Binding Factor 2
TSS	transcriptional start site
vvd	vivid
WCC	White Collar Complex
WT	wild-type
XIST	*X-inactive specific transcript*

CHAPTER 1

1 Chapter 1：Review of literature

(Chapter 1 published in *Journal of Molecular Biology,* 2020 Jan 16. pii: S0022-2836(20)30039-5. Review.PMID:31954735)

1.1 Abstract

Circadian rhythms are generated by transcriptional negative feedback loops and require histone modifications and chromatin remodeling to ensure appropriate timing and amplitude of clock gene expression. Circadian modifications to histones are important for transcriptional initiation and feedback inhibition serving as signaling platform for chromatin-remodeling enzymes. Current models indicate circadian-regulated facultative heterochromatin (CRFH) is a conserved mechanism at clock genes in *Neurospora*, *Drosophila*, and mice. CRFH consists of anti-phasic rhythms in activating and repressive modifications generating chromatin states that cycle between transcriptionally permissive and non-permissive. There are rhythms in histone H3 lysine 9 and 27 acetylation (H3K9ac and H3K27ac) and histone H3 lysine 4 methylation (H3K4me) during activation; while deacetylation, histone H3 lysine 9 methylation (H3K9me) and heterochromatin protein 1 (HP1) are hallmarks of repression. ATP-dependent chromatin-remodeling enzymes control accessibility, nucleosome positioning/occupancy and nuclear organization. In *Neurospora*, the rhythm in facultative heterochromatin is mediated by the *frequency* (*frq*) natural antisense transcript (NAT) *qrf.* While in mammals, histone deacetylases (HDACs), histone H3 lysine 9 methyltransferase (KMT1/SUV39) and components of Nucleosome Remodeling and Deacetylase (NuRD)

are part of the nuclear PERIOD complex (PER complex). Genomics efforts have found relationships among rhythmic chromatin modifications at *clock-controlled genes* (*ccg*) revealing circadian control of genome-wide chromatin states. There are also circadian clock-regulated lncRNAs with an emerging function that includes assisting in chromatin dynamics. The purpose of this literature review is to explore the connections among the circadian clock, chromatin remodeling, lncRNAs, and CRFH and how these impact rhythmicity, amplitude, period, and phase of circadian clock genes. **(Graphical abstract Figure 1-1)**

Keywords: Circadian regulation, chromatin remodeling, chromatin modifications, lncRNAs, aging and diseases

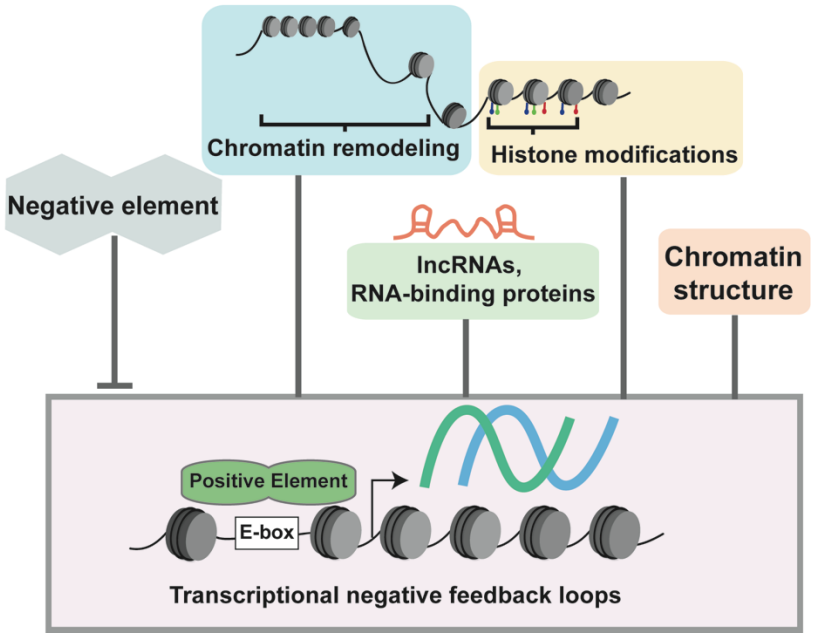


Figure 1-1 A subset of mechanisms involved in circadian clock regulation with a focus on chromatin-associated events.

The schematic depicts generalized factors required for proper circadian clock function that are the subject of this Chapter. (RNA, The red solid line; DNA, black solid line; nucleosomes, grey cylinders).

1.2 The Circadian Clock

The circadian clock is an anticipatory system synchronized to daily changes in light and temperature. The term circadian (Latin: circa = about; dies = day) was introduced by Franz Halberg in the 1950s and describes endogenous oscillations that occur with an approximately 24-hour period. The past decade and a half have revealed the circadian clock has profound influence on human health. Presumably, circadian-related diseases are caused in part by global mis-regulation of circadian-regulated gene expression, where approximately 40% of protein-coding genes are under clock control and participate in various biological functions, such as sleep-wake cycles, body

temperature regulation, blood pressure and hormone release [1]. The circadian clock also regulates behavioral and metabolic rhythms along with a host of molecular processes ensuring gene expression is appropriately timed based on cellular needs [2-6]. At its core, the circadian rhythm is generated by a cell-autonomous molecular oscillator arranged as a transcriptional negative feedback loop [7-9]. The core clock proteins lack strong sequence conservation in *Neurospora*, *Drosophila* and mammals, but the underlying mechanisms, including many of the accessory clock components have highly conserved functions supporting the notion that evolution has commandeered a unifying mechanism to enable 24-hr timed gene expression.

Neurospora crassa and *Drosophila melanogaster* are widely used non-vertebrate models used to study how circadian rhythms function. In fact, the first two clock genes, *Period (per)* and *frequency (frq)* were discovered in *Drosophila* and *Neurospora* [10] [11, 12]. Work on these to organisms were instrumental in describing the mechanisms of how cellular clock functions and laid the ground work for our understanding of the mammalian clock. The basic mechanism of circadian negative feedback in eukaryotes, metazoans and vertebrates consists of a dimeric transcriptional activator that drives expression of the negative elements, which feedback to inhibit their own expression [7]. In vertebrates, the master circadian clock is located in the suprachiasmatic nuclei (SCN) of the hypothalamus [13] , but there are also peripheral clocks in all cells types examined [14, 15].The dimeric transcriptional activator driving the positive arm of the cycle consists of CLOCK and BMAL1 in mammals [16-19], CLK and CYC in *Drosophila* [20, 21], and the White Collar Complex (WCC) in *Neurospora* [22] **(Figure 1-2A.)**. These activators drive expression of the *Period (Per1, Per2, Per3)* [23, 24] and *Cryptochrome*

(*Cry1, Cry2*) in mammals **(Figure 1-2B.)** [9, 25, 26], *dPer* and *Timeless* (*tim*) in *Drosophila* [27-29], and *frequency* (*frq*) in *Neurospora* [30]. RORα and REV-ERBα form a second feedback loop in vertebrates and have a positive and negative role, respectively, in regulating rhythms in *Bmal1* expression [31, 32]. In all 3 model systems, the clock proteins associate with kinases (in particular Casein Kinase I which is tightly associated with the negative elements), chromatin-remodeling and chromatin-modifying enzymes, and RNA-binding proteins [33-39] indicating key elements in feedback repression involve mechanistically conserved chromatin regulation and RNA metabolism.

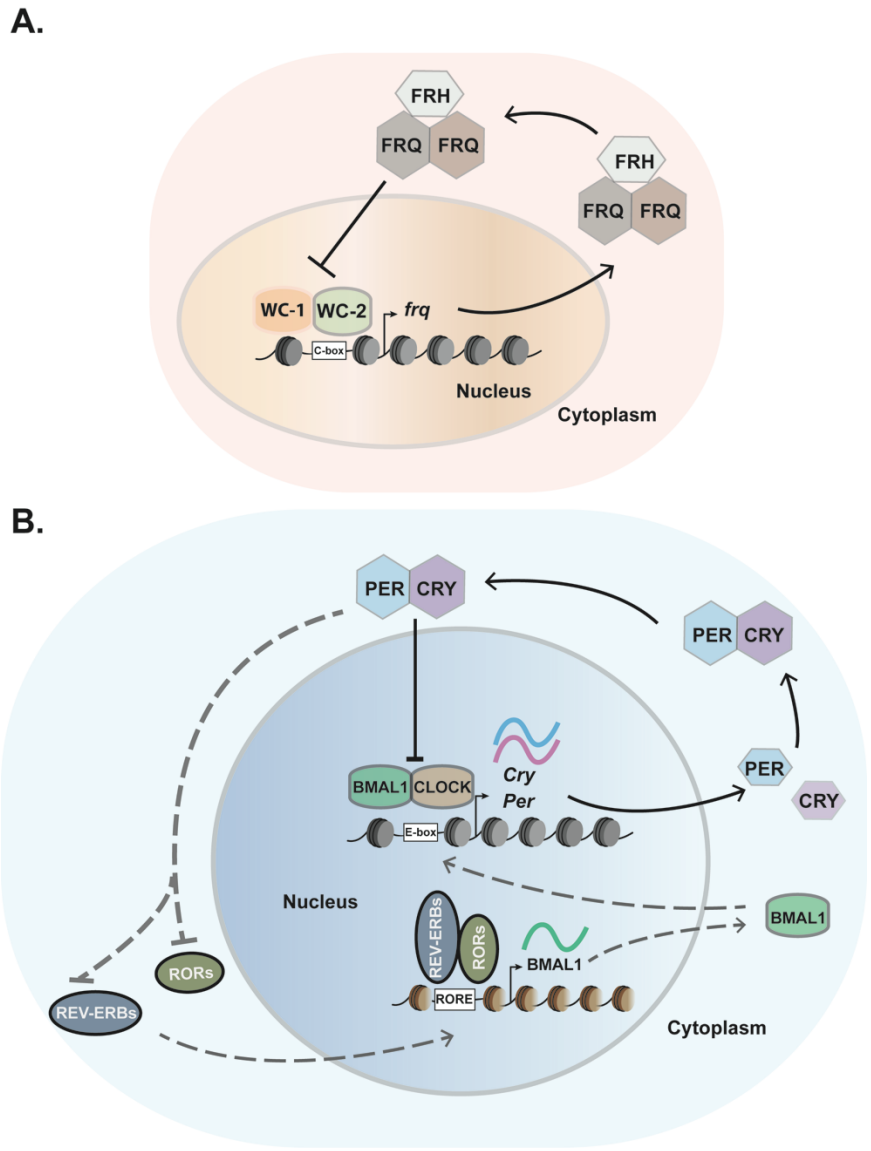


Figure 1-2 Simplified cartoon graphic of negative feedback controlling the circadian clock.

(A) The *Neurospora* transcriptional negative feedback loop. (FRQ, FREQUENCY; FRH, FRQ-interacting RNA Helicase) (B) The mammalian transcriptional feedback loop. (PER, PERIOD; CRY, CRYPYOCHROME; CLOCK, Circadian Locomotor Output Cycles Kaput; BMAL1, Brain and Muscle ARNT-Like 1)

Early studies on negative feedback focused on the direct interaction between the positive elements and the negative elements. PER:CRY associate with CLOCK:BMAL1, dPER:TIM associate with CLK:CYC, and FRQ:FRH associate with the WCC and all interactions are essential for feedback inhibition [40-47]. Also, posttranslational modifications, such as the phosphorylation is an important part of circadian rhythm regulation. There are a conserved set of kinases in all systems and members of the Casein kinase 1 (CK1) family play a predominantly role in most of the phosphorylation events [48]. FRQ get phosphates by casein kinases 1 and 2 (CK1a and CK2), Calcium/Calmodulin-dependent kinase (CAMK-1) and the *Neurospora* homolog of checkpoint kinase-2 (PRD-4) [49-51]. FRQ regulates WCC by changing WCC phosphorylation when they are complexed [52]. In *Drosophila*, DOUBLE-TIME(DBT), the casein kinase Iε promotes the turnover of PER/TIM complexes [53]. The casein kinase I (CKIε and CKIδ) phosphorylate the PER and regulate the stability of PER:CRY complexes in vertebrates [54-56]. The PER proteins, such as PER2, undergo rhythmic phosphorylation. In all three systems phosphorylation controls PER transport between the nucleus and cytoplasm, as well as protein turnover [57]. Current models on phosphorylation indicate it is critical to nuclear import and turnover causing electrostatic repulsion that exposes a degradation signal sequence [54, 58]. There are also global phosphorylation cycles where approximately one fourth of the quantified phosphorylated peptides display circadian oscillations indicating crosstalk between the circadian clock and metabolism [59]. Phosphorylation of PER:CRY in mammals, dPER:TIM in *Drosophila* and FRQ:FRH in *Neurospora* leads to ubiquitination and degradation by the proteasome [60, 61]. However, it appears the actual turnover of negative elements is

ancillary to clock function and instead, it is more a process of cellular recycling because rhythms requiring negative feedback persist without the turnover in *Neurospora* [62]. In other words, once the negative elements have become fully phosphorylated and have successfully guided all the timed events needed for feedback, they are cellular garbage and sent to the recycling center.

The timed phosphorylation by shared kinases led to the phosphorylation-dependent inactivation model where the negative elements guide kinases to post-translationally modify the activators [52, 63-71] and help titrate them off DNA [72, 73]. Guided inactivation via phosphorylation has many appealing features, foremost being that phosphates on the positive elements cause electrostatic repulsion with the phosphate backbone of DNA. However, it is far more complicated because WCC is still phosphorylated in the presumptive null frq^9, which means only some of the phosphorylation is dependent on FRQ, and there are over 100 phosphorylation sites in WCC. Plus arrhythmicity only occurs when all the phosphorylation sites were mutated in WC-1 or WC-2 or within subsets of sites in one, while the other was largely ablated of phosphorylation [74]. Thus, only a small combination of phosphorylation events, which must occur on both proteins, is required for feedback so most of the phosphorylation is either ancillary or there is redundancy with a select few critical sites. Likewise, in *Drosophila*, the majority of phosphorylation occurs after PER-dependent titration of CLK:CYC off the DNA [72]. Therefore, a simple model of phosphorylation-dependent inactivation only scratches the surface of the complex post-translational modifications that occur on the clock transcription factors. Plus, the circadian transcription factors

undergo regulated turnover [75-77] and there are a host of chromatin associated events for which phosphorylation is a possible que.

Complicating matters further, the central clock genes in *Neurospora* and mammals both have antisense transcripts that are expressed 180^{o} out-of-phase [78-80]. Although not fully proven, RNA-seq and ChIP-seq, along with other data indicate that the clock antisense transcripts are controlled by the clock transcription factors yet arise from independent promoters. Thus, phosphorylation-dependent inactivation does not fully account for models involving alternating cycles in the sense and antisense expression. Instead antiphase expression of natural antisense transcripts stipulates there is likely a directionality cue that may involve distinct chromatin states and/or specific phosphorylation events. This creates significant complexities because clock gene expression is no longer a simple case of on or off mediated by phosphorylation, but instead, there may be phosphorylation switches on the transcription factors guiding the transition from sense to antisense expression and back again and this is accompanied by changes in chromatin and/or gene looping, which may also be dependent on the phosphorylation state. Support for this comes from evidence that the *frq* antisense transcript, *qrf*, is controlled by the WCC and *Per2* antisense transcript, *Per2AS* may undergo similar regulation because there is a BMAL1 binding site downstream of *Per2* [79, 81]. In the case with *qrf,* it is clear it is a major regulator of circadian chromatin and assists activation and repression. How circadian chromatin is controlled and how it affects circadian gene expression is discussed below, plus I discuss regulatory events mediated by long non-coding RNA (lncRNA) to serve as an education guide to better understand working models.

1.3 Circadian Chromatin

Chromatin is the DNA protein complex that packages DNA in the nucleus and is predominantly arranged into repeated array of canonical nucleosomes composed of histone octamers containing, 2 H2A, 2 H2B, 2 H3 and 2 H4, further condensed and organized into 30 nm fibers and beyond [82, 83]. Post-translational modifications to chromatin serve as signaling scaffold ensuring genome elements are either accessible or compacted depending on cellular needs. Compaction and accessibility are controlled by combinations of acetylation, phosphorylation, methylation and ubiquitination that are interpreted by chromatin effectors capable of reorganizing the chromatin landscape [84-91]. The last decade and a half have revealed dynamic histone modifications at clock genes affecting the timing and amplitude of oscillations. And although there are some minor differences among circadian model systems, largely due to incomplete information, competing models, or species-specific differences, a mechanism conserved from *Neurospora* to mammals is beginning to emerge. The mechanism consists of circadian-regulated facultative heterochromatin (CRFH) where rhythms in deacetylation, histone H3 lysine 9 methylation (H3K9me) and HP1 binding assist feedback inhibition.

The concept of CRFH first surfaced from studies on the mouse *ccg*, *albumin D element-binding protein* (*Dbp*). An exploratory study on chromatin state changes at *Dbp* revealed anti-phasic oscillations in activating and repressive histone modifications [92]. During the transcriptionally active phase, there were elevated levels of Histone H3 lysine 4 tri-methylation (H3K4me3) and Histone H3 lysine 9 acetylation (H3K9ac), while the repressive phase contained elevated levels of H3K9 di-methylation (H3K9me2) and Heterochromatin Protein 1 (HP1) **(Figure 1-3A.)**. H3K9me2/3 and HP1 are a hallmark of

heterochromatin while H3K9ac and H3K4me3 are typically associated with actively transcribed genes [93, 94]. Although *Dbp* isn't a core clock gene, the rhythm in facultative heterochromatin required the E-box indicating clock-mediated transcription was a requisite for oscillations in permissive and non-permissive chromatin [92]. Shortly thereafter, a similar study on *dper* and *tim* in *Drosophila* also found anti-phasic rhythms in H3K4me3 and H3K9me2 [95]. Separate experiments on the ATP-dependent remodeling enzyme Chromodomain Helicase DNA-binding 1 (CHD1) in *Neurospora* revealed dynamic DNA methylation at a core clock gene and this demonstrated facultative heterochromatin was also conserved from *Neurospora* to mammals [96] because DNA methylation is dependent on H3K9me2/3 and HP1 in *Neurospora* **(Figure 1-3B.)** [97].

Biochemical purification of the nuclear PER-complex provided further, and perhaps some of the most convincing, evidence of CRFH in clock control. The nuclear PER complex is an approximately 1.9 MDa complex that contains around 20-30 proteins, including RNA helicases DDX5 and DHX9, RNA binding proteins NONO and PSF (polypyrimidine tract-binding protein-associated splicing factor), histone lysine 9 methyltransferase KMT1/SUV39H and HP1γ, SIN3-histone deacetylase complex, and the nucleosome remodeling and deacetylase complex (NuRD) **(Figure 1-3C.)** [34, 35, 37, 39, 98]. Thus, the nuclear PER complex guides deacetylation and recruitment of KMT1/SUV39H1 to *Per1* and *Per2* promoters causing circadian di- and trimethylation of H3K9 [34, 37]. For some useful historical context, *Su (Var)3-9* (KMT1) and *Su(var)2-5* (HP1) were first discovered in *Drosophila* as suppressors of position-effect variegation (PEV) of the inverted *White* gene and function in chromatin repression [99, 100].

Biochemical characterization revealed HP1 binds to KMT1-mediated H3K9me2 and H3K9me3 [101-104] creating condensed chromatin. Thus, based on role of H3K9me3 and HP1, and their enrichment during the repressive phase, there is overwhelming support these modifications assist in feedback inhibition.

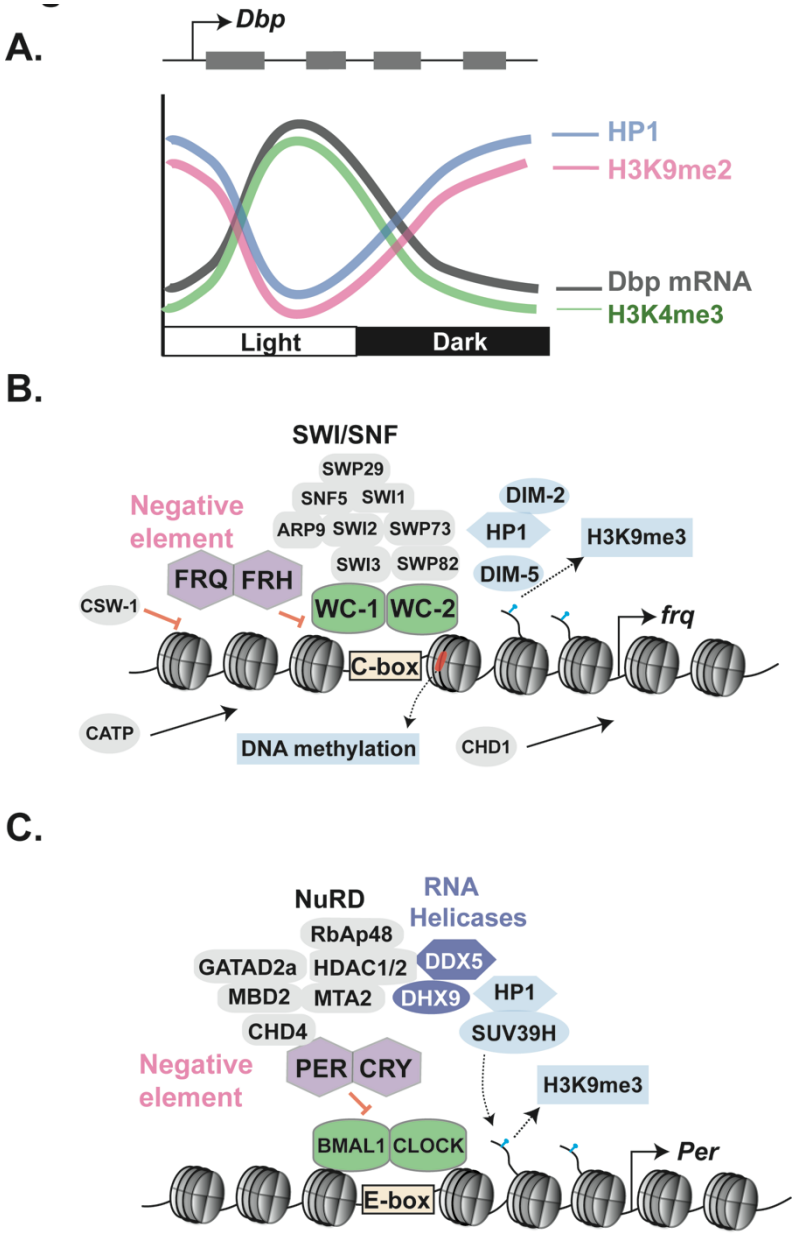


Figure 1-3 Circadian regulated facultative heterochromatin and other associated factors.

(A) A schematic of rhythms showing *Dbp* expression (grey), H3K4me3 (green), H3K9me2 (pink) and HP1 binding (blue). (B) A subset of chromatin-remodeling and modifying enzymes in *Neurospora* clock. (C) A subset of chromatin remodeling and modifying enzymes in the mammalian clock.

Molecular genetic approaches in *Neurospora* designed to understand dynamic DNA methylation at *frq* further supports CRFH is a conserved clock mechanism of fungi to humans. There are rhythms in H3K9me3 and HP1 that occur at *frq* [105] and loss of *kmt1* in *Neurospora* results in a phase shift, and a period decrease. Furthermore, there was also de-repression of circadian output and circadian conidia formation occurs independent of the *Band* mutation [106] suggesting that like *Dbp* in mammals, the amplitude of some *ccgs* involves facultative heterochromatin [105]. Increased output implies increased amplitude of *ccgs* suggesting genome-wide rhythms in H3K9me2/3 assist in controlling transitions to non-permissive states. Unfortunately, a genome-wide study of CRFH looking at H3K9me2/H3K9me3 and HP1 has yet to be performed, but a limited work in *zebrafish* revealed diurnal changes in facultative heterochromatin accompany diurnal changes in transcription, including long non-coding RNAs (lncRNAs) [107]. Whether RNA-mediated heterochromatin serves as a mechanism for circadian chromatin compaction is still unknown, but some correlations between lncRNAs and H3K9me3 are present.

The general model for RNAi-mediated heterochromatin is well-established and merits review. In *S. pombe*, double-stranded RNA (dsRNA) is generated by RdRp (RNA-dependent RNA polymerase) from RNA polymerase II (PolII) derived nascent transcripts. Then Dicer1 cleaves the dsRNA and to produce duplex siRNAs bound by Argonaute. The siRNA is loaded onto RITS (RNA-induced transcriptional silencing) and RITS recruits KMT1/CLR4 complex leading to H3K9me3-mediated gene silencing by forming heterochromatin at pericentric regions and target genes [108-111]. In *Neurospora* CRFH at *frq* shares some similarities with RNAi-mediated heterochromatin

but there are also major differences. CRFH is independent of Dicer and RdRp, but the antisense transcript *qrf* that originates from *frq* gives rise to Argonaute-associated Dicer-independent small interfering RNAs (disiRNAs) [112, 113] and *qrf* is needed for H3K9me3 and DNA methylation at *frq* [114]. In *Neurospora*, duplex RNAs that arise from convergent transcripts are processed by the exonuclease ERI-1 (enhanced RNAi-1) into disiRNAs that trigger DNA methylation [115]. ERI-1 binds nascent transcripts and is associated with DDB1, which is required for H3K9me3 and DNA methylation in *Neurospora*, so it performs a somewhat analogous function to RITS in recruiting KMT1 **(Figure 1-4.)**[115].

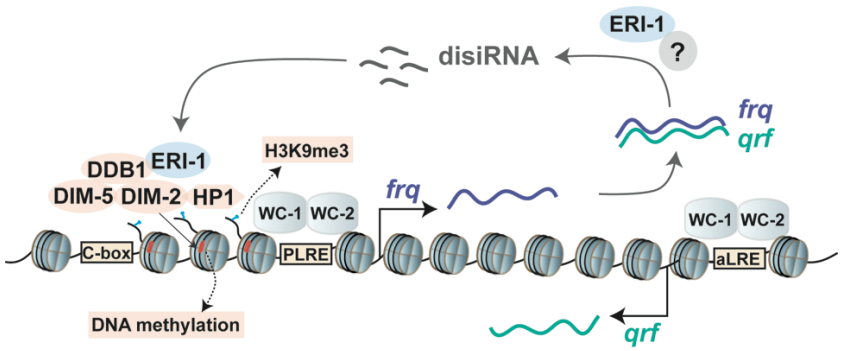


Figure 1-4 The molecular mechanism of NAT-mediated facultative heterochromatin at *frq* in *Neurospora*.

The convergent transcripts *frq* (sense) and *qrf* (antisense) are processed into disiRNA. ERI-1 then helps guide H3K9me3, HP1 and DNA methylation at *frq*.

Oscillations inherent to CRFH requires removal of the repressive modifications and addition of activating modifications, so it is not surprising the clock genes collect a distinct set of other histone modifications during activation phase. Rhythms in histone acetylation and phosphorylation are hallmarks of the activation phase and depending on

the system, H3 lysine 4 tri-methylation (H3K4me3) is also involved. The histone acetyltransferase p300 and PCAF (p300/CBP-associated factor) interact with CLOCK and the rhythm in *Per1*, *Per2* and *Cry1* transcription coincides with a rhythm in RNA polymerase II (PolII) binding and H3 acetylation [116, 117]. In addition, the histone H3 lysine 4 methyltransferase KMT2/MLL1 (Mixed lineage leukemia), interacts with CLOCK:BMAL1 and is needed for proper regulation of circadian transcription [118]. Research on H3K4 methylation in the mammalian clock describe it as an activating mark for clock gene expression [118, 119] and original work on CRFH found H3K4me3 occurs opposite to H3K9me3 [92]. Interestingly, circadian ChIP-seq of H3K4me3 indicate it accompanies circadian gene activation [80] yet global levels of H3K4me3 appear to peak at CT16, the start of the repressive phase of the circadian cycle [79] or roughly 1 hour after the peak in PolII recruitment [120]. In *Neurospora*, the H3K4 methyltransferase, KMT2/SET-1 is necessary for feedback inhibition, DNA methylation and H3K9me3 at *frq* [121, 122]. The role of H3K4me3 in mammals and *Neurospora* is somewhat confounding and highlighting the need for more research, especially when one considers whether H3K4me3 and K3K9me3 contribute to the switch from sense to antisense expression.

In addition to H3K9me2/3 and CRFH, H3K27me3-mediated heterochromatin is also implicated in clock function. The Polycomb group enzyme KMT6A/EZH2 co-precipitates with CLOCK: BMAL1 in liver extracts and methylates H3K27. Knockdown of *Kmt6a/Ezh2* disrupts the *Per2*-luc and *Bmal1*-luc rhythms [123]; Although, in some instances, ChIP of KMT6A/EZH2, H3K27me2, and H3K27me3 at *Per1* and *Per2* promoters did not have statistically significant rhythm [123]. The implications of

KMT6A/EZH2 and H3K27 methylation in the context of CRFH and the mechanism of KMT6A/EZH2 recruitment are not fully understood.

Removal of the methyl-modifications on H3K4, H3K9 and H3K27 is part and parcel of the oscillations; however, the data surrounding the histone lysine (K) demethylases (KMDs) is perplexing. Lysine-specific demethylase 1 (KMD1A/LSD1) is an FAD+-dependent histone demethylase conserved from *S. pombe* to human and demethylates histone H3 lysine 4 and lysine 9 [124]. KMD1A/LSD1 has a rhythm in phosphorylation mediated by PCKα (protein kinase Cα) when complexed with CLOCK:BMAL1 to promote activation of *Per2* [125]. Interestingly, loss of circadian phosphorylation on KMD1A/LSD1 did not affect the methylation status on H3K4 or H3K9, but H3K9ac was reduced [125]. KMD1A/LSD1 is also associated with SIRT1 and additional HDACs [126]. Therefore, loss of circadian phosphorylation may somehow create constitutively active histone deacetylation or alter upstream acetylation. Another histone demethylase implicated in clock function is JumonjiC and ARID domain-containing histone lysine demethylase 1a (KMD5A/JARID1a). KMD5A/JARID1a also associates with CLOCK::BMAL1 to facilitate *Per2* transcription, but like KMD1A/LSD1, it does so independent of demethylation and instead increases histone acetylation by inhibiting HDAC1 [127]. Furthermore, KMD2A/FBXL11 is also required for normal circadian periodicity and assists the negative limb of feedback inhibition [128]; however, like KMD1A/LSD1 and KMD5A/JARID1a, the role of KMD2A/FBXL11 is independent of demethylase activity. Thus, how the histones are catalytically demethylated at clock genes remains an enigma that is further complicated

by crosstalk among other chromatin modifying enzymes and possibly multivalent chromatin readers.

The lack of conclusive evidence of histone demethylation suggests a speculative model where demethylation may involve replacement of methylated nucleosomes or the exchange of individual histones. Likely candidates for nucleosome/histone exchange are ATP-dependent chromatin remodeling enzymes (Possibly members of the chromodomain helicase DNA binding (CHD) family; see below) and/or histone chaperones. Support for histone exchange comes from studies showing the WCC and CLOCK:BMAL1 function as pioneering transcription factors [129-131]. WCC induces nucleosome eviction [130], while CLOCK:BMAL1 incorporates the histone variant H2A.Z to promote the chromatin opening allowing other transcription factors, such as hepatocyte nuclear factor 6 (HNF6) to bind [129]. FACT (Facilitates Chromatin Transcription) -dependent H2A.Z incorporation at *frq* C-box in *Neurospora* supports rhythmic binding of WCC and disruption of DNA replication affects H2A.Z deposition, WCC binding, and clock regulation [132]. Collectively, these findings indicate routine histone exchange, or cell cycle-dependent DNA replication exchange, assists replacement of histones with variant counterparts and reveal histone dynamics are critical for clock regulation. Therefore, removal of heterochromatic histones, instead of demethylation, is not far-fetched, especially in lieu of no specific H3K9 demethylase activity.

In addition to promoter or promoter proximal methylation, there is also a rhythm in H3K36 methylation in the gene body of clock and clock-controlled genes (typically peaking in and around the 3' end). In general, a common theme shared between *Neurospora* and mammals is that H3K36 tri-methylation peaks at the beginning of the

repressive phase with slight variation on the timing [120, 133, 134]. Presumably, this should not come as a huge surprise because KMT3/Set2, the lysine methyltransferase that adds methyl groups to H3K36, is needed for repression in yeast and H3K36me3 serves to recruit the histone deacetylase complex, Rpd3S in order to deacetylate histones after PolII elongation [135-138]. So even though H3K36me3 associates with PolII and H3K36me3 gets deposited during elongation [139, 140], the mark may linger until subsequent circadian cycles.

1.4 Histone acetylation and deacetylation

Circadian oscillations in chromatin modifications also include acetylation and deacetylation. Thus it is not surprising that histone lysine acetyltransferase (KATs) and histone deacetylases (HDACs) generate a rhythm in histone acetylation [117, 141]. Since the discovery Gcn5 is a KAT, the role of histone acetylation as a mark for actively transcribed genes is well-established [142-144] while deacetylation is a requisite for heterochromatin [145]. Much of our understanding of acetylation and deacetylation in the clock has come from studies in mammals. Histone H3 acetylation is rhythmic in *Per1* and *Per2* in fibroblasts, heart and liver and the rhythm coincides with PolII association and transcription [116, 117]. There are also genome-wide rhythms in H3K9ac and H3K27ac [79, 80]. CLOCK is reported to have intrinsic KAT activity [146], but CLOCK is also associated with histone acetyltransferases, p300/CBP and PCAF (p300/CBP-associated factor) [117], so clear indications of CLOCK catalytic KAT activity remains equivocal and would benefit from further studies. Acetylation is reversed by HDACs and the number of HDACs implicated in clock function is relatively large and include

HDAC1, HDAC2 (part of SIN3A, SIN3B and NuRD) HDAC3, SIRT1 and SIRT6 and they play vital, conserved, and sometimes overlapping roles in transcription repression. For example, CRY1 is associated with SIN3B, HDAC1 and HDAC2 [141], and the PER-complex contains deacetylation complexes SIN3A containing HDAC1 [34] and NuRD, which contains HDAC1 and HDAC2 [39]. Thus, HDAC-mediated deacetylation is a key step in reversing the activation phase. Histone deacetylase 3 (HDAC3) is also recruited by Rev-Erbα and is needed for rhythms in hepatic lipid metabolism [147]. The NAD+-dependent deacetylase, SIRT1 remains somewhat enigmatic, in part because differences in competing reports preclude a unifying functional model. SIRT1 is associated with CLOCK:BMAL1 and depending on the report, deacetylates BMAL1 or PER2 [148, 149]. Loss of SIRT1 does not appear to affect the circadian period, only the amplitude, raising the specter of secondary effects. Differences may be compounded by individually isolated *Sirt1* knockout mouse embryonic fibroblasts (MEFs) preventing decisive conclusions on the exact phenotype. Circadian regulation of SIRT1 is also equivocal; however, this may be a moot point because it is clear *nicotinamide phosphoribosyl-transferase* (*Nampt*) expression is rhythmic [150] and this has been independently verified in numerous RNA-seq and microarray experiments by different groups. Thus, at the very least, NAD+ synthesis and the NAD+ salvage pathway are controlled by the clock [151]. Based on this, one can surmise SIRT1-mediated deacetylation, if by no other means than NAD+ cycles, is presumably circadian regulated. And this links to cellular metabolism because SIRT1 deacetylates several transcriptional regulatory proteins involved in glucose homeostasis in liver [152], impairs insulin secretion [153], and affects neuronal function in the brain [154]. SIRT1 also has implications in aging which I

elaborate on below. Other Sirtuins, SIRT3 and SIRT6, are also involved in clock regulation. SIRT6 interacts with CLOCK:BMAL1, but interestingly SIRT6 functions through different mechanism other than histone deacetylation and instead is involved in recruitment of the SREBP-1 [155]. While SIRT3 is involved in circadian clock regulation of mitochondria metabolism [156]. Collectively, HDACs, including Sirtuins play an instrumental role in circadian clock and metabolism and this is supported by a multitude of studies **(Figure 1-5.).** Acetylation and deacetylation studies in *Neurospora* and *Drosophila* are extremely limited, but it is safe to conclude the *ngf-1* (*Neurospora Gcn five*) is the major KAT affecting the clock in *Neurospora* because WC-1 dependent transcription requires *ngf-1* [157] and rhythms in acetylation occur at *dPer* in flies [95].

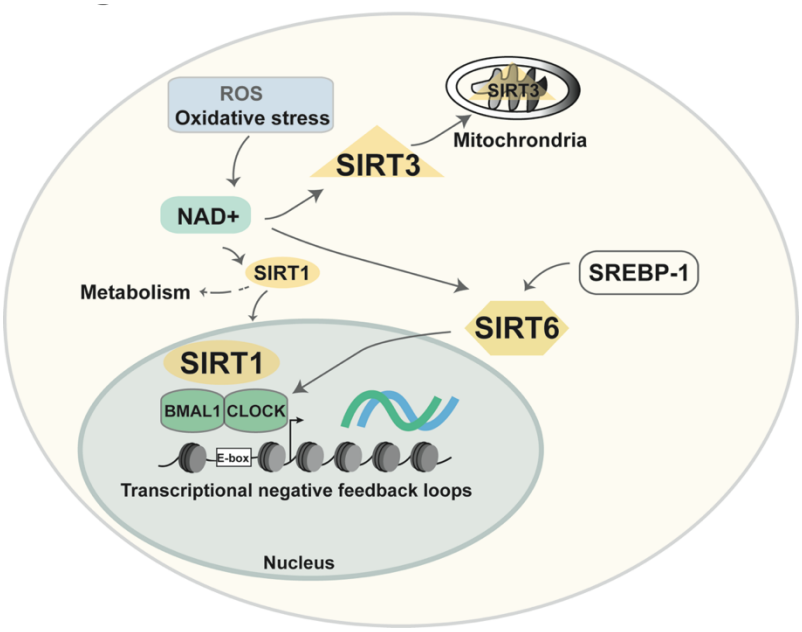


Figure 1-5 Schematic representation of mammalian Sirtuins (SIRT1, SIRT3 and SIRT6) implicated in clock function.

SIRT1, SIRT3 and SIRT6 activity rely on the NAD^+. The subcellular location is different among different SIRTs. SIRT6 associates with CLOCK:BMAL1 through an interaction with SREBP-1.

1.5 ATP-dependent Chromatin-remodeling Enzymes

Circadian post-transcriptional modifications to chromatin are interpreted by effectors including ATP-dependent remodeling enzymes; although, what modifications recruit which remodeling enzymes is sorely lacking. The list of ATP-dependent remodeling enzymes is relatively large and include Clockswitch (CSW-1), Chromodomain Helicase DNA binding 1 (CHD1), Clock ATPase (CATP), SWI/SNF (SWItch/Sucrose NonFermentable) and Ino80 in *Neurospora*, KISMET and BRAHMA in *Drosophila,* and CHD3/CHD4 and SWI/SNF in mammals. Thus far the conserved remodeling enzymes are the SWI/SNF and CHD families and both are needed to support

normal clock gene expression. SWI/SNF interact with WCC and is needed for *frq* expression in *Neurospora* [38]. While, *Brahma (Brm)*, the *Drosophila* SWI/SNF, changes the chromatin landscape at *dper* and *tim* to limit PolII recruitment [36]. In mammals, the SWI/SNF subunit BAF60a is associated with hepatic circadian clock and is necessary for normal energy metabolism [158]. In all instances, and based on work in other systems, the SWI/SNF complex alters chromatin surrounding the transcriptional start site to control the accessibility of PolII and general transcription factors [38, 159].

The role of the CHD enzymes, from work in *Neurospora* and mammals, is far more complicated but a unifying model, based in part on extrapolation, can be discerned. The closest homologues to CHD1 in *Neurospora* is CHD3 and CHD4 in mammals. CHD3/CHD4 is part of the nucleosome remodeling and deacetylates (NuRD) complex that functions as a corepressor [160, 161]. In *Neurospora*, loss of CHD1 results in a reduction in *frq* expression and increases DNaseI inaccessible heterochromatin complete with extensive DNA methylation [96]. This indicates CHD1 is necessary to create a more open chromatin state and reverse the heterochromatin for subsequent cycles of clock gene expression. In mammals, CHD4 is associated with CLOCK:BMAL1 and is necessary for the activation phase. siRNA against CHD4 causes significantly damped expression and result in less PolII at *Per1* and *Per2* suggesting less-accessible chromatin [39]. Collectively, this creates a unifying model where the CHD family is likely involved in reversing the circadian heterochromatin to promote expression. Interestingly, during the repressive phase, the PER complex delivers the remaining NuRD subunits to support bi-modal repressive activity that includes deacetylation and re-establishment of heterochromatin. Considering that *Per2* also has dynamic DNA methylation [162], and

NuRD contains the methyl-DNA binding protein MBD2, it would be interesting to determine whether *Per2* is hypermethylated or packaged into constitutive heterochromatin in cells lacking CHD4.

The remaining chromatin remodeling enzymes have only been found in individual systems; either *Neurospora* or *Drosophila*. In *Neurospora*, loss of *csw-1* causes a defect in negative feedback and is specifically involved in moving a nucleosome adjacent to the c-box to block WCC binding [163]. In contrast, Clock ATPase (CATP) is needed to reposition that same c-box nucleosome to promote *frq* expression [164]. The C2H2 finger domain-containing protein IEC-1, which is associated chromatin remodeling complex INO80 is needed for transcriptional repression of *frq* by creating dense chromatin environment in the *frq* promoter [165]. Whether any of these affect variant exchange would be useful to know. *Drosophila* KISMET (KIS) is a bit unique and is required for normal circadian photo-responses [166].

Collectively, the large array of nucleosome remodeling factors functioning at clock genes indicate there is important catalytic activity that opens, closes and reorganizes the chromatin to support rhythms, but how these interact with modified nucleosome and their exact catalytic activity is still largely a mystery.

1.6 Chromatin Structure

In addition to chromatin remodeling at the nucleosome level, the circadian clock controls 24-hour cycles of gene transcription in part by regulating the 3-D chromatin architecture. Current models on nuclear structure indicate chromosomes are organized into nonrandom polymers that undergo dynamic changes to support metabolic status,

transcription profiles and development [167, 168]. Thus, it is perhaps not surprising there are changes to genome structure on the circadian timescale beyond the localized changes discussed thus far. Chromosome conformation capture (3C) on chip (4C) revealed circadian interactions between *Dbp* with other circadian loci on separate chromosomes and these interactions were dependent on BMAL1 [169]. In circadian entrained MEFs, PAR polymerase 1 (PARP1) and CCCTC-binding factor (CTCF) mediate circadian regulation of high-order chromatin organization by mobilizing *ccg* to lamina-associated domains (LADs) [170]. Consistent with circadian heterochromatin, H3K9me2/3 was a critical component in repositioning circadian genes to the LADs [170]. Cohesion-mediated enhancer-promoter domains are contained within larger CTCF-Cohesin (CCCTC-binding factor) domains [171]. Rev-erbα, a repressor affecting *Bmal1* expression, assists in preventing looping needed for expression, in part by recruiting NCoR (nuclear receptor co-repressor) and HDAC3 (histone deacetylase 3) [172]. These data indicate plasticity in genome-wide chromatin interactions are prevalent and support the importance of clock-controlled nuclear architecture **(Figure 1-6).** Collectively, these reports reveal daily changes in promoter-enhancer looping help modulate rhythms in *ccg* expression governing the temporal dynamics of the circadian regulatory network which likely affect numerous metabolic processes [173]. Future studies will provide more insight into clock-dependent nuclear architecture and how dynamic changes affect metabolic status at the cellular level. In addition, the extent to which facultative heterochromatin guides translocation of clock genes to the nuclear periphery is an open question. Although, it is interesting to speculate that rhythms in facultative heterochromatin are the key determinant in driving clock genes to re-localized to the

nuclear periphery, which is enriched with heterochromatin-packaged telomere DNA (see below).

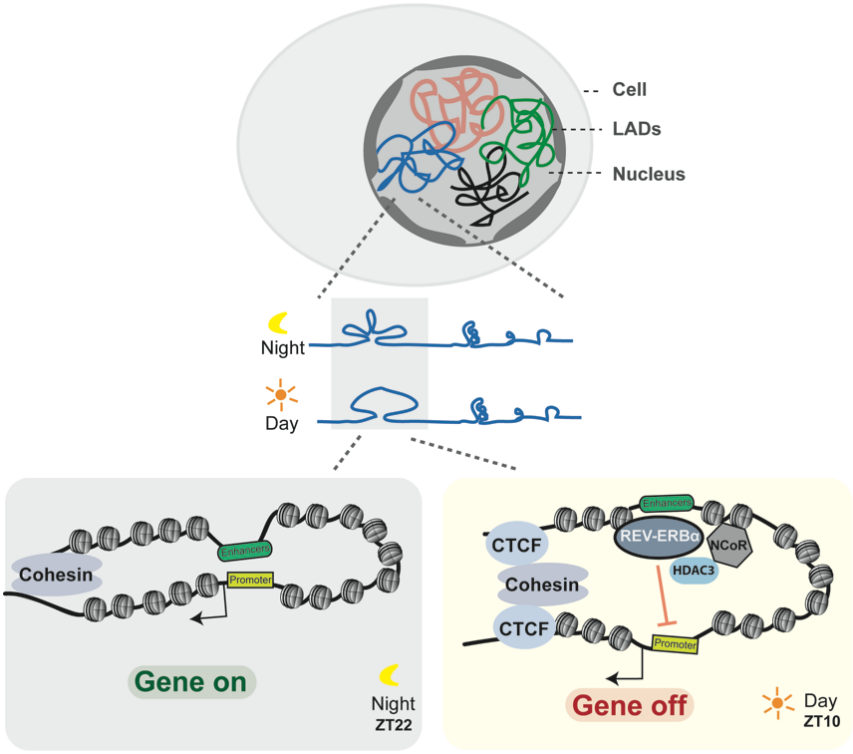


Figure 1-6 An example of circadian plasticity in higher-order chromosome organization over the circadian cycle.

Cohesion supports enhancer and promoter co-localization to help regulate gene activation. Cohesion-CTCF interaction and Rev-erbα disrupt enhancer-promoter looping to assist in transcriptional repression by recruiting nuclear receptor co-repressor (NCoR) and histone deacetylase 3 (HDAC3).

1.7 Long non-coding RNAs and chromatin organization

There is a substantial body of research showing long non-coding RNAs (lncRNAs) regulate chromatin and growing evidence indicates this holds true, in some instances, for circadian lncRNAs. lncRNAs and small RNAs arise from intergenic or antisense transcription and are key regulators of chromatin structure at non-circadian

regulated loci [174, 175]. High-throughput, strand-specific RNA sequencing has revealed that antisense transcripts are widespread throughout the genome in many species [176] and there are circadian lncRNAs and micro RNA (miRNA) [79, 80]. In humans, more than 30% of protein coding genes have antisense transcripts [177] and some of these are important regulators of chromatin architecture and gene regulation [178, 179]. lncRNAs are defined as transcripts larger than 200 nucleotides and are in a separate category from other small RNAs such as microRNAs, small nucleolar RNAs (snoRNAs) and small interfering RNAs (siRNAs). lncRNAs are defined into 5 groups based on their relative position to coding genes. These categories are intergenic, intronic, divergent, antisense or enhancer RNAs (eRNA). Natural Antisense Transcripts (NATs) are specific lncRNAs and are widespread in eukaryotic genomes [180, 181]. NATs overlap with and are transcribed independently from sense RNAs. NATs are involved in different chromatin regulatory processes, including cell fate determination, X chromosome inactivation, imprinting and the circadian clock. One of the best-characterized lncRNAs is *X-inactive specific transcript* (*XIST*) involved in X chromosome inactivation which serves as a scaffold for proteins to alter the chromosomal architecture [182-185]. Another well-known function of lncRNAs is to establish genomic imprinting composed of H3K9me3 and DNA methylation leading to epigenetic silencing of parental-specific genes [186]. For example, H19 is a lncRNA highly expressed in embryogenesis and is involved with *Igf2* imprinting [187]. Although H19 was one of the earliest lncRNAs identified, the mechanism is still not fully understood but does involve CTCF. Currently, the function of most lncRNAs is not defined but techniques such as crosslinking and immunoprecipitation (CLIP) [188], RNA immunoprecipitation (RIP) [189], RNA pull-

down, chromatin isolation by RNA purification (ChIRP) [190], capture hybridization analysis of RNA target (CHART) [191], and RNA antisense purification (RAP) [192] are beginning to provide insights. Use of these techniques have contributed to understanding the function of lncRNAs, *Xist*, and *RepA,* which are known to recruit Polycomb complex to the target genomic loci leading to H3K27me3-dependent heterochromatin formation and repressed gene expression [190, 193]**.**

In addition to the over 1000 circadian lncRNAs in mammals [1], clock genes themselves have NATs in *Neurospora* and mammals [78-80]. In mammals there is the *Per2* antisense transcript (*Per2AS*) and in *Neurospora* the *frq* antisense transcript *qrf.* As mentioned above, *qrf* is required for dynamic DNA methylation at *frq* [96] giving rise to disiRNAs [113] that cause DNA methylation [112]. *qrf* also has a bimodal effect on the clock; *qrf* first promotes *frq* expression by creating a more permissive state before switching to a repressing function where it guides H3K9me3 [114]. Constitutive low-level expression of *qrf* ablates circadian gene expression due to the promoting effects and because a threshold level of *qrf* is necessary to establish H3K9me3 [114]. How *qrf* promotes *frq* expression is unknown, but convergent transcripts can form R-loops to promote transcriptional activation [194]. Alternative models of *frq-qrf* function suggest convergent transcription causes PolII collisions on DNA thus interfering with *frq* expression [195] while the nascent transcripts serve as a docking site for ERI-1-dependent recruitment of KMT1 [115]. Whole-transcriptome RNA sequencing revealed that mammals also have *Per2* natural antisense transcripts *Per2AS* that oscillates anti-phasic to *Per2* in mouse liver but its role, if any, is entirely unknown [79, 80]. There are also enhancers RNA (eRNA), 30% of which are circadian regulated and these correlate

with nucleosomes marked with H3K27ac and H3K4me1 [196]. Enhancer-associated lncRNAs, such as lncCrot is involved in creating long range circadian enhancer interactions to help support *ccg* expression [197]. How the vast array of lncRNAs (including *Per2AS*) and eRNAs function to support proper timing, via changes to chromatin states, remains an important and open question that goes beyond current models of clock function.

1.8 Circadian Clock, Aging and Disease

Circadian rhythms have profound influence on human health and disruption of the circadian clock contributes to cancer, cardiovascular disease, neuroendocrine disorders, metabolic syndromes and others **(Figure 1-7)** [198-200]. A major driver of circadian-related disease undoubtedly resides in global misregulation of *ccg* expression because the clock controls more than 40% of protein coding transcripts either directly or indirectly [1, 201]. Altering expression of just a few critical genes may create homeostatic deficiencies that contribute to disease. Take central metabolism as an example, the circadian clock is a major regulator of metabolism and in turn, transcription of *ccg*s also influences numerous cellular and physiological functions, such as food intake, hormonal synthesis and release, and DNA repair [202, 203]. Thus, circadian disruption in the form of light-at-night or shiftwork, simultaneously perturbs food intake and metabolism, which can impact normal levels of reactive oxygen species (ROS). Increased ROS can serve as a DNA-damaging agent while clock-regulated DNA repair is being negatively affected. Thus, one could envision a scenario where the molecular connection between the higher incidence of cancer in patients with obesity lies in circadian disruption that increase ROS at the same

time it affects expression of DNA repair enzymes such as circadian-regulated *Ogg1* [204], which would increase the likelihood of acquiring a cancer-driving mutation.

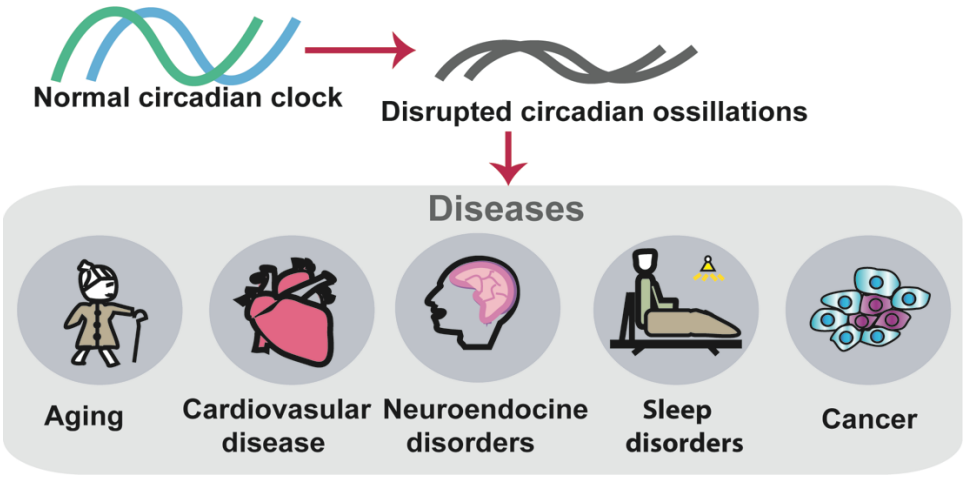


Figure 1-7 Schematic representation showing disruption of the circadian clock gives rise to a host of circadian-related diseases.

The phenomena is perhaps best seen in *Bmal1*$^{-/-}$ mice that have a severe premature aging phenotype complete with increased senescence and increased levels of ROS [205, 206]. *Bmal1* directly regulates NRF2 expression controlling antioxidant proteins in macrophages [207] causing *Bmal1* deficient macrophages to have higher level of ROS and proinflammatory cytokines [208]. Of note, N-acetylcysteine, which is a synthetic precursor of intracellular cysteine wildly used as antioxidant [209], can alleviate the advanced aging phenotype in *Bmal1*$^{-/-}$ [206].

One cannot talk about the clock and aging without noting circadian regulation of telomeres. Telomeres are repeated DNA sequences at the distal tips of eukaryotic chromosomes composed of nucleoprotein structure that protects linear chromosomes [210, 211]. Telomeres are packaged into heterochromatin containing H3K9me3,

H3K27me3 and H4K20me3 (trimethylated histone H4 at lysine 20) catalyzed by KMT1a/SUV39H1, KMT1b/SUV39H2, KMT6A/EZH2 and KMT5B/SUV4-20H [212] [213]. Telomeres shorten when cells divide due to the end-replication problem and telomere shorting is a hallmark of cell senescence and aging [214]. Telomerase is a ribonucleoprotein composed of TERT (telomerase reverse transcriptase) and TERC (Telomerase RNA component) that adds the telomere repeat sequence to help solve the end-replication problem [215]. The activity of telomerase has strong connections to aging and cancer [216] and TERT is reportedly regulated by CLOCK:BMAL1 [217]. Clock deficient animals have shortened telomeres and disrupted expression of TERC and TERT [217]. However, circadian regulation of telomeres goes beyond mere circadian control of *TERT* and *TERC* and is more direct. WC-2 in *Neurospora* and BMAL1 in vertebrates is associated with the telomere repeat and supports a rhythm in telomeric repeat-containing RNA (*TERRA*) [218]. *TERRA* is a long non-coding RNA that is transcribed from the subtelomeric region toward the telomere repeat [219] and is implicated (among others) in heterochromatin formation at the telomeres [220, 221]. Circadian regulation of *TERRA* and heterochromatin at the telomere indicates CRFH is not restricted to clock and clock-controlled genes **(Figure 1-8)** and further demonstrates important connections among the circadian-regulated heterochromatin, lncRNA expression and aging.

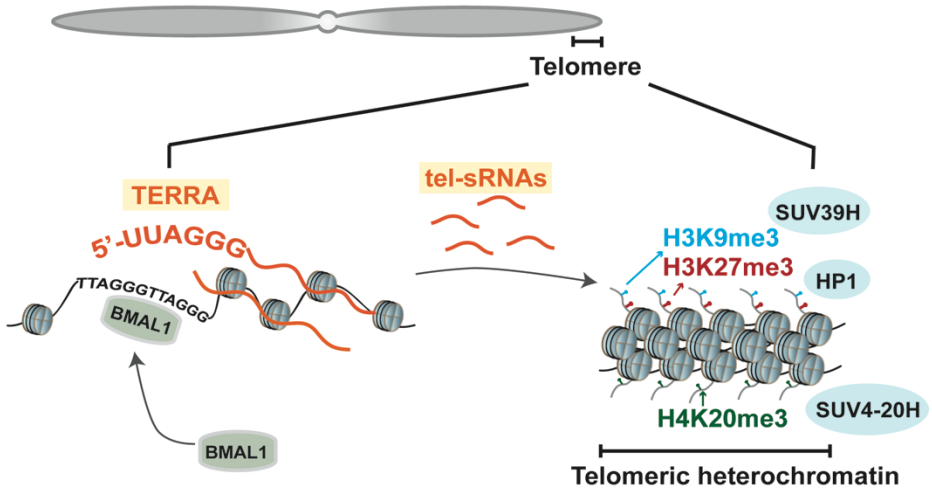


Figure 1-8 Heterochromatin at telomeres in mammals.

BMAL1 binds to telomeres and controls a rhythm in *TERRA*. *TERRA* is converted to Dicer-independent tel-sRNAs and these are implicated in KMT1a/SUV39H1 and HP1-dependent heterochromatin formation at the telomeres.

Another aspect linking the circadian clock, aging and chromatin, lies in the NAD+-dependent deacetylases SIRT1. SIRT1 not only functions in metabolic sensing, but also affects aging and is needed for heterochromatin formation at the telomeres [222, 223]. SIRT1 also associates with PGC-1α and activates the central clock genes in young animals, yet declines with age in the brain [224]. The same concept occurs in aged mice where there is a decrease in acetylation that can be rescued by caloric restriction [225]. These findings encourage further studies to understand the connection between the circadian clock and aging. Interestingly though, age has little effect on the core clock itself, but major changes occur in circadian output [107, 225, 226], supporting a hierarchy in clock regulation that involve age-related epigenetic changes to circadian output and/or circadian output reprogramming, for which SIRT1 is presumably integral [225].

1.9 Concluding remarks and objectives

It is now unequivocal that a major function of the clock is to control chromatin states and genome structure to facilitate proper amplitude and timing of circadian regulated gene expression. A conserved model of circadian heterochromatin involving deacetylation, H3K9me2/3, HP1 binding and dynamic DNA methylation is part and parcel of negative feedback inhibition. Also, the circadian clock is involved in regulating the dynamics of nuclear architecture. The connection between circadian chromatin state changes and circadian lncRNAs is beginning to emerge and evidence from other systems suggest these are important for chromatin regulation. Finally, disruption to the circadian clock has a profound influence on human health and understanding the molecular mechanisms connecting the circadian clock and aging is key to determining the root causes of disease manifestation.

The following dissertation describe my research. I took a multi-organism approach to address three novel questions to understanding connections among the circadian clock, long non-coding RNAs and heterochromatin formation. In Chapter 2, I examined the crosstalk between H3K4me3 and H3K9me3 on genome-wide gene expression. For this, I used *Neurospora crassa* because it was the only circadian system where one can completely obliterate all H3K4 and H3K9 methylation. In Chapter 3, I examined how perturbations to the clock affected CRFH at the telomeres in zebrafish. To accomplish this, I used the commonly used drug alcohol, which is known to induce shortened telomere, to alter diurnal rhythms at the telomere. In Chapter 4, I examined the regulation and functional role of *Per2* NAT in mammalian cells.

CHAPTER 2

2 CHAPTER 2: Histone H3 Lysine 4 methylation is required for facultative heterochromatin at distinct loci

(Chapter 2 published in *BMC Genomics* 20, 350 (2019) doi:10.1186/s12864-019-5729-7[122])

2.1 Preamble to Chapter 2

The filamentous fungus, *Neurospora crassa* is an excellent model eukaryote that has been widely in circadian clock research for decades [227]. In fact, negative feedback inhibition was first proven using *Neurospora* before any of the mammalian clock components were even cloned [228]. The White Collar-1 (WC1) is a blue-light photoreceptor and regulates the clock gene *frequency* (*frq*), and this is a great system to study light entrainment and phase resetting. In this chapter, I used RNA-seq and ChIP-seq to study how H3K4me3 and H3K9me3 affect gene transcription using two mutant strains, *Δkmt1/*Δ*dim-5* which obliterates H3K9me3 and Δ*kmt2/*Δ*set-1* which removes all H3K4 methylation. The choice of *Neurospora* instead of vertebrates (*zebrafish* or mouse) was simple. *Neurospora* is the only system where you can completely block all H3K4me3 and H3K9me3. In vertebrates, there are 6 histone lysine 4 methyltransferase (KMT2) enzymes including SET1A, SET1B, MLL1, MLL2, MLL3, and MLL4 [229] and 6 histone lysine 9 methyltransferase (KMT1) including, G9a, SETDB1, SUV39H1, SUV39H2, EHMT1, and EHMT2 [230]. Therefore, it is impossible to mutate all of them and removal of any individual KMT1 or KMT2 will cause paralog-compensation and/or

chromatin redistribution, supporting erroneous conclusions. Therefore, *Neurospora* is most ideal model instead of vertebrates due to redundancy of methyltransferases.

2.2 Abstract

Background: Light-activated gene expression in *Neurospora crassa* (*Neurospora*) is mediated by transcriptional and post-transcriptional events to control early- and late-light responses and photo-adaptation. Despite extensive studies examining the light response, very little is known about histone H3 methylation and how it contributes to light-regulated gene expression.

Results: In this report, I performed a multi-dimensional genomic analysis to understand the role of 2 opposing modifications; histone H3 lysine 4 tri-methylation (H3K4me3) and histone H3 lysine 9 tri-methylation (H3K9me3) using the *Neurospora* light response as the system. RNA-seq on strains lacking H3 lysine 4 methyltransferase (KMT2/SET-1) and histone H3 lysine 9 methyltransferase (KMT1/DIM-5) revealed many light-activated genes had altered expression. Comparing these 2 mutants to wild-type (WT), I found that roughly equal numbers of genes showed elevated and reduced expression in the dark and the light making the environmental stimulus somewhat ancillary to the genome-wide effects. ChIP-seq experiments revealed H3K4me3 and H3K9me3 had only minor changes in response to light in WT, but there were notable alterations in H3K4me3 in *Δkmt1/*Δ*dim-5* and H3K9me3 in Δ*kmt2/*Δ*set-1* indicating crosstalk and redistribution between the modifications. Integrated analysis of the RNA-seq and ChIP-seq highlighted context-dependent roles for KMT2/SET1 and KMT1/DIM-5 as either co-activators or co-repressors with some overlap as co-regulators. At a small

subset of loci, H3K4 methylation is required for H3K9me3-mediated facultative heterochromatin including, the central clock gene *frequency* (*frq*). Moreover, using re-ChIP I confirm *Neurospora* contains K4/K9 bivalent domains.

Conclusions: Collectively, these data indicate there are obfuscated regulatory roles for H3K4 methylation and H3K9 methylation depending on location with some minor overlap and co-dependency.

Keywords: Histone lysine methyltransferase, heterochromatin, histone H3 lysine 4 methylation, histone H3 lysine 9 methylation, long non-coding RNA, Light-activated gene expression

2.3 Introduction

The *Neurospora* light response is a transcriptional cascade that responds to environmental cues and controls circadian entrainment and developmental programs [231-234]. White Collar-1 (WC-1) and WC-2 are GATA-type transcription factors that drive light-mediated expression and associate via PAS (Per, Arnt and SIM) domains to form the White collar complex (WCC) [235-237]. WC-1 servers as the photoreceptor sensing blue light via its LOV (light, oxygen and voltage) domain in a photochemical reaction involving a cysteinyl-flavin adenine dinucleotide (FAD) adduct [238, 239]. Microarray and RNA-seq experiments indicate that approximately 3-20% of all genes change expression in response to light and fall into 1 of 3 categories; early light-responsive that require the WCC, late light-responsive that require SUB-1, and light-repressed [130, 240-242]. There are also light-activated changes to chromatin structure and nucleosome position mediated by ATP-dependent chromatin-remodeling enzymes

[38, 130, 163]. The nucleosome movement and remodeling are likely initiated by the *Neurospora* homologue of GCN5 (NGF-1), which acetylates histone H3 on lysine 14 and is needed for proper light-activated gene expression [157]. Despite our knowledge of WCC, SUB-1, NGF-1, chromatin remodeling and RNA polymerase II (PolII) activity, the role of histone modifications in the *Neurospora* light response is still very limited.

Previously, it was demonstrated that KMT2/SET-1-dependent histone H3 lysine 4 methylation and KMT1/DIM-5-dependent histone H3 lysine 9 methylation are involved in light- and clock-regulated gene expression playing a supporting role in photoadaptation and negative feedback inhibition [121, 243]. Histone H3 lysine 4 tri-methylation (H3K4me3) is generally viewed as a modification supporting transcriptional activation because it is enriched in active genes [244] even though H3K4me3 was originally found to be involved in repression [245-247]. The role of KMT2/SET-1 in repression is not fully understood, but studies in *S. pombe* indicate it is needed for repression at subtelomeric genes and retrotransposons, possibly independent of H3K4 methylation [248-250]. Moreover, recent models suggest that H3K4me3 may be less of mark for activation, and more reflective of cell-specific transcriptional states [251, 252]. Thus, even though KMT2/SET-1-dependent H3K4 methylation is generally associated with euchromatin, KMT2/SET-1 appears to have a context dependent role in silencing. In contrast, H3K9me3, which is typically bound by Heterochromatin Protein 1 (HP1), is entrenched as a repressive modification underlying condensed heterochromatin [253-255]. Yet evidence suggests H3K9me3 can also be found in actively transcribed genes in mammals [256] and heterochromatic regions (e.g. centromeres and telomeres) are often transcriptionally active due to the requirement of RNAi-mediated heterochromatin [257].

In *Neurospora*, it is widely believed there is little, if any cross talk between KMT2/SET-1 and KMT1/DIM-5 because H3K9me3 and H3K4me2 appear to be mutually exclusive [258] and *Neurospora* centromeres lack H3K4 methylation [259]. However, both KMT2/SET-1 and KMT1/DIM-5 assist circadian negative feedback of the clock *frequency* (*frq*), and both are required for DNA methylation ($5^{m}C$) at *frq*. In addition, the deletion strains appear to have some effect on photoadaptation and H3K4me3 and H3K9me3 appear to peak within *frq* approximately 30 minutes after exposure to light [121, 243], a time when VIVID (VVD)-mediated down-regulation and adaptation has commenced [260, 261]. Specifically, the loss of DNA methylation at *frq* in Δ*kmt2/Δset-1* implicates H3K4me3 in KMT2/DIM-5-dependent facultative heterochromatin and counters prevailing views on H3K4 methylation.

The dynamic DNA methylation and facultative heterochromatin at *frq* also require coordinated expression of a light-activated long non-coding natural antisense transcript (NAT) *qrf* [78, 112, 114]. Strains with constitutive low-level expression of *qrf* have a localized defective in DNA methylation and heterochromatin [96, 114]. As a side, the initial burst of light-activated *qrf* appears to contribute to bimodal regulation of *frq* by initially creating a more permissive state for expression prior to heterochromatin-mediated silencing [114]. The role of *qrf* in establishing heterochromatin may not be unique because many sense-antisense pairs and convergent transcripts gives rise to Dicer-independent small interfering RNA (disiRNA) and subsequent DNA methylation [112, 262].

In order to further understand the role of KMT2/SET-1, KMT1/DIM-5 and long noncoding RNA (lncRNA), I took a genome-wide approach and examined the

connections between H3K4me3 and H3K9me3 and how these impact gene expression in the *Neurospora* light response. Part of the premise is that, at least at *frq,* KMT2/SET-1 and KMT1/DIM-5 along with an antisense transcript are all needed for heterochromatin. Therefore, I coupled RNA-seq and transcript discovery with H3K4me3 and H3K9me ChIP-seq. My results reveal conventional paradigms surrounding H3K4me3 and H3K9me3 are not entirely universal and point to a context, inter-dependent nature for both modifications. I have found that at some loci, KMT2/SET1 and KMT1/DIM-5 are involved in repression, and at other loci, they aid in co-activation. In addition, KMT1/DIM-5 and KMT2/SET-1 can function alone or in combination to help modulate gene regulation. The data reveal that H3K4me3 and H3K9me3 are not solely confined to activation or repression and instead point to complex combinatorial inter- and context-dependent relationship between these modifications and highlight the need for more mechanistic studies.

2.4 Materials and Methods:

2.4.1 Strains and Growth conditions

The strains used in this report are WT (FGSC 2489), Δ*kmt2/*Δ*set*-1 (XB140-10) and Δ*kmt1/*Δ*dim5* (XB18-11) and were described previously [121, 243]. Briefly, Δ*kmt2/*Δ*set*-1 and Δ*kmt1/*Δ*dim5* were made by the *Neurospora* Knockout Consortium and obtained from the Fungal Genetics Stock center (FGSC) then backcrossed to FGSC2489 or FGSC4200. Neurospora was grown as follows; conidia were inoculated in 2% liquid culture media (2% LCM) consisting of 1X Vogel's salts, 2% glucose, 0.17% arginine and grown in 100 mm Petri dish overnight at 30 °C to generate mycelia mats. Plugs were cut

from the mycelia and used to inoculate flasks containing 100 ml of 2% LCM and grown for a total of 2 days at 25 °C. For DD24, cultures were grown in the light for 24 hours then transferred to constant dark for 24 hours and harvested. The LP30 samples were grown in a similar fashion but just prior to harvesting, the cultures were placed in constant light for 30 min. Depending on the experiment the tissue was either immediately snap frozen or crosslinked with 1% formaldehyde.

2.4.2 RNA-seq Sample preparation

Frozen tissue from WT, Δ*kmt2/*Δ*set*-1 and Δ*kmt1/*Δ*dim-5* grown in constant darkness for 24 hours or subject to a 30-minute light pulse was ground to a fine powder in the presence of liquid nitrogen. A small fraction of the ground mycelia was used to extract total RNA using Trizol (Invitrogen) following the manufacturer's instructions. The total RNA from 2 biological replicates was sent to the Columbia Genome Center for library preparation and RNA-seq. Ribosomal RNAs were depleted and the remaining RNA was for cDNA library preparation using the TrueSeq library preparation kit version 2. A total of 60 million 100bp paired-ended reads were sequenced on the Illumina 2500 instrument. The RNA-seq is deposited in Gene Expression Omnibus (GEO) with the accession number, GSE121356.

2.4.3 RNA-seq Analysis

I used HISAT2 (version 2.1.0) and StringTie (version 1.3.3b) [263] for mapping and transcripts discovery. The 100bp pair-end reads were mapped to the *Neurospora* NC12 reference genome using the existing GTF file as a guide with the default settings [264]. The resulting *Neurospora* transcriptome for each sample and time point was

assembled and merged with the existing GTF file using StringTie to generate a new GTF file containing novel transcripts. Gffcompare (version v0.10.1) was used with the 2 gtf files to determine how many assembled transcripts matched annotated genes. NATs and lincRNAs were determined based on class code; NATs were called if the class code was x or s, then filtered by i and o determine overlap and direction; lincRNA were called if the transcript had a class code:u. Next, I used Cuffdiff (version v2.2.1) and cummeRbund (version 2.20.0) [265, 266] for differential expression and statistical analysis. Genes that had a change in expression between light and dark or between strains under identical conditions were determined using the getSig function in CummeRbund ($q<0.05$) [267]. Transcripts visualization was done using Integrative Genomics Viewer (IGV)(version 2.3.94) [268].

2.4.4 Chromatin Immunoprecipitation

The ChIP experiments were performed on WT, Δ*kmt2/*Δ*set-1* and Δ*kmt1/*Δ*dim-5* at DD24 and LP30 using two biological replicates. The ChIP experiments followed our general laboratory protocol. Tissue was crosslinked with 1% formaldehyde for 10 min then quenched with 0.1 M Glycine for 10 min. Crosslinked tissue was ground with a mortar and pestle, and the lysates were suspended in 10 mL of ChIP lysis buffer (0.05 M Hepes (pH 7.4), 0.15 M NaCl, 0.001 M EDTA, 1% Triton X-100, 0.1% SDS) containing protease inhibitors (2.0 µg/mL leupeptin, 2.0 µg/mL pepstatin A, 1.0 mM PMSF). Complete cell lysis and gross chromatin shearing was achieved by sonicating 2 times at 20% power with a microtip. The resulting lysates were cleared of cellular debris by centrifugation at 2,000 × g for 10 min. 200 µl of supernatant was transferred to a 1.5 ml polystyrene tube and the chromatin was sheared to an average size of 500 bp using a

Misonix cup sonicator 6-8 times at 20% power. The ChIPs were performed with antibodies specific to H3K9me3 (Abcam #ab8898) and H3K4me3 (Abcam #ab8580). Each ChIP sample contained approximately 10 µg of lysate and 0.025 µg antibody prebound to Protein A magnetic beads. The antibody, lysates and beads were incubated overnight at 4ºC and washed 4 times with RIPA buffer then eluted with 0.1M sodium bicarbonate containing 1% SDS. The crosslinks were reversed by heating at 65 ºC and the protein was removed by the addition of proteinase K. The DNA was further purified by phenol chloroform extraction and ethanol precipitation. The purified DNA was sent to Beijing Genome Institute (BGI) for library preparation and sequencing. The samples were sequenced to a depth of 50 million 50-bp single-end read. The ChIP-seq data is deposited in GEO and can be found under GSE121356.

2.4.5 ChIP-seq data analysis

ChIP-seq reads were mapped to NC12 using Burrows-Wheeler Aligner (BWA) (version 0.7.10) [269]. The resulting Bam files were processed using macs2 call peak function [270]. For H3K9me3, I used the broad peaks options with the H3K9me3 ChIP from Δ*kmt1/*Δ*dim-5* DD as the control file for background subtraction. For H3K4me3 ChIP, we used the default setting and the H3K4me3 ChIP on Δ*kmt2/*Δ*set-1* as the control file. Next, I used Diffbind (version 2.6.6) [271]to determine H3K9me3 and H3K4me3 enrichment between the different strains and conditions and visualized the data with the MA plot (dba.plotMA) function ($p<0.05$). ChIPseeker (version 1.14.1) [272] was used to examine coverage of H3K9me3 and H3K4me3 over chromosomes using the *covplot* function. The *TagHeatmap* function was used to generate the heatmap relative to the TSS plus or minus 2000bp. In addition, I used deepTools2 [273] plotHeatmap function to

create additional heatmaps. RnaChipIntegrator was used to determine genes that are close to ChIP-seq regions using the -edge=both function within 2 kb of the TSS and TES. I used *bedtools intersect* (version v.2.25.0) [274] to determine the overlap of H3K4me3 and H3K9me3.

2.4.6 Sequential ChIP

The sequential ChIP followed the standard ChIP protocol described above with minor modifications. First, the H3K4me3 and H3K9me3 antibodies were covalently attached to BioMag carboxyl microparticles (Bangs Laboratories) using EDAC following manufacturer's guidelines. After coupling, beads were titrated to find the optimal antibody-bead to lysate ratio. The primary ChIP used 10 times the normal amount used in a standard ChIP (approximately 50 µg of lysate) and 50 µl beads. After the initial ChIP, the chromatin was eluted with 2 x 25 µl (0.1 M $NaHCO_3$, 1.0% SDS) by heating at 37^{o}C for 10 min. The pH of eluates was adjusted to near neutral pH by the addition of 6 µl of 1.0 M Tris pH 6.5. The SDS was lowered to 0.1% for the second ChIP by the addition of 10 times the eluate volume of ChIP Lysis Buffer with no SDS (500 µl). At this stage, 1/10 the initial ChIP was removed and the remaining was subjected to a second ChIP with 10 µl beads containing the second antibody. For the sequential ChIP I performed reciprocal reactions to fully validate the findings; In one re-CHIP, H3K4me3 was used first and H3K9me3 second. Simultaneously, separate ChIP was performed using H3K9me3 first and H3K4me3 second. In total, the sequential ChIP was performed 3 independent times on 3 separate biological replicates in reciprocal duplicates.

2.5 Results

2.5.1 Role of KMT1/DIM-5 and KMT2/SET-1 in transcriptional regulation.

In order to understand the role of KMT1 and KMT2 in the Neurospora gene regulation, I performed a comprehensive RNA-sequencing (RNA-seq) on *Δkmt1/Δdim-5* and *Δkmt2/Δset-1* compared to WT with RNA isolated from mycelia grown in the dark for 24 hours (DD24) and after a 30-minute light exposure (LP30). The RNA-seq was performed on ribo-depleted RNA and the DNA strand was preserved during library preparation to identify lncRNA, NATs and to monitor potential spurious transcription in the mutants. I used HISAT2 and StringTie for mapping and transcripts identification [263] **(Supplemental Figure 2-1a)**. HISAT2 and StringTie were chosen because other mapping and transcript discovery algorithms were unable to identify the *frq* NAT, *qrf,* and this was a minimum requisite to ensure confidence in transcript discovery. In total, I identified 21,475 isoform level transcripts; 10,784 are annotated protein-coding genes and 10,692 were transcripts that are not present in the NC12 annotation. Further predictions indicated there are 2444 NATs and 2268 long intergenic noncoding RNA (lincRNA) (**Figure 2-1a)** from the unannotated transcripts. The absolute number of lincRNAs and NATs is slightly larger, but consistent with previous reports and the small differences are likely due to different stringency constraints used during identification [275, 276]. I also performed multidimensional scaling (MDS) on the replicates to gage the variance among the strains and conditions and found the strain background contributed more to the variance than changes in environmental conditions **(Supplemental Figure 2-1b)**. Next, I examined the expression level distribution among

the different transcript types in WT under DD and LP30 and found that NATs and lincRNA are typically expressed at a much lower level than annotated protein-coding genes **(Figure 2-1b)**. Similar results were found for *Δkmt1/Δdim-5* **(Supplemental Figure 2-2a)** and *Δkmt2/Δset-1* **(Supplemental Figure 2-2b)**, and I did not detect major global differences among the strains (data not shown).

To examine the defects in transcription arising from loss of H3K4 and/or H3K9 methylation in the light, I compared the expression differences of *Δkmt1/Δdim-5* and *Δkmt2/Δset-1* to WT. Hierarchical clustering of light-regulated genes in each strain revealed many were not activated in *Δkmt1/Δdim-5* or *Δkmt2/Δset-1* while many light-repressed genes became activated ($q<0.05$) **(Figure 2-1c**). Independent clustering of *Δkmt1/Δdim-5* and *Δkmt2/Δset-1* revealed analogous result where subsets of genes that became unresponsive while different sets became responsive (**Supplemental Figure 2-3a & b**). Comparison of light-activated genes in WT, *Δkmt1/Δdim-5* and *Δkmt2/Δset*-1 revealed that 144 transcripts required both KMT1/DIM-5 and KMT2/SET-1, while 43 and 19 became light-activated in *Δkmt1/Δdim-5* and *Δkmt2/Δset-1,* respectively **(Figure 2-1d)**. In contrast, 170 of 220 light-repressed genes in WT required both KMT1/DIM-5 and KMT2/SET-1 **(Figure 2-1e)**. Among the light-activated genes in WT, 12.5% are predicted to be lincRNA and 3% are light-activated NATs **(Supplemental Figure 2-3c)**. Collectively, the data indicate that although the light response is still largely intact, a subset of genes lose proper regulation in the absence of KMT1/DIM-5 or KMT2/SET-1 while other genes become light-activated or light-repressed in *Δkmt1/Δdim-5* or *Δkmt2/Δset-1*.

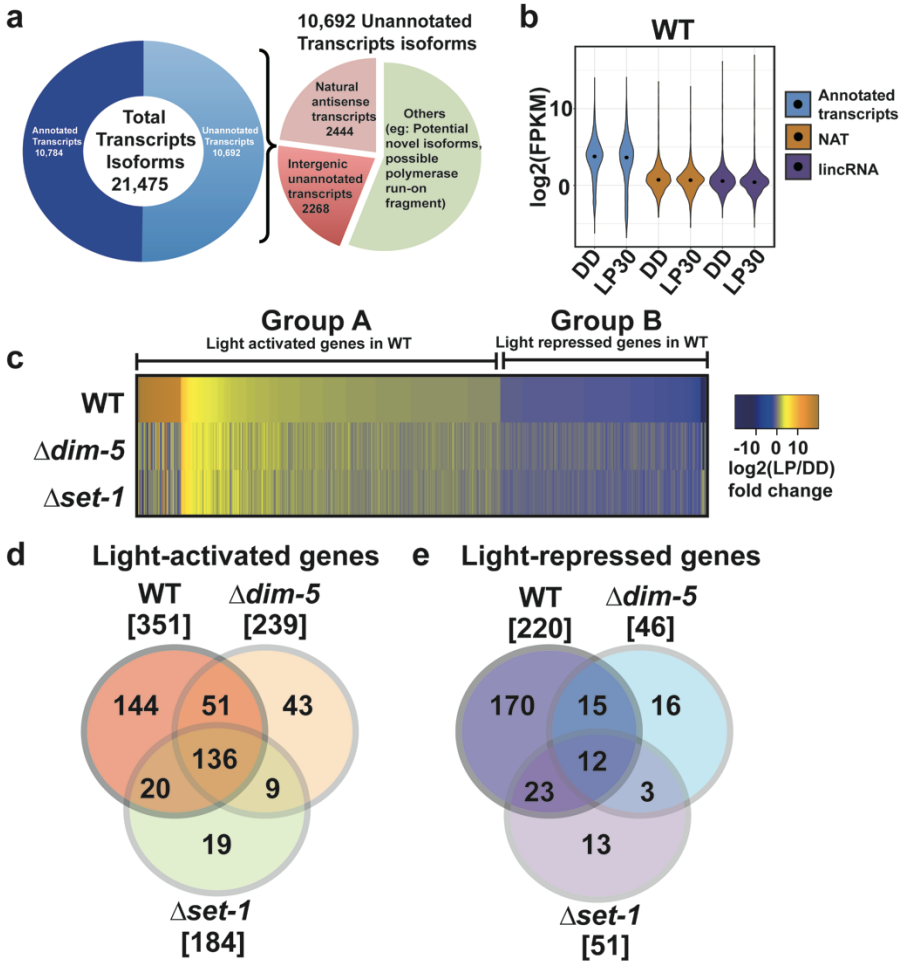


Figure 2-1 RNA-seq analysis in WT, Δ*kmt1*/Δ*dim-5* and Δ*kmt2*/Δ*set-1*.

(a) Graphical representation of identified transcripts. In total, there were 21,475 unique transcripts classified as 10,784 previously annotated protein-coding transcripts and 10,692 unannotated transcripts, that were further subdivided into 2444 natural antisense transcripts and 2268 predicted lincRNAs. (b) Violin plot depicting the expression levels (FPKM, log2) of transcripts belonging to annotated transcripts, NATs and lincRNAs. The black dots indicate median. (c) Heatmap showing the clustering of light-activated and light-repressed genes in WT ($q<0.05$) and compared to the same genes in Δ*kmt1*/Δ*dim-5* and Δ*kmt2*/Δ*set-1*. Group A includes genes that were light-activated in WT while Group B represents genes that were light-repressed. Relative expression levels are shown as log2 fold change of LP vs. DD as indicated in the range bar. The overlap among (d) light-activated ($q<0.05$) or (e) light-repressed ($q<0.05$) genes are represented as Venn diagrams

2.5.2 KMT1/DIM-5 and KMT2/SET1 are co-activators and co-repressors.

Next, I sought to better define how loss of H3K4 or H3K9 methylation changed the underlying steady-state expression in the dark and after 30 min light. I did differential analysis to find genes that had expression differences ($q<0.05$) in both the dark and light for Δ*kmt1/*Δ*dim-5* and Δ*kmt2/*Δ*set-1* relative to WT. The changes in gene expression relative to WT in DD24 and LP30 are pronounced for both Δ*kmt1/*Δ*dim-5* **(Figure 2-2 a & b)** and Δ*kmt2/*Δ*set-1* **(Figure 2-2c & d)**. Overall, I found that for some genes, KMT1/DIM-5 and KMT2/SET1 are needed for repression, and at other loci, they are needed for co-activation. Next, I explored the extent of overlap among the misregulated transcripts in Δ*kmt1/*Δ*dim-5* and Δ*kmt2/*Δ*set-1* and found that 137 transcripts were elevated in both Δ*kmt1/*Δ*dim-5* and Δ*kmt2/*Δ*set-1* relative to WT in the dark **(Figure 2-2e)**, while 140 transcripts are elevated in the light **(Figure 2-2f)**. In contrast, 43 transcripts had reduced expression in both Δ*kmt1/*Δ*dim-5* and Δ*kmt2/*Δ*set-1* in the dark **(Figure 2-2g)**, while 61 were reduced at LP30 **(Figure 2-2h)**. To further define the gene expression pattern among all conditions and within the mutants, I used K-means clustering. I arbitrarily clustered into 12 groups (determined empirically) reasoning this gave a good a cross-comparison of the 3 strains under both conditions **(Supplemental Figure 2-4)**. Overall, many genes in Δ*kmt1/*Δ*dim-5* and Δ*kmt2/*Δ*set-1* maintained a similar pattern to WT (Clusters 1, 2, 4, 5, & 7), while others were marginally or dramatically affected in Δ*kmt1/*Δ*dim-5* or Δ*kmt2/*Δ*set-1* (Clusters 3, 6, 8-12). Specifically, cluster 8 contained light-activated transcripts that are de-repressed in both Δ*kmt1/*Δ*dim-5* and Δ*kmt2/*Δ*set-1* but remain light responsive. Cluster 9 was de-repressed in Δ*kmt1/*Δ*dim-5* while clusters 3 and10 were de-repressed in both strains. In contrast,

cluster 11 required *kmt1/dim-5* for normal expression, while cluster 6 required *kmt2/set-1* for normal expression. Of note, cluster 12 contained genes that were de-repressed in Δ*kmt2*/Δ*set-1*.

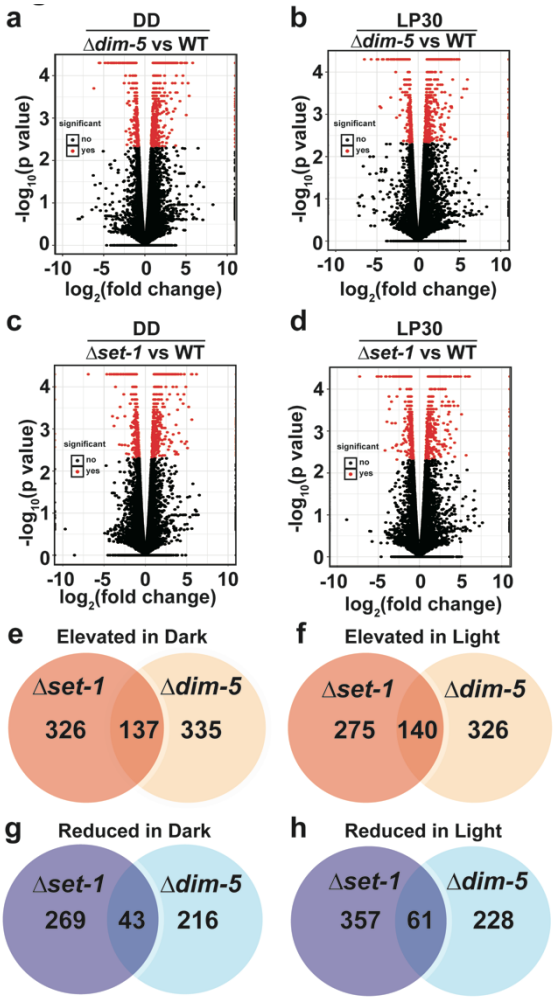


Figure 2-2 KMT1 and KMT2 are needed for activation and repression.

Volcano plots indicating the expression changes in (a) Δ*kmt1*/Δ*dim-5* grown in the dark (DD) or (b) in the light. The red dots represent genes that have altered expression in the mutant relative to WT ($q<0.05$). Same as in a and b, except I examined expression changes in Δ*kmt2*/Δ*set-1* relative to WT in DD (c) or LP30 (d). (e-h) Venn diagrams show the amount of overlap among the misregulated genes. (e) Genes that are elevated in the dark in both Δ*kmt2*/Δ*set-1* (orange) or Δ*kmt1*/Δ*dim-5* (yellow) in DD. (f) Genes elevated in the light in in Δ*kmt2*/Δ*set-1* (Orange) or Δ*kmt1*/Δ*dim-5* (yellow). (g) Genes that have reduced expression in the dark *kmt2*/Δ*set-1* (dark blue) and Δ*kmt1*/Δ*dim-5* (light blue). (h) Genes that have reduced expression in the light in Δ*kmt2*/Δ*set-1* (dark blue) and Δ*kmt1*/Δ*dim-5* (light blue).

2.5.3 Genome-wide distribution of H3K9me3 and H3K4me3.

To further examine the connection between H3K9me3 and H3K4me3, and how these modifications influence expression, I performed H3K9me3 ChIP-seq in WT (DD and LP30) and Δ*kmt2/*Δ*set-1* (DD and LP30) using Δ*kmt1/*Δ*dim-5* DD for background subtraction. Likewise, I performed H3K4me3 ChIP-seq in WT (DD and LP30) and Δ*kmt1/*Δ*dim-5* (DD and LP30) using Δ*kmt2/*Δ*set-1* DD as background. I identified broad peaks enriched with H3K9me3 and narrow peaks with H3K4me3 using macs2 ($p<0.1$, the macs2 cutoff used by ECNODE). The broad peak option was chosen for H3K9me3 because the vast majority of H3K9me3 occurs in constitutive heterochromatin domains that are relics of repeat induced point mutations (RIP) [258, 277]. In contrast, H3K4me3 is typically restricted to nucleosomes near the transcriptional start site and deposited in more localized regions [244, 251]. Consistent with this, I found H3K9me3 was enriched at centromeres and constitutive heterochromatin, but H3K4me4 was absent from those same regions in all conditions (data not shown).

To further examine the distribution of H3K9me3 and H3K4me3, ChIP-seq peak intensities were clustered for all transcripts identified including the lincRNAs and NATs and spliced isoforms. I clustered H3K9me3 and H3K4me3 into 4 groups (determined empirically) to represent our findings and included 2 kb of DNA upstream and downstream of the transcriptional start site (TSS) and transcriptional end site (TES), respectively. As expected, H3K9me3 was largely absent from expressed genes and 2 of the clusters had no associated peaks (data not shown). However, in the other 2 clusters, H3K9me3 was present in the gene body of a small subset of loci **(Figure 2-3a, black arrow)** and there were also peaks upstream of the TSS. In total the cluster I contained

1348 transcripts. We also found H3K9me3 downstream of the TES **(Supplemental Figure 2-5)**. The upstream and downstream peaks represent transcripts adjacent to constitutive heterochromatin. In addition, there was also a small subset of genes that contained H3K9me3 in the gene body **(Figure 2-3a)**. Consistent with a different distribution, H3K4me3 tended to cluster at nucleosomes proximal to the TSS **(Figure 2-3b,** cluster I, 6076 transcripts**)**. There were also instances where H3K4me3 was enriched upstream of the TSS (cluster II, 2604 transcripts) and downstream of the TES (cluster III, 2936 transcripts). The second 2 clusters have H3K4me3 in upstream genes (cluster II) and downstream genes (cluster III) due to the compact nature of the Neurospora genome where most genes are within 2 kb of an adjacent gene. The remaining H3K4me3 cluster was largely devoid of H3K4me3 (data not shown). I next sought to examine the extent of H3K9me3 and H3K4me3 spreading in either *Δkmt2/Δset-1* or *Δkmt1/Δdim-5*. Alternative clustering revealed a minor trend where H3K9me3 spread into gene bodies in *Δkmt2/Δset-1*, but there was no evidence of H3K4me3 spreading in Δkmt1/Δdim-5 **(Supplemental Figure 2-6)**. The spreading of H3K9me3 into euchromatin regions in the absence of H3K4 methylation supports the notion that K4 methylation helps establish chromatin boundary elements and is consistent with the role of KMT2/SET-1 as an antisilencing factor [247].

2.5.4 Interdependent relationships between H3K4me3 and H3K9me3.

I next examined if there were any interdependent relationships between H3K4me3 and H3K9me3, which if present, would presumably be opposing effects. To accomplish this, I examined how H3K9me3 changed in Δ*kmt2/*Δ*set-1*, and how H3K4me3 changed in Δ*kmt1/*Δ*dim-5* relative to WT in the dark and light. I began by examining H3K9me3 in

WT compared to Δ*kmt2/*Δ*set-1* in DD and LP30 **(Figure 2-3c & d)**. In the dark, I found 440 H3K9me3 peaks changed between WT and Δ*kmt2/*Δ*set-1* **(Figure 2-3c)**. Among them, 199 peaks have higher H3K9me3 levels in the absence of K4 methylation, whereas 241 peaks have lower H3K9me3 levels in Δ*kmt2/*Δ*set-1*. Similar results were found in the light where 210 peaks changed between WT and Δ*kmt2/*Δ*set-1* **(Figure 2-3d)**. Among those, 165 peaks have higher H3K9me3 density in Δ*kmt2/*Δ*set-1*, compared to 44 peaks that have lower H3K9me3 density. The decrease or loss of H3K9me3 in the absence of KMT2/SET1 indicates that H3K4me3 is a requisite modification (either direct or indirect) for facultative heterochromatin including *frq* (see below). However, much of the increase may be due to spreading of H3K9me3 into euchromatin in the absence of H3K4 methylation as determined above.

Next, I examined how H3K4me3 changed when H3K9 methylation was absent. I did similar differential analysis comparing Δ*kmt1/*Δ*dim-5* relative to WT in the dark and light. I found 1530 H3K4me3 peaks changed between WT and Δ*kmt1/*Δ*dim-5* in the dark. Among those, 537 peaks have higher H3K4me3 levels in Δ*kmt1/dim5* DD, whereas 993 peaks have lower H3K4me3 levels **(Figure 2-3e)**. In the light, a total of 2227 peaks changed intensity between WT and Δ*kmt1/*Δ*dim-5*. Among those, 1415 peaks have higher H3K4me3 density in Δ*kmt1/*Δ*dim-5* and 812 peaks had a decrease in H3K4me3 **(Figure 2-3f)**.

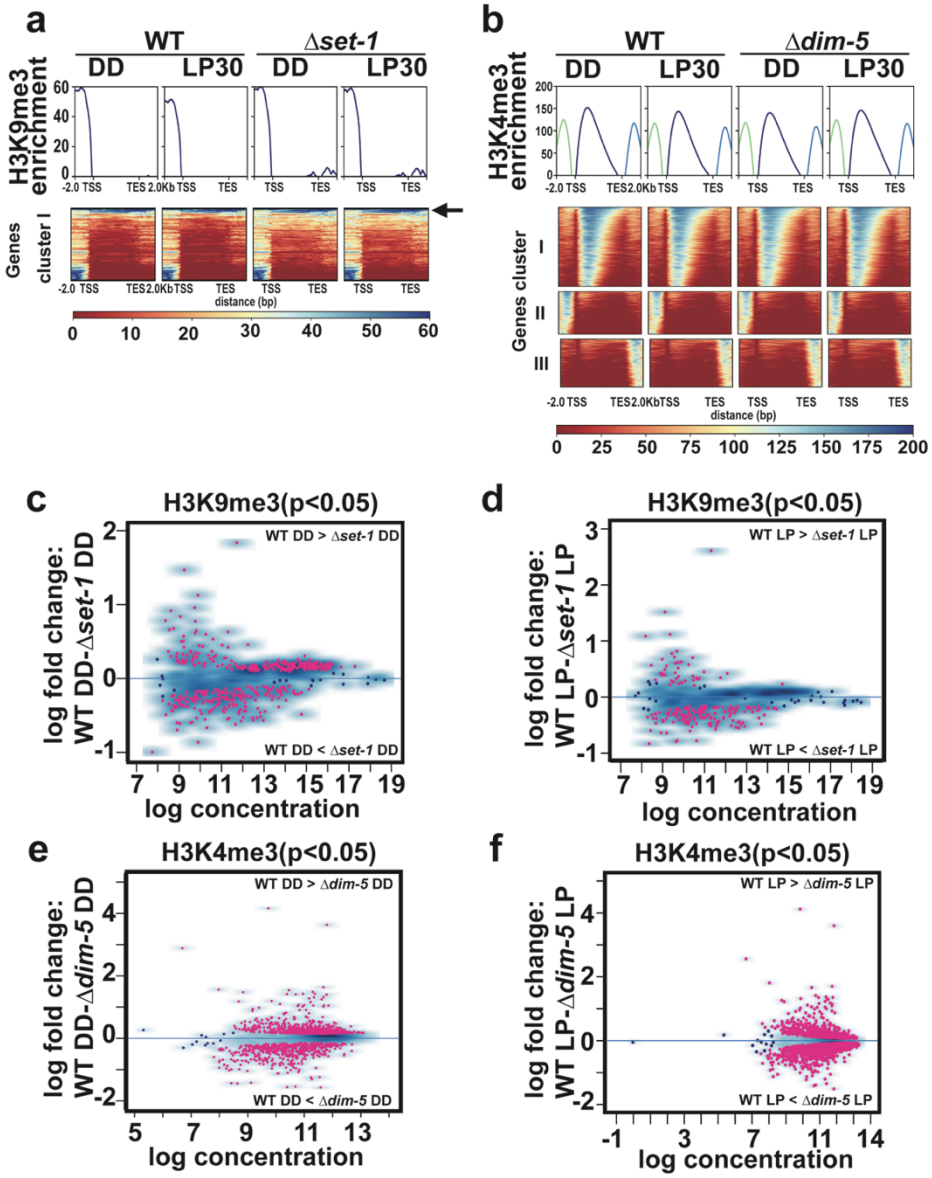


Figure 2-3 Global profile of H3K9me3 and H3K4me3.

The heatmaps display ChIP enrichment of (a) H3K9me3 for a small cluster of genes in WT and Δ*kmt2/*Δ*set-1* and (b) H3K4me3 of three independent clusters of genes in WT and Δ*kmt1/*Δ*dim-5*. Regions included 2kb upstream of TSS and downstream of TES. The MAplots show quantitative difference in H3K9me3 peak densities between WT and Δ*kmt2/*Δ*set-1* in (c) DD and (d) LP30. (e-f) Same as in c & d except differences in H3K4me3 peak densities were between WT and Δ*kmt1/*Δ*dim-5*. The red spots on the MA plots indicate a log fold change ($p < 0.05$).

In order to examine the expression pattern of transcripts that had a decrease in H3K9me3 in Δ*kmt2/*Δ*set-1,* I extracted the peaks with a reduced H3K9me3 (log fold change ≥0.3) in Δ*kmt2/*Δ*set-1* and identified the corresponding transcripts based on proximity to the H3K9me3 peak. Overall there were 141 transcripts in the dark and 119 transcripts in the light that had a decrease in H3K9me3 when H3K4 methylation was missing **(Figure 2-4a)**. Of the 260 transcripts that had a decrease in H3K9me3 in Δ*kmt2/*Δ*set-1*, the majority had no change in expression; however, there was a small subset that appeared to be de-repressed due to loss of heterochromatin in Δ*kmt2/*Δ*set-1* (19 in the light and 21 in the dark). I classified these as genes having KMT2/SET-1-dependent heterochromatin. Also, six genes (2 in the DD and 4 in the LP) appear to have reduced expression when H3K9me3 is lost in Δ*kmt2/*Δ*set-1* **(Supplemental Figure 2-7a-c)**.

We previously showed H3K4me3 functions as a repressive modification at *frq* and KMT2/SET-1 is needed for circadian negative feedback and DNA methylation at *frq* [121]. To gain insight into a possible mechanism of the repression, I focused on the spatial distribution of both modifications within *frq* and examined how the marks changed in the mutants relative to WT. Examination of the RNA-seq, showing the 3 known transcripts that originate from *frq* (*frq, qrf* and a small upstream transcript that spans the c-box coined *frq-mini*) and the ChIP-seq showing H3K4me3 and H3K9me3 is contained in **Figure 2-4b**. Also included are WC-2 ChIP-seq (SRX015820) and disiRNA-seq (GSE21175) to provide more context to the chromatin regulation and spatial distribution [262, 278]. In WT, H3K9me3 was largely confined to nucleosomes near the c-box within the *frq-mini* gene body. In response to light, there was an increase

in H3K9me3 that corresponded with light-activated expression of *frq-mini.* In contrast, H3K4me3 appeared restricted to the nucleosome(s) proximal to the *qrf* TSS and was not contained within the nucleosome proximal to *frq*. As predicted, I also found that H3K9me3 is completely absent in Δ*kmt2/*Δ*set-1* ($p<0.005$). In contrast, H3K4me3 at the *qrf* TSS appeared to increase in the absence of Δ*kmt1/*Δ*dim5,* and in response to light ($p<0.005$). Interestingly, H3K4me3 and H3K9me3 appear spatially separated by approximately 5 kb suggesting more complex mechanism than just H3K4me3-dependent H3K9me3, unless there is a long-range chromatin contact due to gene looping. Similar to a previous report, I did not find any change in *qrf* expression; however, *frq-mini* appeared to be elevated in both Δ*kmt2/*Δ*set*-1 and Δ*kmt1/*Δ*dim-5* **(Supplemental Figure 2-8)**.

In addition to *frq,* the divergently transcribed genes NCU02913 and NCU16472 also had KMT2/SET-1-dependent H3K9me3 **(Figure 2-4c)**. In WT, there were minor hints that TSS proximal nucleosomes are K4/K9 bivalent domains because H3K9me3 and H3K4me3 are both present and overlap defined regions. Further support can be found in the observation that in the absence of H3K4 methylation, H3K9me3 was lost and in the absence of H3K9 methylation, the nucleosome(s) proximal to one of the genes (NCU16472) had a substantial increase in H3K4me3 ($p<0.005$). In this instance, both genes had increased expression in Δ*kmt2/*Δ*set*-1 and Δ*kmt1/*Δ*dim-5* revealing that loss of H3K9me3, either through KMT2/SET-1-dependent H3K9me3 or by deleting *kmt1/dim-5* caused de-repression **(Figure 2-4d)**.

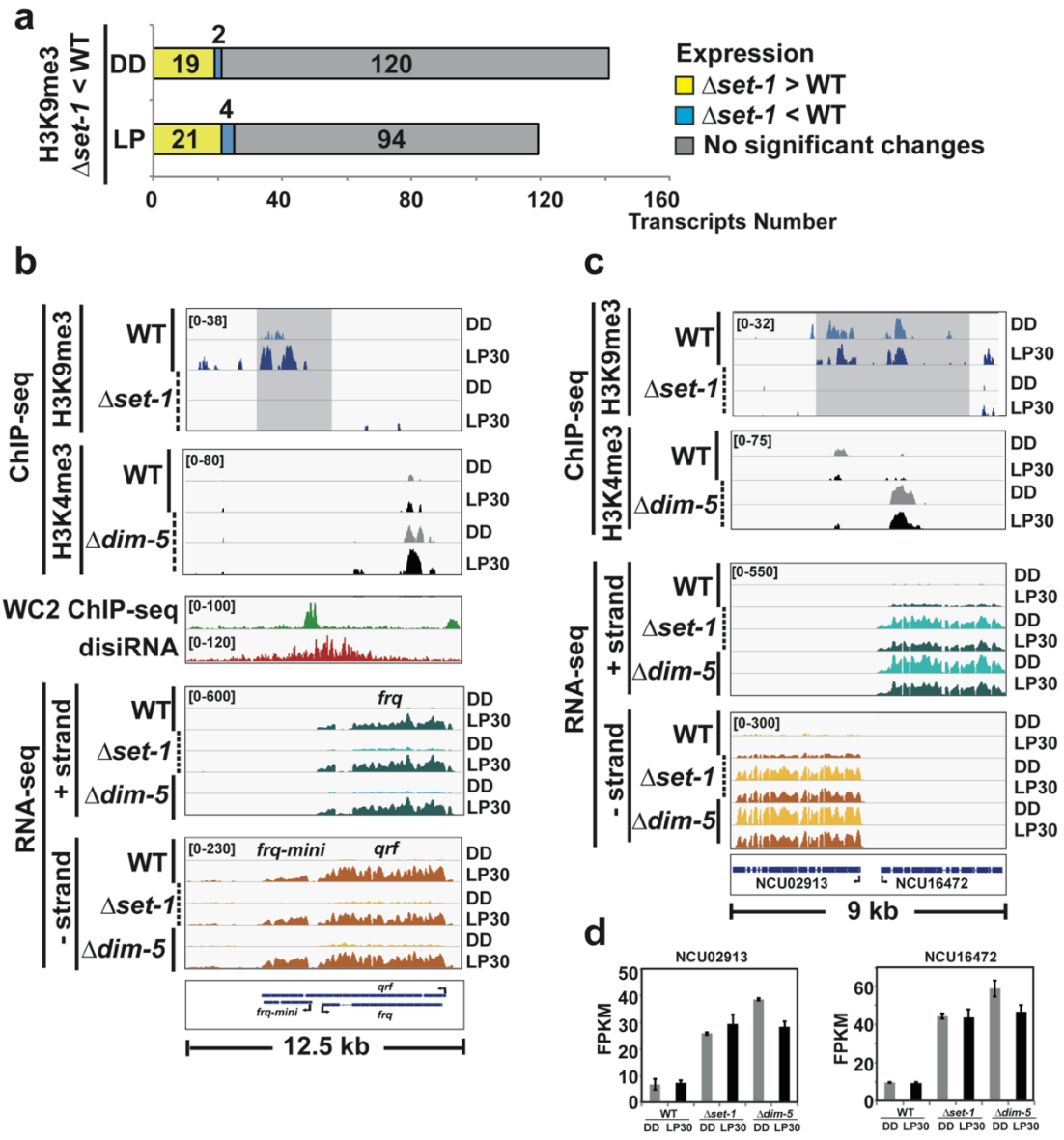


Figure 2-4 Examples of KMT2/SET1 are needed for H3K9me3.

(a) Stacked bar plot shows the number (x-axis) of transcripts in proximity to an H3K9me3 peak that had a decreased H3K9me3 in Δ*kmt2*/Δ*set-1* ($p<0.05$). Corresponding transcripts were catalogued based on whether there was a change in expression relative to WT ($q<0.05$). Yellow represents elevated expression in Δ*kmt2*/Δ*set-1,* blue indicates a lower expression and grey are transcripts with no significant change. (b) Gene-level plot of ChIP-seq and RNA-seq of the *frq* locus. H3K9me3 peaks from the ChIP-seq done in WT and Δ*kmt2*/Δ*set-1* along with H3K4me3 peaks from ChIP performed in WT and Δ*kmt1*/Δ*dim-5*. The growth conditions (DD or LP30) are shown to the right. RNA-seq data is

separated to display transcripts that originate from the plus (+) or minus (-) strand and are separated for the 3 different strains under each condition. As an added reference WC-2 ChIP-seq and disiRNA at *frq* are also shown. (c) Same as in B except the diagram shows the divergently transcribed genes NCU02913 and NCU16472. (d) Bar plot showing the expression changes of NCU02913 and NCU16472 in WT, Δ*kmt2/*Δ*set-1* and Δ*kmt1/*Δ*dim-5* in DD and LP30.

I next examined how the reduction or loss of H3K4me3 in Δ*kmt1/*Δ*dim-5* affected expression. Overall, there were minor effects on genes that had a decrease in H3K4me3 upon loss of H3K9 methylation, but the vast majority showed no significant change in expression **(Figure 2-5a)**. Nevertheless, there were 62 transcripts in the dark and 66 in the light that had reduced H3K4me3 levels and lower expression in Δ*kmt1/*Δ*dim-5.* Transcripts that appeared to require KMT1/DIM-5 for normal elevated expression suggest KMT1/DIM-5 can theoretically function as a co-activator. However, this comes with major caveats, because loss of H3K9me3 can cause redistribution of chromatin modifications including Histone H3 lysine 27 methylation and/or other pleiotropic effects [277, 279]; plus, in general, genes that required KMT1/DIM-5 for expression tended to fall between regions of constitutive heterochromatin, so I cannot rule out other mechanisms affecting expression. A typical example of this occurred in a region containing three genes (NCU10058, NCU10059 and NCU16377) located between 2 constitutive heterochromatin domains **(Figure 2-5b)**. In this instance, all 3 genes had a reduction in H3K4me3 levels in Δ*kmt1/*Δ*dim5* and all had a corresponding reduction in expression. These results are notable for a couple of reasons; first, expression appears lower in Δ*kmt2/*Δ*set*-1 but lowest in Δ*kmt1/*Δ*dim-5* **(Figure 2-5c)** and second, two of the genes are convergent transcripts, but do not appear to give rise to disiRNA (not shown). Thus, I surmise the likely conclusion is that reduction in H3K4me3 and loss of

expression may be due to spreading or redistribution of other chromatin modifications such as H3K27 methylation, and/or that flanking constitutive heterochromatin is necessary to maintain localized chromatin states.

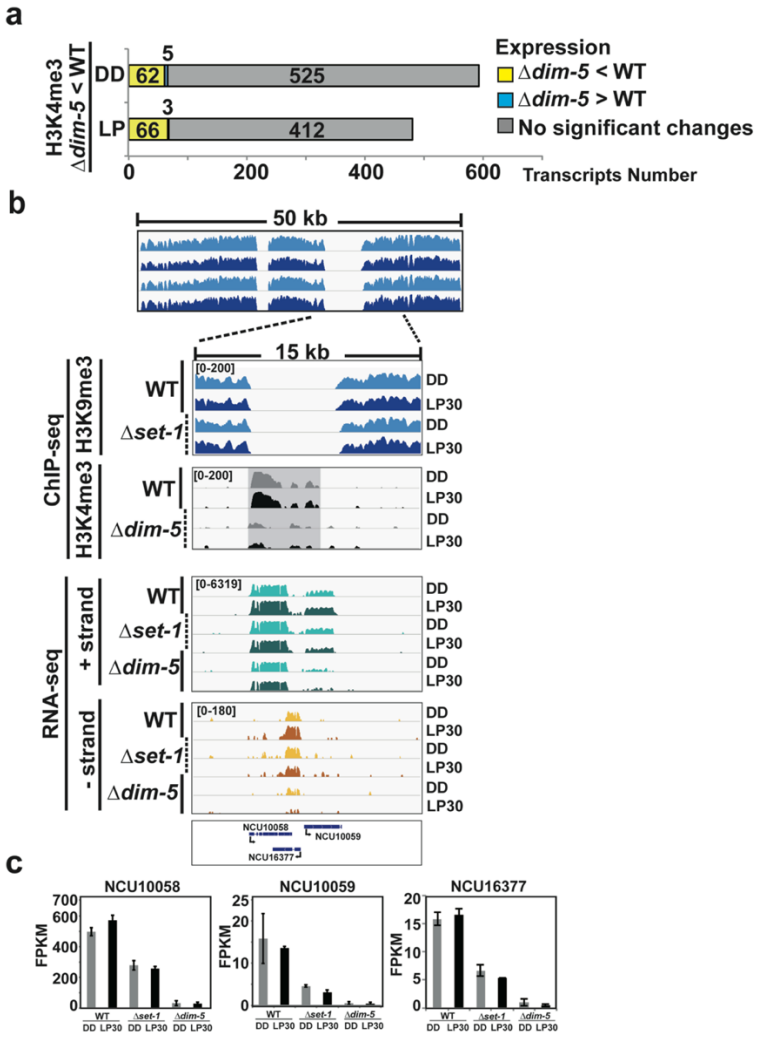


Figure 2-5 Loss of KMT1/DIM5 affects H3K4me3 distribution.

(a) Stacked bar plot showing the number (x-axis) of H3K4me3-containing transcripts that have a decreased in H3K4me3 levels in Δ*kmt1/*Δ*dim-5* ($p<0.05$). Transcripts are catalogued by expression changes in WT versus Δ*kmt1/*Δ*dim-5* ($q<0.05$). (b) Representative gene-level plot of a locus that has a reduction in H3K4me3 in Δ*kmt1/*Δ*dim-5*. H3K9me3 and H3K4me3 ChIP-seq data is shown for WT and the corresponding knockouts. Transcripts originating for either the + or – strand were separated and plotted individually. (c) The expression level of the three genes is represented as bar plots for NCU10058, NCU10059 and NCU16377 in WT, Δ*kmt2/*Δ*set-1* and *kmt1/*Δ*dim-5* comparing expression in DD (grey) versus LP30 (black).

I next focused on genes with a co-dependent and overlapping relationship between H3K9me3 and H3K4me3. In total, I found 56 loci that contained both modifications in the dark and 77 after LP30. 36 of these loci were shared under both environmental conditions **(Figure 2-6a)**. In certain instances, H3K4me3 was lost without H3K9 methylation and H3K9me3 was lost without H3K4 methylation. One example occurred in a transposon containing NCU09968 and *sly*-1 (NCU09969), which incidentally also appears to be a clock-controlled ADV-1 target gene **(Figure 2-6b)** [131, 280, 281]. Both H3K9me3 and H3K4me3 are contained within the gene body of NCU09968 and spread into *sly-1* and both modifications are dependent on one another. Interestingly, the expression of the two divergently transcribed transcripts was only affected in the absence of KMT2/SET-1 **(Figure 2-6& d)**. To confirm the existence of K4/K9 bivalent domains, I performed sequential ChIP (re-ChIP) under reciprocal conditions using either H3K4me3 or H3K9me3 antibodies first (1^{o}), followed by the opposing antibody in the re-ChIP (2^{o}) **(Figure 2-6e)**. As an added control, I took 1/10 the first ChIP to monitor recovery **(Figure 2-6f)**. Another instance occurred at a centromeric gene on supercontig 12.2 (chromosome II) **(Supplemental Figure 2-9)**. In this instance, only the heterochromatin in the gene body that contained both H3K9me3 and H3K4me3 had the co-dependent relationship and the nearby constitutive heterochromatin was not affected. As with NCU09969 and *sly-1*, expression of NCU16628 was dependent on KMT2/SET-1.

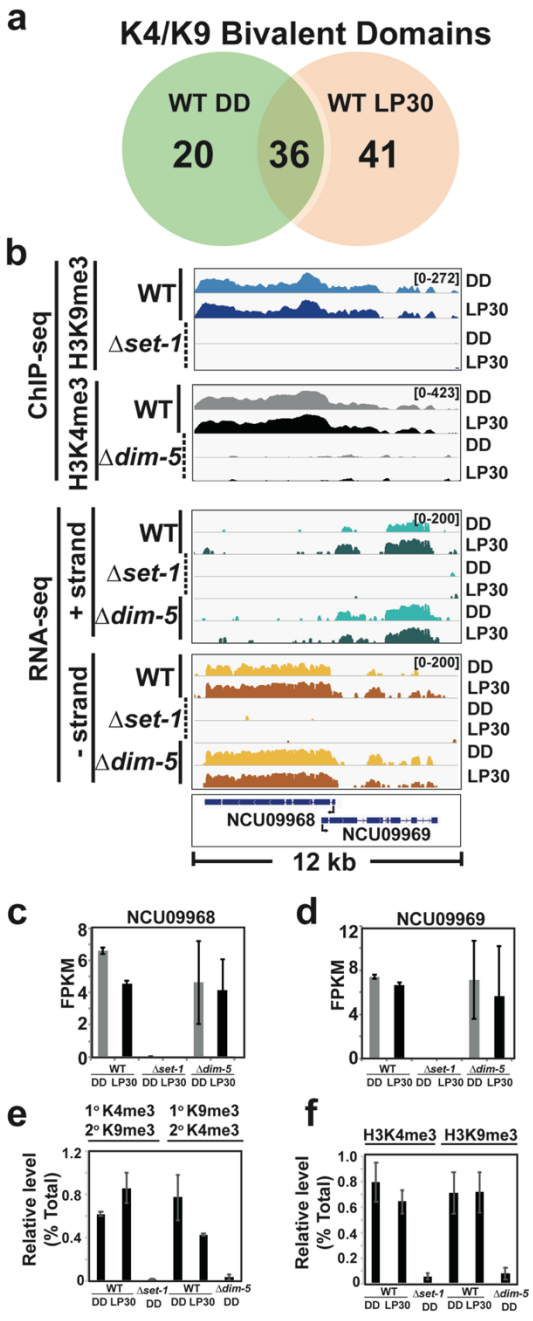


Figure 2-6 H3K4me3 and H3K9me3 can be co-dependent.

(a). Venn diagram show presumptive K4/K9 bivalent domains in WT that contain both H3K4me3 and H3K9me3 and the extent of overlap in DD and LP30. (b) The gene level plot displays H3K9me3 ChIP-seq (DD, Blue and LP30, Navy) in WT and Δ*kmt2*/Δ*set-1* and H3K4me3 ChIP-seq (DD, Grey and LP30, black) in WT and Δ*kmt1*/Δ*dim-5* at divergently transcribed NCU09968 and NCU09969. Relative expression of transcripts originating from either the plus

(Green) or minus (Orange) strands is shown for the 3 strains. NCU09969 also has an antisense transcript(s) identified by SringTie (not drawn). Expression level of (c) NCU09968 and (d) NCU09969 are shown as bar plots in WT, *kmt2/*Δ*set-1* and *kmt1/*Δ*dim-5* for DD (grey) and LP30 (black). (e) Reciprocal sequential ChIP (re-ChIP) in WT with first antibody (1^o) and second antibody (2^o) are shown above the bar plot. (f) One tenth of the primary ChIP from e.

Loss of H3K9me3 causes heightened circadian-regulated conidia formation similar to the *ras-1*^{bd} mutant and Δ*kmt1/*Δ*dim*-5 has a synthetic effect with *ras-1*^{bd} [243]. In *Neurospora*, conidia development is a light-activated process, so I wanted to examine changes in H3K9me3 in response to light in WT**.** I found that 244 loci had a reduction in H3K9me3 after 30 min light, whereas only 4 peaks (including *frq*) had an increase in H3K9me3 after 30min light in WT **(Figure 2-7a).** A specific example of light-activated loss in H3K9me3 occurred at *vvd,* which is a light-activated gene necessary for photoadaptation **(Figure 2-7b)**. Inspection of *vvd* revealed some unexpected findings, including the observation that *vvd* contains a light-activated NAT and there are two upstream overlapping lincRNAs that form a sense-antisense pair. As reported previously, *vvd* appeared to have an increase in light-activated expression in the absence of Δ*kmt2/*Δ*set*-1 or Δ*kmt1/*Δ*dim*-5 **(Figure 2-7c & d)**. Upon light activation, H3K9me3 is lost and there is a small increase in H3K4me3. Of note, H3K9me3 was unaffected by Δ*kmt2/*Δ*set-1* at *vvd*. In addition, there were 5 loci where H3K9me3 changed in response to light in Δ*kmt2/*Δ*set*-1 **(Supplemental Figure 2-10a)**. In these instances, 4 peaks had a decrease in H3K9me3 **(Supplemental Figure 2-10b)** while 1 peak had an increase in H3K9me3 **(Supplemental Figure 2-10c).**

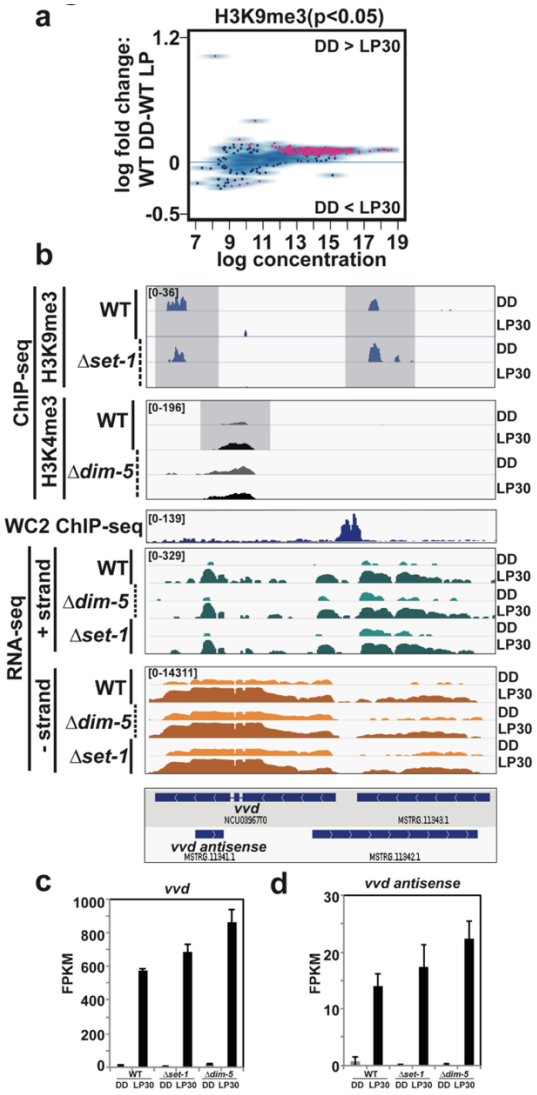


Figure 2-7 Effects of light on H3K9me3.

(a) The MA plots show quantitative difference ($p<0.05$) in H3K9me3 peak densities between WT DD and WT LP30. (b) Gene-level plot of ChIP-seq and RNA-seq of the *vvd* locus. H3K9me3 peaks from the ChIP-seq done in WT and Δ*kmt2*/Δ*set-1* along with H3K4me3 peaks from ChIP performed in WT and Δ*kmt1*/Δ*dim-5*. The growth conditions (DD or LP30) are shown to the right. RNA-seq data indicating transcripts that originate from the plus (+) or minus (-) strand are separated for the 3 different strains under the two conditions. As an aided reference WC-2 ChIP-seq is also shown. (c) Bar plot showing the expression changes of *vvd* and *vvd antisense* in WT, Δ*kmt2*/Δ*set-1* and Δ*kmt1*/Δ*dim-5* in DD and LP30.

Finally, I examined how light affects H3K4me3 in WT and found 760 H3K4me3 peaks showed a reduction in response to light whereas 31 H3K4me3 peaks increased **(Supplemental Figure 2-11)**. The finding that more transcripts lost H3K4me3 in response to light versus transcripts that gained H3K4me3 is counter to the notion that H3K4me3 is solely a mark for actively transcribed genes and may be more reflective of transcriptional memory. However, I cannot rule out that H3K4me3 nucleosomes are being disassembled to make room for transcriptional machinery. Regardless of the cause, it is clear H3K4me3, and for that matter H3K9me3, play more complex roles in transcriptional regulation than current paradigms suggest.

2.6 Discussion:

In this project, I performed a comprehensive analysis to understand how H3K9me3 and H3K4me3 impact the *Neurospora* light response. Initially, there were some *a priori* assumptions, which didn't entirely pan out as expected and instead there was a more complicated interplay between the modifications. For example, at the start, it seemed reasonable to assume that loss of KMT2/SET-1 would predominantly manifest as a defect in light-activated gene expression due to the widely-accepted premise that H3K4me3 is a mark for activation (or at the very least a mark for actively transcribed genes in euchromatin). Instead, I found that SET1 is needed for both expression and repression. In addition, I found that H3K4me3 tended to decrease after 30 minutes in the light (data not shown); a time when adaptation has solidly commenced, so it is possible the activating H3K4me3 modifications are being removed. It is also plausible that the reduction in H3K4me3 follows a general trend where the nucleosome density is decreasing at actively transcribed genes, which is consistent with the finding that nucleosome occupancy proximal to regions of light-activated genes decrease in response to light [130]. In contrast to KMT2/SET-1-dependent H3K4 methylation, I expected to find mostly de-repression in the absence of H3K9me3 because although H3K9me3 is widespread in *Neurospora*, it is largely restricted to constitutive heterochromatin domains devoid of protein coding genes, but a number of genes are de-repressed in Δ*kmt1/*Δ*dim-5,* possibly due to the redistribution of H3K27 methylation [279]. Consistent with this, I found that a number of genes are de-repressed in Δ*kmt1/*Δ*dim-5,* but I also found that KMT1/DIM-5 appears to be needed for light-activated expression **(Figure 2-1c)**. However, the extent to which these changes are pleiotropic need further examination,

because even though I found numerous genes with H3K9me3, only a small subset had altered expression in the corresponding deletion strain. Thus, the misregulation that occurs in Δ*kmt1/*Δ*dim-5* is potentially due, at least in some instances, to redistribution of other modifications like re-localization of H3K27me3 [277, 279].

The initial premise of this work was rooted in the finding that H3K4 methylation, which is normally excluded from regions containing methylated DNA, is required for DNA methylation at *frq* [282]. I was hopeful that these experiments would illuminate why H3K4me3 is needed for DNA methylation at *frq,* and potentially find other loci, but after analysis of the data, a definitive conclusion is still lacking. However, the data did reveal that H3K9me3 was completely absent from *frq* in Δ*kmt2/*Δ*set-1,* confirming the requirement for KMT2/SET-1-dependent heterochromatin. The actual mechanism of how this occurs is confounding in part due to the spatial distribution of H3K4me3 and H3K9me3 within *frq*. I found that H3K4me3 was predominantly localized in a region proximal to the TSS in *qrf*, while H3K9me3 was restricted to nucleosome surrounding the c-box **(Figure 2-4b)**. Based on the need for *qrf* in facultative heterochromatin at *frq* [114], the obvious answer would suggest that loss of H3K4me3 at *qrf* might effect *qrf* expression; however, as previously reported, *qrf* expression was not altered in *kmt1/set-1* knockout [121]. It is also possible the KMT2/SET-1-dependent heterochromatin at *frq* occurs analogous to silencing of retrotransposons in *S. pombe*, which requires SET1, but appears to be independent of H3K4 methylation [248, 249]. Some ancillary support for this can be found in the region containing *Sly-1* and *NCU09968*, which I identified as the K4/K9 bivalent domain **(Figure 2-6)**. This domain is believed to be a transposon in the WT parent strain (FGSC2489) used in this [280], but it also appears to be circadian

regulated [131] and a target of ADV-1 [281]. However, in this instance both modifications are present and contained on the same nucleosome, but expression is only dependent on H3K4me3. Another possible explanation is that the KMT2/SET-1-dependent facultative heterochromatin may be caused by secondary affects. However, when I examined the expression of known components of DCDC, DRDM, or RNAi components, I did not find any major defects, nor did I observe any other global changes in constitutive heterochromatin. Support for a direct effect comes from findings that other loci displayed similar co-dependent requirements and thus *frq* is not entirely unique. Whether or not mono- or di-methylation of H3K4, and not H3K4me3, are the requisite modifications for H3K9me3 needs to be examined further. For example, one could speculate that H3K4me1 and H3K4me2 (or other modifications dependent on H3K4 methylation) are missing in Δ*kmt2/*Δ*set-1* and these modification(s) overlap with the peak in H3K9me3. An example of this hypothesis would posit that H3K4me1, which tends to accumulate at the 3' end of genes, would be enriched at the 3' end of *qrf*, and this is close to the H3K9me3 peaks. However, this is pure speculation at this juncture and more comprehensive series of studies are needed, because even if H3K4me1 is co-localized and missing; it does not reveal the mechanism. Regardless, is remains clear that at *frq,* and 260 other loci, KMT2/SET-1 influences H3K9me3-mediated facultative heterochromatin and in some instances H3K4me3 and H3K9me3 appear as bivalent chromatin domains [283].

As it pertains to K4/K9 bivalent chromatin domains, *Neurospora* is uniquely suited to further define their role in regulation because KMT1/DIM-5 and KMT2/SET-1 are the only H3K9 and H3K4 methyltransferases and neither is essential for viability so

in theory, dissecting the importance of each modification in bivalent domains should be straightforward. However, it is far more complicated than meets the eye, because in certain instances, these modifications were dependent on one another and in many cases, loss of either didn't dramatically affect the expression of the underlying transcripts. When one considers there is also redistribution of other modifications, the system becomes sufficiently complex and would likely yield inappropriate conclusions. Thus, the only thing I can infer is that H3K4me3/H3K9me3 bivalent domains represent functional states where the modifications may be reliant on one another to establish the appropriate transcriptional program and may not necessarily be in a poised state ready for activation or repression as proposed for K4/K27 bivalent domains [284]. In addition, other combinations of modifications are likely contained within the nucleosome, beyond just H3K4me3 and H3K9me3, and these will likely contribute to the transcriptional program. For example, multivalent chromatin readers may not necessarily discern if the full-compliment of modifications are present to maintain the gene in a repressed or activated state. Whether or not this extends to classical bivalent domains that contain H3K4me3 and H3K27me3 should be considered, because the data presented here seems to suggest it is more complicated than initially perceived.

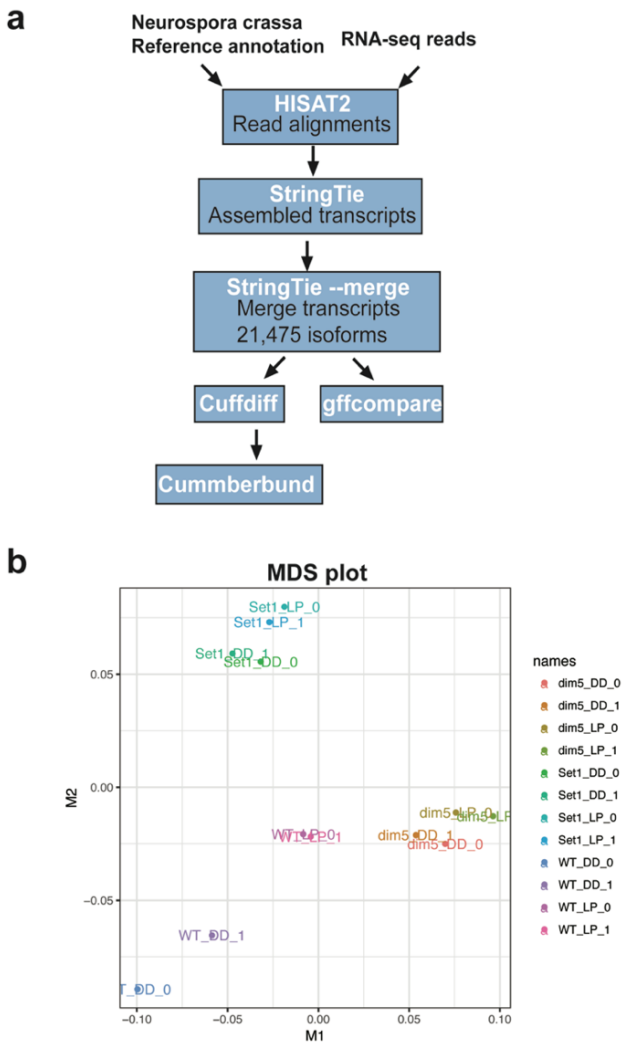
As a final note, there are less than 50 convergent loci in Neurospora that contain disiRNAs and DNA methylation [112, 262] and in this report I found 2,444 transcripts that contain NATs, which greatly increases the number of convergent overlapping transcripts. However, only a small subset contains DNA methylation and H3K9me3. Given that H3K9me3 is needed for DNA methylation [285], the majority of convergent transcripts do not give rise to disiRNA-mediated DNA methylation. So ultimately,

convergent transcripts are not universal in their ability to generate heterochromatin and in this instance, *frq* and the other disiRNA loci, appear to be special cases with unique epigenetic regulation.

2.7 Conclusions:

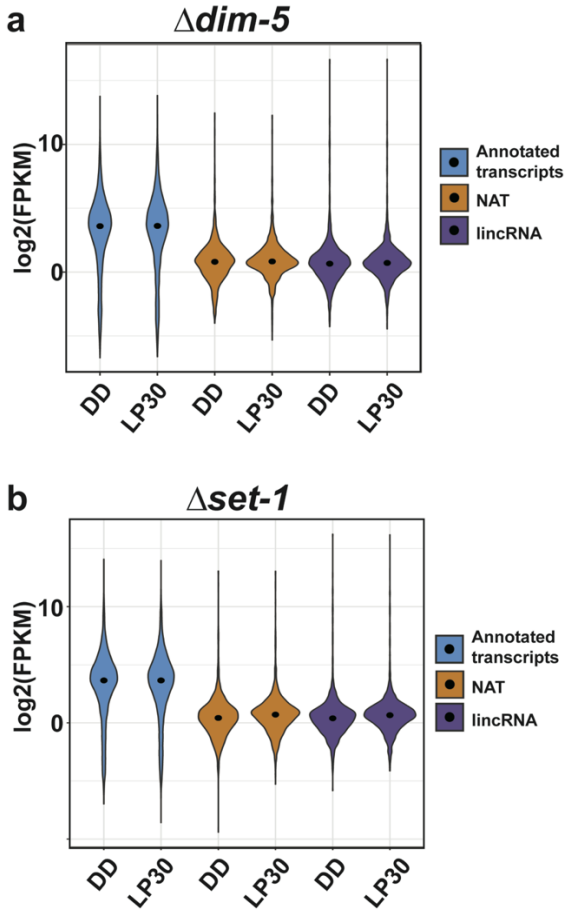
Sensing and responding to light provide organisms an adaptive advantage, in part by altering gene expression. The complement of light-activated genes in model organisms is largely known, and mechanisms of activation and photoadaptation are well-defined. To further understand how light alters post translation modifications to chromatin, I performed a comprehensive analysis looking at H3K4me3 and H3K9me3. Using a combination of RNA-seq and ChIP-seq, I found that H3K4 methylation and H3K9 methylation play an important but restricted role in light-activated gene expression. Both modifying enzymes need to be present for normal compliment of genes to be light-activated and repressed. In addition, I found KMT2/SET-1-dependent heterochromatin and H3K4me3/H3K9me3 bivalent domains that appear to function in a complex intertwined manner that merits further research but not necessarily part of the light response.

2.8 Supplemental Figures



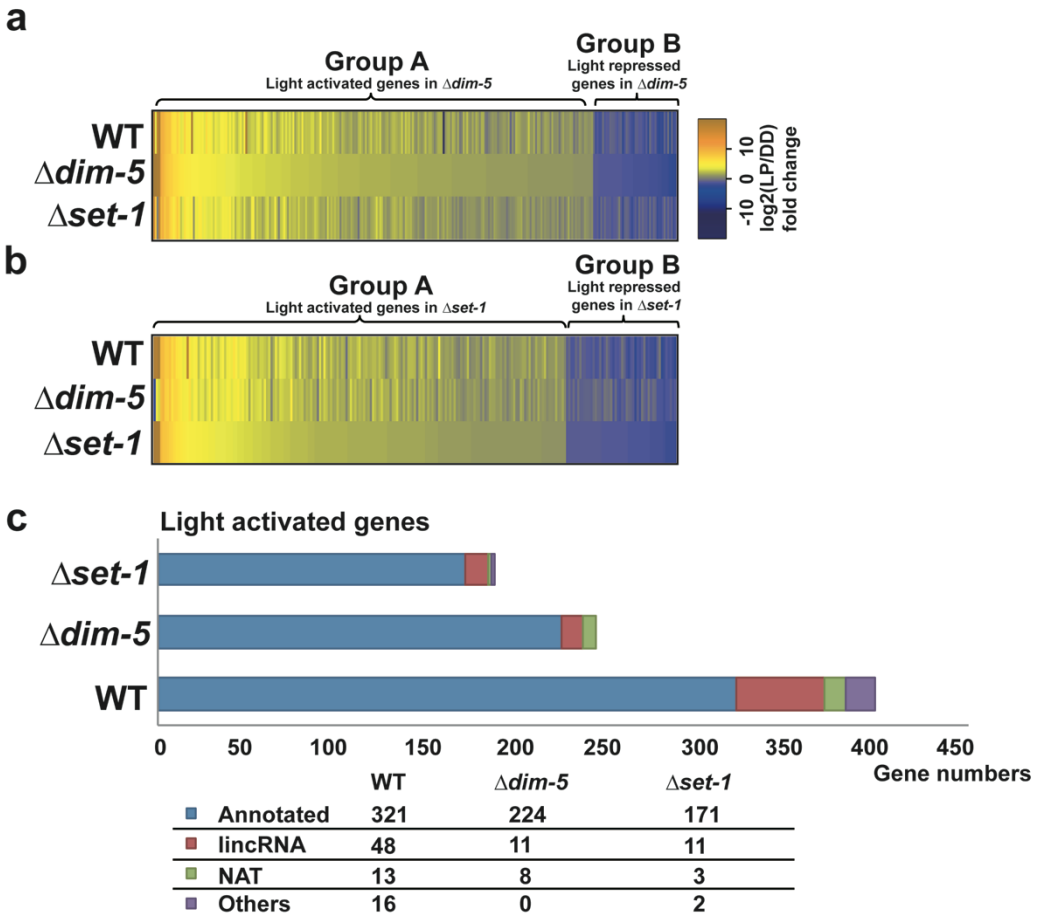
Supplemental Figure 2-1

(a) Schematic representation of the transcript discovery pipeline used in this study. After StringTie merge, we identified 21,475 transcripts at the isoforms level using the default settings in HISAT2 and StringTie. Transcripts expression differences were identified using Cuffdiff and further analysis was performed using CummeRbund. Gffcompare was used to classify newly identified transcripts relative to the reference annotation NC12. (b) Multidimensional scaling of the RNA-seq samples reveal that the underlying mutations have a larger effect than light treatment. M1: Dimension 1, M2: Dimension 2.



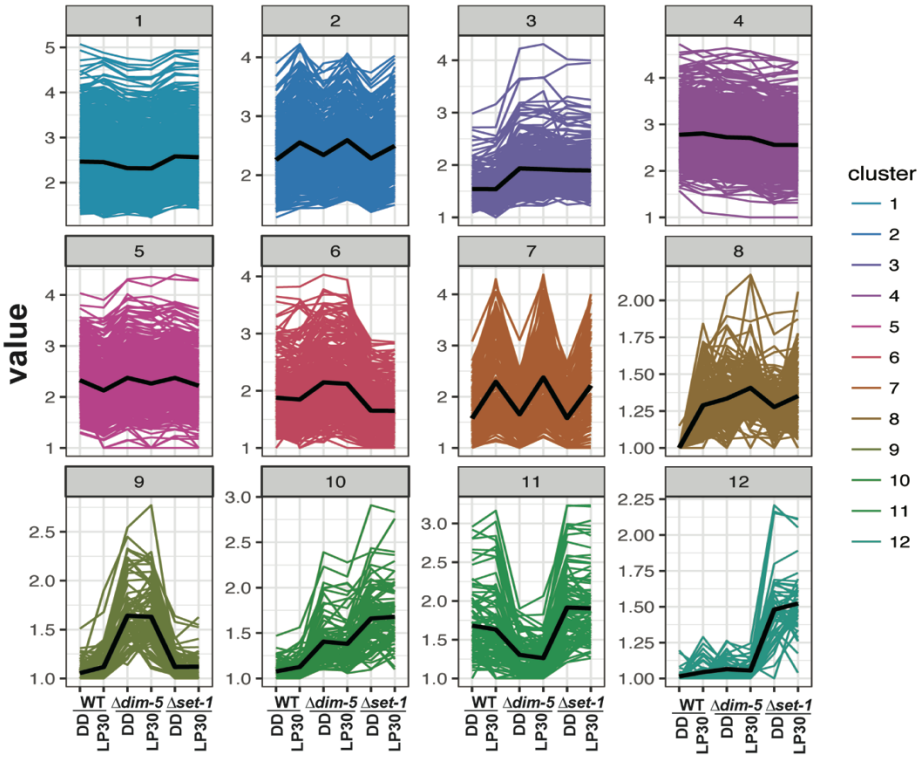
Supplemental Figure 2-2 Variation among transcript types in Δ*kmt2/*Δ*set-1* and Δ*kmt1/*Δ*dim-5*.

Violin plot depicting the expression levels (FPKM, log2) of transcripts belonging to existing annotated transcripts, NATs and lincRNAs in (a) *kmt1/*Δ*dim-5* (b) *kmt2/*Δ*set-1*.



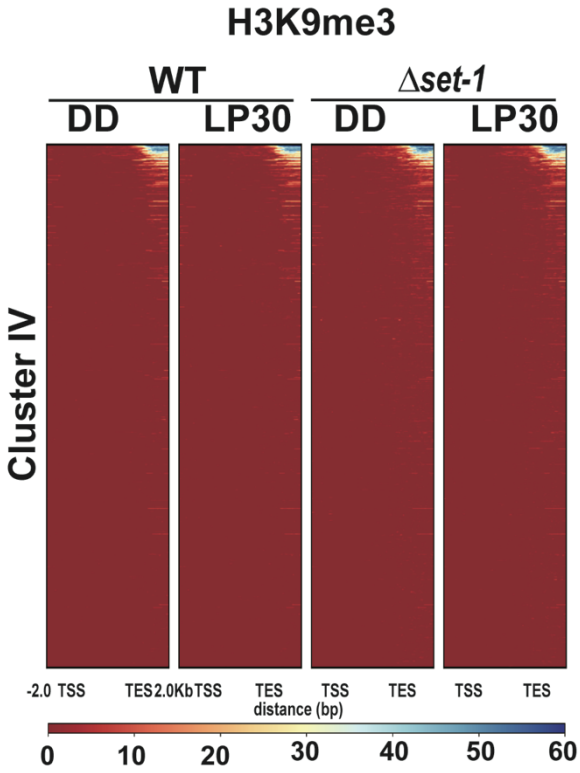
Supplemental Figure 2-3 Hierarchical clustering of light regulated genes in Δ*kmt2/*Δ*set-1* and Δ*kmt1/*Δ*dim-5*.

(a) Heatmap showing the clustering of genes that are differentially expressed in Δ*kmt1/*Δ*dim-5* DD versus LP30 and compared to same genes in WT and Δ*kmt2/*Δ*set-1*. The expression levels are log2 fold change (q<0.05). (b) Same as in A except clustering was done in Δ*kmt2/*Δ*set-1* DD versus LP30. Genes in Group A are light activated and Group B includes genes that are light repressed. Changes in expression between the two conditions are displayed with range [-10,10], with levels above and below the mean shown in yellow or blue. (c) Bar plot depicting the number of different transcripts categorized as an existing annotated gene, lincRNA, NAT or other that were light activated in WT, Δ*kmt1/*Δ*dim-5* and Δ*kmt2/*Δ*set-1*. The x-axis is the gene number and the y-axis is the strain.



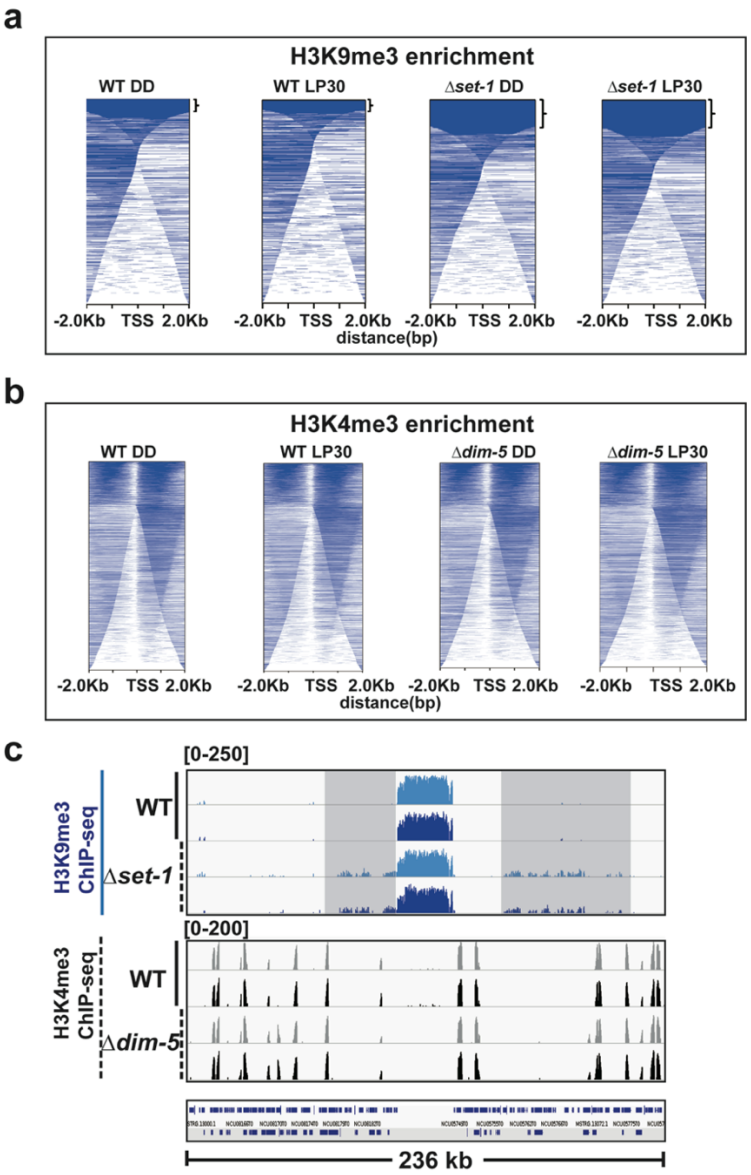
Supplemental Figure 2-4 Gene expression clusters of WT, Δ*kmt1*/Δ*dim-5* and Δ*kmt2*/Δ*set-1*.

Clustering gene expression into 12 groups based on the expression profile. The colored lines represent the pattern for each gene, and the black lines represent the median of all the genes in a given cluster.



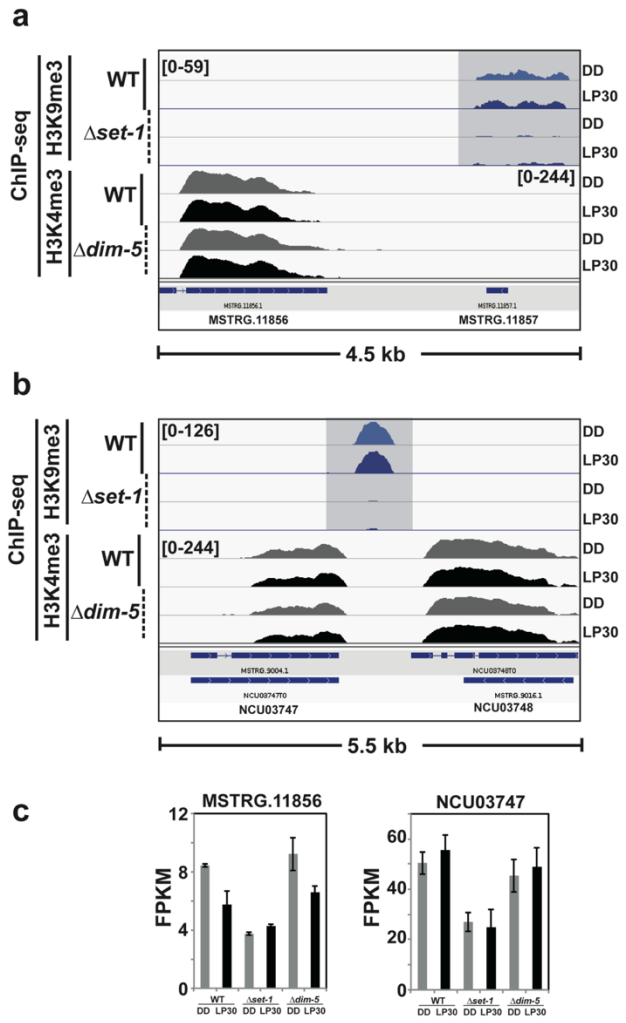
Supplemental Figure 2-5 The heatmaps display ChIP enrichment of H3K9me3 in WT and Δ*kmt2/*Δ*set-1*.

Regions included 2kb upstream of TSS and downstream of TES. In this specific cluster, H3K9me3 is found downstream of the TES in genes adjacent to constitutive heterochromatin.



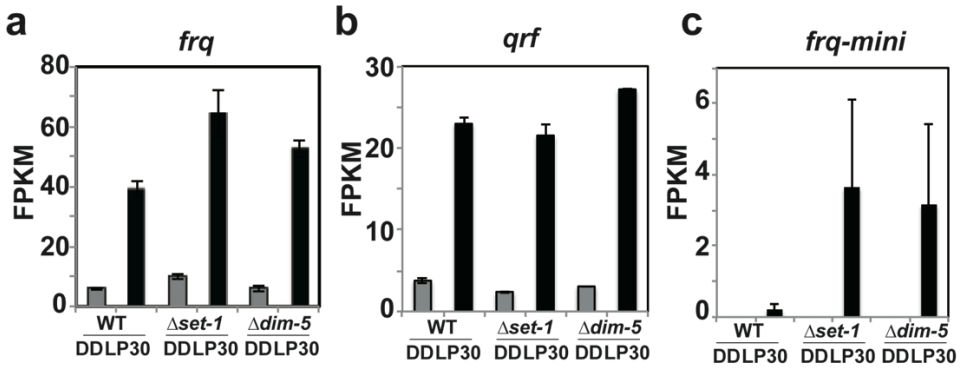
Supplemental Figure 2-6 H3K9me3 spreading in *kmt2/Δset-1* strain.

Heatmap display signal distribution for (a) H3K9me3 (b) H4K4me3 density plotted in a 2-kb windows centered on the TSS. The curly bracket(s) in panel A indicate the extent of H3K9me3 spreading in Δ*kmt2/*Δ*set-1*. (c) Gene-level plot of a 236 kb region on chromosome VII (supercontig 12.7) showing H3K9me3 ChIP-seq (DD Blue and LP30 Navy) for the WT and Δ*kmt2/*Δ*set-1* and H3K4me3 ChIP-seq (DD Grey and LP30 black) for the WT and Δ*kmt1/*Δ*dim-5*. The shaded boxes highlight representative examples of H3K9me3 spreading into euchromatic regions in Δ*kmt2/*Δ*set-1*.



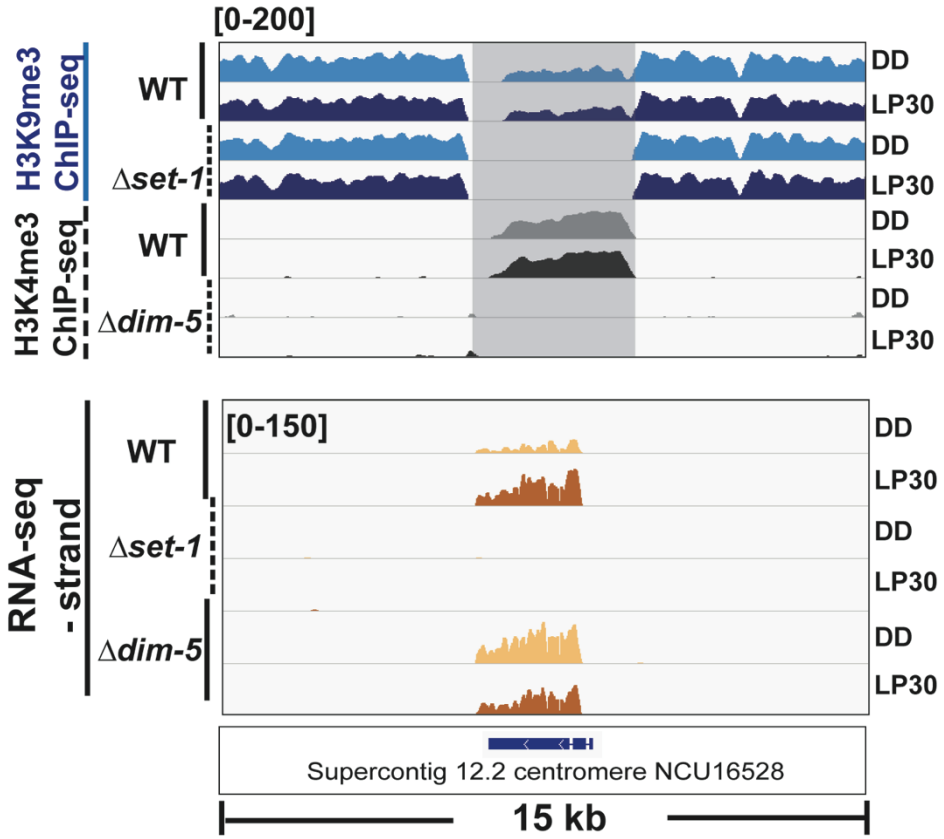
Supplemental Figure 2-7 Two additional loci that have KMT2/SET-1-dependent heterochromatin.

Gene-level diagram of H3K9me3 ChIP-seq (DD Blue and LP30 Navy) in WT and Δ*kmt2/*Δ*set-1,* and H3K4me3 ChIP-seq (DD Grey and LP30 black) in WT and Δ*kmt1/*Δ*dim-5* for a presumptive lincRNA, (a) MSTRG.11857 and (b) NCU03747. It is clear from the traces that both have a significant decrease in H3K9me3 ($p<0.05$) in Δ*kmt2/*Δ*set-1*. (c) Expression bar plot of MSTRG.11856 and NCU03747, which have a decrease in expression when KMT2/SET-1-dependent heterochromatin is lost.



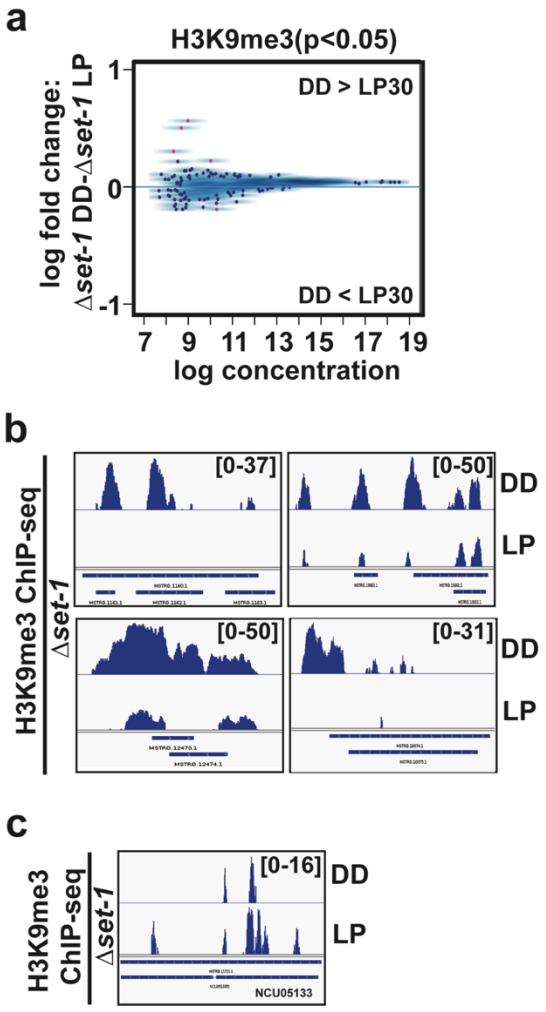
Supplemental Figure 2-8 Expression of transcripts arising from the *frq* locus.

FPKM values for (a) *frq*, (b) *qrf* and (c) *frq-mini* are shown as bar plots in WT, *kmt2/*Δ*set-1* and *kmt1/*Δ*dim-5* for DD (grey) and LP30 (black).



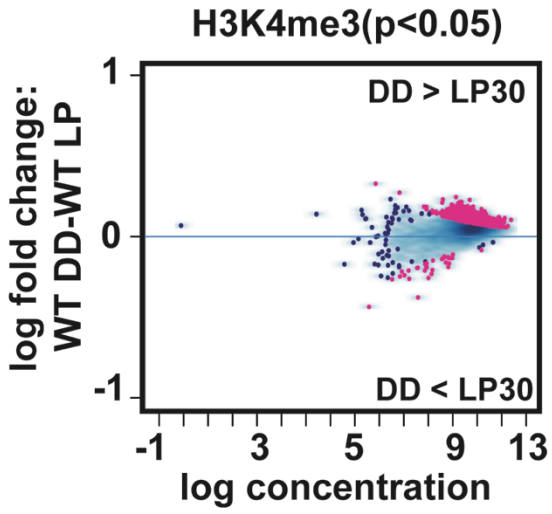
Supplemental Figure 2-9 Reciprocal dependence of H3K9me3 and H3K4me3.

Gene level plot shows a centromeric gene on Chromosome II (supercontig 12.2). The IGV diagram displays H3K9me3 ChIP-seq (DD Blue and LP30 Navy) in WT and Δ*kmt2*/Δ*set-1* and H3K4me3 ChIP-seq (DD Grey and LP30 black) in WT and Δ*kmt1*/Δ*dim-5*. The corresponding RNA-seq traces are also shown for NCU16528.



Supplemental Figure 2-10 Changed in H3K9me3 in response to light in Δ*kmt2*/Δ*set-1*

(a) Quantitative difference in H3K9me3 levels from the ChIP-seq in Δ*kmt2*/Δ*set-1* upon light exposure. The MA plot shows log fold change ($p < 0.05$) in enrichment in DD (log fold change > 0) or LP (log fold change < 0). (b) IGV diagrams of 4 genes that had a decrease in H3K9me3 in Δ*kmt2*/Δ*set-1* in response to light ($p<0.05$) (c) IGV diagram of NCU05133 which had an increase in H3K9me3 in Δ*kmt2*/Δ*set-1* ($p<0.05$) in response to light.



Supplemental Figure 2-11 Light induced changed in H3K4me3 in WT

Quantitative difference in H3K4me3 levels in WT DD versus WT LP30. The MA plot shows increase in H3K4me3 in DD (log fold change > 0) or in response to light (log fold change < 0). Spots shown in red have a $p < 0.05$.

CHAPTER 3

3 Chapter 3: Alcohol negatively impacts normal diurnal regulation of *TERRA* and heterochromatin at the telomeres

3.1 Abstract

Alcohol metabolism produces the DNA damaging agents acetaldehyde and reactive oxygen species. In addition, alcohol has a negative effect on the circadian clock, sleep and cell physiology. Chronic alcohol consumption causes telomere shortening, but other than the DNA damaging effects of alcohol, the mechanism of alcohol-induced telomere shortening is largely unknown. Recently we determined the circadian clock protein BMAL1 associates with telomeres and is needed for rhythms in telomere heterochromatin and *TERRA*, a long noncoding RNA that originates from telomeres. To determine if alcohol affects diurnal regulation at the telomere, which would in theory make it more susceptible to DNA damage, I establish a zebrafish model of short-term alcohol consumption to ascertain how alcohol impacts normal oscillations at the telomere. I found that short-term alcohol exposure causes a phase-shift in the normal *TERRA* rhythm and this corresponded with phase shift in *per2*. In addition, alcohol consumption disrupted the diurnal rhythm in heterochromatin at the telomeres. My data show that alcohol intake alters normal telomere homeostasis on multiple levels, including altering the diurnal rhythm in *TERRA* expression and heterochromatin. These findings reveal multifarious mechanisms involved in alcohol-induced telomere disruption.

Key words: Zebrafish, Ethanol exposure, Telomere, *TERRA*, H3K9me3

3.2 Introduction

Alcohol is one of the most widely and commonly abused drugs and causes many adverse health consequences [286, 287]. For example, alcohol consumption increases cellular senescence in metabolic and inflammatory diseases reminiscent of accelerated aging [288]. Alcohol breakdown in the liver increases reactive oxygen species (ROS) leading to oxidative damage in the liver [289, 290]. Increased oxidative stress induces cell senescence and DNA damage [291]. The breakdown of alcohol also generates acetaldehyde, a reactive metabolite that causes DNA double-strand breaks and abnormal chromosome rearrangement [292, 293]. Alcohol also causes destruction of the telomeres and chronic drinkers have shorter telomeres relative to non-drinkers [294]. Shortened telomeres are observed in many pathologies such as cardiovascular disease and cancer; although, the extent of causation versus correlation remains shrouded [295-298].

As a reminder, telomeres are specialized domains at the distal tips of linear chromosomes composed of the DNA repeat, TTAGGG, bound by a protein complex called shelterin. Linear chromosomes present 2 problems for cells, the end-replication problem and the end-protection problem. Shelterin and the T-loop solve the end-protection problem and prevent the chromosome ends from appearing as double strand breaks [210, 211]. The end-replication problem is solved by telomerase, a ribonucleoprotein composed of telomerase reverse transcriptase (TERT) and *telomerase RNA component* (*TERC*), which adds the telomeric repeat to the chromosome ends [215]. A long noncoding RNA called *TERRA* (telomeric repeat-containing RNA) is expressed from telomeres and is implicated in both end-replication and end-protection. [299]. *TERRA* is critical in the control of telomeric chromatin structure and is thought to be

important telomere heterochromatin formation. *TERRA* is also associated with the shelterin components TRF1 and TRF2, which are important for telomere structure and stability maintenance [220, 300]. In addition, *TERRA* interact with TERT and TERC [301, 302], depleting *TERRA* increases telomerase activity and DNA damage at telomere in embryonic stem cells. Thus, TERRA plays many different roles as a structural regulator of telomeres. Recently, we have shown a diurnal rhythm in *TERRA* that is regulated by the circadian clock transcription factor, Brain Muscle Arnt-like 1 (BMAL1) [218].

The circadian clock governs circa 24-hour gene expression integrating environment cues that adapt to cellular needs [2, 4, 5]. The circadian clock controls metabolism and alcohol intake is affected by *Per1* and *Per2*, and the mutation of *Per2* and *Clock* increased alcohol consumption [303-305]. On the other hand, chronic alcohol intake decrease CLOCK expression in ventral tegmental area [305]. Moreover, the chronic and heavy ethanol intake could alter the free-running period in mice, and potential lead to hepatic steatosis [306-308].

Despite connections between the circadian clock and alcohol, how chronic alcohol consumption negatively impacts circadian cell physiology is only partially understood. The discovery that BMAL1 can associate with telomere repeat controlling TERRA expression and shortened telomeres are observed in chronic drinkers I sought to explore whether alcohol interfered with circadian control of telomere homeostasis. To examine this, I develop a short-term alcohol consumption model in zebrafish designed to mimic binge drinking and examined if alcohol intake disrupts normal telomere homeostasis. I found short-term ethanol exposure causes a defect in telomere regulation

by disrupting normal diurnal rhythms in *TERRA* and heterochromatin formation. These findings provide the molecular basis on how ethanol affects normal telomere regulation, and how chronic drinking causes disruption of diurnal TERRA expression and heterochromatin formation at the telomere. Extrapolation suggests loss of diurnal control at the telomeres plays a pivotal role in alcohol-related diseases that arise from increased cellular senescence brought on by alcohol-induced telomere shortening.

3.3 Materials and methods

3.3.1 Animal Care

All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of Rutgers University. Wild-type zebrafish were obtained from ZIRC (Zebrafish International Resource Center, Oregon) and housed according to procedures approved by the Policy on the Housing of Vertebrate Animals Outside of Animal Facilities as described previously [218]. Zebrafish were kept in system water (conductivity 800 ± 200 μS and pH 7.5), fed twice daily ad libitum, and maintained under a 14 hour(h):10h light:dark cycle for breeding, or 12h:12h light-dark cycle for experiments done under diurnal conditions. Alcohol exposure was administered to the animals as follow. Wild-type adult zebrafish (approximately 12-14 months old) representing both sexes were housed in an aquarium equipped with a programmable automatic fish feeder set to feed twice daily (one hour after lights on (Zeitgeber time ZT1) and two hours before the light off (ZT10)) and kept on a 12h:12h light:dark cycle. For alcohol exposure, half the zebrafish were transferred to tank containing 1% (volume/volume) ethanol for 1 hour between ZT2 and ZT3. After the 1-hour ethanol

exposure, zebrafish were transferred to the original aquarium. After 4 days of this binge drinking model, zebrafish were euthanatized in Tricaine (MS-222, 0.04 mg/ml, Western Chemical), decapitated and dissected under PBS. Fish were harvested every 4 hour over a 24-hour period [309]. Liver samples were obtained from each zebrafish, snap frozen in liquid nitrogen and store at -80°C for later use. Nighttime samples were obtained under red light in a dark room.

3.3.2 Chromatin immunoprecipitation (ChIP)

ChIP experiments followed the general procedure described previously [114, 122, 218]. Isolated zebrafish tissue was cross-linked with 1% formaldehyde for 10 min at room temperature then quenched with 0.1 M glycine for 10 min. The cross-linked tissue was snap-frozen in liquid nitrogen and stored at -80 °C. The tissue was homogenized with a micro pestle in the presence of 100 µl ChIP lysis buffer (0.05 M Hepes pH 7.4, 0.15 M NaCl, 0.001 µM EDTA, 1% triton X-100, 0.1% SDS) 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 an additional 6 times for 20 s at 20% power. The resulting lysates were cleared of cellular debris by centrifugation at 5000 × g for 10 min. The sonication consistently yielded chromatin sheared to an average size of 500 bp. 200 µg of clarified lysate was mixed with histone H3 lysine 9 tri-methylated (H3K9me3) antibody prebound to protein A magnetic beads (Dynabeads, Invitrogen). The ChIP was washed 1 times with RIPA-150 (50mM Tris-HCl pH8, 0.15 M NaCl, 1mM EDTA pH8, 0.1% SDS, 1% Triton

X-100, 0.1% sodium deoxycholate) buffer, 2 times with RIPA-500 (same as RIPA-150 except using 0.5M NaCl), and 2 times with RIPA-LiCl (50mM Tris-HCl pH8, 1mM EDTA pH8, 1% Nonidet P-40, 0.7% sodium deoxycholate, 0.5M LiCl). Chromatin was eluted from the beads by treatment with 2 x 50 µl 0.1 M sodium bicarbonate, 1.0% SDS at 37 °C for 10 min. The cross-links were reversed by adding 2 µl of 5M NaCl and incubating the samples for a minimum of 4 hours at 65 °C. Protein was removed by the addition of 1 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 55 °C for 1-hour. DNA was purified by a phenol-chloroform extraction. The relative levels of H3K9me3 was determined by qPCR using oligonucleotides contained in **Table 3-1**. The H3K9me3 antibody was obtained from Abcam (#ab8898).

3.3.3 RNA isolation and RT-PCR

Total liver RNA was isolated by Trizol (Invitrogen) following the manufacturer's guidelines. Precipitated RNA was resuspended in RNase-free water and the concentration was determined using a Nanodrop spectrophotometer. cDNA was reverse transcribed from 1µg of total RNA using High-Capacity cDNA reverse transcription kit (Applied Biosystems) following manufacturer's guidelines. RT-PCR data were analyzed using comparative Ct method (2-[delta][delta]Ct) method, using *hmbs* as internal control. Oligonucleotides for RT-PCR are listed in **Table 3-1.**

Table 3-1 Primers used for RT-PCR and ChIP q-PCR

*adh8a*_F	CCCTCTTCCTCTCTCAGTGTG
*adh8a*_R	CTTGTAGGTTCAGCCATAATGTCAT
*adh8b*_F	GACAGACAATAAAGGTTTTCCCAC
*adh8b*_R	GTTTTGGGATATGACATATATTCAA
*per2*_exon17_F	CTCATCTGAAGGAGCAGTC

*per2*_exon17_R	CTGTAGCTGCACTGGCTGG
subtelomere_ChIP_F	TGTGTTCGTGGAGTGGATCTTG
subtelomere_ChIP_R	GCGTGATCGTGGTTATGTTCA
*tert*_F	CCCTGGTGCGGCTTGA
*tert*_R	TAGGCCGTCATACCGTGAGTAGT
*terc*_F	CGGCTTGTATGCTGTTTTGTCT
*terc*_R	AAACCGCGGAAAGCTTCA
hmbs_F	GAACGAGGAAAAGCCAGTTG
hmbs_R	TGTGGACATGGCAACAATCT

3.3.4 Northern blot

TERRA Northern blot was performed under general protocol described previously [218]. 3-5 μg of total RNA was incubated for 15 min at 65 °C 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 in a 1.3% agarose gel made in 1 x MOPS, 1.85 % formaldehyde for 2-3 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* oligonucleotide probe containing 5 TTAGGG repeats (IDT DNA) at 42°C overnight in EZ hyb buffer (Roche). Membranes were washed two times with 2X SSC, 0.1% SDS at 42°C for 10 min each and then three times in 0.1% SSC, 0.1% SDS at 65 °C for 15 min. The membrane was placed in DIG blocking buffer for 2-hours then incubated with anti-digoxigenin Fab fragments for 30 min. The membranes were washed 5 x 10 min with 1x maleic acid, 0.3% tween and visualized using CDP-star (Roche).

3.3.5 Data analysis

Rhythmicity of all relevant data was determine using Cosinor with DiscoRhythm [310]. Other statistical analyses shown are one-way ANOVA of peak to trough determined with Bonferroni Post Hoc test.

3.4 Results

3.4.1 Alcohol dehydrogenase (Adh) Adh8 is diurnally regulated

I set out to test whether alcohol interfered with normal diurnal regulation at the telomere. To examine this question, I used zebrafish, which is an ideal model for studying alcohol exposure [311, 312]. Alcohol dehydrogenases in zebrafish are highly conserved with humans making it an ideal system to examine the effect of alcohol on human health [313, 314]. As a first step, I sought to determine the phase of zebrafish alcohol dehydrogenase (*adh*) genes and use this information, compared with the known phase of *TERRA*, to identify a time likely to generate an easily observable response to alcohol. *Zebrafish* have 2 *adh* genes, *adh8a* and *adh8b* so I measured diurnal expression of both by RT-qPCR. In zebrafish, Adh8a is the dominant dehydrogenase that breaks down ethanol, while Adh8b metabolizes longer chain primary alcohol substrates [314]. Cosinor analysis using DiscoRhythm indicated *adh8a* and *adh8b* both had a diurnal rhythm ($p<0.05$) with *adh8a* peaking at approximately ZT 20 while *adh8b* appeared to peak at roughly ZT 0 **(Figure 3-1A.)**. Based on these results, I designed the alcohol exposure to coincide with the trough in *adh8a* expression reasoning this would likely reveal the greatest effect. Therefore, I subjected zebrafish to 1% alcohol for 1 hour over consecutive four days at

ZT2. After the four days ethanol treatment, I sacrifice the fish every 4 hours for a 24h cycle and used the resulting tissue for the following experiments **(Figure 3-1B)**.

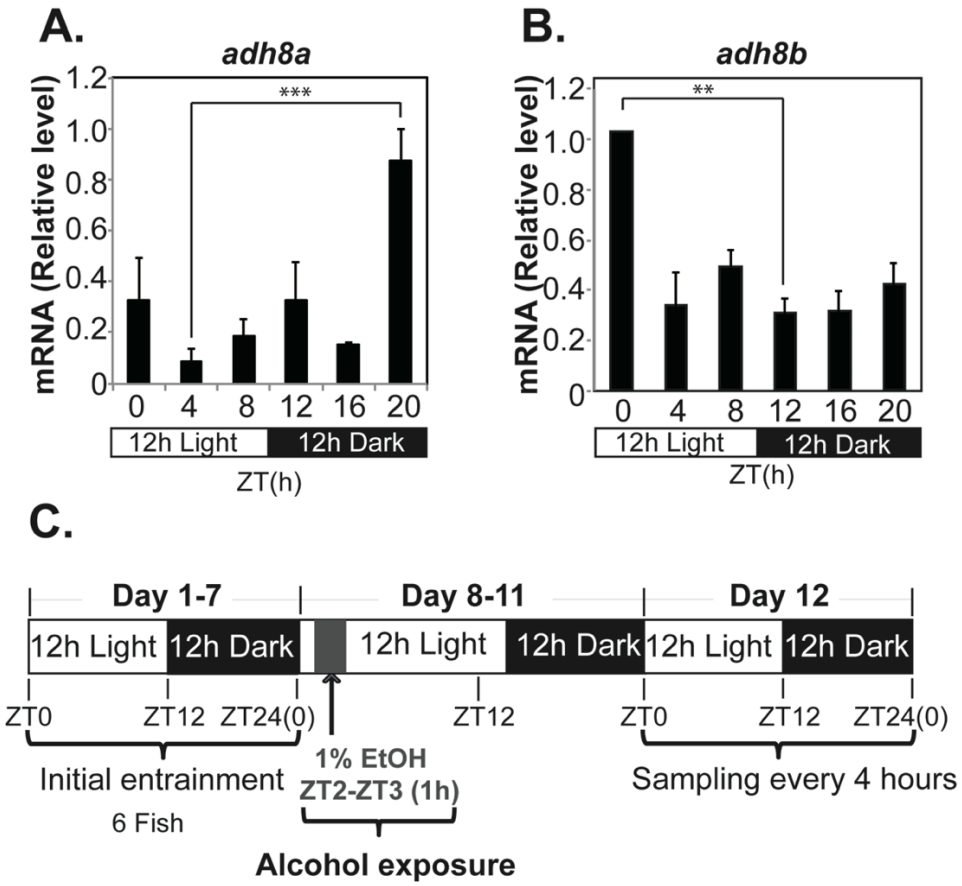


Figure 3-1 Diurnal rhythm of alcohol dehydrogenase in zebrafish liver and experimental design.

Real-time PCR was used to determine mRNA levels of *adh8a* (n=5, n represents the biological replicates) (A) *adh8b* (n=3) (B) Zebrafish alcohol consumption model (C). The white and black bar indicate the light dark cycle. The first week (Day 1 to 7) indicates the initial entrainment. Alcohol exposure last 4 days (Day 8-11), and one hour a day from ZT2-ZT3 as indicated as grey bar and arrow. After exposure, zebrafish were sampled every four hours. The error bar represents the SEM and the cosinor analysis showed $p<0.05$. The statistical analysis of peak to trough was performed by one-way ANOVA with Bonferroni Post Hoc test showing **$p\leq0.005$ and ***$p\leq0.0005$.

3.4.2 Short-term Ethanol exposure causes a phase shift in *per2*

To begin, I sought to confirm on *per2* gene expression, As a reminder, the circadian clock plays a critical role in maintain health and physiology [315] and chronic alcohol consumption affects the circadian clock by altering expression the circadian clock gene *Per2* both in clinical studies and mice [308, 316-318]. Based on this I tested whether our short-term ethanol treatment altered *per2* expression using RT-PCR **(Figure 3-2A)**. I found short-term ethanol exposure caused an approximate 2-hour phase shift, but that overall rhythmicity was intact. This indicates that short-term ethanol exposure causes a phase shift in *per2* expression, but otherwise the clock had a relatively normal pattern in expression.

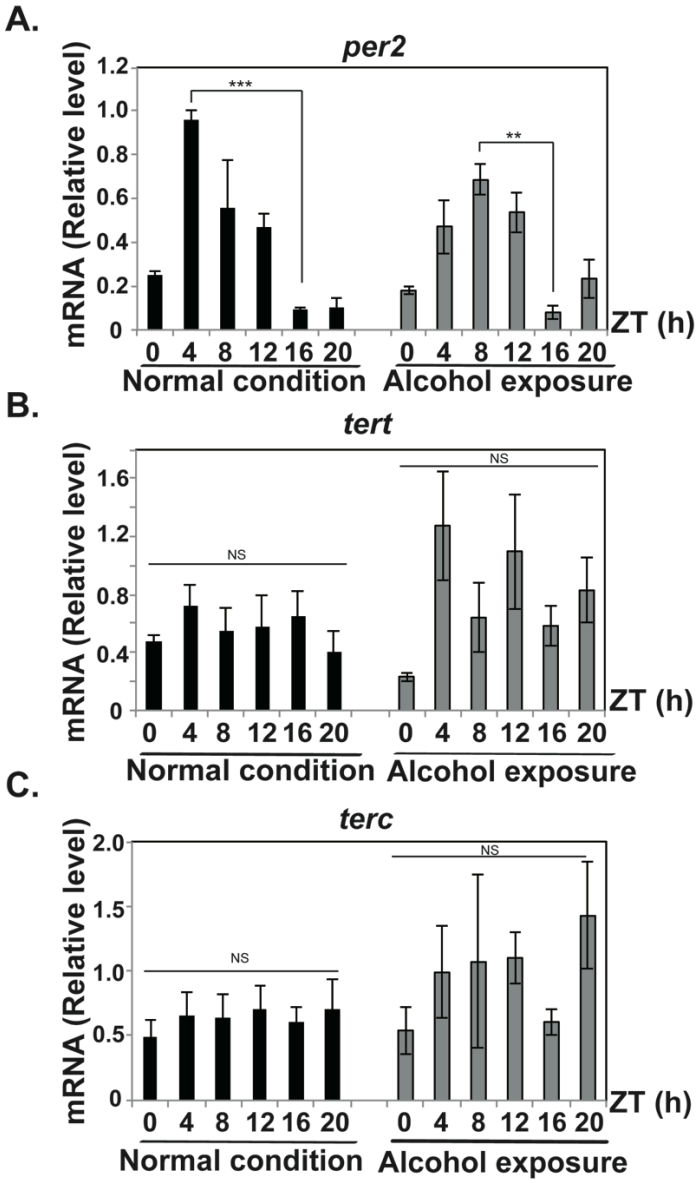


Figure 3-2 The rhythm of *per2* in zebrafish after exposed to 1% ethanol for 1h.

The bar chart indicates the zebrafish (n=3, n represents the biological replicates) mRNA level of *per2* (A) tert (Cosinor, $p>0.5$) (B) and terc (Cosinor, $p>0.5$) (C) over 24 hours under 12:h-12:h light and dark cycle. (Black bar: normal condition. Grey bar: 1 hour 1% alcohol exposure from ZT2-ZT3) The error bar represents the SEM and the cosinor analysis showed $p<0.001$. The statistical analysis of peak to trough was performed by one-way ANOVA with Bonferroni Post Hoc test showing **p≤0.005 and ***p≤0.0005, and the NS represents $p>0.05$.

3.4.3 Ethanol exposure disrupted TERRA rhythmicity

I next sought to examine whether alcohol affected the diurnal rhythm in TERRA expression because telomere shorting is a biomarker of cirrhosis formation along with an increase in hepatocyte senescence, which are both hallmarks alcohol-induced liver disease [319, 320]. Based on this, I sought to determine whether the negative impact of alcohol on telomeres arise because alcohol affects the normal diurnal regulation of *TERRA*. To accomplish this, I performed Northern blots on total zebrafish RNA from liver samples isolated over 24h cycle sampling every 4h using a probe that recognizes *TERRA*. Cosinor analysis of 6 independent biological replicates indicated ethanol exposure altered the diurnal rhythm in *TERRA* **(Figure 3-3)**. In WT, I confirmed the *TERRA* rhythm ($p<0.05$), while alcohol treated animals displayed pseudo-constitutive expression with no discernable rhythm (Cosinor, $p>0.05$) when averaged over 6 animals.

Next, I wanted to see whether alcohol also affected telomerase, which if true might contribute to telomere shortening. Reports indicate telomerase expression and activity are under circadian control in human and mice [217]. Therefore, I tested whether short-term ethanol exposure affects telomerase in zebrafish. I was unable to confirm the rhythms in *terc* or *tert* in WT zebrafish liver **(Figure 3-2 B & C)** and there appeared to be no statistically significant change in expression due to alcohol. However, alcohol seemed to dramatically increase the variability in *terc* or *tert* making strong conclusions elusive. The lack of a rhythm in *tert* and *terc* may due to the limited expression of telomerase in adult zebrafish somatic liver cells [321], making the liver a less-than-ideal tissue to detect a telomerase rhythm. In addition, the age of the fish I chose (adult fish, 12-14 month) may

be too old to detect a rhythm in *terc* and *tert*, due to increase variability of telomerase activity in older fish [321, 322].

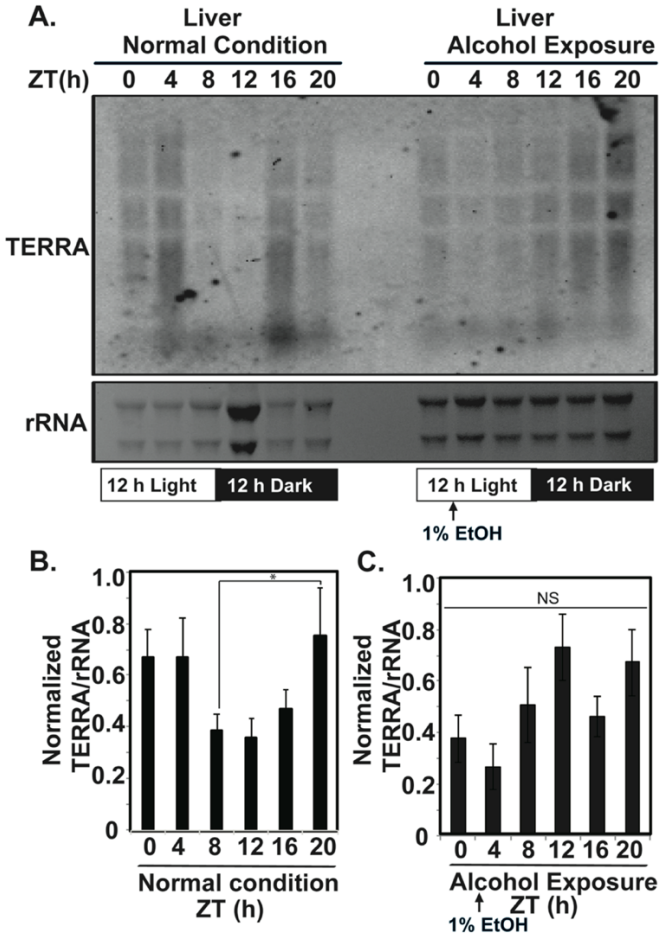


Figure 3-3 Effect of ethanol on Zebrafish *TERRA* rhythmic pattern.

A representative northern blot data (A) showed the rhythmic pattern of TERRA over 24 hours both in normal condition (left) and ethanol exposure (right). The ribosomal RNA bands used for normalization. The quantification bar plot for normal condition (n=6, n represents the biological replicates) (B) and alcohol exposure (n=6) (C). The white and black bar at the bottom indicate the 12:12 h light dark cycle**.** The error bar represents the SEM and the rhythmicity by cosinor analysis showed $p<0.001$. The statistical analysis of peak to trough was performed by one-way ANOVA showing *$p\leq0.05$, and the NS represents $p>0.05$.

3.4.4 Ethanol affects H3K9me3 rhythm at telomeres

The loss of diurnal TERRA expression led me to test how that impacted heterochromatin formation at the telomeres. Epigenetic regulation of telomeres is essential for telomere stability and homeostasis [323] and telomere was widely thought to be packaged into constitutive heterochromatin containing histone H3 lysine 9 trimethylation (H3K9me3). Telomere heterochromatin it is required for repression of telomere recombination and is necessary to maintain telomere length in mammals [212] and *TERRA,* through its association with the shelterin protein TRF2, plays a critical role in telomere heterochromatin formation including heterochromatin protein 1 (HP1) binding and H3K9me3 [220]. Because the rhythm in *TERRA* is accompanied by a rhythm in H3K9me3 at telomere [218], I sought to examine whether alcohol also disrupted the normal heterochromatin rhythm. To accomplish this, I examined the H3K9me3 at the telomere using ChIP. I confirmed the rhythm in H3K9me3 at telomeres in zebrafish liver under normal conditions, (Cosinor, $p<0.01$) **(Figure 3-4A)** and found that alcohol caused a disruption in the H3K9me3 rhythm (Cosinor, $p>0.1$) **(Figure 3-4B)**. Based on this, I conclude that alcohol causes a disruption of diurnal regulation of *TERRA* and heterochromatin at the telomere and this may increase telomere erosion observed in chromic drinkers.

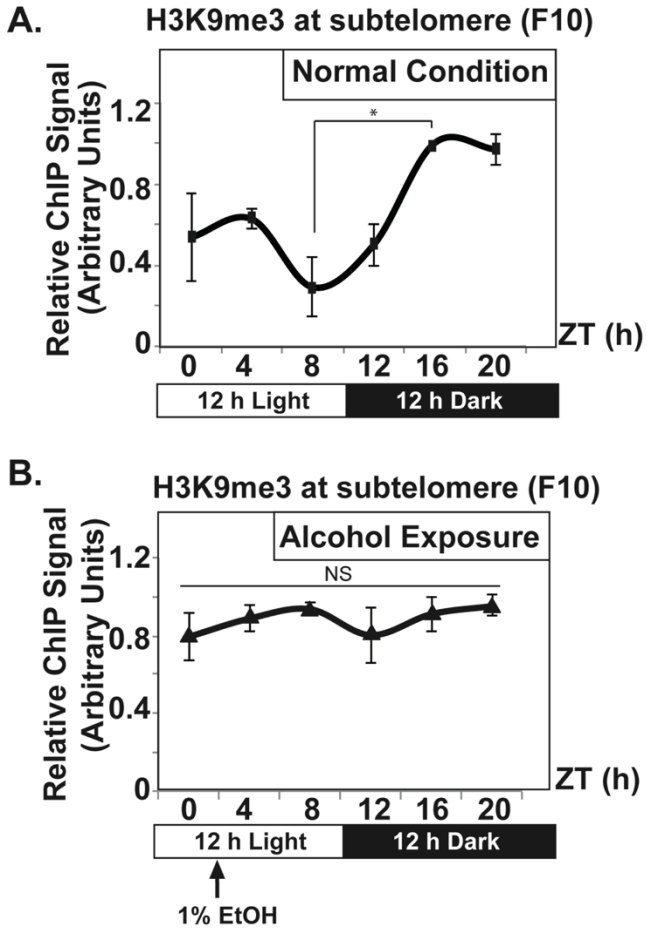


Figure 3-4 Ethanol affect H3K9me3 rhythm at subtelomere region.

H3K9me3 is rhythmic under normal diurnal condition (A) but is changed after alcohol exposure (B) at telomere. The data are displayed as mean ± SEM of three independent experiments (n=3). the Cosinor analysis showed $p<0.01$. The statistical analysis of peak to trough was performed by one-way ANOVA with Bonferroni Post Hoc test showing *$p\leq0.05$, and the NS represents $p>0.05$.

3.5 Discussion

In this Chapter, I found that short-term ethanol exposure in zebrafish causes a phase shift in *per2* and disrupts normal telomere homeostasis. Specifically, alcohol

exposure altered the rhythm in *TERRA* expression and the rhythm in H3K9me3. The consequences of these findings are of significant interest. *TERRA* it thought to play a critical role in maintaining telomere stability and interacts with proteins important for telomere signaling, DNA replication, cell cycle regulation and DNA repair [302]. Thus, changes in *TERRA* expression likely impacts many of those processes and leads to abnormal regulation of telomere-specific events, such as DNA repair and telomere signaling. I also showed that alcohol exposure disrupts the H3K9me3 rhythm at telomere. The epigenetic status of telomere and subtelomeres is strongly connected with telomere length, and shorten telomeres correlate with, and may be a driver of, diseases such as cancer and premature aging [323, 324]. The alcohol-induced defects in heterochromatin could theoretically lead to more rapid telomere erosion. For example, the increase in alcohol-induced double-strand breaks may cause more rapid telomere erosion because *TERRA* expression and heterochromatin are mis-regulated in response to alcohol [325]. In other words, the lack of heterochromatin may cause the telomeres to be more accessible to DNA damaging agents at the same time the DNA repair signaling is affected. Support for this comes from studies that show ethanol exposure increases the telomeric DNA protection protein, TRF2 (Telomeric Repeat Binding Factor 2) and other telomere-associated proteins in cultured cells without altering telomerase activity. This suggests telomere are more susceptible to DNA damage and DNA protection proteins are upregulated to help prevent telomere shortening [326, 327].

In humans, alcohol intake causes alcoholic liver disease including fatty liver, alcoholic hepatitis, liver cirrhosis and the hepatocellular carcinoma [328]. Chronic alcohol exposure in zebrafish also causes alcoholic liver disease [329] making it a good

system to study the effects of alcohol on the liver. Based on the work presented here one cell physiological responses to alcohol is loss of diurnal regulation at the telomere and this presumably would occur in humans and may be one of the causative events linking alcohol consumption and liver disease.

Alcohol abuse also results in accelerated aging and is a risk factors for different types cancers such as the liver, colon, rectum and female breast cancer [330, 331]. Despite many studies showing alcohol is a factor in human carcinogenesis the mechanism of alcohol-induced cancer is still not fully determined [332]. One of the critical keys that associated with aging and cancers is telomeres. Cancer cells tend to have shorter telomeres compare with normal cells, even with reactivation telomerase that required for cancer growth [216, 333]. Telomere homeostasis is maintained under complexity of protein interacting networks, long non-coding RNAs, and epigenetic regulation [334]. The defect of epigenetic regulation have strong connection with human diseases including cancer [335, 336].

In this report I also found the normal BMAL1-dependent diurnal rhythm in *TERRA* and heterochromatin at telomeres is altered by alcohol consumption [218]. Because H3K9me3 is a critical factor in telomere heterochromatin, it is interesting to speculate alcohol-induced disruption of the diurnal rhythm at telomeres is a potential factor contributing to alcohol-associated cancers. Our results showing short-term alcohol exposure disrupts the normal rhythm in *TERRA* and H3K9me3 give direct molecular evidence that alcohol interrupts normal telomere homeostasis. This provides insight into clinical implications of binge drinking and suggest it could disrupt telomere regulation and

lead to shorter telomeres, and this could one of the direct factors that leading to accelerate aging and potential causes of cancers.

In conclusion, my data show short-term ethanol exposure alters the normal diurnal rhythm in *TERRA* and heterochromatin and suggest alcohol effects on the telomere might be a contributor to shorten telomere seen in alcoholic liver diseases, cancers and advanced aging. Whether long-term alcohol exposure has a more severe effect on telomeres remains to be determined.

CHAPTER 4

4 Chapter 4: hnRNP M interact with the *Per2* natural antisense transcript and assists heterochromatin formation at *Per2*

4.1 Abstract

Circadian oscillations require histone modifications and chromatin remodeling to ensure appropriate amplitude, period, and phase of clock gene expression. The *Period2* (*Per2*) locus in vertebrates has a non-coding natural antisense transcript (NAT) *Per2AS*, whose function is completely unknown. In this Chapter, I determined the diurnal rhythm of *Per2AS* is dependent on BMAL1. Comprehensive identification of RNA-binding proteins by mass spectrometry (ChIRP-MS) revealed heterogenous ribonuclear proteins and other RNA binding proteins that are part of the nuclear PER2-complex associate with *Per2AS*. Phenotypic characterization of a subset of *Per2AS*-interacting protein indicated hnRNP M is needed for normal circadian regulation of *Per2* and depletion of hnRNP M caused a reduction in amplitude and a small period increase. Knockdown of hnRNP M also showed hnRNP M is necessary to maintain the normal levels of both H3K9me3 and H3K27me3 at *Per2*. Artificial induction of *Per2AS* using dCas9-VPR showed a minor enhancement in *Per2* expression, but the mechanism of this was elusive. Collectively, these data show hnRNP M is associated with *Per2AS* and plays an assistive role in circadian clock regulation and suggests a model where *Per2AS* serves as scaffold for hnRNP M and other RNA-binding proteins to efficiently coordinate heterochromatin formation at *Per2*.

Key Words: Circadian clock, Natural antisense transcript, *Per2AS*, hnRNP M, Heterochromatin

4.2 Introduction

Circadian rhythm is controlled by transcriptional negative feedback loop [2, 7-9] and proper control of chromatin is a key aspect in activation and feedback inhibition [79, 337, 338]. In vertebrates and mammals, the positive arm is controlled by CLOCK:BMAL1 heterodimer driving the expressions of *Period (Per1, Per2 and Per3*) and *Cryptochrome (Cry1 and Cry2)* genes [16-18, 339]. PER:CRY form the negative element that inhibit CLOCK:BMAL1-mediated transcription via direct protein-protein interactions and chromatin associated events [40-42]. Timed chromatin remodeling and histone modifications serve to establish the proper timing, phasing and amplitude of the clock gene expression via circadian-regulated facultative heterochromatin (CRFH) [337]. CRFH is orchestrated oscillations in chromatin that changes from a transcriptionally permissive to non-permissive states on the 24-hr time scale. The activation phase is initiated by CLOCK:BMAL1 binding to E-box elements, accompanied by acetylation and PolII binding, while the repressive phase consists of deacetylation, histone H3 lysine 9 methylation (H3K9me), and Heterochromatin Protein 1 (HP1) binding [34, 37, 79].

Chromatin oscillations during feedback repression is mediated in part by the nuclear PER complex which is an approximate 1.9 MDa complex that contains numerous RNA binding proteins NONO and PSF (polypyrimidine tract-binding protein-associated splicing factor) the RNA helicases DDX5 and DHX9. In addition, the histone lysine 9 methyltransferase KMT1/SUV39H1, HP1γ, SIN3-histone deacetylase complex, and the

nucleosome remodeling and deacetylase complex NuRD are also found within the PER-complex [34, 35, 37, 39, 98]. The PER complex helps guide KMT1/Suv39H1 to *Per1* and *Per2* promoters [37]. During repression *Per2* is first deacetylated and then collects a set of repressive histone modifications that include histone H3 lysine 9 di- and tri-methylation (H3K9me2 and H3K9me3), and histone H3 lysine 27 tri-methylation (H3K27me3) [37, 118, 123].

Natural antisense transcripts (NATs) are specific long non-coding RNAs (lncRNAs) that overlap with, but are transcribed independently from sense RNAs [174, 175]. NATs haven been implicated in epigenetic regulation such as X chromosome inactivation [340] and RNAi mediated DNA methylation or heterochromatin formation [341, 342]. The *frequency* gene in *Neurospora* and *Per2* in mammals each contains NATs. In *Neurospora*, the NAT *qrf* is needed for CRFH and dynamic DNA methylation at *frq* in a manner somewhat analogous to RNAi-mediated heterochromatin, except the duplex RNAs that arise from convergent transcripts are processed by the exonuclease ERI-1 (enhanced RNAi-1) [115] to form Argonaut-associated Dicer-independent small interfering RNAs (disiRNAs) [112, 113]. The ERI-1 binds nascent transcripts, interacts with DDB1(a core component of the histone methyltransferase complex DCDC), and facilitates H3K9me3 and DNA methylation in *Neurospora* [115] [114].

Despite a fair amount of knowledge on *qrf* in *Neurospora*, the function of *Per2*AS, if any, is still unknown. For example, it is still not known whether *Per2AS* transcription is dependent on BMAL1. Nor is it known whether *Per2AS* functions analogous to *qrf* transcription and participates in establishing heterochromatin at *Per2*. Given the conserved mechanism of CRFH in clock function, I sought to elucidate how *Per2AS* is

regulated and how it contributes to clock regulation. To address these questions, I examined *Per2AS* expression in *Bmal1*$^{-/-}$ mice and found *Per2AS* expression was dependent on BMAL1. I also identified proteins associated with *Per2AS* by ChIRP-MS (Comprehensive identification of RNA-binding proteins by mass spectrometry) and found factors that are also shared in PER-complex. In addition, *Per2AS* is associated with heterogenous nuclear ribonucleoproteins (hnRNP M). Specifically, hnRNP M alters normal circadian clock transcription and heterochromatin formation at *Per2*. Based on these data and emerging models in other systems, my findings provide evidence that *Per2AS* may act as scaffold to recruit various functional proteins that coordinate *Per2* transcription and CRFH at *Per2*.

4.3 Materials and Methods

4.3.1 Animal Care and Cell culture

All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of Rutgers University. *Bmal1* mutant mice (B6.129-*Arntl*^{tm1Bra}/J) were obtained from Jackson Laboratories and bred to generate homozygous mutant mice and compared C57BL/6J wild-type (WT) controls [343]. Mice were genotyped by PCR using genomic DNA of ear tissues from 8-week-old mice [343]. Both male and female WT and homozygous *Bmal1*$^{-/-}$ mice were sorted and used for experiments. For time-course sampling, mice were maintained under a diurnal light cycle (12-hour light:12-hour dark) and sacrificed at 4-hour intervals over 24 hours. Rhythms in the relevant data was analyzed using Cosinor with DiscoRhythm [310].

NIH 3T3 mouse fibroblasts cells (ATCC, #CRL-1658) and *Per2*::*Luciferase* (*Per2*::*Luc*) mouse embryonic fibroblasts (MEFs) were originally isolated from *Per2*::*Luc* transgenic mice [344]. Cells were cultured and passaged in DMEM (Gibco) with phenol red, 10% FBS, 1% penicillin/streptomycin at 37 °C with 5% CO2. For time-course sampling, the cells are grown to 80% confluency, entrained with 100 nM dexamethasone for 2 hours and sampled at designated times.

4.3.2 RNA isolation and Real-time PCR

RNA was isolated by Trizol (Invitrogen) following the manufacturer's guidelines from cells and mouse liver. Precipitated RNA was resuspended in RNase-free water and the concentration was determined using a Nanodrop spectrophotometer. cDNA was reverse transcribed from 1μg of total RNA with High-Capacity cDNA reverse transcription kit (Applied Biosystems) using with strand-specific oligonucleotides oligo-m*Per2*_RT and oligo-m*Per2*_intron5_AS_RT (**Table 4-1**) for *Per2* and *Per2AS*. Quantitative PCR of the cDNA was done using oligos-m*Per2*_exon6_F4/R4, which detects both *Per2* and *Per2AS*, along with Power SYBR green PCR Master Mix (Applied Biosystems). The absolute copy number of *Per2* and *Per2AS* was determined using a plasmid containing *Per2* gene. The relative copy number of each transcript was normalized to the average copy number between different replicates.

4.3.3 Chromatin immunoprecipitation (ChIP)

ChIP experiments followed the general procedure described previously [114, 122, 218]. Cells were detached by trypsinization and quenched by cultured media. The cells were pelleted and washed 2 times with PBS. Then the washed cells were cross-

linked with 1% formaldehyde for 10 min at room temperature, then quenched with 0.1 M Glycine for 10 min. The cross-linked cells were snap-frozen in liquid nitrogen and stored at -80 °C. To generate the chromatin lysate, the cells were homogenized with a micro pestle in the presence of 100 µl ChIP lysis buffer (0.05 M Hepes (pH 7.4), 0.15 M NaCl, 0.001 M EDTA, 1% Triton X-100, 0.1% SDS) 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 s at 20% power). The resulting lysates were cleared of cellular debris by centrifugation at 5000 × g for 10 min. The sonication yielded chromatin sheared to an average size of 500 bp and 200 µg of clarified lysate was used in each ChIP. Prior to the ChIP, antibody was prebound to protein A-conjugated magnetic beads (Dynabeads, Invitrogen). The ChIP was washed five times with ChIP lysis buffer and then eluted twice with 50 µl 0.1 M sodium bicarbonate, 1.0% SDS for 10 min at 37 °C. The cross-links were reversed by heating at 65 °C for a minimum of 4 hours after the addition of 2 µl of 5M NaCl. After the cross-links were reversed, chromatin protein was removed with 1 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 55 °C for 1-hour. DNA was then purified by a phenol-chloroform extraction and used in a quantitiative PCR reaction using oligonucleotides (oligos-m*Per2*_ChIP_Upstream_F/R and m*Per2*_ChIP_qPCR_F/R) listed in **Table 4-1**. Oligonucleotides for *Per2* E-box (oligos- m*Per2*_ChIP_qPCR_F/R)

were identical to ones reported previously [37]. The H3K9me3 (#ab8898) and H3K27me3 (#ab6002)) antibody was purchased from Abcam.

4.3.4 Western blotting

The cells were lysed in RIPA buffer (150mM NaCl, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS, 25mM Tris pH 7.4) with protease inhibitors (1mM PMSF, 1µg/ml leupeptin, 1µg/ml pepstain). The lysate was cleared of cellular debris by centrifuge at 5000g for 10 min at 4 °C. The protein sample 2.5 × sample buffer and then heat-denaturing 95ºC for 5 min then separated on 10% SDS-polyacrylamide gels and transferred to PVDF. Proteins were detected with primary antibody at room temperature for 1 hour, followed by 5 washes with TBST (1X Tris-buffered saline with 0.1% Tween), then incubated with secondary antibody for 1 hour at room temperature. and detected with ECL Plus western blotting substrate (Thermo Scientific). hnRNP antibodies were obtained from Abcam, hnRNP M (#ab177957), hnRNP F (#ab50982), and hnRNP L(#ab6106).

4.3.5 Real-time Bioluminescence monitoring

The real-time bioluminescence monitoring assay was performed as previously described [345]. Briefly, the synchronous *Per2*::*Luc* MEF cells were washed with PBS and replaced with 2 ml of assay medium (phenol red free L15 (Life Technologies), 10% FBS, 1% Pen Strep and 0.1 mM luciferin). Tissue culture dishes were sealed with 38 circle #1 glass coverslips and silicone grease and transferred to the LumiCycle (Actimetrics, Wilmette, IL) at 36°C. Data analysis was performed with the LumiCycle Analysis software (Actimetrics).

4.3.6 shRNA Transfections

The shRNA constructs were made by first annealing two DNA oligonucleotides (IDT) then the annealed oligos were cloned into pSUPERIOR (Oligoengine) by standard cloning methods using BglII and XhoI and T4 DNA ligase. The shRNAs contained 19-nt of DNA matching the mRNA transcript of the targeted gene driven by H1 promoter. The negative control was a scrambled shRNA sequence used previously [346]. Cells were transfected by reverse-transfection using Lipofectamine 3000 (Invitrogen L3000001) as described [347]. Briefly, for 6 well plates, I prepared 250 µl of DNA-lipid mixture composed of 125 µl Opti-MEM media containing 7.5 µl lipofectamine 3000 and 125 µl Opti-MEM containing 5 µg pSUPERIOR-shRNA and 10 µl P3000 reagent for each well and then incubated the mixture at room temperature for 20-30 min. Confluent cells from a 100 mm plate were incubated with 1 ml of 0.25% trypsin then incubated at 37^{o}C for 5 minutes and quenched 5 ml DMEM 10% FBS. Cells were spun at 150 x g 5 min and resuspended to 125,000-150,000 cells/ml in a 12 ml volume. 2 mls of cell and media were added to a six-well plate and the DNA-lipofectamine mix was added to the cells. After 24 hours, the media with transfection reagent was changed to DMEM 10% FBS and the cells were allowed to grow for an addtional 48-72 hours, then harvested in RIPA buffer and used to check knockdown efficiency. The shRNA sequences of the target genes are listed in **Table 4-1.**

4.3.7 Comprehensive identification of RNA-binding proteins by mass spectrometry (ChIRP-MS)

The ChIRP-MS was performed following published methods [190, 348]. Ten 100 x 20 mm dishes of NIH 3T3 cells were used per ChIRP-MS experiment. Cells were harvest by treating with trypsin and quenched with cultured media. The suspended cells were pellet and washed 3 times with PBS. Cells were cross-linked with 3% formaldehyde for 30 min and quenched by incubating in 0.125 M glycine for 5 min. Cells were lysed in lysis buffer (50mM Tris-Cl pH 7.0, 10mM EDTA, 1% SDS) containing protease inhibitors (1mM PMSF, 1μg/ml leupeptin, 1μg/ml pepstain) and RNase inhibitor (Invitrogen AM2694, Superase-in) (1 mL lysis buffer per 100 mg cell pellet). After lysis, DNA was sheared to 100-500 bp by sonication and the lysate was incubated with biotinylated anti-*Per2AS* DNA probes. The lysate and DNA probes (100 pmol probes per 1ml lysate) were hybridized overnight in hybridization buffer (750 mM NaCl, 1% SDS, 50 mM Tris-Cl pH 7.0, 1 mM EDTA, 15% formamide contain PMSF, P.I. and Superase-in). In total, 10 anti-*Per2AS* biotinylated DNA probes spaced approximately 200-500 bp apart were used in the hybridization. The negative control contained biotinylated oligos (Biotin_FrqP12F and Biotin_FrqP7R) that matched the *frequency* gene from *Neurospora*. The biotinylated oligo probes (Biotin_m*Per2*_1-10) are listed in **Table 4-1.** The DNA probes and the associated RNA-binding proteins were recovered by adding 100 μl Streptavidin C-1 magnetic beads (Invitrogen) and incubating for 1 hour at 37°C. The beads were washed five times for 5 min at 37°C with wash buffer (2x NaCl and Sodium citrate (SSC), 0.5% SDS, 1mM PMSF). After the final wash, one tenth of the fraction was removed and used to check *Per2AS* enrichment by RT-PCR after RNA isolation.

Finally, the protein was eluted two times with biotin elution buffer (12.5 mM biotin (Invitrogen), 7.5 mM HEPES (pH 7.5), 75 mM NaCl, 1.5 mM EDTA, 0.15% SDS, 0.075% sarkosyl, and 0.02% Na-Deoxycholate), first at room temperature for 20 min and then at 65°C for 10 min. The eluted protein was precipitated with 10% trichloroacetic acid (TCA) at 4°C overnight, pelleted at 16, 000 × g at 4°C for 30 min and washed with ice cold acetone. Proteins were solubilized in 1× sample buffer and boiled at 95°C for 30 min to reverse-crosslinking. The identity of eluted proteins was determined by mass spectrometry at the Rutgers Biological Mass Spectrometry Core Facility.

4.3.8 dCas9-VPR induction

Per2AS was induced with transfected SP-dCas9-VPR as described [349]. Briefly, the plasmid containing SP-dCas9-VPR (Addgene Plasmid #63798) along with 4 gRNA (localized to intron 6 of *Per2*) constructs under control of U6 promoter were transfected into cells as described above. The gRNA target sequences (m*Per2*_AS_gRNA 1-4) were listed in **Table 4-1.** The transfection was performed using 7μg of dCas9-VPR and 750 ng of each individual gRNA (total 3μg) per well in a 6-well plate. The lipofectamine transformation mix was replaced with fresh media after 24 hours and the cultures were grown for an additional 24 hours then processed for real time PCR to check *Per2AS* induction or for ChIP.

4.3.9 R-loop formation

Detection of *in vitro* R-loop formation was measured by gel mobility following published methods [350]. A 1 kb DNA fragment of *Per2AS* was cloned into pCR4-TOPO plasmid (Invitrogen), with T3 promoter driving *Per2AS* transcription and T7 promoter

driving *Per2* transcription (**Supplemental Figure 4-5A**). The oligonucleotide sequences used for TOPO cloning are listed in **Table 4-1**. The *in vitro* transcription was performed with the MAXIscript T7/T3 transcription kit (Invitrogen) following manufacture's guidelines. Briefly, 1 μg of plasmid was used in a 20 μl reaction mix containing transcription buffer and 1μl of each of 10mM ATP, CTP, GTP, and UTP, and T7 or T3 RNA polymerase. The transcription mixture was incubated for 1 h at 37° and terminated by incubation at 70° for 15 min. Free RNA was degraded by incubating at 37° for 30 min in the presence of 1 μg of RNase A. The control reaction was done under identical condition, except 0.5 U of Ribonuclease H (Invitrogen) was added along with the RNase A. The nucleic acids were purified by phenol and chloroform extraction followed by ethanol precipitation and dissolved in 10 μl of TE buffer (10 mM Tris-Cl, pH 8.0, 1mM EDTA). The purified nucleic acids were then resolved on a 1% agarose gel run in 1×Tris-Borate-EDTA buffer (without ethidium bromide) for 2–3 h at 70 volts, stained with ethidium bromide, and visualized under UV light.

Table 4-1 Sequences of Oligos and probes

Oligo name	Sequences
m*Per2*_RT	TGCTCTTGCACCTTGACCAGG
m*Per2*_intron5_AS_RT	CCTGCTAGTGTTGCATTGGGC
m*Per2*_exon6_F4	CCAAGTGGCCTCCATCTTT
m*Per2*_exon6_R4	AGGGGTGGTGTAGCTGTGG
Biotin_m*Per2*_1	5'Biotin-CCAGCCACGAGTCCAAAGACC
Biotin_m*Per2*_2	5'Biotin-GCACAGGTGGCTGGGATGCTTC
Biotin_m*Per2*_3	5'Biotin-CCAGTGAGAGCCAGCCCTGC
Biotin_m*Per2*_4	5'Biotin-CATCTAAGGGTCTGGGAGAG
Biotin_m*Per2*_5	5'Biotin-CCTGCTAGTGTTGCATTGGG
Biotin_m*Per2*_6	5'Biotin-CATCTCTAACCAAGTGGCCTC
Biotin_m*Per2*_7	5'Biotin-GGCAGAGGAGTGTTCGGTGG

Biotin_m*Per2*_8	5'Biotin-GTCCTCTCCAAAACTCCAGTTGG
Biotin_m*Per2*_9	5'Biotin-GGAAGCAGCTCACCCTCCAGG
Biotin_m*Per2*_10	5'Biotin-GGCTGCAATCTAGGCAGCGT
Biotin_FrqP12F	5'Biotin-GGCGCCGACATTGTCGGCCC
Biotin_FrqP7R	5'Biotin-CCTTAACAGGCCAGTGTTGG
sh-hnRNP M-1	GCAATCGCTTTGAGCCATA
sh-hnRNP M-2	GGAAGATCCTGATGGTGAA
sh-hnRNP M-3	GCACAGTATTTGTAGCAAA
sh-hnRNP F-1	GCATGGGACACCGGTATAT
sh-hnRNP F-2	GGCACAGGTATATTGAAGT
sh-hnRNP F-3	GGGAGAGACCTCAGCTATT
sh-hnRNP L-1	CCAAGCAACCAGCCATTAT
sh-hnRNP L-2	CCTGCAGCTATAAAGACTT
sh-hnRNP L-3	GCAATGTACTGCACTTCTT
sh-Control	CAGTCGCGTTTGCGACTGG
m*Per*2_ChIP_Upstream_F	CAGGGCCCTCAAAGATATGG
m*Per2*_ChIP_Upstream_R	CGCCTCTCTGACTGCTTCCT
m*Per2*_ChIP_qPCR_F	AAGAGCGCGCAGCATCTTCATT
m*Per2*_ChIP_qPCR_R	ATTGGTCGGAGTGCCACCTCATTT
m*Per2*_AS_gRNA1	AGTAAGCCATGGGGCAGAGC
m*Per2*_AS_gRNA2	ACGTGCACCAAGTATGTAGC
m*Per2*_AS_gRNA3	TCTCAAACTCACGCTGAATT
m*Per2*_AS_gRNA4	TGACCATTATGTAACTCCAC
R_loop_TOPO_F	GAACTTGTTGCTCCTGCT
R_loop_TOPO_R	CCCACCGAACACTCCTCTG

4.4 Results

4.4.1 *Per2AS* rhythm is dependent on BMAL1.

Circadian whole-transcriptome RNA-sequencing revealed *Per2* has an antisense transcript (*Per2AS*) that oscillates anti-phasic to *Per2* in mouse liver [79, 80]. This pattern is similar to *frequency* NAT *qrf* in *Neurospora* [78, 79]. The *qrf* transcript is also light induced and dependent on WCC [351]. To understand the function and regulation of

Per2AS, I first investigated how BMAL1 impacts *Per2AS* expression. To accomplish this, I performed strand-specific real-time PCR on RNA isolated for WT and *Bmal1*$^{-/-}$ liver. I confirmed the *Per2* (**Figure 4-1A.**) and *Per2AS* (**Figure 4-1B.**) rhythms are anti-phasic in WT and are consistent with previous findings [79, 80]. However, loss of BMAL1 caused *Per2AS* to lose rhythmicity (Cosinor, $p > 0.05$) and only basal level of *Per2AS* was detectable. These data indicate the expression and diurnal rhythm of *Per2AS* is dependent on BMAL1, either directly or indirectly.

The circadian clock is a cell-autonomous oscillator and rhythms can be detected in entrained NIH 3T3 cells by treating with high concentrations of serum or chemicals [352]. Therefore, we sought to examine the *Per2AS* rhythm in NIH 3T3 cells between 0-24 hours and 48-72 hours after entraining with dexamethasone. We found *Per2AS* was rhythmic ($n=6$, Cosinor $p<0.05$) in the first 24 hours after synchronization; however, at 48-72 hours after synchronization the rhythm in *Per2AS* was dampened to point where it was insignificant (**Supplemental Figure 4-1)** ($n=5$, Cosinor $p> 0.05$). The loss of the rhythm is likely due to the limited nutrients for cell growth. Interestingly, in the first 24 hours, the *Per2AS* rhythm was not anti-phasic to *Per2* as in liver tissue. Instead, the *Per2AS* expression pattern of cells under chemical synchronization was similar to light-induced *qrf* in *Neurospora*, where the external zeitgeber induced an increase in both the sense and antisense expression at same time. This suggest that one potential function of *Per2AS* may be to assist in entrainment.

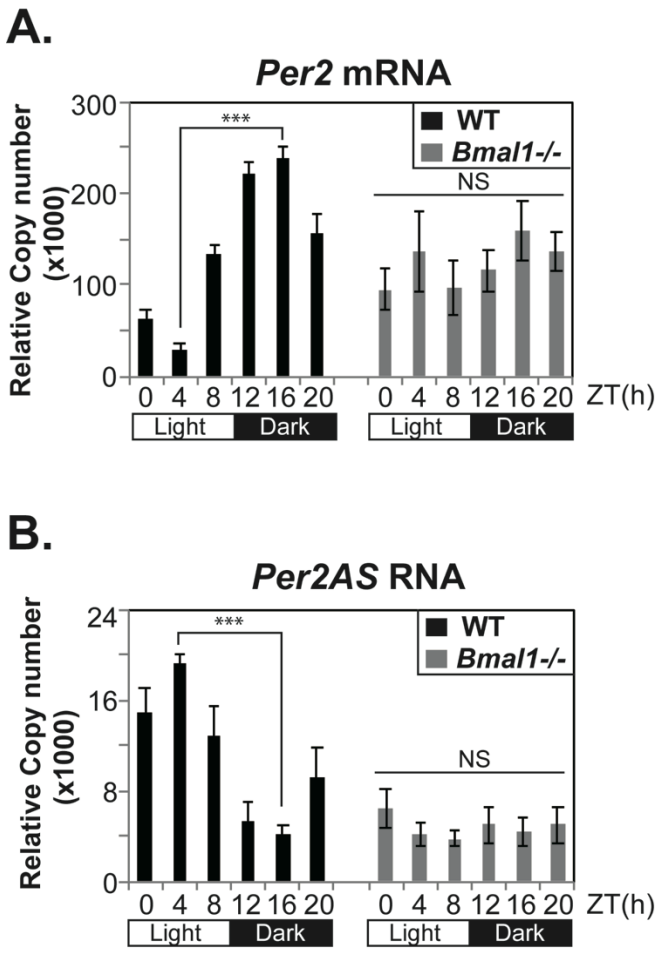


Figure 4-1 BMAL1 regulates *Per2AS*.

Strand-specific real-time PCR was used to determine the levels of *Per2* mRNA (A) and *Per2AS* RNA (B). The y axis represents the relative copy number of both transcripts in WT (black, Cosinor $p<0.005$) and Bmal1-/- (grey, Cosinor $p> 0.05$) over 24 hours under 12:h-12:h light-dark cycle. The white and black bars indicate the light-dark cycle. The error bar represents the mean +/- standard error of the mean (SEM) for 4 independent biological replicates (n=4). Peak to trough statistics was performed by one-way ANOVA with Tukey Post Hoc test (***$p\leq0.0005$ and NS (not significant) $p>0.05$).

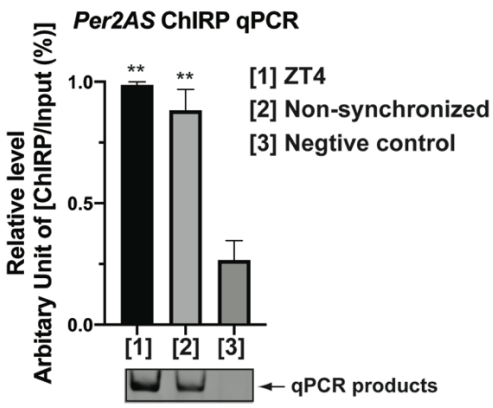
4.4.2 Identification of *Per2AS*-interacting proteins

In an effort to understand the role of *Per2AS* (if any) in clock function, I took a biochemical approach and identify *Per2AS*-associated proteins by comprehensive identification of RNA-binding proteins by mass spectrometry (ChIRP-MS) [190]. This method was used to identify 81 proteins associated *Xist* , including proteins involved in chromatin modification, nuclear matrix, and RNA remodeling pathways [190]. To identify *Per2AS* interacting proteins, we performed ChIRP-MS on NIH 3T3 lysates using non-synchronized NIH 3T3 cells, or cells isolated 4 hours after entrainment with 100 nM dexamethasone (Dex) (ZT4). To isolate *Per2AS*, we used 10 biotin-labeled oligonucleotides capable of hybridizing to *Per2AS*. As a control, I used scrambled oligonucleotides that do not match the mouse genome. The initial design used samples at ZT4 and unsynchronized cells with the intention to find *Per2AS*-associated proteins during the peak of expression and determine if they differed. To validate the ChIRP, I checked the enrichment of *Per2A*S by strand-specific RT-PCR. The *Per2AS* antisense DNA probes retrieve significantly higher *Per2AS* RNA (3-fold more) in both the ZT4 samples and non-synchronized cells compared to the negative control (**Figure 4-2A.**).

After confirming the ChIRP selectively enriched *Per2AS*, I performed mass-spectrometry on the eluted proteins. The table in **Figure 4-2B** shows peptide counts of enriched proteins. Proteins shaded black had peptide counts $\geqq$ 5 and 2-fold more peptides in both ZT 4 and non-synchronized cells compare to the negative control. Proteins shaded grey did not meet that cut-off, but may be significant because the PER2-associated RNA-binding protein NONO fell in that group [353]. The most abundant *Per2AS*-associated proteins include members of the heterogeneous nuclear

ribonucleoproteins (hnRNPs), hnRNP M, hnRNP U, hnRNP F and hnRNP L. In addition, the RNA helicase, DDX5 and splicing factor proline and glutamine rich (SFPQ) were highly enriched. DDX5 is a component of the nuclear Per-complex and is involved in transcriptional termination [98]. SFPQ has also been shown to be associate with *Per2* and is thought to facilitate transcriptional inhibition [37]. The finding that *Per2AS* interacts with both DDX5 and SFPQ suggest it may play an important role in negative feedback. The presence of known components of the PER complex, such as DDX5 and SFPQ provided further confidence in the fidelity of the ChIRP. Given that DDX5 and SFPQ have already been characterized in clock function, we focused our attention on understanding potential roles for the hnRNP proteins.

A.



B.

Protein	Peptide		
	ZT4	Non-synchronized cell	Negative control
hnRNP M	26	24	3
hnRNP U	17	10	4
DDX5	12	8	2
hnRNP F	12	7	3
hnRNP L	7	5	1
SFPQ	8	8	2
hnRNP K	19	10	6
J3QNY1	18	12	7
NONO	6	4	3
DDX3X	7	6	4
PABP1	7	4	3
RBM14	6	4	0
FUBP2	4	3	1
UBP5	5	3	0

Figure 4-2 ChIRP-MS of *Per2AS*.

(A) The bar plot shows *Per2AS* enrichment from the ChIRP measured by qPCR. Data were normalized with total input samples (n=3, t-test, two tailed. **p≤0.005). Products from the *Per2AS* ChIRP were resolved on a 10% TBE gel to analyze specificity and background (B) Values indicate the number of different peptides (Peptide Counts) detected by mass spectrometry in different ChIRP samples (Non-synchronized cells, ZT4 sampled at 4 hours after dexamethasone treatment, and negative control). Proteins shaded black represent peptide counts ≧5, and more than 2-fold enrichment in both samples relative to background seen in the negative control. The grey color proteins also appeared to be enriched in the ChIRP-MS but did not pass the cutoff, which may provide insights for future study.

4.4.3 hnRNP M assists clock regulation

hnRNPs are a large family of RNA-binding proteins mostly localized in the nucleus [354, 355]. Among other functions, hnRNPs are involved in regulating gene transcription [354]. Accumulating evidence indicate hnRNPs can interact with lncRNAs and are involved in assisting gene expression, maintaining RNA stability, and aid in controlling chromatin [354, 355]. To investigate whether hnRNPs affect clock function, we measured luciferase rhythms in *Per2*::*Luc* MEFs using shRNA directed against hnRNP M, hnRNP F and hnRNP L. First, I confirmed knockdown of hnRNP M by western blot using three independent shRNA (**Figure 4-3A**). The knockdown experiment indicated there was a sizable reduction in the amount of hnRNP M protein. Therefore, I proceeded to tested whether knock down of hnRNP M affected *Per2*. In collaboration with Dr. Joshua Gamsby at University of South Florida, I determined that hnRNP M is needed to maintain the normal *Per2* amplitude and a reduction in hnRNP M caused a slight increase period length (**Figure 4-3B-D.**). Similar experiments examining hnRNP F or hnRNP L did not affect the underlying period (**Supplemental Figure 4-2A-F.**). Thus, these data indicated that hnRNP M, via its association with *Per2AS* is important for normal *Per2* expression and clock function.

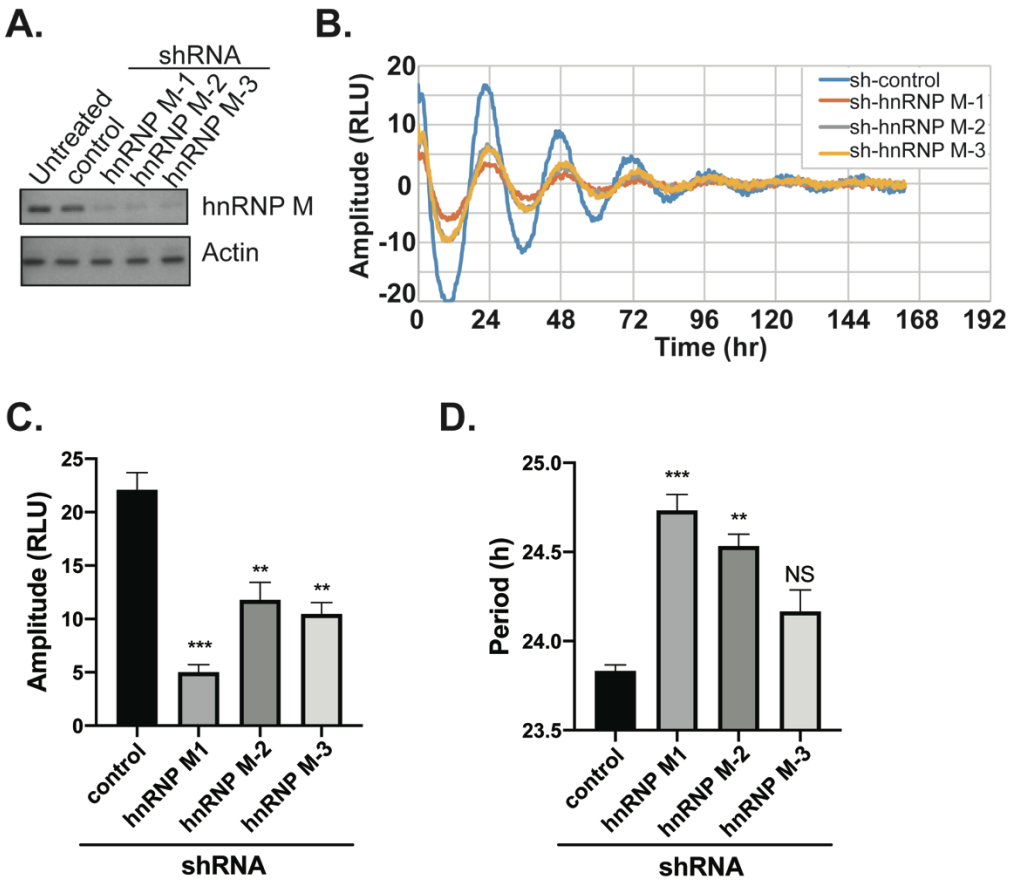


Figure 4-3 Knockdown of hnRNP M affects the *Per2* rhythm.

(A) Western blot showing the abundance of hnRNP M in cultured unsynchronized mouse fibroblasts after delivery of three different shRNA targeting hnRNP M relative to a scramble shRNA control. ACTIN was used as a loading control. Untreated cells lacked shRNA and lipofectamine treatment **(B)** Circadian oscillations of bioluminescence in synchronized reporter line after delivery of scramble control shRNA (blue) or three different hnRNP M sh-RNAs (Orange, grey, yellow). The bar plot shows the quantification data in B examining amplitude **(C)** and period **(D)** (RLU, relative light unit. x-axis, hours after synchronization.). The bar plots are the mean +/- SEM of 3 independent (n=3) experiments. The statistical analysis was performed by one-way ANOVA with Tukey Post Hoc test showing ***$p \leq 0.0005$, **$p \leq 0.005$ and the NS represents $p > 0.05$.

4.4.4 hnRNP M affect the H3K9me3 and H3K27me3 at *Per2*.

I next sought to understand the function hnRNP M plays via its interaction with *Per2AS*. In general, lncRNAs, including antisense transcripts, can serve as scaffolds to recruit heterochromatin assembly factors [356-359], a function that is shared by *qrf* in *Neurospora* [114]. Specifically, evidence indicates that hnRNPs interact with non-coding RNAs and impact nuclear architecture. For example, the lncRNA *Firre* interacts with hnRNP U and affects chromatin organization [354]. Therefore, I investigated whether hnRNP M affected H3K9me3 at *Per2*. I performed H3K9me3 ChIP examining both the *Per2* E-box and an upstream region on *Per2,* with and without hnRNP M depletion. I found that depletion of hnRNP M caused a reduction in H3K9me3 at *Per2* **(Figure 4-4A. and Supplemental Figure 4-3B.).** This indicates hnRNP M, presumably through its interaction with *Per2AS,* contributes H3K9me3 formation at *Per2*.

In addition to H3K9me2/3 at *Per2*, KMT6A/EZH2-mediated H3K27me3 is also implicated in clock function. KMT6A/EZH2 is a Polycomb group enzyme that methylates H3K27 and co-precipitates with CLOCK:BMAL1 in liver extracts [123]. Depletion of KMT6A/EZH2 alters luciferase rhythm of both *Per2* and *Bmal1* [123]. Data indicate hnRNP K interacts with *Xist* and contributes to H3K27me3 on the inactive X chromosome [190, 360]. Based on similar function of hnRNP K, H3K27me3 and *Xist*, I tested whether hnRNP M is involved assisting H3K27me3 at *Per2.* I determined that depletion of hnRNP M led to a reduction in H3K27me3 at *Per2* **(Figure 4-4B. and Supplemental Figure 4-3C.)**. For all the ChIP experiments, I assayed the knockdown efficiency hnRNP M and consistently obtained reduced levels (**Figure 4-4C**.). Overall, I found that depletion of hnRNP M caused a reduction in both H3K9me3 and H3K27me3

at *Per2* indicating hnRNP M, presumably through its association with *Per2AS,* assists heterochromatin formation at *Per2*.

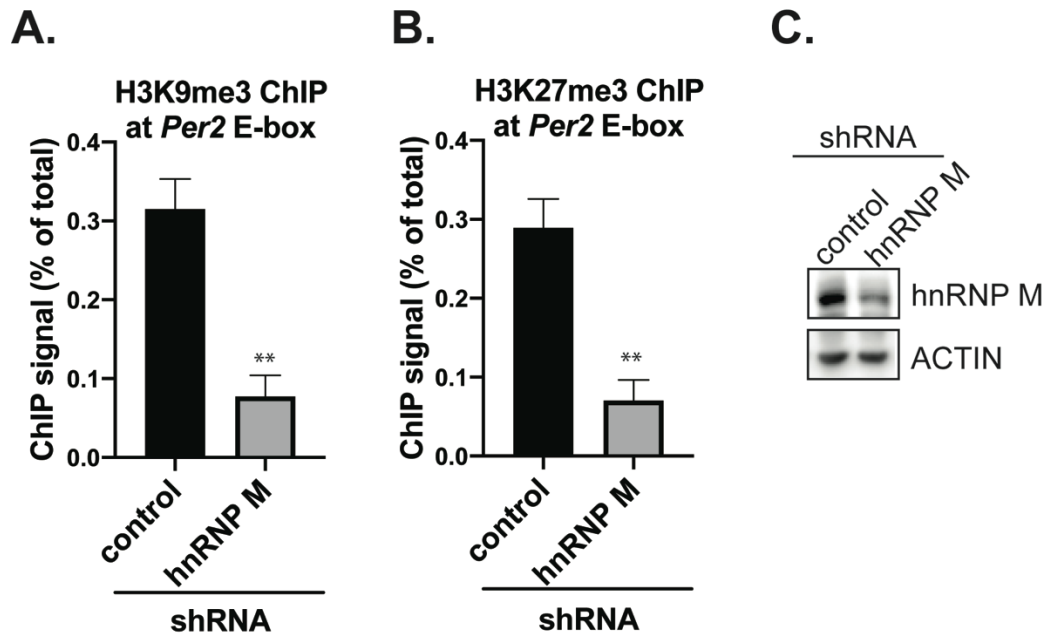


Figure 4-4 Depletion of hnRNP M causes a reduction in H3K9me3 and H3K27me3 at *Per2*

The ChIP-qPCR (mean +/-SEM of 4 independent experiments. t-test, two tailed. **p≤0.005) showed a substantial reduction in H3K9me3 **(A)** and H3K27me3**(B)** at *Per2* E-box site after knockdown of hnRNP M compared to unsynchronized mouse fibroblasts cells. **(C)** Representative Western blot showing of lysates from cells used in the ChIP measure hnRNP M abundance after shRNA knockdown using ACTIN as loading control.

4.4.5 Induction of *Per2AS* enhances *Per2* expression

To further investigate the role of *Per2AS*, I induced expression using defective RNA-guided nuclease-null dCas9 fused to VP64-p65-Rta (dCas9-VPR) trans-activation domain. Using this system, I routinely achieved a 3-fold increase in *Per2AS* expression (**Figure 4-5A**). Accompanying the induction was a slight, but consistent increase in

endogenous *Per2* expression (**Figure 4-5B**). Next, I examined rhythms in *Per2* using *Per2*::*Luc* MEF cells over 72 hours and repeatedly, but not consistently, saw slightly elevated levels in *Per2* expression **(Figure 4-5C**). I am currently repeating these experiments with our collaborator (Dr. Joshua Gamsby) because our system was not designed to faithfully measure *Per2::Luc* rhythm.

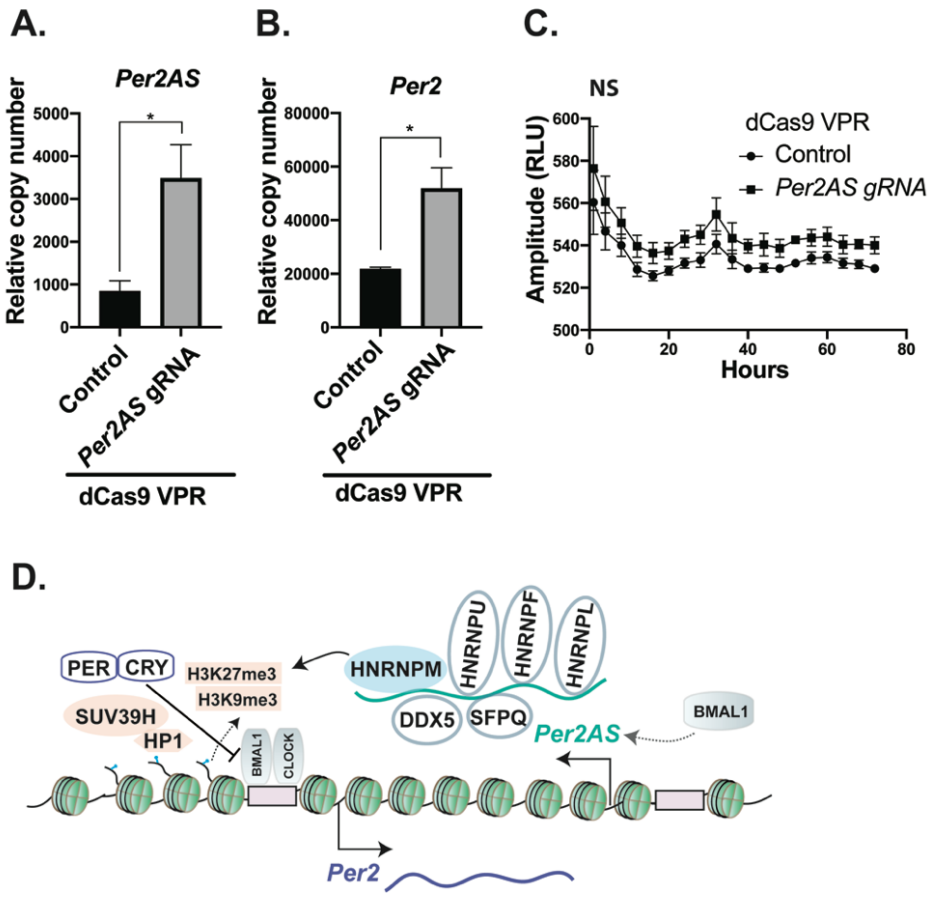


Figure 4-5 Induction of *Per2AS* causes an increase in *Per2* expression.

The bar plot shows the levels of *Per2AS* (A) and *Per2* (B) RNA after inducing *Per2AS* expression using dCas9-VPR with *Per2*, intron 6-specific gRNAs and control (no gRNAs) (n=4, t-test, two tailed, *p≤0.05.). (C) Circadian oscillations were measured using bioluminescence of *Per2*::*Luc* MEF cells over 72 hours with and without induction of *Per2AS* (n=3, NS represents p>0.05). (RLU, relative light unit, x-axis, hours after synchronization). (D) Potential model of *Per2AS.* Schematic representation of *Per2AS* function whereby it helps establish H3K9me3 and H3K27me3 facultative heterochromatin formation at *Per2* E-box through an association with hnRNP M. (Blue line: *Per2* mRNA. Green line: *Per2AS* RNA).

The increase in *Per2* expression with *Per2AS* induction is reminiscent of *qrf's* effect on *frq* [114]. Recent reports indicate that antisense transcripts can form R-loops to

promote sense transcription and hnRNP M associates with RNA-DNA hybrids [194, 361]. Therefore, I sought to examine whether induction of *Per2AS* could form R-loops. To accomplish this, I cloned a 1kb fragment of the *Per2* locus encompassing *Per2AS* and performed an in vitro transcription assay to see if *Per2AS* could form a DNA-RNA hybrid. Unfortunately, the assay proved to be problematic and I was not convinced of any R-loop formation induced by *Per2AS* **(Supplemental Figure 4-5B)**.

Because reductions in hnRNP M caused a defect in H3K9me3 and H3K27me3 and because *Per2AS* appeared to increase *Per2*, I also examined if levels of H3K9me3 and H3K27me3 were altered at *Per2* after induction of *Per2AS* (**Supplemental Figure 4-4**). H3K9me3 and H3K27me3 ChIP revealed there was no significant difference in H3K9me3 or H3K27me3 levels after induction of *Per2AS* with dCas9-VPR. These experiments reveal that hnRNP M is necessary for normal levels of H3K9me3 and H3K27me3, but artificial induction of *Per2AS* is not sufficient to induce their addition or removal.

4.5 Discussion

In this chapter, I set out to understand the role of *Per2AS* and how it is regulated. I began by examining *Per2AS* expression and found it was dependent on BMAL1. I also identified *Per2AS*-associated proteins by ChIRP-MS and examined if the top candidates were involved in clock regulation. Among them, the one with the most pronounced phenotype was hnRNP M. Knockdown of hnRNP M revealed that it affects the amplitude and period of *Per2*. The phenotype of hnRNP M contrasts KMT1/SUV39H1 and KMT6A/EZH2, where depletion of KMT1/SUV39H1 shortens the period length of *Per2* and depletion of KMT6A/EZH2 reduced the *Per2* amplitude in mouse fibroblast cells [37, 123]. Based on the growing evidence that hnRNP M assists in establishing epigenetic states and the role of *qrf* in *Neurospora*, I examined whether H3K9me3 and H3K27me3 levels were altered at *Per2* in hnRNP M knockdown. I found depletion of hnRNP M led to a reduction in both H3K9me3 and H3K27me3 at *Per2*. These results indicate that hnRNP M contributes to facultative heterochromatin formation at *Per2* and presumably this occurs via its associations with *Per2AS*. These findings can be extrapolated into a model (**Figure 4-5D**) in which *Per2AS* could potentially serve as a scaffold to recruit the regulatory proteins that coordinate heterochromatin formation at *Per2*.

It is interesting, but not unexpected that deletion of BMAL1 disrupt *Per2AS* expression. At the very least, I expected a loss in rhythmicity at *Per2AS,* but I also found the levels of *Per2AS* only achieved basal levels of expression and this is similar to WCC regulating both *frq* and *qrf* in *Neurospora* [351]. The regulation of *Per2AS* expression, like *qrf* is complicated, as is coordinate regulation of most all sense-antisense pairs when

viewed at the level of chromatin. RNA-seq data indicate there are oscillations of H3K4me3 at the potential transcription start site (TSS) of *Per2AS*, which is anti-phasic with the H3K4me3 at *Per2* [80]. This indicates histone modifications, along with chromatin state changes, are critical for sense-antisense transcription. Consider the anti-phasic expression of both *Per2* and *Per2AS* where they share the same transcription factors (e.g. BMAL1, and presumably CLOCK or NPAS2 (Neuronal PAS Domain Protein 2)). This anti-phasic, convergent transcription means there are regulatory switches, presumably at the level of chromatin, that provide a directionality que. Thus, specific modifications (possibly H3K9me3) are presumably laid down by PER-complex, which is then read by a chromatin-effectors to control accessibility of the E-box elements at both *Per2* and *Per2AS*. This somewhat speculative model can serve as a foundation for future research on the circadian directionality cue and should be factored into circadian chromatin studies in the future.

The availability of RNA-centric approaches is helping to uncover lncRNA-interacting proteins, which is critical to understanding their function. Techniques to interrogate lncRNA-protein interactions, include cross-linking and immunoprecipitation (CLIP) [188], RNA immunoprecipitation (nRIP) [189], Comprehensive identification of RNA-binding proteins by mass spectrometry (ChIRP-ms) [190], capture hybridization analysis of RNA target (CHART) [191], and RNA antisense purification (RAP) [192]. These techniques are starting to provide insight into the composition and regulatory roles of lncRNAs. Well-studied examples included, *Xist*, the long non-coding RNA for X-chromosome inactivation and heterochromatin formation [362], and Telomeric Repeat containing RNA (*TERRA),* a non-coding RNA that originates at the telomeres that are

essential for establishment of heterochromatin makers at telomere [221]. Proteins that interact with *Xist* include subunits of PRC1 (Polycomb repressive complex 1), HDAC (histone deacetylase) complexes, and hnRNP U and hnRNP K [190]. Proteins that interact with *TERRA* can be subdivided in multiple groups based on their functions including cell cycle, shelterin complex, chromatin, and nuclear matrix proteins including hnRNPs [302]. To better understanding the function of *Per2AS*, we isolated *Per2AS*-associated proteins using the ChIRP-MS and found several hnRNPs family members are associated with *Per2AS*. Interestingly, our top candidates that interact with *Per2AS* are the same nuclear matrix proteins hnRNPs, which also associate with *Xist* [190] and *TERRA* [363] and both of these are implicated in heterochromatin formation. In the future studies, it will be interesting to know whether hnRNPs binding to lncRNAs, especially NATs, is a genome-wide event, and to determine whether hnRNPs have specific or global functions.

hnRNP are somewhat difficult to study because deletion of some hnRNPs (hnRNP A1, hnRNP L and hnRNP U) is embryonic lethal, and lethality causes pleiotropic effects [364-366]. There are however, studies linking hnRNPs to the circadian clock and loss of one copy of hnRNP U is sufficient to change mouse locomotor activity and metabolic rhythms [367]. Because hnRNP U was already somewhat characterized in clock function, I focused on the other hnRNPs that are associated with *Per2AS*. I showed that hnRNP M interacts with *Per2AS* and knockdown of hnRNP M alters the amplitude and period of *Per2*. In addition to the hnRNPs, I also identified proteins that are also part of the nuclear PER complex, such as DDX5. An a priori hunch was that *Per2AS* may be contained in the nuclear PER complex serving as an intermediate or scaffold to assist

clock function; however, besides for DDX5 and possibly NONO, we didn't detect any core clock component associated with *Per2AS*. That said, ChIRP-MS enrichment of a low-abundant non-coding RNA such as *Per2A*S was significantly difficult and I believe I missed a number of interacting proteins. Whether or not *Per2AS* is part of nuclear PER-complex associated with DDX5, NONO and SPFQ remains to be determined, as does experiments examining whether hnRNPs associate with the nuclear PER complex.

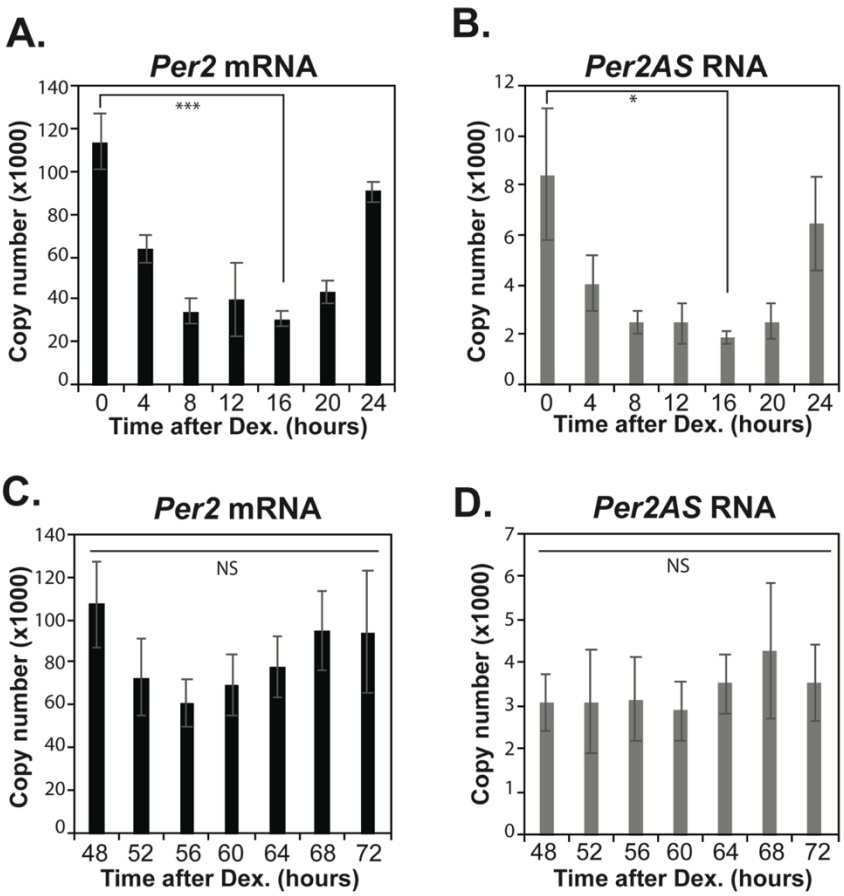
LncRNAs studied to date indicate they usually associate with specific proteins to affect structural and regulatory outcomes. For instance, the lncRNAs, *Xist*, and *RepA* can recruit Polycomb complex to the target genomic loci, which leads to H3K27me3-dependent heterochromatin formation and represses gene expression [190]**.** The long non-coding RNA, *H19* is associated with *Igf2* imprinting [368]. H19 interacts with hnRNP U, which is required for H19-mediated transcription repression [369]. I showed that *Per2AS* interact with hnRNP M, which can affect heterochromatin at *Per2* as evidenced by reduction in H3K9me3 and H3K27me3 levels. This suggests that hnRNP M is needed to recruitment methyltransferase(s) or HP1. Consistent with this notion, hnRNP U has been shown to interact with HP1 and needed for transcriptional repression [370] Whether hnRNP M or other hnRNPs interact with HP1 and assist global transcription remains an open question.

We also tested how induction of *Per2AS* alters the system and found elevated levels of *Per2,* suggesting induction creates a more permissive chromatin state. In order to try to understand the mechanism, I tested whether the nascent *Per2AS* transcript invades the *Per2* DNA duplex to form an R-loop because antisense transcripts are capable of forming R-loops with the sense DNA [194, 371]. Unfortunately, this assay

proved challenging and conclusive results were elusive, possibly due to the partial sequence I cloned, and it did not behave *in vitro* as it does *in vivo.* Or the R-loop only forms at the beginning of entrainment and it was difficult to detect. As all transcription have the tendency to form the RNA:DNA hybrids structure, cells evolved different mechanism to avoid it, including RNase H-mediated degradation of RNA:DNA hybrids [372]. With these uncertainty, future study still needs to focus on both *in vitr*o and *in vivo* studies using different assays. For the *in vitro* studies, more precise sequence of the *Per2AS* may be needed to detect R-loop formation of gel mobility experiments. The *in vivo* studies could focus on detecting RNA:DNA hybrids with the S 9.6 or expression of exogenous catalytically inactive RNase H followed by ChIP to check the RNase H enrichment as a sign of R-loop formation [373].

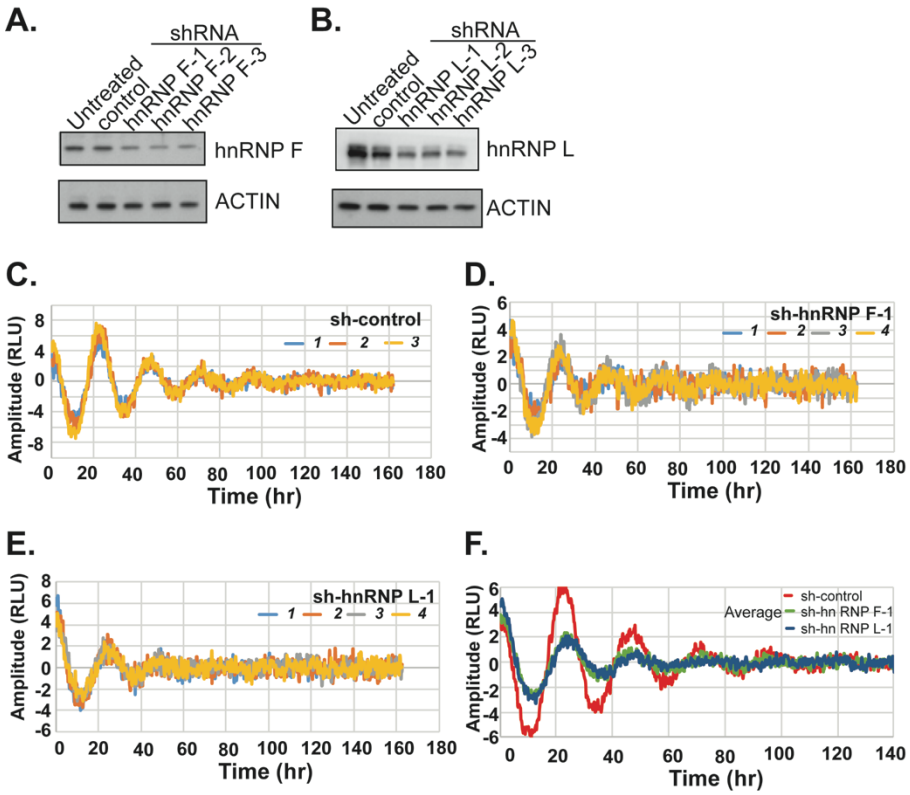
In conclusion, I set out to understand crosstalk between the circadian clock, *Per2AS*, and circadian-regulated facultative heterochromatin. My study found *Per2AS* is regulated by BMAL1 and associates with hnRNPs. Reduction in hnRNP M caused a minor clock phenotype and altered heterochromatin at *Per2*. The work presented here increases our understanding of *Per2AS* incrementally, from nothing to something, and I am in full acknowledgement that much more work needs to be done to fully understand the molecular mechanism of circadian-regulated heterochromatin and the role of antisense transcripts.

4.6 Supplemental Figures:



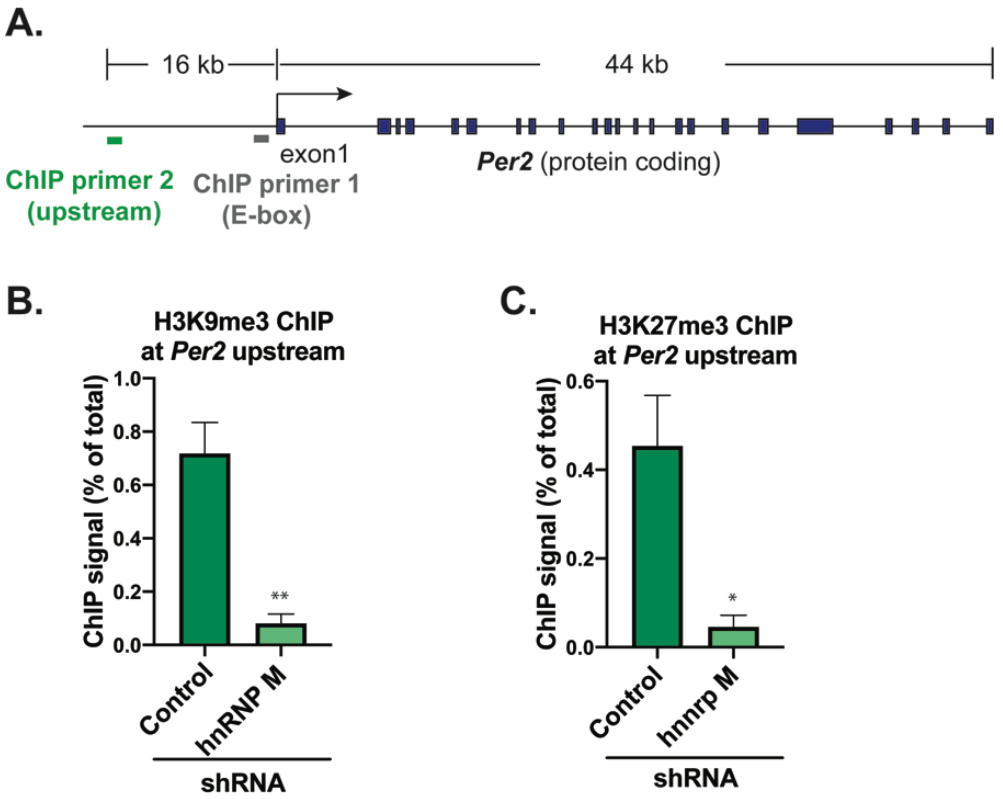
Supplemental Figure 4-1 The rhythm in *Per2* and *Per2AS* in synchronized NIH 3T3 cells.

Strand-specific real-time PCR was used to determine *Per2* mRNA (A) and *Per2AS* RNA (B) levels (y axis- copy number) from 0 to 24 hours (x-axis) (n=6) after dexamethasone (Dex) synchronization. (C and D) same as in A and B except levels were measured from 48 to 72 hours (x-axis) (n=5) after Dex synchronization. The statistical analysis of peak to trough was performed by one-way ANOVA with Tukey Post Hoc test showing ***$p < 0.0005$, *$p \leq 0.05$ and the NS represents $p > 0.05$.



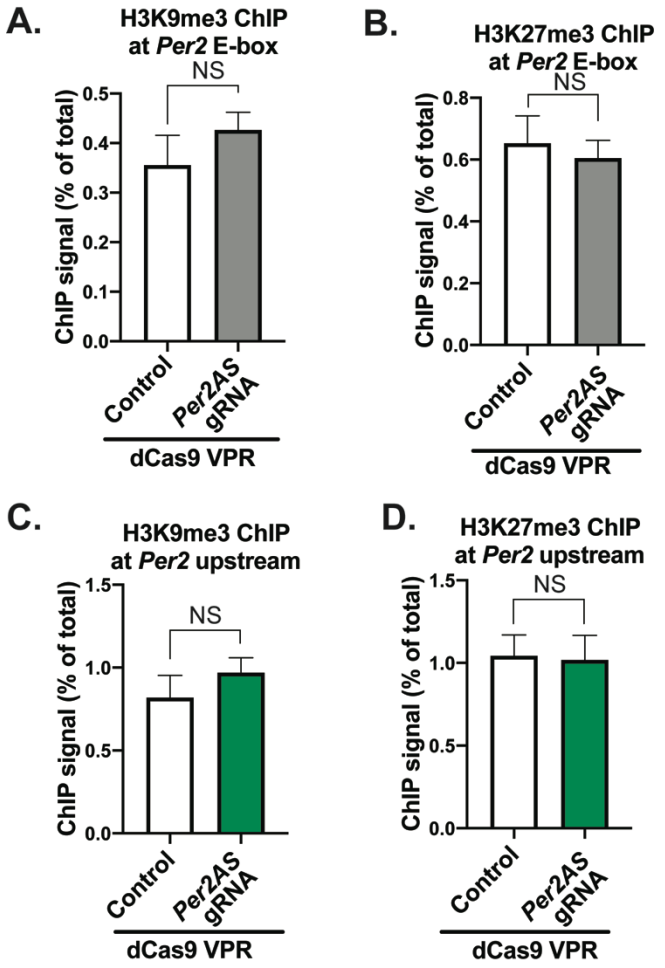
Supplemental Figure 4-2 hnRNP F and hnRNP L affect the amplitude of *Per2*.

Western blot showing the abundance of hnRNP F (A) and hnRNP L (B) in cultured unsynchronized mouse fibroblasts (untreated) and scramble shRNA (control) relative to three independent shRNA targeting hnRNP F or hnRNP L. ACTIN showed as loading control. Circadian oscillations of bioluminescence in synchronized *Per2::Luc* MEFs after delivery of scramble control shRNA (n=3) (C) or shRNA specific to hnRNP F-1 (n=4) (D) and shRNA-hnRNP L-1 (n=4) (E). The Circadian oscillations of bioluminescence (F) is the average of the independent data of (C)(D)(E). (RLU, relative light unit. n, numbers of independent experiments. x-axis, hours after synchronization).



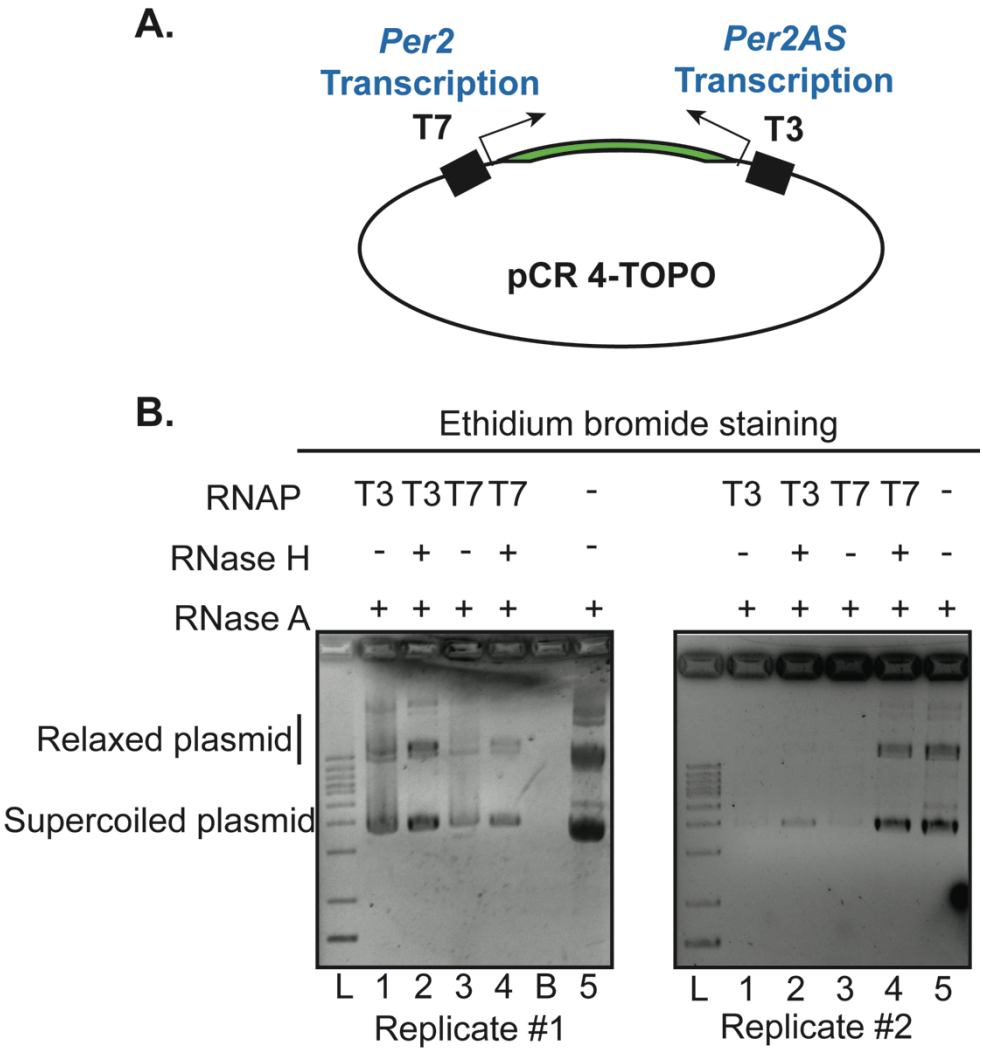
Supplemental Figure 4-3 Depletion of hnRNP M causes a reduction in H3K9me3 and H3K27me3 upstream of *Per2*.

(A) The ChIP oligos position relative to *Per2*. The ChIP-qPCR (mean +/-SEM of 4 independent experiments. t-test, two tailed. **p≤0.005, *p≤0.05) showed a reduction in H3K9me3 (B) and H3K27me3(C) upstream of *Per2* in unsynchronized cells upon knockdown of hnRNP M.



Supplemental Figure 4-4 Induction of *Per2AS* does not affect H3K9me3 and H3K27me3 at *Per2*.

The ChIP qPCR (mean +/-SEM, t-test, two tailed, NS represents $p>0.05$.) of 5 independent experiments showing H3K9me3 and H3K27me3 levels at *Per2 E*-box (A, B) and *Per2* upstream (C, D) with and without induction of *Per2AS.* The ChIP levels were normalized to percentage of total input (y-axis).



Supplemental Figure 4-5 *In vitro* detection of R-loop formation at *Per2*.

(A) Schematic of the 1kb *Per2AS* and *Per2* overlapping sequence cloned in pCR 4-TOPO vector driven by the T7 promoter (*Per2*) and T3 promoter (*Per2AS*). (B) Representative agarose gel stained with ethidium bromide. There was no obvious shift in supercoiled relative to relaxed DNA after RNase H treatment with either in *Per2AS* and Per2 induction.

Chapter 5

5 Chapter 5. Conclusions and Future Directions

The circadian clock governs 24-hour gene expression controlling cell physiological events, and if it is disrupted, it can have a negative influence on human health [315]; thus, it is critical to understand mechanisms of circadian clock regulation so appropriate interventions can be taken. At the molecular level, circadian clock gene expression and chromatin regulation is strongly connected, including the conserved mechanism of the circadian-regulated facultative heterochromatin (CRFH). CRFH involves rhythms in deacetylation, H3K9me3 and HP1 binding at the central clock gene in *Neurospora*, *Drosophila*, zebrafish and mammals. Recent findings by multiple groups have shown that the antisense transcript *qrf* is involved in heterochromatin formation at *frq* and the regulation is somewhat analogous to RNAi-mediated heterochromatin. PolII recruitment and chromatin remodeling occur along with and facilitates circadian-regulated transcription [79]. Negative feedback inhibition mediated by the negative elements shut down the transcription and create non-permissive chromatin. In this dissertation, I used a combination of techniques including genomic analyses (ChIP-seq and RNA-seq), molecular biology and biochemistry to further understand the crosstalk among the circadian clock, non-coding RNAs and heterochromatin regulation.

In chapter 2, I performed a genome-wide study to understand the effect and crosstalk between H3K4me3 and H3K9me3. Specifically, I performed RNA-seq and ChIP-seq on strains unable to establish histone H3 lysine 4 (Δ*kmt2/*Δ*set-1*) or histone H3 lysine 9 methylation (*Δkmt1/*Δ*dim-5*) compared to a WT, I determined nearly equal

numbers of genes had elevated and reduced expression in the dark and the light in the two mutants compare to WT, but the overall light response was not dramatically impaired. This indicates SET1 and DIM5 play critical roles in genome-wide regulation by assisting both transcriptional activation and repression, but it is not universal across light-activated genes. The ChIP-seq data revealed there are profound changes in H3K4me3 in *Δkmt1/*Δ*dim-5* and H3K9me3 in Δ*kmt2/*Δ*set-1* more so than the changes in H3K4me3 and H3K9me3 in response to light. The lack of major changes in H3K4me3 in response to light demonstrates that H3K4me3, a mark widely believed to accompany gene activation, was dispensable for the light response. I also found that H3K9me3-mediated facultative heterochromatin is dependent on the KMT1/SET-1, not only at the central clock gene *frequency* (*frq*), but also at a small subset of other loci. Future studies will still need to address the detailed mechanism of how H3K4me3 affect DNA methylation and H3K9me3 at *frq* (and at other loci), and how the two chromatin signals interplay with each other. I hypothesize this effect may be due to SET1 interactions with other chromatin regulators or the resulting histone marks crosstalk with each other through chromatin readers, writers and erasers. Therefore, it would be necessary to determine other SET1 interacting proteins or find out other regulatory components. In addition, my genome-wide study also found overlapping and co-dependency further highlighting the complicated regulatory roles for H3K4 methylation and H3K9 methylation and shows combination of modifications is dependent on specific locations. Future work still needs to address whether other modifications and regulators interfere or facilitate the crosstalk between H3K4me3 and H3K9me3.

The research topic of the chapter 3 was inspired by the discovery that circadian clock transcription factor BMAL1 associates with the telomere repeat DNA and, regulates the rhythm in *TERRA* and heterochromatin at telomere. To further understand how perturbing circadian regulation at telomeres affects telomere homeostasis, I chose to use the commonly abused drug, alcohol. Alcohol has been indicated in altering the circadian clock function [374]. In addition, chromic alcohol consumption leads to shortened telomere [294] and the molecular mechanism is not fully understood. My data indicates that alcohol consumption changed normal telomere homeostasis altering the diurnal rhythm in *TERRA* and heterochromatin. These findings have profound influence on our understanding of the serious health effects cause by drug abuse. In the future, it would be interesting to see whether alcohol consumption at different time of day causes a different effect to determine if there a "best" time-of-day to consume alcohol.

Because the shelterin complex is critical to maintain the telomere structures [334]more studies are needed to determine whether alcohol also perturbs the protein network at telomere. In addition, this work would benefit from a long-term study to mimic chronic drinking. Models for *TERRA* indicate it function in *cis* at telomere, but can also serve as epigenomic modulator in *trans* [302]. Therefore, it would be interesting to see whether alcohol alters the genome-wide *TERRA* pattern and to understand whether this has the potential to disease.

In chapter 4, I set out to understand the *Per2* NAT, *Per2AS*. I found the diurnal rhythm of *Per2AS* is dependent on BMAL1. Using the ChIRP-MS, I identified *Per2AS*-interacting proteins. I found hnRNP M interact with *Per2A*S and is required to maintain the normal amplitude and period of *Per2* expression. Furthermore, I demonstrate that

hnRNP M is involved in regulating H3K9me3 and H3K27me3 at *Per2*. Using artificial induction of *Per2AS* by dCas9-VPR, I found *Per2AS* appeared to enhance *Per2* expression. and I tested if this was due to R-loop formation or a reduction in heterochromatin marks. Unfortunately, neither result was conclusive, and the mechanism is still unknown. Regardless, my data support a model where *Per2AS* may serve as scaffold for hnRNP M and other associated proteins, which are needed for heterochromatin formation at *Per2*. Further work should focus on investigating whether hnRNPs are in PER complex. Also, it would be interesting to know the genome-wide crosstalk between hnRNPs, lncRNAs and heterochromatin formation.

In conclusion, this dissertation investigated the crosstalk of histone modifications on genome-wide gene expression and increased our understanding of the circadian clock, non-coding RNAs, and the heterochromatin formation. Using functionally characterized lncRNAs I was able to further our understanding of regulatory mechanisms and how altering the expression of lncRNA can alter heterochromatin formation; a least at the telomere. I also found *Per2AS* interacting proteins and determined hnRNP M is need for CRFH at *Per2*.

6 References:

1. Zhang R, Lahens NF, Ballance HI, Hughes ME, Hogenesch JB: **A circadian gene expression atlas in mammals: implications for biology and medicine**. *Proceedings of the National Academy of Sciences* 2014, **111**(45):16219-16224.
2. Dunlap JC, Loros JJ, Liu Y, Crosthwaite SK: **Eukaryotic circadian systems: cycles in common**. *Genes to Cells* 1999, **4**(1):01-10.
3. Takahashi JS: **Transcriptional architecture of the mammalian circadian clock**. *Nature Reviews Genetics* 2017, **18**(3):164.
4. Bass J: **Circadian topology of metabolism**. *Nature* 2012, **491**(7424):348.
5. Roenneberg T, Merrow M: **The network of time: understanding the molecular circadian system**. *Current Biology* 2003, **13**(5):R198-R207.
6. Partch CL, Green CB, Takahashi JS: **Molecular architecture of the mammalian circadian clock**. *Trends in cell biology* 2014, **24**(2):90-99.
7. Dunlap JC: **Molecular bases for circadian clocks**. *Cell* 1999, **96**(2):271-290.
8. Darlington TK, Wager-Smith K, Ceriani MF, Staknis D, Gekakis N, Steeves TD, Weitz CJ, Takahashi JS, Kay SA: **Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim**. *Science* 1998, **280**(5369):1599-1603.
9. Shearman LP, Sriram S, Weaver DR, Maywood ES, Chaves I, Zheng B, Kume K, Lee CC, Hastings MH, Reppert SM: **Interacting molecular loops in the mammalian circadian clock**. *Science* 2000, **288**(5468):1013-1019.
10. Konopka RJ, Benzer S: **Clock mutants of Drosophila melanogaster**. *Proceedings of the National Academy of Sciences* 1971, **68**(9):2112-2116.
11. Bargiello TA, Jackson FR, Young MW: **Restoration of circadian behavioural rhythms by gene transfer in Drosophila**. *Nature* 1984, **312**(5996):752.
12. Reddy P, Zehring WA, Wheeler DA, Pirrotta V, Hadfield C, Hall JC, Rosbash M: **Molecular analysis of the period locus in Drosophila melanogaster and identification of a transcript involved in biological rhythms**. *Cell* 1984, **38**(3):701-710.
13. Klein DC, Moore RY: **Suprachiasmatic nucleus: the mind's clock**: Oxford University Press, USA; 1991.
14. Herzog ED, Takahashi JS, Block GD: **Clock controls circadian period in isolated suprachiasmatic nucleus neurons**. *Nature neuroscience* 1998, **1**(8):708-713.
15. Mohawk JA, Green CB, Takahashi JS: **Central and peripheral circadian clocks in mammals**. *Annual review of neuroscience* 2012, **35**:445-462.
16. Antoch MP, Song EJ, Chang AM, Vitaterna MH, Zhao Y, Wilsbacher LD, Sangoram AM, King DP, Pinto LH, Takahashi JS: **Functional identification of the mouse circadian Clock gene by transgenic BAC rescue**. *Cell* 1997, **89**(4):655-667.
17. King DP, Zhao Y, Sangoram AM, Wilsbacher LD, Tanaka M, Antoch MP, Steeves TD, Vitaterna MH, Kornhauser JM, Lowrey PL *et al*: **Positional cloning of the mouse circadian clock gene**. *Cell* 1997, **89**(4):641-653.

18. Vitaterna MH, King DP, Chang AM, Kornhauser JM, Lowrey PL, McDonald JD, Dove WF, Pinto LH, Turek FW, Takahashi JS: **Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior**. *Science* 1994, **264**(5159):719-725.
19. Hogenesch JB, Gu YZ, Jain S, Bradfield CA: **The basic-helix-loop-helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors**. *Proc Natl Acad Sci U S A* 1998, **95**(10):5474-5479.
20. Allada R, White NE, So WV, Hall JC, Rosbash M: **A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless**. *Cell* 1998, **93**(5):791-804.
21. Rutila JE, Suri V, Le M, So WV, Rosbash M, Hall JC: **CYCLE is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of Drosophila period and timeless**. *Cell* 1998, **93**(5):805-814.
22. Crosthwaite SK, Dunlap JC, Loros JJ: ***Neurospora wc-1* and *wc-2*: transcription, photoresponses, and the origins of circadian rhythmicity**. *Science* 1997, **276**(5313):763-769.
23. Zheng B, Larkin DW, Albrecht U, Sun ZS, Sage M, Eichele G, Lee CC, Bradley A: **The mPer2 gene encodes a functional component of the mammalian circadian clock**. *Nature* 1999, **400**(6740):169-173.
24. Cermakian N, Monaco L, Pando MP, Dierich A, Sassone-Corsi P: **Altered behavioral rhythms and clock gene expression in mice with a targeted mutation in the Period1 gene**. *EMBO J* 2001, **20**(15):3967-3974.
25. Thresher RJ, Vitaterna MH, Miyamoto Y, Kazantsev A, Hsu DS, Petit C, Selby CP, Dawut L, Smithies O, Takahashi JS *et al*: **Role of mouse cryptochrome blue-light photoreceptor in circadian photoresponses**. *Science* 1998, **282**(5393):1490-1494.
26. van der Horst GT, Muijtjens M, Kobayashi K, Takano R, Kanno S, Takao M, de Wit J, Verkerk A, Eker AP, van Leenen D *et al*: **Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms**. *Nature* 1999, **398**(6728):627-630.
27. Sehgal A, Price JL, Man B, Young MW: **Loss of circadian behavioral rhythms and per RNA oscillations in the Drosophila mutant timeless**. *Science* 1994, **263**(5153):1603-1606.
28. Baylies MK, Bargiello TA, Jackson FR, Young MW: **Changes in abundance or structure of the per gene product can alter periodicity of the Drosophila clock**. *Nature* 1987, **326**(6111):390-392.
29. Konopka RJ, Benzer S: **Clock mutants of Drosophila melanogaster**. *Proc Natl Acad Sci U S A* 1971, **68**(9):2112-2116.
30. Aronson BD, Johnson KA, Loros JJ, Dunlap JC: **Negative feedback defining a circadian clock: autoregulation of the clock gene *frequency***. *Science* 1994, **263**(5153):1578-1584.
31. Sato TK, Panda S, Miraglia LJ, Reyes TM, Rudic RD, McNamara P, Naik KA, FitzGerald GA, Kay SA, Hogenesch JB: **A functional genomics strategy**

reveals Rora as a component of the mammalian circadian clock. *Neuron* 2004, **43**(4):527-537.
32. Preitner N, Damiola F, Lopez-Molina L, Zakany J, Duboule D, Albrecht U, Schibler U: **The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator**. *Cell* 2002, **110**(2):251-260.
33. Cheng P, He Q, Wang L, Liu Y: **Regulation of the *Neurospora* circadian clock by an RNA helicase**. *Genes Dev* 2005, **19**(2):234-241.
34. Duong HA, Robles MS, Knutti D, Weitz CJ: **A molecular mechanism for circadian clock negative feedback**. *Science* 2011, **332**(6036):1436-1439.
35. Brown SA, Ripperger J, Kadener S, Fleury-Olela F, Vilbois F, Rosbash M, Schibler U: **PERIOD1-associated proteins modulate the negative limb of the mammalian circadian oscillator**. *Science* 2005, **308**(5722):693-696.
36. Kwok RS, Li YH, Lei AJ, Edery I, Chiu JC: **The catalytic and non-catalytic functions of the Brahma chromatin-remodeling protein collaborate to fine-tune circadian transcription in Drosophila**. *PLoS genetics* 2015, **11**(7):e1005307.
37. Duong HA, Weitz CJ: **Temporal orchestration of repressive chromatin modifiers by circadian clock Period complexes**. *Nat Struct Mol Biol* 2014, **21**(2):126-132.
38. Wang B, Kettenbach AN, Gerber SA, Loros JJ, Dunlap JC: **Neurospora WC-1 recruits SWI/SNF to remodel frequency and initiate a circadian cycle**. *PLoS genetics* 2014, **10**(9):e1004599.
39. Kim JY, Kwak PB, Weitz CJ: **Specificity in circadian clock feedback from targeted reconstitution of the NuRD corepressor**. *Molecular cell* 2014, **56**(6):738-748.
40. King DP, Takahashi JS: **Molecular genetics of circadian rhythms in mammals**. *Annual review of neuroscience* 2000, **23**(1):713-742.
41. Panda S, Antoch MP, Miller BH, Su AI, Schook AB, Straume M, Schultz PG, Kay SA, Takahashi JS, Hogenesch JB: **Coordinated transcription of key pathways in the mouse by the circadian clock**. *Cell* 2002, **109**(3):307-320.
42. Griffin EA, Staknis D, Weitz CJ: **Light-independent role of CRY1 and CRY2 in the mammalian circadian clock**. *Science* 1999, **286**(5440):768-771.
43. Sangoram AM, Saez L, Antoch MP, Gekakis N, Staknis D, Whiteley A, Fruechte EM, Vitaterna MH, Shimomura K, King DP: **Mammalian circadian autoregulatory loop: a timeless ortholog and mPer1 interact and negatively regulate CLOCK-BMAL1-induced transcription**. *Neuron* 1998, **21**(5):1101-1113.
44. Sato TK, Yamada RG, Ukai H, Baggs JE, Miraglia LJ, Kobayashi TJ, Welsh DK, Kay SA, Ueda HR, Hogenesch JB: **Feedback repression is required for mammalian circadian clock function**. *Nature genetics* 2006, **38**(3):312.
45. Gekakis N, Saez L, Delahaye-Brown A-M, Myers MP, Sehgal A, Young MW, Weitz CJ: **Isolation of timeless by PER protein interaction: defective**

interaction between timeless protein and long-period mutant PERL. *Science* 1995, **270**(5237):811-815.
46. Zeng H, Qian Z, Myers MP, Rosbash M: **A light-entrainment mechanism for the Drosophila circadian clock**. *Nature* 1996, **380**(6570):129.
47. Denault DL, Loros JJ, Dunlap JC: **WC-2 mediates WC-1–FRQ interaction within the PAS protein-linked circadian feedback loop of Neurospora**. *The EMBO journal* 2001, **20**(1-2):109-117.
48. Knippschild U, Gocht A, Wolff S, Huber N, Löhler J, Stöter M: **The casein kinase 1 family: participation in multiple cellular processes in eukaryotes**. *Cellular signalling* 2005, **17**(6):675-689.
49. Yang Y, Cheng P, Zhi G, Liu Y: **Identification of a calcium/calmodulin-dependent protein kinase that phosphorylates the Neurospora circadian clock protein FREQUENCY**. *Journal of Biological Chemistry* 2001, **276**(44):41064-41072.
50. Görl M, Merrow M, Huttner B, Johnson J, Roenneberg T, Brunner M: **A PEST-like element in FREQUENCY determines the length of the circadian period in Neurospora crassa**. *The EMBO Journal* 2001, **20**(24):7074-7084.
51. Pregueiro AM, Liu Q, Baker CL, Dunlap JC, Loros JJ: **The Neurospora checkpoint kinase 2: a regulatory link between the circadian and cell cycles**. *Science* 2006, **313**(5787):644-649.
52. Schafmeier T, Haase A, Kaldi K, Scholz J, Fuchs M, Brunner M: **Transcriptional feedback of *Neurospora* circadian clock gene by phosphorylation-dependent inactivation of its transcription factor**. *Cell* 2005, **122**(2):235-246.
53. Kloss B, Rothenfluh A, Young MW, Saez L: **Phosphorylation of period is influenced by cycling physical associations of double-time, period, and timeless in the Drosophila clock**. *Neuron* 2001, **30**(3):699-706.
54. Vielhaber E, Eide E, Rivers A, Gao Z-H, Virshup DM: **Nuclear entry of the circadian regulator mPER1 is controlled by mammalian casein kinase I ε**. *Molecular and cellular biology* 2000, **20**(13):4888-4899.
55. Keesler GA, Camacho F, Guo Y, Virshup D, Mondadori C, Yao Z: **Phosphorylation and destabilization of human period I clock protein by human casein kinase Iε**. *Neuroreport* 2000, **11**(5):951-955.
56. Camacho F, Cilio M, Guo Y, Virshup D, Patel K, Khorkova O, Styren S, Morse B, Yao Z, Keesler G: **Human casein kinase Iδ phosphorylation of human circadian clock proteins period 1 and 2**. *FEBS letters* 2001, **489**(2-3):159-165.
57. Nakajima M, Imai K, Ito H, Nishiwaki T, Murayama Y, Iwasaki H, Oyama T, Kondo T: **Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro**. *Science* 2005, **308**(5720):414-415.
58. Querfurth C, Diernfellner AC, Gin E, Malzahn E, Höfer T, Brunner M: **Circadian conformational change of the Neurospora clock protein FREQUENCY triggered by clustered hyperphosphorylation of a basic domain**. *Molecular cell* 2011, **43**(5):713-722.

59. Robles MS, Humphrey SJ, Mann M: **Phosphorylation is a central mechanism for circadian control of metabolism and physiology**. *Cell metabolism* 2017, **25**(1):118-127.
60. Busino L, Bassermann F, Maiolica A, Lee C, Nolan PM, Godinho SI, Draetta GF, Pagano M: **SCFFbxl3 controls the oscillation of the circadian clock by directing the degradation of cryptochrome proteins**. *science* 2007, **316**(5826):900-904.
61. Hirano A, Yumimoto K, Tsunematsu R, Matsumoto M, Oyama M, Kozuka-Hata H, Nakagawa T, Lanjakornsiripan D, Nakayama KI, Fukada Y: **FBXL21 regulates oscillation of the circadian clock through ubiquitination and stabilization of cryptochromes**. *Cell* 2013, **152**(5):1106-1118.
62. Larrondo LF, Olivares-Yañez C, Baker CL, Loros JJ, Dunlap JC: **Decoupling circadian clock protein turnover from circadian period determination**. *Science* 2015, **347**(6221):1257277.
63. Yu W, Zheng H, Price JL, Hardin PE: **DOUBLETIME plays a noncatalytic role to mediate CLOCK phosphorylation and repress CLOCK-dependent transcription within the Drosophila circadian clock**. *Mol Cell Biol* 2009, **29**(6):1452-1458.
64. Kim EY, Ko HW, Yu W, Hardin PE, Edery I: **A DOUBLETIME kinase binding domain on the Drosophila PERIOD protein is essential for its hyperphosphorylation, transcriptional repression, and circadian clock function**. *Mol Cell Biol* 2007, **27**(13):5014-5028.
65. Yu W, Zheng H, Houl JH, Dauwalder B, Hardin PE: **PER-dependent rhythms in CLK phosphorylation and E-box binding regulate circadian transcription**. *Genes Dev* 2006, **20**(6):723-733.
66. He Q, Cha J, He Q, Lee HC, Yang Y, Liu Y: **CKI and CKII mediate the FREQUENCY-dependent phosphorylation of the WHITE COLLAR complex to close the *Neurospora* circadian negative feedback loop**. *Genes Dev* 2006, **20**(18):2552-2565.
67. He Q, Shu H, Cheng P, Chen S, Wang L, Liu Y: **Light-independent phosphorylation of WHITE COLLAR-1 regulates its function in the Neurospora circadian negative feedback loop**. *J Biol Chem* 2005, **280**(17):17526-17532.
68. Tamaru T, Hirayama J, Isojima Y, Nagai K, Norioka S, Takamatsu K, Sassone-Corsi P: **CK2alpha phosphorylates BMAL1 to regulate the mammalian clock**. *Nat Struct Mol Biol* 2009, **16**(4):446-448.
69. Kondratov RV, Chernov MV, Kondratova AA, Gorbacheva VY, Gudkov AV, Antoch MP: **BMAL1-dependent circadian oscillation of nuclear CLOCK: posttranslational events induced by dimerization of transcriptional activators of the mammalian clock system**. *Genes Dev* 2003, **17**(15):1921-1932.
70. Yoshitane H, Takao T, Satomi Y, Du NH, Okano T, Fukada Y: **Roles of CLOCK phosphorylation in suppression of E-box-dependent transcription**. *Mol Cell Biol* 2009, **29**(13):3675-3686.

71. Dardente H, Fortier EE, Martineau V, Cermakian N: **Cryptochromes impair phosphorylation of transcriptional activators in the clock: a general mechanism for circadian repression**. *Biochem J* 2007, **402**(3):525-536.
72. Menet JS, Abruzzi KC, Desrochers J, Rodriguez J, Rosbash M: **Dynamic PER repression mechanisms in the Drosophila circadian clock: from on-DNA to off-DNA**. *Genes Dev* 2010, **24**(4):358-367.
73. Chiou YY, Yang Y, Rashid N, Ye R, Selby CP, Sancar A: **Mammalian Period represses and de-represses transcription by displacing CLOCK-BMAL1 from promoters in a Cryptochrome-dependent manner**. *Proc Natl Acad Sci U S A* 2016, **113**(41):E6072-E6079.
74. Wang B, Kettenbach AN, Zhou X, Loros JJ, Dunlap JC: **The Phospho-Code Determining Circadian Feedback Loop Closure and Output in Neurospora**. *Molecular cell* 2019, **74**(4):771-784. e773.
75. Talora C, Franchi L, Linden H, Ballario P, Macino G: **Role of a white collar-1-white collar-2 complex in blue-light signal transduction**. *EMBO J* 1999, **18**(18):4961-4968.
76. Stratmann M, Suter DM, Molina N, Naef F, Schibler U: **Circadian Dbp Transcription Relies on Highly Dynamic BMAL1-CLOCK Interaction with E Boxes and Requires the Proteasome**. *Mol Cell* 2012, **48**(2):277-287.
77. Schafmeier T, Diernfellner A, Schäfer A, Dintsis O, Neiss A, Brunner M: **Circadian activity and abundance rhythms of the Neurospora clock transcription factor WCC associated with rapid nucleo–cytoplasmic shuttling**. *Genes & development* 2008, **22**(24):3397-3402.
78. Kramer C, Loros JJ, Dunlap JC, Crosthwaite SK: **Role for antisense RNA in regulating circadian clock function in Neurospora crassa**. *Nature* 2003, **421**(6926):948-952.
79. Koike N, Yoo S-H, Huang H-C, Kumar V, Lee C, Kim T-K, Takahashi JS: **Transcriptional architecture and chromatin landscape of the core circadian clock in mammals**. *Science* 2012, **338**(6105):349-354.
80. Vollmers C, Schmitz RJ, Nathanson J, Yeo G, Ecker JR, Panda S: **Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome**. *Cell metabolism* 2012, **16**(6):833-845.
81. Rey G, Cesbron F, Rougemont J, Reinke H, Brunner M, Naef F: **Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver**. *PLoS Biol* 2011, **9**(2):e1000595.
82. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ: **Crystal structure of the nucleosome core particle at 2.8 Å resolution**. *Nature* 1997, **389**(6648):251-260.
83. Kornberg RD: **Chromatin structure: a repeating unit of histones and DNA**. *Science* 1974, **184**(4139):868-871.
84. Durrin LK, Mann RK, Grunstein M: **Nucleosome loss activates CUP1 and HIS3 promoters to fully induced levels in the yeast Saccharomyces cerevisiae**. *Molecular and Cellular Biology* 1992, **12**(4):1621-1629.

85. Workman J, Kingston R: **Alteration of nucleosome structure as a mechanism of transcriptional regulation**. In.: Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA; 1998.
86. Bannister AJ, Kouzarides T: **Regulation of chromatin by histone modifications**. *Cell research* 2011, **21**(3):381-395.
87. Smith E, Shilatifard A: **The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes**. *Molecular cell* 2010, **40**(5):689-701.
88. Li G, Reinberg D: **Chromatin higher-order structures and gene regulation**. *Current opinion in genetics & development* 2011, **21**(2):175-186.
89. Fischle W, Wang Y, Allis CD: **Histone and chromatin cross-talk**. *Current opinion in cell biology* 2003, **15**(2):172-183.
90. Narlikar GJ, Fan H-Y, Kingston RE: **Cooperation between complexes that regulate chromatin structure and transcription**. *Cell* 2002, **108**(4):475-487.
91. Ivana L, Ohkawa Y, Imbalzano AN: **Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers**. *Nature Reviews Genetics* 2006, **7**(6):461.
92. Ripperger JA, Schibler U: **Rhythmic CLOCK-BMAL1 binding to multiple E-box motifs drives circadian Dbp transcription and chromatin transitions**. *Nature genetics* 2006, **38**(3):369.
93. Becker JS, Nicetto D, Zaret KS: **H3K9me3-dependent heterochromatin: barrier to cell fate changes**. *Trends in Genetics* 2016, **32**(1):29-41.
94. Trojer P, Reinberg D: **Facultative heterochromatin: is there a distinctive molecular signature?** *Molecular cell* 2007, **28**(1):1-13.
95. Taylor P, Hardin PE: **Rhythmic E-box binding by CLK-CYC controls daily cycles in per and tim transcription and chromatin modifications**. *Mol Cell Biol* 2008, **28**(14):4642-4652.
96. Belden WJ, Lewis ZA, Selker EU, Loros JJ, Dunlap JC: **CHD1 remodels chromatin and influences transient DNA methylation at the clock gene frequency**. *PLoS genetics* 2011, **7**(7):e1002166.
97. Tamaru H, Selker EU: **A histone H3 methyltransferase controls DNA methylation in Neurospora crassa**. *Nature* 2001, **414**(6861):277.
98. Padmanabhan K, Robles MS, Westerling T, Weitz CJ: **Feedback regulation of transcriptional termination by the mammalian circadian clock PERIOD complex**. *Science* 2012, **337**(6094):599-602.
99. Tschiersch B, Hofmann A, Krauss V, Dorn R, Korge G, Reuter G: **The protein encoded by the Drosophila position-effect variegation suppressor gene Su (var) 3-9 combines domains of antagonistic regulators of homeotic gene complexes**. *The EMBO journal* 1994, **13**(16):3822-3831.
100. Saunders WS, Chue C, Goebl M, Craig C, Clark RF, Powers JA, Eissenberg JC, Elgin S, Rothfield NF, Earnshaw WC: **Molecular cloning of a human homologue of Drosophila heterochromatin protein HP1 using anti-**

centromere autoantibodies with anti-chromo specificity. *Journal of Cell Science* 1993, **104**(2):573-582.
101. Jacobs SA, Taverna SD, Zhang Y, Briggs SD, Li J, Eissenberg JC, Allis CD, Khorasanizadeh S: **Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3**. *The EMBO journal* 2001, **20**(18):5232-5241.
102. Jacobs SA, Khorasanizadeh S: **Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail**. *Science* 2002, **295**(5562):2080-2083.
103. Aagaard L, Laible G, Selenko P, Schmid M, Dorn R, Schotta G, Kuhfittig S, Wolf A, Lebersorger A, Singh PB: **Functional mammalian homologues of the Drosophila PEV-modifier Su (var) 3-9 encode centromere-associated proteins which complex with the heterochromatin component M31**. *The EMBO journal* 1999, **18**(7):1923-1938.
104. Ivanova AV, Bonaduce MJ, Ivanov SV, Klar AJ: **The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast**. *Nature genetics* 1998, **19**(2):192.
105. Ruesch CE, Ramakrishnan M, Park J, Li N, Chong HS, Zaman R, Joska TM, Belden WJ: **The histone H3 lysine 9 methyltransferase DIM-5 modifies chromatin at frequency and represses light-activated gene expression**. *G3: Genes, Genomes, Genetics* 2015, **5**(1):93-101.
106. Belden WJ, Larrondo LF, Froehlich AC, Shi M, Chen C-H, Loros JJ, Dunlap JC: **The band mutation in Neurospora crassa is a dominant allele of ras-1 implicating RAS signaling in circadian output**. *Genes & development* 2007, **21**(12):1494-1505.
107. Park J, Belden WJ: **Long non-coding RNAs have age-dependent diurnal expression that coincides with age-related changes in genome-wide facultative heterochromatin**. *BMC genomics* 2018, **19**(1):777.
108. Montgomery MK, Xu S, Fire A: **RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans**. *Proceedings of the national academy of sciences* 1998, **95**(26):15502-15507.
109. Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A: **On the role of RNA amplification in dsRNA-triggered gene silencing**. *Cell* 2001, **107**(4):465-476.
110. Wassenegger M: **The role of the RNAi machinery in heterochromatin formation**. *Cell* 2005, **122**(1):13-16.
111. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA: **Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi**. *science* 2002, **297**(5588):1833-1837.
112. Dang Y, Li L, Guo W, Xue Z, Liu Y: **Convergent transcription induces dynamic DNA methylation at disiRNA loci**. *PLoS genetics* 2013, **9**(9):e1003761.
113. Lee H-C, Li L, Gu W, Xue Z, Crosthwaite SK, Pertsemlidis A, Lewis ZA, Freitag M, Selker EU, Mello CC: **Diverse pathways generate microRNA-like RNAs**

and Dicer-independent small interfering RNAs in fungi. *Molecular cell* 2010, **38**(6):803-814.
114. Li N, Joska TM, Ruesch CE, Coster SJ, Belden WJ: **The frequency natural antisense transcript first promotes, then represses, frequency gene expression via facultative heterochromatin**. *Proceedings of the National Academy of Sciences* 2015, **112**(14):4357-4362.
115. Dang Y, Cheng J, Sun X, Zhou Z, Liu Y: **Antisense transcription licenses nascent transcripts to mediate transcriptional gene silencing**. *Genes & development* 2016, **30**(21):2417-2432.
116. Etchegaray J-P, Lee C, Wade PA, Reppert SM: **Rhythmic histone acetylation underlies transcription in the mammalian circadian clock**. *Nature* 2003, **421**(6919):177.
117. Curtis AM, Seo S-b, Westgate EJ, Rudic RD, Smyth EM, Chakravarti D, FitzGerald GA, McNamara P: **Histone acetyltransferase-dependent chromatin remodeling and the vascular clock**. *Journal of Biological Chemistry* 2004, **279**(8):7091-7097.
118. Katada S, Sassone-Corsi P: **The histone methyltransferase MLL1 permits the oscillation of circadian gene expression**. *Nature structural & molecular biology* 2010, **17**(12):1414.
119. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NT, Schreiber SL, Mellor J, Kouzarides T: **Active genes are tri-methylated at K4 of histone H3**. *Nature* 2002, **419**(6905):407.
120. Le Martelot G, Canella D, Symul L, Migliavacca E, Gilardi F, Liechti R, Martin O, Harshman K, Delorenzi M, Desvergne B *et al*: **Genome-wide RNA polymerase II profiles and RNA accumulation reveal kinetics of transcription and associated epigenetic changes during diurnal cycles**. *PLoS Biol* 2012, **10**(11):e1001442.
121. Raduwan H, Isola AL, Belden WJ: **Methylation of histone H3 on lysine 4 by the lysine methyltransferase SET1 protein is needed for normal clock gene expression**. *Journal of Biological Chemistry* 2013, **288**(12):8380-8390.
122. Zhu Q, Ramakrishnan M, Park J, Belden WJ: **Histone H3 lysine 4 methyltransferase is required for facultative heterochromatin at specific loci**. *BMC genomics* 2019, **20**(1):350.
123. Etchegaray J-P, Yang X, DeBruyne JP, Peters AH, Weaver DR, Jenuwein T, Reppert SM: **The polycomb group protein EZH2 is required for mammalian circadian clock function**. *Journal of Biological Chemistry* 2006, **281**(30):21209-21215.
124. Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y: **Histone demethylation mediated by the nuclear amine oxidase homolog LSD1**. *Cell* 2004, **119**(7):941-953.
125. Nam HJ, Boo K, Kim D, Han D-H, Choe HK, Kim CR, Sun W, Kim H, Kim K, Lee H: **Phosphorylation of LSD1 by PKCα is crucial for circadian rhythmicity and phase resetting**. *Molecular cell* 2014, **53**(5):791-805.

126. Mulligan P, Yang F, Di Stefano L, Ji J-Y, Ouyang J, Nishikawa JL, Toiber D, Kulkarni M, Wang Q, Najafi-Shoushtari SH: **A SIRT1-LSD1 corepressor complex regulates Notch target gene expression and development**. *Molecular cell* 2011, **42**(5):689-699.
127. DiTacchio L, Le HD, Vollmers C, Hatori M, Witcher M, Secombe J, Panda S: **Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock**. *Science* 2011, **333**(6051):1881-1885.
128. Reischl S, Kramer A: **Fbxl11 Is a Novel Negative Element of the Mammalian Circadian Clock**. *J Biol Rhythms* 2015, **30**(4):291-301.
129. Menet JS, Pescatore S, Rosbash M: **CLOCK: BMAL1 is a pioneer-like transcription factor**. *Genes & development* 2014, **28**(1):8-13.
130. Sancar C, Ha N, Yilmaz R, Tesorero R, Fisher T, Brunner M, Sancar G: **Combinatorial control of light induced chromatin remodeling and gene activation in Neurospora**. *PLoS Genet* 2015, **11**(3):e1005105.
131. Hurley JM, Dasgupta A, Emerson JM, Zhou X, Ringelberg CS, Knabe N, Lipzen AM, Lindquist EA, Daum CG, Barry KW *et al*: **Analysis of clock-regulated genes in Neurospora reveals widespread posttranscriptional control of metabolic potential**. *Proc Natl Acad Sci U S A* 2014, **111**(48):16995-17002.
132. Liu X, Dang Y, Matsu-ura T, He Y, He Q, Hong CI, Liu Y: **DNA replication is required for circadian clock function by regulating rhythmic nucleosome composition**. *Molecular cell* 2017, **67**(2):203-213. e204.
133. Koike N, Yoo SH, Huang HC, Kumar V, Lee C, Kim TK, Takahashi JS: **Transcriptional architecture and chromatin landscape of the core circadian clock in mammals**. *Science* 2012, **338**(6105):349-354.
134. Zhou Z, Liu X, Hu Q, Zhang N, Sun G, Cha J, Wang Y, Liu Y, He Q: **Suppression of WC-independent frequency transcription by RCO-1 is essential for Neurospora circadian clock**. *Proc Natl Acad Sci U S A* 2013, **110**(50):E4867-4874.
135. Strahl BD, Grant PA, Briggs SD, Sun ZW, Bone JR, Caldwell JA, Mollah S, Cook RG, Shabanowitz J, Hunt DF *et al*: **Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression**. *Mol Cell Biol* 2002, **22**(5):1298-1306.
136. Li B, Jackson J, Simon MD, Fleharty B, Gogol M, Seidel C, Workman JL, Shilatifard A: **Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription**. *J Biol Chem* 2009, **284**(12):7970-7976.
137. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP *et al*: **Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription**. *Cell* 2005, **123**(4):581-592.
138. Joshi AA, Struhl K: **Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation**. *Mol Cell* 2005, **20**(6):971-978.

139. Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD: **A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation**. *Mol Cell Biol* 2005, **25**(8):3305-3316.
140. Li J, Moazed D, Gygi SP: **Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation**. *J Biol Chem* 2002, **277**(51):49383-49388.
141. Naruse Y, Oh-hashi K, Iijima N, Naruse M, Yoshioka H, Tanaka M: **Circadian and light-induced transcription of clock gene Per1 depends on histone acetylation and deacetylation**. *Molecular and cellular biology* 2004, **24**(14):6278-6287.
142. Kuo MH, Allis CD: **Roles of histone acetyltransferases and deacetylases in gene regulation**. *Bioessays* 1998, **20**(8):615-626.
143. Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondson DG, Roth SY, Allis CD: **Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation**. *Cell* 1996, **84**(6):843-851.
144. Hebbes TR, Thorne AW, Crane-Robinson C: **A direct link between core histone acetylation and transcriptionally active chromatin**. *The EMBO journal* 1988, **7**(5):1395-1402.
145. Grewal SI, Jia S: **Heterochromatin revisited**. *Nature Reviews Genetics* 2007, **8**(1):35.
146. Doi M, Hirayama J, Sassone-Corsi P: **Circadian regulator CLOCK is a histone acetyltransferase**. *Cell* 2006, **125**(3):497-508.
147. Sun Z, Feng D, Everett LJ, Bugge A, Lazar MA: **Circadian Epigenomic Remodeling and Hepatic Lipogenesis: Lessons from HDAC3**. *Cold Spring Harb Symp Quant Biol* 2011.
148. Nakahata Y, Kaluzova M, Grimaldi B, Sahar S, Hirayama J, Chen D, Guarente LP, Sassone-Corsi P: **The NAD+-dependent deacetylase SIRT1 modulates CLOCK-mediated chromatin remodeling and circadian control**. *Cell* 2008, **134**(2):329-340.
149. Asher G, Gatfield D, Stratmann M, Reinke H, Dibner C, Kreppel F, Mostoslavsky R, Alt FW, Schibler U: **SIRT1 regulates circadian clock gene expression through PER2 deacetylation**. *Cell* 2008, **134**(2):317-328.
150. Nakahata Y, Sahar S, Astarita G, Kaluzova M, Sassone-Corsi P: **Circadian control of the NAD+ salvage pathway by CLOCK-SIRT1**. *Science* 2009, **324**(5927):654-657.
151. Sassone-Corsi P: **Minireview: NAD+, a circadian metabolite with an epigenetic twist**. *Endocrinology* 2012, **153**(1):1-5.
152. Rodgers JT, Lerin C, Haas W, Gygi SP, Spiegelman BM, Puigserver P: **Nutrient control of glucose homeostasis through a complex of PGC-1α and SIRT1**. *Nature* 2005, **434**(7029):113.
153. Bordone L, Motta MC, Picard F, Robinson A, Jhala US, Apfeld J, McDonagh T, Lemieux M, McBurney M, Szilvasi A: **Sirt1 regulates insulin secretion by repressing UCP2 in pancreatic β cells**. *PLoS biology* 2005, **4**(2):e31.

154. Araki T, Sasaki Y, Milbrandt J: **Increased nuclear NAD biosynthesis and SIRT1 activation prevent axonal degeneration**. *Science* 2004, **305**(5686):1010-1013.
155. Masri S, Rigor P, Cervantes M, Ceglia N, Sebastian C, Xiao C, Roqueta-Rivera M, Deng C, Osborne TF, Mostoslavsky R: **Partitioning circadian transcription by SIRT6 leads to segregated control of cellular metabolism**. *Cell* 2014, **158**(3):659-672.
156. Peek CB, Affinati AH, Ramsey KM, Kuo H-Y, Yu W, Sena LA, Ilkayeva O, Marcheva B, Kobayashi Y, Omura C: **Circadian clock NAD+ cycle drives mitochondrial oxidative metabolism in mice**. *Science* 2013, **342**(6158):1243417.
157. Grimaldi B, Coiro P, Filetici P, Berge E, Dobosy JR, Freitag M, Selker EU, Ballario P: **The Neurospora crassa White Collar-1 dependent blue light response requires acetylation of histone H3 lysine 14 by NGF-1**. *Molecular biology of the cell* 2006, **17**(10):4576-4583.
158. Tao W, Chen S, Shi G, Guo J, Xu Y, Liu C: **SWItch/sucrose nonfermentable (SWI/SNF) complex subunit BAF60a integrates hepatic circadian clock and energy metabolism**. *Hepatology* 2011, **54**(4):1410-1420.
159. Yudkovsky N, Logie C, Hahn S, Peterson CL: **Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators**. *Genes & development* 1999, **13**(18):2369-2374.
160. Denslow S, Wade P: **The human Mi-2/NuRD complex and gene regulation**. *Oncogene* 2007, **26**(37):5433.
161. Xue Y, Wong J, Moreno GT, Young MK, Côté J, Wang W: **NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities**. *Molecular cell* 1998, **2**(6):851-861.
162. Azzi A, Dallmann R, Casserly A, Rehrauer H, Patrignani A, Maier B, Kramer A, Brown SA: **Circadian behavior is light-reprogrammed by plastic DNA methylation**. *Nature neuroscience* 2014, **17**(3):377.
163. Belden WJ, Loros JJ, Dunlap JC: **Execution of the circadian negative feedback loop in Neurospora requires the ATP-dependent chromatin-remodeling enzyme CLOCKSWITCH**. *Molecular cell* 2007, **25**(4):587-600.
164. Cha J, Zhou M, Liu Y: **CATP is a critical component of the Neurospora circadian clock by regulating the nucleosome occupancy rhythm at the frequency locus**. *EMBO reports* 2013, **14**(10):923-930.
165. Gai K, Cao X, Dong Q, Ding Z, Wei Y, Liu Y, Liu X, He Q: **Transcriptional repression of frequency by the IEC-1-INO80 complex is required for normal Neurospora circadian clock function**. *PLoS genetics* 2017, **13**(4):e1006732.
166. Dubruille R, Murad A, Rosbash M, Emery P: **A constant light-genetic screen identifies KISMET as a regulator of circadian photoresponses**. *PLoS genetics* 2009, **5**(12):e1000787.
167. Cavalli G, Misteli T: **Functional implications of genome topology**. *Nature structural & molecular biology* 2013, **20**(3):290.

168. Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *science* 2009, **326**(5950):289-293.
169. Aguilar-Arnal L, Hakim O, Patel VR, Baldi P, Hager GL, Sassone-Corsi P: **Cycles in spatial and temporal chromosomal organization driven by the circadian clock**. *Nature structural & molecular biology* 2013, **20**(10):1206.
170. Zhao H, Sifakis EG, Sumida N, Millán-Ariño L, Scholz BA, Svensson JP, Chen X, Ronnegren AL, de Lima CDM, Varnoosfaderani FS: **PARP1-and CTCF-mediated interactions between active and repressed chromatin at the lamina promote oscillating transcription**. *Molecular cell* 2015, **59**(6):984-997.
171. Xu Y, Guo W, Li P, Zhang Y, Zhao M, Fan Z, Zhao Z, Yan J: **Long-range chromosome interactions mediated by cohesin shape circadian gene expression**. *PLoS genetics* 2016, **12**(5):e1005992.
172. Kim YH, Marhon SA, Zhang Y, Steger DJ, Won K-J, Lazar MA: **Rev-erbα dynamically modulates chromatin looping to control circadian gene transcription**. *Science* 2018, **359**(6381):1274-1277.
173. Mermet J, Yeung J, Hurni C, Mauvoisin D, Gustafson K, Jouffe C, Nicolas D, Emmenegger Y, Gobet C, Franken P: **Clock-dependent chromatin topology modulates circadian transcription and behavior**. *Genes & development* 2018, **32**(5-6):347-358.
174. Moazed D: **Small RNAs in transcriptional gene silencing and genome defence**. *Nature* 2009, **457**(7228):413-420.
175. Cech TR, Steitz JA: **The noncoding RNA revolution—trashing old rules to forge new ones**. *Cell* 2014, **157**(1):77-94.
176. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap C, Suzuki M, Kawai J: **Antisense transcription in the mammalian transcriptome**. *Science* 2005, **309**(5740):1564-1566.
177. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F: **Landscape of transcription in human cells**. *Nature* 2012, **489**(7414):101-108.
178. Mercer TR, Mattick JS: **Structure and function of long noncoding RNAs in epigenetic regulation**. *Nature structural & molecular biology* 2013, **20**(3):300-307.
179. Mattick JS: **The genetic signatures of noncoding RNAs**. *PLoS Genet* 2009, **5**(4):e1000459.
180. Mattick JS, Makunin IV: **Non-coding RNA**. *Human molecular genetics* 2006, **15**(suppl_1):R17-R29.
181. Deveson IW, Hardwick SA, Mercer TR, Mattick JS: **The dimensions, dynamics, and relevance of the mammalian noncoding transcriptome**. *Trends in Genetics* 2017, **33**(7):464-478.
182. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S: **The product of the mouse Xist gene is a 15 kb inactive X-**

specific transcript containing no conserved ORF and located in the nucleus. *Cell* 1992, **71**(3):515-526.
183. Brown CJ, Hendrich BD, Rupert JL, Lafreniere RG, Xing Y, Lawrence J, Willard HF: **The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus**. *Cell* 1992, **71**(3):527-542.
184. Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES: **The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome**. *Science* 2013, **341**(6147):1237973.
185. Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B: **Xist RNA and the mechanism of X chromosome inactivation**. *Annual review of genetics* 2002, **36**(1):233-278.
186. Nakao M, Sasaki H: **Genomic imprinting: significance in development and diseases and the molecular mechanisms**. *The Journal of Biochemistry* 1996, **120**(3):467-473.
187. Bartolomei MS, Zemel S, Tilghman SM: **Parental imprinting of the mouse H19 gene**. *Nature* 1991, **351**(6322):153.
188. Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J: **iCLIP-transcriptome-wide mapping of protein-RNA interactions with individual nucleotide resolution**. *JoVE (Journal of Visualized Experiments)* 2011(50):e2638-e2638.
189. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E: **Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs**. *Cell* 2007, **129**(7):1311-1323.
190. Chu C, Zhang QC, Da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, Chang HY: **Systematic discovery of Xist RNA binding proteins**. *Cell* 2015, **161**(2):404-416.
191. Simon MD: **Capture hybridization analysis of RNA targets (CHART)**. *Current protocols in molecular biology* 2013:21.25. 21-21.25. 16.
192. McHugh CA, Chen C-K, Chow A, Surka CF, Tran C, McDonel P, Pandya-Jones A, Blanco M, Burghard C, Moradian A: **The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3**. *Nature* 2015, **521**(7551):232-236.
193. Zhao J, Sun BK, Erwin JA, Song J-J, Lee JT: **Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome**. *Science* 2008, **322**(5902):750-756.
194. Boque-Sastre R, Soler M, Oliveira-Mateos C, Portela A, Moutinho C, Sayols S, Villanueva A, Esteller M, Guil S: **Head-to-head antisense transcription and R-loop formation promotes transcriptional activation**. *Proceedings of the National Academy of Sciences* 2015, **112**(18):5785-5790.

195. Xue Z, Ye Q, Anson SR, Yang J, Xiao G, Kowbel D, Glass NL, Crosthwaite SK, Liu Y: **Transcriptional interference by antisense RNA is required for circadian clock function**. *Nature* 2014, **514**(7524):650.
196. Fang B, Everett LJ, Jager J, Briggs E, Armour SM, Feng D, Roy A, Gerhart-Hines Z, Sun Z, Lazar MA: **Circadian enhancers coordinate multiple phases of rhythmic gene transcription in vivo**. *Cell* 2014, **159**(5):1140-1152.
197. Fan Z, Zhao M, Joshi PD, Li P, Zhang Y, Guo W, Xu Y, Wang H, Zhao Z, Yan J: **A class of circadian long non-coding RNAs mark enhancers modulating long-range circadian gene regulation**. *Nucleic acids research* 2017, **45**(10):5720-5738.
198. Fu L, Lee CC: **The circadian clock: pacemaker and tumour suppressor**. *Nature Reviews Cancer* 2003, **3**(5):350-361.
199. Willis GL: **Parkinson's disease as a neuroendocrine disorder of circadian function: dopamine-melatonin imbalance and the visual system in the genesis and progression of the degenerative process**. *Reviews in the Neurosciences* 2008, **19**(4):245.
200. Sahar S, Sassone-Corsi P: **Metabolism and cancer: the circadian clock connection**. *Nature Reviews Cancer* 2009, **9**(12):886.
201. Ptitsyn AA, Zvonic S, Conrad SA, Scott LK, Mynatt RL, Gimble JM: **Circadian clocks are resounding in peripheral tissues**. *PLoS computational biology* 2006, **2**(3):e16.
202. Wijnen H, Young MW: **Interplay of circadian clocks and metabolic rhythms**. *Annu Rev Genet* 2006, **40**:409-448.
203. Eckel-Mahan K, Sassone-Corsi P: **Metabolism control by the circadian clock and vice versa**. *Nature structural & molecular biology* 2009, **16**(5):462-467.
204. Manzella N, Bracci M, Strafella E, Staffolani S, Ciarapica V, Copertaro A, Rapisarda V, Ledda C, Amati M, Valentino M: **Circadian modulation of 8-oxoguanine DNA damage repair**. *Scientific reports* 2015, **5**:13752.
205. Kondratov RV, Kondratova AA, Gorbacheva VY, Vykhovanets OV, Antoch MP: **Early aging and age-related pathologies in mice deficient in BMAL1, the core componentof the circadian clock**. *Genes & development* 2006, **20**(14):1868-1873.
206. Kondratov RV, Vykhovanets O, Kondratova AA, Antoch MP: **Antioxidant N-acetyl-L-cysteine ameliorates symptoms of premature aging associated with the deficiency of the circadian protein BMAL1**. *Aging (Albany NY)* 2009, **1**(12):979.
207. Pekovic-Vaughan V, Gibbs J, Yoshitane H, Yang N, Pathiranage D, Guo B, Sagami A, Taguchi K, Bechtold D, Loudon A: **The circadian clock regulates rhythmic activation of the NRF2/glutathione-mediated antioxidant defense pathway to modulate pulmonary fibrosis**. *Genes & development* 2014, **28**(6):548-560.
208. Early JO, Menon D, Wyse CA, Cervantes-Silva MP, Zaslona Z, Carroll RG, Palsson-McDermott EM, Angiari S, Ryan DG, Corcoran SE: **Circadian clock**

protein BMAL1 regulates IL-1β in macrophages via NRF2. *Proceedings of the National Academy of Sciences* 2018, **115**(36):E8460-E8468.
209. Zafarullah M, Li W, Sylvester J, Ahmad M: **Molecular mechanisms of N-acetylcysteine actions**. *Cellular and Molecular Life Sciences CMLS* 2003, **60**(1):6-20.
210. Greider CW, Blackburn EH: **Identification of a specific telomere terminal transferase activity in Tetrahymena extracts**. *cell* 1985, **43**(2):405-413.
211. Chan SR, Blackburn EH: **Telomeres and telomerase**. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 2004, **359**(1441):109-122.
212. García-Cao M, O'Sullivan R, Peters AH, Jenuwein T, Blasco MA: **Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases**. *Nature genetics* 2004, **36**(1):94.
213. Benetti R, Gonzalo S, Jaco I, Schotta G, Klatt P, Jenuwein T, Blasco MA: **Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination**. *The Journal of cell biology* 2007, **178**(6):925-936.
214. Hemann MT, Strong MA, Hao L-Y, Greider CW: **The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability**. *Cell* 2001, **107**(1):67-77.
215. Greider CW, Blackburn EH: **The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity**. *Cell* 1987, **51**(6):887-898.
216. Blasco MA: **Telomeres and human disease: ageing, cancer and beyond**. *Nature Reviews Genetics* 2005, **6**(8):611.
217. Chen W-D, Wen M-S, Shie S-S, Lo Y-L, Wo H-T, Wang C-C, Hsieh I-C, Lee T-H, Wang C-Y: **The circadian rhythm controls telomeres and telomerase activity**. *Biochemical and biophysical research communications* 2014, **451**(3):408-414.
218. Park J, Zhu Q, Mirek E, Na L, Raduwan H, Anthony TG, Belden WJ: **BMAL1 associates with chromosome ends to control rhythms in TERRA and telomeric heterochromatin**. *PLOS ONE* 2019, **14**(10):e0223803.
219. Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E, Lingner J: **Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends**. *Science* 2007, **318**(5851):798-801.
220. Deng Z, Norseen J, Wiedmer A, Riethman H, Lieberman PM: **TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres**. *Molecular cell* 2009, **35**(4):403-413.
221. Montero JJ, López-Silanes I, Megías D, Fraga MF, Castells-García Á, Blasco MA: **TERRA recruitment of polycomb to telomeres is essential for histone trymethylation marks at telomeric heterochromatin**. *Nature communications* 2018, **9**(1):1548.
222. Palacios JA, Herranz D, De Bonis ML, Velasco S, Serrano M, Blasco MA: **SIRT1 contributes to telomere maintenance and augments global homologous recombination**. *The Journal of cell biology* 2010, **191**(7):1299-1313.

223. Vaquero A, Scher M, Lee D, Erdjument-Bromage H, Tempst P, Reinberg D: **Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin**. *Molecular cell* 2004, **16**(1):93-105.
224. Chang H-C, Guarente L: **SIRT1 mediates central circadian control in the SCN by a mechanism that decays with aging**. *Cell* 2013, **153**(7):1448-1460.
225. Sato S, Solanas G, Peixoto FO, Bee L, Symeonidi A, Schmidt MS, Brenner C, Masri S, Benitah SA, Sassone-Corsi P: **Circadian reprogramming in the liver identifies metabolic pathways of aging**. *Cell* 2017, **170**(4):664-677. e611.
226. Solanas G, Peixoto FO, Perdiguero E, Jardí M, Ruiz-Bonilla V, Datta D, Symeonidi A, Castellanos A, Welz P-S, Caballero JM: **Aged stem cells reprogram their daily rhythmic functions to adapt to stress**. *Cell* 2017, **170**(4):678-692. e620.
227. Baker CL, Loros JJ, Dunlap JC: **The circadian clock of Neurospora crassa**. *FEMS microbiology reviews* 2012, **36**(1):95-110.
228. Aronson BD, Johnson KA, Dunlap JC: **Circadian clock locus frequency: protein encoded by a single open reading frame defines period length and temperature compensation**. *Proceedings of the National Academy of Sciences* 1994, **91**(16):7683-7687.
229. Shilatifard A: **The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis**. *Annual review of biochemistry* 2012, **81**:65-95.
230. Mozzetta C, Boyarchuk E, Pontis J, Ait-Si-Ali S: **Sound of silence: the properties and functions of repressive Lys methyltransferases**. *Nature reviews Molecular cell biology* 2015, **16**(8):499-513.
231. Chen CH, Dunlap JC, Loros JJ: **Neurospora illuminates fungal photoreception**. *Fungal Genet Biol* 2010, **47**(11):922-929.
232. He Q, Liu Y: **Molecular mechanism of light responses in Neurospora: from light-induced transcription to photoadaptation**. *Genes Dev* 2005, **19**(23):2888-2899.
233. Proietto M, Bianchi MM, Ballario P, Brenna A: **Epigenetic and Posttranslational Modifications in Light Signal Transduction and the Circadian Clock in Neurospora crassa**. *Int J Mol Sci* 2015, **16**(7):15347-15383.
234. Schafmeier T, Diernfellner AC: **Light input and processing in the circadian clock of Neurospora**. *FEBS Lett* 2011, **585**(10):1467-1473.
235. Ballario P, Vittorioso P, Magrelli A, Talora C, Cabibbo A, Macino G: **White collar-1, a central regulator of blue light responses in Neurospora, is a zinc finger protein**. *EMBO J* 1996, **15**(7):1650-1657.
236. Ballario P, Talora C, Galli D, Linden H, Macino G: **Roles in dimerization and blue light photoresponse of the PAS and LOV domains of Neurospora crassa white collar proteins**. *Mol Microbiol* 1998, **29**(3):719-729.

237. Linden H, Macino G: **White collar 2, a partner in blue-light signal transduction, controlling expression of light-regulated genes in Neurospora crassa**. *EMBO J* 1997, **16**(1):98-109.
238. Froehlich AC, Liu Y, Loros JJ, Dunlap JC: **White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter**. *Science* 2002, **297**(5582):815-819.
239. He Q, Cheng P, Yang Y, Wang L, Gardner KH, Liu Y: **White collar-1, a DNA binding transcription factor and a light sensor**. *Science* 2002, **297**(5582):840-843.
240. Wu C, Yang F, Smith KM, Peterson M, Dekhang R, Zhang Y, Zucker J, Bredeweg EL, Mallappa C, Zhou X *et al*: **Genome-wide characterization of light-regulated genes in Neurospora crassa**. *G3 (Bethesda)* 2014, **4**(9):1731-1745.
241. Chen CH, Ringelberg CS, Gross RH, Dunlap JC, Loros JJ: **Genome-wide analysis of light-inducible responses reveals hierarchical light signalling in Neurospora**. *EMBO J* 2009, **28**(8):1029-1042.
242. Lewis ZA, Correa A, Schwerdtfeger C, Link KL, Xie X, Gomer RH, Thomas T, Ebbole DJ, Bell-Pedersen D: **Overexpression of White Collar-1 (WC-1) activates circadian clock-associated genes, but is not sufficient to induce most light-regulated gene expression in Neurospora crassa**. *Mol Microbiol* 2002, **45**(4):917-931.
243. Ruesch CE, Ramakrishnan M, Park J, Li N, Chong HS, Zaman R, Joska TM, Belden WJ: **The histone H3 lysine 9 methyltransferase DIM-5 modifies chromatin at frequency and represses light-activated gene expression**. *G3 (Bethesda)* 2014, **5**(1):93-101.
244. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T: **Active genes are tri-methylated at K4 of histone H3**. *Nature* 2002, **419**(6905):407-411.
245. Briggs SD, Bryk M, Strahl BD, Cheung WL, Davie JK, Dent SY, Winston F, Allis CD: **Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae**. *Genes Dev* 2001, **15**(24):3286-3295.
246. Fingerman IM, Wu CL, Wilson BD, Briggs SD: **Global loss of Set1-mediated H3 Lys4 trimethylation is associated with silencing defects in Saccharomyces cerevisiae**. *J Biol Chem* 2005, **280**(31):28761-28765.
247. Venkatasubrahmanyam S, Hwang WW, Meneghini MD, Tong AH, Madhani HD: **Genome-wide, as opposed to local, antisilencing is mediated redundantly by the euchromatic factors Set1 and H2A.Z**. *Proc Natl Acad Sci U S A* 2007, **104**(42):16609-16614.
248. Lorenz DR, Meyer LF, Grady PJ, Meyer MM, Cam HP: **Heterochromatin assembly and transcriptome repression by Set1 in coordination with a class II histone deacetylase**. *Elife* 2014, **3**:e04506.
249. Lorenz DR, Mikheyeva IV, Johansen P, Meyer L, Berg A, Grewal SI, Cam HP: **CENP-B cooperates with Set1 in bidirectional transcriptional silencing**

and genome organization of retrotransposons. *Mol Cell Biol* 2012, **32**(20):4215-4225.
250. Jezek M, Gast A, Choi G, Kulkarni R, Quijote J, Graham-Yooll A, Park D, Green EM: **The histone methyltransferases Set5 and Set1 have overlapping functions in gene silencing and telomere maintenance**. *Epigenetics* 2017, **12**(2):93-104.
251. Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, Devarajan K, Daugherty AC, Kundaje AB, Mancini E *et al*: **H3K4me3 breadth is linked to cell identity and transcriptional consistency**. *Cell* 2014, **158**(3):673-688.
252. Howe FS, Fischl H, Murray SC, Mellor J: **Is H3K4me3 instructive for transcription activation?** *Bioessays* 2017, **39**(1):1-12.
253. Freitag M, Hickey PC, Khlafallah TK, Read ND, Selker EU: **HP1 is essential for DNA methylation in neurospora**. *Mol Cell* 2004, **13**(3):427-434.
254. Honda S, Selker EU: **Direct interaction between DNA methyltransferase DIM-2 and HP1 is required for DNA methylation in Neurospora crassa**. *Mol Cell Biol* 2008, **28**(19):6044-6055.
255. Lewis ZA, Adhvaryu KK, Honda S, Shiver AL, Knip M, Sack R, Selker EU: **DNA methylation and normal chromosome behavior in Neurospora depend on five components of a histone methyltransferase complex, DCDC**. *PLoS Genet* 2010, **6**(11):e1001196.
256. Vakoc CR, Mandat SA, Olenchock BA, Blobel GA: **Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin**. *Mol Cell* 2005, **19**(3):381-391.
257. Buhler M, Moazed D: **Transcription and RNAi in heterochromatic gene silencing**. *Nat Struct Mol Biol* 2007, **14**(11):1041-1048.
258. Lewis ZA, Honda S, Khlafallah TK, Jeffress JK, Freitag M, Mohn F, Schubeler D, Selker EU: **Relics of repeat-induced point mutation direct heterochromatin formation in Neurospora crassa**. *Genome Res* 2009, **19**(3):427-437.
259. Smith KM, Phatale PA, Sullivan CM, Pomraning KR, Freitag M: **Heterochromatin is required for normal distribution of Neurospora crassa CenH3**. *Mol Cell Biol* 2011, **31**(12):2528-2542.
260. Heintzen C, Loros JJ, Dunlap JC: **The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting**. *Cell* 2001, **104**(3):453-464.
261. Schwerdtfeger C, Linden H: **VIVID is a flavoprotein and serves as a fungal blue light photoreceptor for photoadaptation**. *The EMBO journal* 2003, **22**(18):4846-4855.
262. Lee HC, Li L, Gu W, Xue Z, Crosthwaite SK, Pertsemlidis A, Lewis ZA, Freitag M, Selker EU, Mello CC *et al*: **Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi**. *Mol Cell* 2010, **38**(6):803-814.

263. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL: **Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown**. *Nature protocols* 2016, **11**(9):1650.
264. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S *et al*: **The genome sequence of the filamentous fungus Neurospora crassa**. *Nature* 2003, **422**(6934):859-868.
265. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L: **Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks**. *Nat Protoc* 2012, **7**(3):562-578.
266. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, Salzberg SL, Wold BJ, Pachter L: **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation**. *Nature biotechnology* 2010, **28**(5):511.
267. Goff L, Trapnell C, Kelley D: **cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequening data.** *R package version 2200* 2013.
268. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP: **Integrative genomics viewer**. *Nat Biotechnol* 2011, **29**(1):24-26.
269. Li H, Durbin R: **Fast and accurate long-read alignment with Burrows-Wheeler transform**. *Bioinformatics* 2010, **26**(5):589-595.
270. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W *et al*: **Model-based analysis of ChIP-Seq (MACS)**. *Genome Biol* 2008, **9**(9):R137.
271. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T *et al*: **Orchestrating high-throughput genomic analysis with Bioconductor**. *Nat Methods* 2015, **12**(2):115-121.
272. Yu G, Wang LG, He QY: **ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization**. *Bioinformatics* 2015, **31**(14):2382-2383.
273. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T: **deepTools2: a next generation web server for deep-sequencing data analysis**. *Nucleic Acids Res* 2016, **44**(W1):W160-165.
274. Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features**. *Bioinformatics* 2010, **26**(6):841-842.
275. Arthanari Y, Heintzen C, Griffiths-Jones S, Crosthwaite SK: **Natural antisense transcripts and long non-coding RNA in Neurospora crassa**. *PLoS One* 2014, **9**(3):e91353.
276. Cemel IA, Ha N, Schermann G, Yonekawa S, Brunner M: **The coding and noncoding transcriptome of Neurospora crassa**. *BMC Genomics* 2017, **18**(1):978.
277. Jamieson K, Wiles ET, McNaught KJ, Sidoli S, Leggett N, Shao Y, Garcia BA, Selker EU: **Loss of HP1 causes depletion of H3K27me3 from facultative**

heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Res* 2016, **26**(1):97-107.
278. Smith KM, Sancar G, Dekhang R, Sullivan CM, Li S, Tag AG, Sancar C, Bredeweg EL, Priest HD, McCormick RF *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* 2010, **9**(10):1549-1556.
279. Basenko EY, Sasaki T, Ji L, Prybol CJ, Burckhardt RM, Schmitz RJ, Lewis ZA: **Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth**. *Proc Natl Acad Sci U S A* 2015, **112**(46):E6339-6348.
280. Wang Y, Smith KM, Taylor JW, Freitag M, Stajich JE: **Endogenous Small RNA Mediates Meiotic Silencing of a Novel DNA Transposon**. *G3 (Bethesda)* 2015, **5**(10):1949-1960.
281. Dekhang R, Wu C, Smith KM, Lamb TM, Peterson M, Bredeweg EL, Ibarra O, Emerson JM, Karunarathna N, Lyubetskaya A *et al*: **The Neurospora Transcription Factor ADV-1 Transduces Light Signals and Temporal Information to Control Rhythmic Expression of Genes Involved in Cell Fusion**. *G3 (Bethesda)* 2017, **7**(1):129-142.
282. Raduwan H, Isola AL, Belden WJ: **Methylation of histone H3 on lysine 4 by the lysine methyltransferase SET1 protein is needed for normal clock gene expression**. *J Biol Chem* 2013, **288**(12):8380-8390.
283. Matsumura Y, Nakaki R, Inagaki T, Yoshida A, Kano Y, Kimura H, Tanaka T, Tsutsumi S, Nakao M, Doi T *et al*: **H3K4/H3K9me3 Bivalent Chromatin Domains Targeted by Lineage-Specific DNA Methylation Pauses Adipocyte Differentiation**. *Mol Cell* 2015, **60**(4):584-596.
284. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K *et al*: **A bivalent chromatin structure marks key developmental genes in embryonic stem cells**. *Cell* 2006, **125**(2):315-326.
285. Tamaru H, Zhang X, McMillen D, Singh PB, Nakayama J, Grewal SI, Allis CD, Cheng X, Selker EU: **Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in Neurospora crassa**. *Nat Genet* 2003, **34**(1):75-79.
286. Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, Bouvard V, Altieri A, Cogliano V: **Carcinogenicity of alcoholic beverages**. *The lancet oncology* 2007, **8**(4):292-293.
287. Graham K, Bernards S, Knibbe R, Kairouz S, Kuntsche S, Wilsnack SC, Greenfield TK, Dietze P, Obot I, Gmel G: **Alcohol-related negative consequences among drinkers around the world**. *Addiction* 2011, **106**(8):1391-1405.
288. Kong CM, Lee XW, Wang X: **Telomere shortening in human diseases**. *The FEBS journal* 2013, **280**(14):3180-3193.
289. Bailey SM, Cunningham CC: **Contribution of mitochondria to oxidative stress associated with alcoholic liver disease**. *Free Radical Biology and Medicine* 2002, **32**(1):11-16.

290. Wu D, Cederbaum AI: **Alcohol, oxidative stress, and free radical damage**. *Alcohol Research and Health* 2003, **27**:277-284.
291. Cadet J, Wagner JR: **DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation**. *Cold Spring Harbor perspectives in biology* 2013, **5**(2):a012559.
292. Garaycoechea JI, Crossan GP, Langevin F, Mulderrig L, Louzada S, Yang F, Guilbaud G, Park N, Roerink S, Nik-Zainal S: **Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells**. *Nature* 2018, **553**(7687):171.
293. Joenje H: **Metabolism: alcohol, DNA and disease**. *Nature* 2011, **475**(7354):45.
294. Pavanello S, Hoxha M, Dioni L, Bertazzi PA, Snenghi R, Nalesso A, Ferrara SD, Montisci M, Baccarelli A: **Shortened telomeres in individuals with abuse in alcohol consumption**. *International Journal of Cancer* 2011, **129**(4):983-992.
295. Panossian L, Porter V, Valenzuela H, Zhu X, Reback E, Masterman D, Cummings J, Effros R: **Telomere shortening in T cells correlates with Alzheimer's disease status**. *Neurobiology of aging* 2003, **24**(1):77-84.
296. Fyhrquist F, Saijonmaa O, Strandberg T: **The roles of senescence and telomere shortening in cardiovascular disease**. *Nature Reviews Cardiology* 2013, **10**(5):274.
297. Calado RT, Young NS: **Telomere diseases**. *New England Journal of Medicine* 2009, **361**(24):2353-2365.
298. Maciejowski J, de Lange T: **Telomeres in cancer: tumour suppression and genome instability**. *Nature reviews Molecular cell biology* 2017, **18**(3):175.
299. Azzalin CM, Reichenbach P, Khoriauli L, Giulotto E, Lingner J: **Telomeric repeat–containing RNA and RNA surveillance factors at mammalian chromosome ends**. *Science* 2007, **318**(5851):798-801.
300. De Lange T: **Shelterin: the protein complex that shapes and safeguards human telomeres**. *Genes & development* 2005, **19**(18):2100-2110.
301. Redon S, Reichenbach P, Lingner J: **The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase**. *Nucleic acids research* 2010, **38**(17):5797-5806.
302. Chu H-P, Cifuentes-Rojas C, Kesner B, Aeby E, Lee H-g, Wei C, Oh HJ, Boukhali M, Haas W, Lee JT: **TERRA RNA antagonizes ATRX and protects telomeres**. *Cell* 2017, **170**(1):86-101. e116.
303. Gamsby J, Templeton E, Bonvini L, Wang W, Loros J, Dunlap J, Green A, Gulick D: **The circadian Per1 and Per2 genes influence alcohol intake, reinforcement, and blood alcohol levels**. *Behavioural brain research* 2013, **249**:15-21.
304. Spanagel R, Pendyala G, Abarca C, Zghoul T, Sanchis-Segura C, Magnone MC, Lascorz J, Depner M, Holzberg D, Soyka M: **The clock gene Per2 influences the glutamatergic system and modulates alcohol consumption**. *Nature medicine* 2005, **11**(1):35.

305. Ozburn AR, Falcon E, Mukherjee S, Gillman A, Arey R, Spencer S, McClung CA: **The role of clock in ethanol-related behaviors**. *Neuropsychopharmacology* 2013, **38**(12):2393.
306. Seggio JA, Fixaris MC, Reed JD, Logan RW, Rosenwasser AM: **Chronic ethanol intake alters circadian phase shifting and free-running period in mice**. *Journal of Biological Rhythms* 2009, **24**(4):304-312.
307. Zhou P, Werner JH, Lee D, Sheppard AD, Liangpunsakul S, Duffield GE: **Dissociation between diurnal cycles in locomotor activity, feeding behavior and hepatic PERIOD2 expression in chronic alcohol-fed mice**. *Alcohol* 2015, **49**(4):399-408.
308. Zhou P, Ross RA, Pywell CM, Liangpunsakul S, Duffield GE: **Disturbances in the murine hepatic circadian clock in alcohol-induced hepatic steatosis**. *Scientific reports* 2014, **4**:3725.
309. Schafmeier T, Haase A, Káldi K, Scholz J, Fuchs M, Brunner M: **Transcriptional feedback of Neurospora circadian clock gene by phosphorylation-dependent inactivation of its transcription factor**. *Cell* 2005, **122**(2):235-246.
310. Carlucci M, Kriščiūnas A, Li H, Gibas P, Koncevičius K, Petronis A, Oh G: **DiscoRhythm: an easy-to-use web application and R package for discovering rhythmicity**. *Bioinformatics* 2019.
311. Gerlai R, Lahav M, Guo S, Rosenthal A: **Drinks like a fish: zebra fish (Danio rerio) as a behavior genetic model to study alcohol effects**. *Pharmacology biochemistry and behavior* 2000, **67**(4):773-782.
312. Echevarria DJ, Toms CN, Jouandot DJ: **Alcohol-induced behavior change in zebrafish models**. *Reviews in the Neurosciences* 2011, **22**(1):85-93.
313. Mathur P, Guo S: **Use of zebrafish as a model to understand mechanisms of addiction and complex neurobehavioral phenotypes**. *Neurobiology of disease* 2010, **40**(1):66-72.
314. Reimers MJ, Hahn ME, Tanguay RL: **Two zebrafish alcohol dehydrogenases share common ancestry with mammalian class I, II, IV, and V alcohol dehydrogenase genes but have distinct functional characteristics**. *Journal of Biological Chemistry* 2004, **279**(37):38303-38312.
315. Roenneberg T, Merrow M: **The circadian clock and human health**. *Current biology* 2016, **26**(10):R432-R443.
316. Huang MC, Ho CW, Chen CH, Liu SC, Chen CC, Leu SJ: **Reduced expression of circadian clock genes in male alcoholic patients**. *Alcoholism: Clinical and Experimental Research* 2010, **34**(11):1899-1904.
317. Filiano AN, Millender-Swain T, Johnson Jr R, Young ME, Gamble KL, Bailey SM: **Chronic ethanol consumption disrupts the core molecular clock and diurnal rhythms of metabolic genes in the liver without affecting the suprachiasmatic nucleus**. *PloS one* 2013, **8**(8):e71684.
318. Sarkar DK: **Circadian genes, the stress axis, and alcoholism**. *Alcohol research: current reviews* 2012, **34**(3):362.

319. Kitada T, Seki S, Kawakita N, Kuroki T, Monna T: **Telomere shortening in chronic liver diseases**. *Biochemical and biophysical research communications* 1995, **211**(1):33-39.
320. Wiemann SU, Satyanarayana A, Tsahuridu M, Tillmann HL, Zender L, Klempnauer J, Flemming P, Franco S, Blasco MA, Manns MP: **Hepatocyte telomere shortening and senescence are general markers of human liver cirrhosis**. *The FASEB journal* 2002, **16**(9):935-942.
321. Anchelin M, Murcia L, Alcaraz-Pérez F, García-Navarro EM, Cayuela ML: **Behaviour of telomere and telomerase during aging and regeneration in zebrafish**. *PloS one* 2011, **6**(2):e16955.
322. Carneiro MC, Henriques CM, Nabais J, Ferreira T, Carvalho T, Ferreira MG: **Short telomeres in key tissues initiate local and systemic aging in zebrafish**. *PLoS genetics* 2016, **12**(1):e1005798.
323. Blasco MA: **The epigenetic regulation of mammalian telomeres**. *Nature Reviews Genetics* 2007, **8**(4):299.
324. Galati A, Micheli E, Cacchione S: **Chromatin structure in telomere dynamics**. *Frontiers in oncology* 2013, **3**:46.
325. Passos JF, Saretzki G, von Zglinicki T: **DNA damage in telomeres and mitochondria during cellular senescence: is there a connection?** *Nucleic acids research* 2007, **35**(22):7505-7513.
326. Harpaz T, Abumock H, Beery E, Edel Y, Lahav M, Rozovski U, Uziel O: **The effect of ethanol on telomere dynamics and regulation in human cells**. *Cells* 2018, **7**(10):169.
327. Denchi EL, de Lange T: **Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1**. *Nature* 2007, **448**(7157):1068.
328. Mann RE, Smart RG, Govoni R: **The epidemiology of alcoholic liver disease**. *Alcohol research and health* 2003, **27**:209-219.
329. Park K-H, Kim S-H: **Low dose of chronic ethanol exposure in adult zebrafish induces hepatic steatosis and injury**. *Biomedicine & Pharmacotherapy* 2019, **117**:109179.
330. Spencer RL, Hutchison KE: **Alcohol, aging, and the stress response**. *Alcohol Research and Health* 1999, **23**(4):272-283.
331. Bagnardi V, Rota M, Botteri E, Tramacere I, Islami F, Fedirko V, Scotti L, Jenab M, Turati F, Pasquali E: **Alcohol consumption and site-specific cancer risk: a comprehensive dose–response meta-analysis**. *British journal of cancer* 2015, **112**(3):580.
332. Boffetta P, Hashibe M: **Alcohol and cancer**. *The lancet oncology* 2006, **7**(2):149-156.
333. Okamoto K, Seimiya H: **Revisiting telomere shortening in cancer**. *Cells* 2019, **8**(2):107.
334. Lu W, Zhang Y, Liu D, Songyang Z, Wan M: **Telomeres—structure, function, and regulation**. *Experimental cell research* 2013, **319**(2):133-141.

335. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K: **Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer**. *Nature genetics* 2005, **37**(4):391.
336. Jones PA, Baylin SB: **The fundamental role of epigenetic events in cancer**. *Nature reviews genetics* 2002, **3**(6):415.
337. Zhu Q, Belden WJ: **Molecular Regulation of Circadian Chromatin**. *Journal of Molecular Biology* 2020.
338. Aguilar-Arnal L, Sassone-Corsi P: **Chromatin landscape and circadian dynamics: Spatial and temporal organization of clock transcription**. *Proceedings of the National Academy of Sciences* 2015, **112**(22):6863-6870.
339. Hogenesch JB, Gu Y-Z, Jain S, Bradfield CA: **The basic-helix–loop–helix-PAS orphan MOP3 forms transcriptionally active complexes with circadian and hypoxia factors**. *Proceedings of the National Academy of Sciences* 1998, **95**(10):5474-5479.
340. Brockdorff N, Turner BM: **Dosage compensation in mammals**. *Cold Spring Harbor perspectives in biology* 2015, **7**(3):a019406.
341. Davis E, Caiment F, Tordoir X, Cavaillé J, Ferguson-Smith A, Cockett N, Georges M, Charlier C: **RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus**. *Current Biology* 2005, **15**(8):743-749.
342. Martienssen R, Moazed D: **RNAi and heterochromatin assembly**. *Cold Spring Harbor perspectives in biology* 2015, **7**(8):a019323.
343. Bunger MK, Wilsbacher LD, Moran SM, Clendenin C, Radcliffe LA, Hogenesch JB, Simon MC, Takahashi JS, Bradfield CA: **Mop3 is an essential component of the master circadian pacemaker in mammals**. *Cell* 2000, **103**(7):1009-1017.
344. Yoo S-H, Yamazaki S, Lowrey PL, Shimomura K, Ko CH, Buhr ED, Siepka SM, Hong H-K, Oh WJ, Yoo OJ: **PERIOD2:: LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues**. *Proceedings of the National Academy of Sciences* 2004, **101**(15):5339-5346.
345. Grant GD, Gamsby J, Martyanov V, Brooks 3rd L, George LK, Mahoney JM, Loros JJ, Dunlap JC, Whitfield ML: **Live-cell monitoring of periodic gene expression in synchronous human cells identifies Forkhead genes involved in cell cycle control**. *Molecular biology of the cell* 2012, **23**(16):3079-3093.
346. Drori S, Girnun GD, Tou L, Szwaya JD, Mueller E, Kia X, Shivdasani RA, Spiegelman BM: **Hic-5 regulates an epithelial program mediated by PPARγ**. *Genes & Development* 2005, **19**(3):362-375.
347. Hacisuleyman E, Goff LA, Trapnell C, Williams A, Henao-Mejia J, Sun L, McClanahan P, Hendrickson DG, Sauvageau M, Kelley DR: **Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre**. *Nature structural & molecular biology* 2014, **21**(2):198.

348. Chu C, Quinn J, Chang HY: **Chromatin isolation by RNA purification (ChIRP)**. *JoVE (Journal of Visualized Experiments)* 2012(61):e3912.
349. Chavez A, Scheiman J, Vora S, Pruitt BW, Tuttle M, Iyer EP, Lin S, Kiani S, Guzman CD, Wiegand DJ: **Highly efficient Cas9-mediated transcriptional programming**. *Nature methods* 2015, **12**(4):326.
350. Yu K, Roy D, Huang FT, Lieber MR: **Detection and Structural Analysis of R-Loops**. *Methods in enzymology* 2006, **409**:316-329.
351. Smith KM, Sancar G, Dekhang R, Sullivan CM, Li S, Tag AG, Sancar C, Bredeweg EL, Priest HD, McCormick RF: **Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for Neurospora white collar complex**. *Eukaryotic cell* 2010, **9**(10):1549-1556.
352. Nagoshi E, Brown SA, Dibner C, Kornmann B, Schibler U: **Circadian gene expression in cultured cells**. In: *Methods in enzymology*. vol. 393: Elsevier; 2005: 543-557.
353. Kowalska E, Ripperger JA, Hoegger DC, Bruegger P, Buch T, Birchler T, Mueller A, Albrecht U, Contaldo C, Brown SA: **NONO couples the circadian clock to the cell cycle**. *Proceedings of the National Academy of Sciences* 2013, **110**(5):1592-1599.
354. Sun X, Haider Ali MSS, Moran M: **The role of interactions of long non-coding RNAs and heterogeneous nuclear ribonucleoproteins in regulating cellular functions**. *Biochemical Journal* 2017, **474**(17):2925-2935.
355. Geuens T, Bouhy D, Timmerman V: **The hnRNP family: insights into their role in health and disease**. *Human genetics* 2016, **135**(8):851-867.
356. Cam HP, Chen ES, Grewal SI: **Transcriptional scaffolds for heterochromatin assembly**. *Cell* 2009, **136**(4):610-614.
357. Cho DH, Thienes CP, Mahoney SE, Analau E, Filippova GN, Tapscott SJ: **Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF**. *Molecular cell* 2005, **20**(3):483-489.
358. Pelechano V, Steinmetz LM: **Gene regulation by antisense transcription**. *Nature Reviews Genetics* 2013, **14**(12):880-893.
359. Tomita S, Abdalla MOA, Fujiwara S, Yamamoto T, Iwase H, Nakao M, Saitoh N: **Roles of long noncoding RNAs in chromosome domains**. *Wiley Interdisciplinary Reviews: RNA* 2017, **8**(2):e1384.
360. Nguyen ED, Balas MM, Griffin AM, Roberts JT, Johnson AM: **Global profiling of hnRNP A2/B1-RNA binding on chromatin highlights LncRNA interactions**. *RNA biology* 2018, **15**(7):901-913.
361. Wang IX, Grunseich C, Fox J, Burdick J, Zhu Z, Ravazian N, Hafner M, Cheung VG: **Human proteins that interact with RNA/DNA hybrids**. *Genome research* 2018, **28**(9):1405-1414.
362. Wutz A: **Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation**. *Nature Reviews Genetics* 2011, **12**(8):542-553.

363. De Silanes IL, d'Alcontres MS, Blasco MA: **TERRA transcripts are bound by a complex array of RNA-binding proteins**. *Nature communications* 2010, **1**:33.
364. Liu T-Y, Chen Y-C, Jong Y-J, Tsai H-J, Lee C-C, Chang Y-S, Chang J-G, Chang Y-F: **Muscle developmental defects in heterogeneous nuclear Ribonucleoprotein A1 knockout mice**. *Open biology* 2017, **7**(1):160303.
365. Gaudreau M-C, Heyd F, Bastien R, Wilhelm B, Möröy T: **Alternative splicing controlled by heterogeneous nuclear ribonucleoprotein L regulates development, proliferation, and migration of thymic pre-T cells**. *The Journal of Immunology* 2012, **188**(11):5377-5388.
366. Roshon MJ, Ruley HE: **Hypomorphic mutation in hnRNP U results in post-implantation lethality**. *Transgenic research* 2005, **14**(2):179-192.
367. Lai B, Zou J, Lin Z, Qu Z, Song A, Xu Y, Gao X: **Haploinsufficiency of hnRNP U Changes Activity Pattern and Metabolic Rhythms**. *The American journal of pathology* 2018, **188**(1):173-183.
368. Sasaki H, Ishihara K, Kato R: **Mechanisms of Igf2/H19 imprinting: DNA methylation, chromatin and long-distance gene regulation**. *The Journal of Biochemistry* 2000, **127**(5):711-715.
369. Bi H-s, Yang X-y, Yuan J-h, Yang F, Xu D, Guo Y-j, Zhang L, Zhou C-c, Wang F, Sun S-h: **H19 inhibits RNA polymerase II-mediated transcription by disrupting the hnRNP U–actin complex**. *Biochimica et Biophysica Acta (BBA)-General Subjects* 2013, **1830**(10):4899-4906.
370. Ameyar-Zazoua M, Souidi M, Fritsch L, Robin P, Thomas A, Hamiche A, Percipalle P, Ait-Si-Ali S, Harel-Bellan A: **Physical and functional interaction between heterochromatin protein 1α and the RNA-binding protein heterogeneous nuclear ribonucleoprotein U**. *Journal of Biological Chemistry* 2009, **284**(41):27974-27979.
371. Tan-Wong SM, Dhir S, Proudfoot NJ: **R-Loops Promote Antisense Transcription across the Mammalian Genome**. *Molecular cell* 2019, **76**(4):600-616. e606.
372. Cerritelli SM, Crouch RJ: **Ribonuclease H: the enzymes in eukaryotes**. *The FEBS journal* 2009, **276**(6):1494-1505.
373. Chen J-Y, Zhang X, Fu X-D, Chen L: **R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1**. *Nature protocols* 2019, **14**(5):1661.
374. Wasielewski JA, Holloway FA: **Alcohol's interactions with circadian rhythms: A focus on body temperature**. *Alcohol Research & Health* 2001, **25**(2):94.