INVESTIGATING THE EVOLUTION OF THE MOLECULAR CLOCK MECHANISM USING THE HOUSEFLY, MUSCA DOMESTICA

A Thesis

by

COURTNEY MARIE CASTER

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

MASTER OF SCIENCE

Chair of Committee,	Paul Hardin
Committee Members,	Christine Merlin
	Vlad Panin
	Mark Zoran
Intercollegiate Faculty Chair,	Dorothy Shippen

August 2017

Major Subject: Genetics

ABSTRACT

Mapping the evolution of the transcriptional feedback loops that regulate the circadian clock will lead to the understanding of how this essential pacemaker allows insects to anticipate changes in their environment. Transcriptional feedback loops were first discovered in *Drosophila*, and while there is extensive conservation to mammals, it is evident that different insects have diverged or adopted different mechanisms to construct their circadian clocks. Here, we investigate the divergence of the clock in dipterans by characterizing the evolution of molecular clock mechanisms from the model organism *Drosophila* to another dipteran, the housefly (*Musca domestica*).

In *Drosophila* the core feedback loop is composed of the core components: *Clock* (*Clk*), *cycle* (*cyc*), *period* (*per*), *timeless* (*tim*), *clockwork orange* (*cwo*), and *cryptochrome* (*Cry1)*. CLK binds CYC to initiate the feedback loop, and together they drive transcription of *per* and *tim*. PER:TIM heterodimers then repress their own transcription by binding to and inhibiting CLK:CYC. In addition, CWO is a transcriptional repressor which synergizes with PER to repress CLK:CYC activation. Insects such as monarch butterflies, however, appear to have an ancestral clock, possessing not only the *Drosophila* clock genes but also mammalian components such as the *cycle* ortholog (*Bmal1*) and a mammalian *cryptochrome, cry2*. In mice and "ancestral-like" insect clocks, BMAL1 contains the main transactivation domain of the CLK:BMAL1 complex.

In these insects CRY2 also takes the place of *Drosophila* PER as the main transcriptional repressor.

Examination of the genomic sequences of *M. domestica* identified both the core *Drosophila* clock genes and the ancestral clock gene, *cry2*. However, further analysis, showed that *mdCry2* is in fact a (6-4) DNA photolyase. Furthermore, characterization of the circadian clock in the housefly and its comparison to the *Drosophila* clock revealed the functional genetic differences in their transcriptional feedback loops. Indeed, in the housefly CWO appears to have taken on the role of *Drosophila* PER as the main transcriptional repressor. This work confirms that the organization of the clock has diverged during insect evolution, and that there are multiple genetic approaches to regulating overt rhythmicity.

DEDICATION

To Mom, Stephen, David, and my large furry support system: Eowyn, Fergus, Seamus, Fiona, Bert, Ellie, Quinn and all those before them.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Hardin, and my committee members, Dr. Merlin, Dr. Panin, and Dr. Zoran, for their guidance and support throughout the course of this research. I definitely would not have been able to make it without the patience, advice, and support of Dr. Merlin. I would also like to give a special thanks to Dr. Meisel and Chelsea Holcomb who took the time to teach me about housefly husbandry.

Thanks also to my lab mates and friends who supported me and helped me through all the tough times of grad school! And I want to also give a special thanks to my mom, my brother, and David Forgacs who never gave up on me.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Dr. Paul Hardin (Chair of Committee), Dr. Christine Merlin, and Dr. Mark Zoran of the Department of Biology and Dr. Vlad Panin of the Department of Biochemistry and Biophysics.

All work for the thesis was completed by the student, in collaboration with Dr. Christine Merlin of the Department of Biology.

Funding Sources

This work was made possible in part by John W. Lyons Jr. ’59 Endowed Chair. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Texas A&M University.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

The Circadian Clock

The study of how circadian clocks keep time is a relatively young field. It was not until the 1960s and 1970s that the field of chronobiology gained momentum. Early work done by Seymour Benzer and his student Ron Konopka with *Drosophila* led to the isolation of three mutants: per^0, per^S, *and* per^L and the discovery of the first clock gene *period*, on the X-chromosome (Konopka et al., 1971). The *per* mutants had altered periods in adult locomotor activity and eclosion rhythms (Konopka et al., 1971). The discovery that all three mutants affected both the pupal eclosion rhythms as well as the adult activity rhythms suggested that the clock is necessary for both the development and the physiology of an organism (Konopka et al., 1971). The use of different species to study the pacemaker by Colin S. Pittendrigh showed that even though the clock controlled different behaviors in different animals the general function of the clock remained "clock-like" among species (Pittendrigh et al., 1976). Their work and work done by other early chronobiologists paved the way to our understanding of the clock today.

The clock in all organisms serves the function of harmonizer, adapting and adjusting an organism's physiology, metabolism, and behavior to reflect changes in the environment (Sandrelli et al., 2008). These changes in the environment include seasonal fluctuations in temperature and the daily light and dark cycles due to the revolution of the earth

around the sun and the daily rotation of the earth on its axis (Meuti et al., 2013). Input from environmental cues such as temperature and length of daylight act as entraining factors for the oscillator (Bell-Pedersen et al, 2005). Clocks use the information gained from these environmental inputs to adjust and regulate their outputs, which include behavioral rhythms and clock-controlled gene (CCG) expression (Bell-Pedersen et al., 2005).

Whether its complex organisms like humans or single-celled cyanobacteria their behavioral rhythms are controlled by at least one circadian oscillator (Bell-Pedersen et al., 2005). All oscillators contain both positive and negative feedback elements, which form feedback loops and cause roughly 24-hour rhythms (Bell-Pedersen et al., 2005). Factors such as phosphorylation, dephosphorylation, light degradation, and the shuttling of clock proteins between the cytoplasm and the nucleus lead to changes in both clock and CCG expression (Zheng et al., 2008). These molecular events are responsible for the ability of an organism to act in coordination with its environment (Zheng et al., 2008). Its also interesting to note that even under constant conditions (e.g. constant darkness), the clock still continues to function in the absence of any environmental inputs (Sandrelli et al., 2008).

Without a functional clock, organisms would not be able to acclimate and effectively survive in their surrounding environment. Therefore, when the first clock gene *period* was discovered in *Drosophila* it was assumed that the *Drosophila* clock would be

representative of the clock in all animals. In fact, for many years *Drosophila* was the only organism, which had the molecular tools necessary for scientists to investigate the clock (Hall, 2003). But the notion that *Drosophila* held the key to our understanding of the clock biased our view on both how the clock functioned, as well as how the clock evolved in different species (Sandrelli et al., 2008). It was not until the mammalian *Clock* and *period* genes were identified that scientists were first able to determine that although vertebrate clocks share a common evolutionary origin with the *Drosophila* clock there were notable differences (Sandrelli et al., 2008). In addition, the sequencing of other insect species as well as the development of new molecular tools have led to the understanding that the clock has evolved in many different ways amongst insects.

The *Drosophila* Clock

The use of *Drosophila* (order Diptera) as a model organism began over 100 years ago and has been used to study many topics including human diseases, drug screening, pest control, development, and behavior (Beckingham et al., 2005; Jennings, 2011). Its use as a model organism first began in the laboratory of William E. Castle in 1901 but work done by T. H. Morgan set *Drosophila* on the scientific map (Beckingham et al., 2005; Jennings, 2011). He chose *Drosophila* for several reasons including its short generation time, small size, and ability to produce large quantities of progeny (Beckingham et al., 2005). Morgan added on to Gregor Mendel's work by redefining the theory of inheritance and establishing that genes were actually found within chromosomes (Jennings, 2011). By using *Drosophila*, Morgan's team also showed recombination

between homologous chromosomes as well as the chromosomal basis of sex determination (Beckingham et al., 2005).

Although outwardly *Drosophila* and mammals share no physical similarities they do in fact share similarities in both their development and behavior linking them evolutionarily (Beckingham et al., 2005; Jennings, 2011). However, the genetic tools and methods that have been generated for *Drosophila* surpass that of mammals, making *Drosophila* a valuable tool in understanding the genetics behind behavior and development (Beckingham et al., 2005; Jennings, 2011). In addition, sequencing of mammal genomes has shown that most genes in mammals have been duplicated (Beckingham et al., 2005; Jennings, 2011). *Drosophila* lacks the genetic redundancy of mammals, thus the complications encountered in unraveling mammalian genomes (Beckingham et al., 2005). *Drosophila* has also been used as a model for human diseases providing the ability to do large scale genetic screening and leading to the identification of genes involved in those diseases (Beckingham et al., 2005). Besides being useful in understanding the inner workings of the genome in mammals, *Drosophila* has also been fundamental in determining common genes found in both pest and beneficial insects, aiding in the development of different forms of pest control that target genes found only in pest species (Roberts, 2006).

The clock in *Drosophila* is responsible for the daily rhythms in many behaviors such as sleep, time of feeding, and courtship (Sakai et al., 2001; Tataroglu et al., 2014). In fact,

the robust rhythms seen in *Drosophila* courtship are caused by not only the clock mechanism but also by females, which generate mating rhythms (Sakai et al., 2001). The *Drosophila* clock also controls physiology, driving rhythms in many peripheral organs including the antennae where olfactory sensitivity rhythms persist (Tataroglu et al., 2014). However, it is *Drosophila's* eclosion rhythms and locomotor activity that have been used to discover the molecular mechanism of the clock and led to the discovery of the first clock gene, *period* (Dubowy et al., 2017; Konopka et al., 1971).

Eclosion is the process by which an adult fly emerges from its pupae and their emergence can be monitored as a rhythm with a population of flies (Dubowy et al., 2017). In wild-type *Drosophila*, the peak of eclosion is usually early in the morning (Dubowy et al., 2017). When adult wild-type *Drosophila* are entrained to a 12:12 hour light:dark (LD) cycle, their activity rhythms show two peaks, one at dawn and one at dusk making *Drosophila* crepuscular (Dubowy et al., 2017; Peschel et al., 2011). Before those peaks occur there is usually an anticipation of the lights turning on and off (Allada et al., 2010; Dubowy et al., 2017). When kept in constant conditions like constant darkness they continue to keep the same activity rhythms as when they were in a 12:12 hour LD cycle with a period of approximately 24 hours (Allada et al., 2010; Peschel et al., 2011).

As is seen in Figure 1, CLK and CYC form heterodimers and bind to the E-box elements (CACGTG) in the promoter regions of *per* and *tim* from the middle of the day to early in

the evening (Hardin, 2005; Rubin et al., 2006). These E-box elements are conserved between most species (Hardin, 2011). In addition, *Clk* and *cyc* are both part of a group of basic-helix-loop-helix-Period-Arnt-Single-minded (bHLH-PAS) transcription factors that are known to bind to E-box elements and activate the transcription of genes (Hardin, 2011). But unlike mammals, *Drosophila* CLK contains the main transactivation domain rather than CYC. And PER, like CLK and CYC, contains a PAS domain, but lacks the bHLH (Rosato et al., 2006).

The binding of CLK and CYC to E-box elements in the promoter regions of *per* and *tim* initiates their transcription (Hardin, 2011). PER and TIM protein levels peak late at night, a delay caused by phosphorylation events (Hardin, 2005). While in the cytoplasm PER and TIM form heterodimers (Figure 1). Once heterodimers, PER and TIM return to the nucleus where PER binds to CLK, thus effectively repressing its own transcription by inhibiting CLK:CYC DNA binding (Figure 1; Hardin, 2005). When exposed to the early morning light TIM is degraded, which leads to the destabilization and eventual degradation of PER (Hardin, 2011). CLK levels then increase and CLK and CYC are once again able to bind to the E-box elements and initiate the transcription of *per* and *tim* (Hardin, 2011). The repression of its own transcription by PER and TIM forms a negative feedback loop.

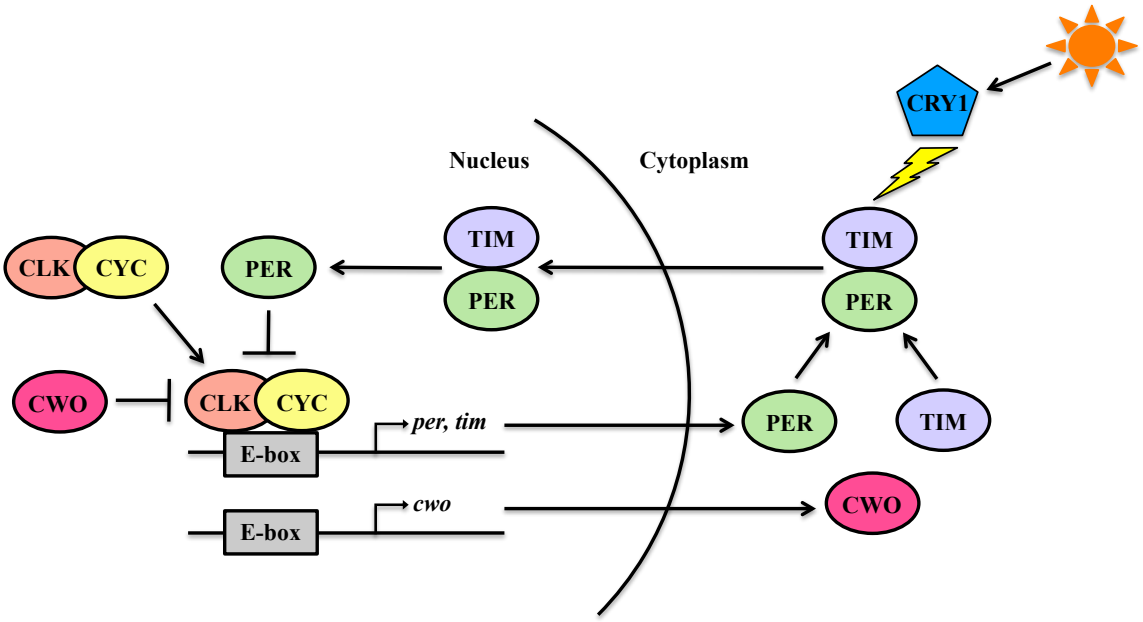

Figure 1: *Drosophila* clock. In *Drosophila* the clock is composed of the core components *Clock* (*Clk*), *cycle* (*cyc*), *period* (*per*), *timeless* (*tim*), *clockwork orange* (*cwo*), and *cryptochrome* (*cry1)*.

In *Drosophila*, light cues from the environment can entrain the circadian oscillator by causing either phase delays or phase advances (Hardin, 2005). *Cry1*, a blue light photoreceptor, is involved in this light entrainment and is known to co-localize in clock cells with both PER and TIM (Yuan et al, 2007; Hardin, 2005). *Cry1* is also a *cryptochrome* related to photolyases, which are DNA-binding proteins involved in repairing DNA damage caused by UV-light (Rosato et al., 2006). When exposed to light, CRY1 interacts with TIM leading to its degradation (Figure 1). This in turn leads to the destabilization and eventual degradation of PER (Rubin et al., 2006).

A recent discovery to the feedback loop in *Drosophila* is *clockwork orange* (*cwo*), a bHLH transcriptional repressor (Kadener et al., 2007; Zhou et al., 2016). Like *per* and

tim, CLK:CYC initiates transcription of *cwo*, and *cwo* transcription is in turn repressed by PER (Figure 1; Kadener et al., 2007). In Schneider (S2) cells, when both PER and CWO are co-transfected they were able to repress CLK activation by five to ten fold (Zhou et al., 2016). This suggests that not only is CWO a strong transcriptional repressor but it also acts in synergy with PER to repress CLK and CYC (Zhou et al., 2016). Further evidence of this synergy was shown in *Drosophila*. It was discovered that when PER binds to CLK and CYC during the early morning it not only reduces the binding affinity of CLK and CYC to the DNA, but it allows CWO to out compete CLK and CYC for binding to the E-boxes of *tim* (Zhou et al., 2016). However, this was not seen during the early night when PER is absent and CLK:CYC are strongly bound to the E-boxes (Zhou et al., 2016). This evidence suggests that together PER and CWO co-repress CLK and CYC by competing for E-box binding (Zhou et al., 2016).

Differences between *Drosophila* and Mammalian Clocks

Although the focus of this thesis is primarily on how the clock has evolved in insects, it is important to define the differences between the mammalian and *Drosophila* clocks because the clock in mammals has distinctive features also found in most insects other than *Drosophila*. The components of the mammalian clock include *brain and muscle Arnt-like protein-1* (*Bmal1*), *clock* (*clk*), *period* (*per1*, *per2*, and *per3*; although *per3* does not play a critical role in the feedback loops of the oscillator) and *cryptochrome* (*cry1* and *cry2*) (Figure 2; Ko et al., 2006). Based on the model shown in Figure 2, we can see that the primary feedback loop includes both positive and negative elements.

BMAL1 and CLK comprise the positive elements of the mammalian clock and are part of the basic helix-loop-helix (bHLH)-PAS (Period-Arnt-Single-minded) transcription factor family (Ko et al., 2006; Levi et al., 2007). Together the CRYs and PERs act as the negative elements of the feedback loop (Ko et al., 2006).

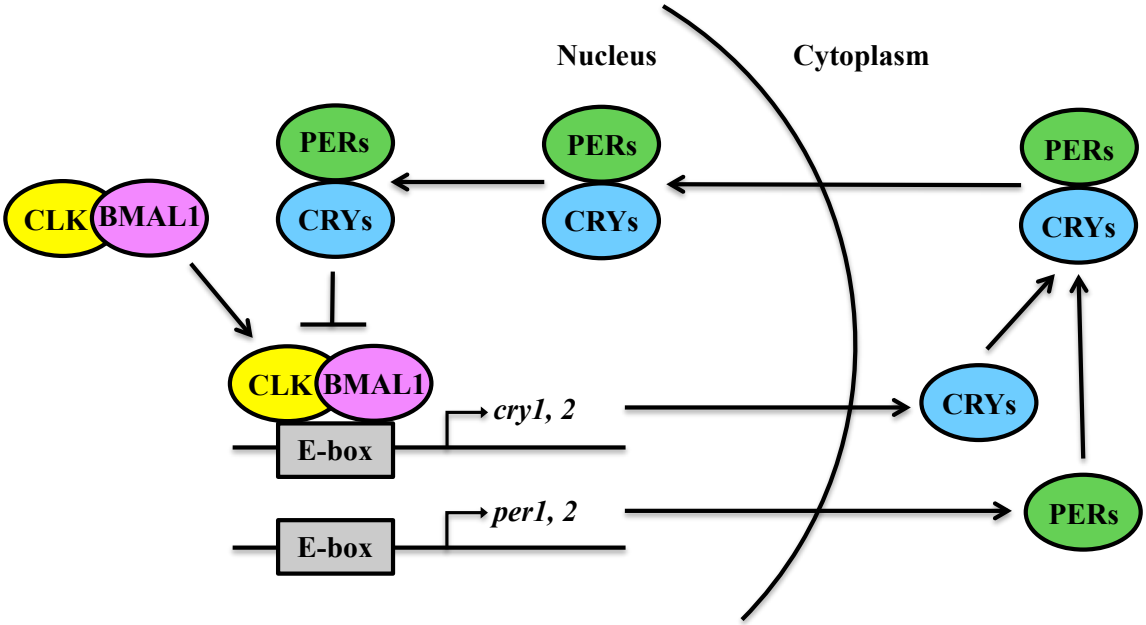

Figure 2: Mammalian clock. In mammals the clock is composed of the core components *clock* (*clk*), *brain and muscle Arnt-like protein-1* (*Bmal1*), *period* (*per1* and *per2*), and *cryptochrome* (*cry1* and *cry2)*.

BMAL1 and CLK form heterodimers, which bind to the E-boxes of *cryptochrome* and *period* (Ko et al., 2006; Levi et al., 2007). The binding by the transcription factors, BMAL1 and CLK drives the transcription of *cryptochrome* and *period*. In the cytoplasm, PERs and CRYs accumulate and form a complex (Levi et al., 2007). They then translocate back into the nucleus where they repress their own transcription by inhibiting the DNA binding activity of BMAL1:CLK (Ko et al., 2006). When the abundance of PERs and CRYs is at its maximum the levels of *per1*, *per2*, *cry1*, and *cry2* mRNAs

decrease (Maywood et al., 2007). Without any transcriptional activation, the amount of PER and CRY proteins also eventually decline (Maywood et al., 2007). When the amount of PERs and CRYs reaches a minimum negative regulation of their own genes is removed and the feedback loop can begin again (Levi et al., 2007; Maywood et al., 2007).

One key difference in the mammalian and *Drosophila* clocks is that in mammals, the negative elements (i.e. the PERs and CRYs) are found in multiple copies (Ko et al., 2006; Clayton et al., 2001). These gene duplications increase not only the complexity of the clock in mammals but also its redundancy (Clayton et al., 2001). Redundancy in the clock mechanism is a form of compensation; when a single *period* or *cryptochrome* gene is mutated the clock still continues to oscillate. However, if both the *cryptochrome* and/or the *period* genes were mutated together, it would cause arrhythmicity (Ko et al., 2006). It is also interesting to note that if one of the *cryptochrome* or *period* genes were mutated by itself, there would be an effect on the period of the animal (Ko et al., 2006). This indicates that each clock component cannot be fully compensated by another component of the clock (Ko et al., 2006).

Another key difference between mammalian and *Drosophila* clocks is that mammalian BMAL1 acts as the main transcriptional activator (Gustafson et al., 2017). In mammals, BMAL1 contains a C-terminal transactivation domain (Gustafson et al., 2017). In *Drosophila*, CYC is a truncated form of mammalian BMAL1 that lacks the C-terminal

transactivation domain (Sandrelli et al., 2008). However, in the C-terminus of *dmClk* there has been an expansion of the poly-Q repeats in comparison to mammalian *clk* (Sandrelli et al., 2008). This expansion of the poly-Q region is thought to provide the activation domain necessary for normal clock function in *Drosophila* (Sandrelli et al., 2008).

In addition, the *cryptochromes* in mammals replace dmPER as the main transcriptional repressor, and replace dmTIM as a partner for mammalian PERs (Sandrelli et al., 2008). Unlike *dmCry1*, none of the mammalian *cryptochromes* are blue light photoreceptors and instead they only have a repressive role (Sandrelli et al., 2008). Mammals also lack an equivalent of *dmTim*, but they do have the paralog to *Drosophila timeout* (Sandrelli et al., 2008). However, it is not known exactly what function *timeout* plays in either *Drosophila* or mammals (Sandrelli et al., 2008).

Despite the differences between the mammalian and *Drosophila* clocks, there is substantial commonality. Most of the clock genes of mammals and *Drosophila* are homologs of each other and both clock mechanisms work via autoregulatory feedback loops. This suggests that both the *Drosophila* and mammalian clocks share a common evolutionary origin, which raises questions about the identity of an ancestral clock for both mammals and *Drosophila*. With the sequencing of more insect species, researchers were able to discover that the ancestral clock can be found in insects as some insects such as the monarch butterfly contain clock components reminiscent of both mammalian

and *Drosophila* clocks. This discovery has offered some clues as to how the clock has evolved in insects.

The Ancestral Insect Clock

Every fall, monarch butterflies (order Lepidoptera) migrate from eastern North America to their overwintering sites in central Mexico (Froy et al., 2003; Shlizerman et al., 2016). Their migratory path is nearly 4000 km, and throughout their long journey they continually correct their flight direction to maintain a southwesterly orientation (Shlizerman et al., 2016). While flying they utilize two compasses: a time-compensated sun compass controlled by the clock and an inclination-based magnetic compass (Reppert et al., 2016). The dual system of the compasses allows monarchs to utilize different environmental cues such as the sun's orientation in the sky as well as the inclination angle of the earth's magnetic field, keeping monarchs on the correct migratory trajectory (Reppert et al., 2016). In the spring, the migratory monarch butterflies return north, but because of the long distances traveled by the migrants it takes at least two subsequent generations during the spring and summer before monarch butterflies can reach numbers great enough to repopulate their historic range (Reppert et al., 2016). In addition, during the summer, monarch butterflies are considered to be non-migratory and do not display oriented flight (Reppert et al., 2016). They utilize their clock only for non-navigational purposes such as adult eclosion, sleep-wake cycles, and metabolic rhythms (Reppert et al., 2016).

Since the sun's position varies throughout the day, migrant monarchs utilize their clock to adjust their flight orientation in relation to the sun's shifting position making their sun compass time-compensated (Reppert et al., 2016). They use their time-compensated sun compass during their fall migration southward as well as during their spring migration northward (Reppert et al., 2016). Sunlight sensed by the eyes and relayed to the brain is not time-compensated by clocks in the brain, but by clocks found within the antennae (Reppert et al., 2016). This was discovered when monarchs that either had both antennae removed or painted black, could no longer display a southward orientation during flight simulations. However, when a monarch had at least one functional antenna they were able to display correct flight orientation and time-compensation (Reppert et al., 2016).

The discovery of mammalian clock components and sequencing of non-drosophilid insects have given substantial proof that the clock mechanism in most species is not *Drosophila*-like. In fact, the monarch butterfly clock appears to represent an ancestral clock as it contains not only components from the mammalian clock but also components from the *Drosophila* clock (Figure 3). Like in mammals, monarch butterfly *Bmal1* has the main transactivation domain instead of *clk* (Sandrelli et al., 2008). The monarch butterfly has two *cryptochromes*: *cry1,* a blue light photoreceptor similar to the one found in *Drosophila* and *cry2*, which functions as the main transcriptional repressor, a function similar to the mammalian *cryptochromes* (Reppert, 2007; Zhu et al, 2005). Monarch butterflies also have *tim* similar to the one found in *Drosophila* (Sandrelli et al., 2008).

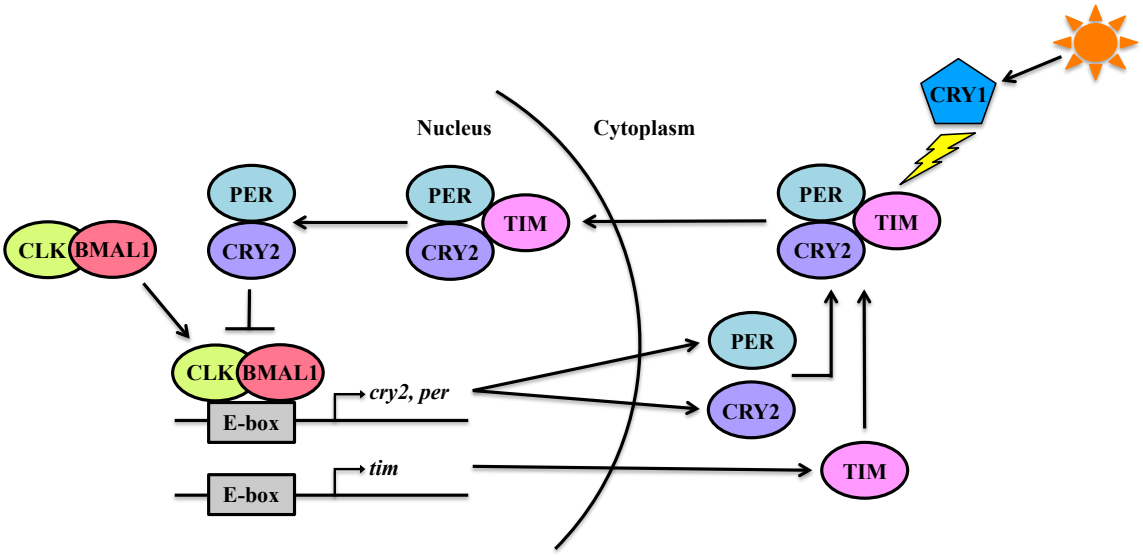

Figure 3: Ancestral insect (monarch butterfly) clock. The monarch butterfly contains not only the core *Drosophila* clock genes but it also contains mammalian components: the mammalian-like *cycle* ortholog (*Bmal1*) and a mammalian-like *cryptochrome 2* (*cry2*).

In the ancestral clock, BMAL1 and CLK form heterodimers and bind E-boxes at the *cry2*, *per*, and *tim* loci (Figure 3; Reppert et al., 2016). This in turn initiates transcription of *cry2*, *per*, and *tim* (Reppert et al., 2016). In the cytoplasm CRY2, PER, and TIM form a complex, and then translocate back into nucleus where CRY2 binds to BMAL1 and CLK (Figure 3; Reppert et al., 2016). This inhibits CLK:BMAL1 activity and E-box binding, and thus effectively inhibits the transcription of *cry2*, *per*, and *tim*. Just like in mammals and *Drosophila*, the monarch clock acts as a transcriptional feedback loop (Shlizerman et al., 2016).

Like *Drosophila*, the monarch butterfly has *cry1* which is a blue light photoreceptor involved in light entrainment in circadian oscillator cells (Zhu et al, 2008). Previous

studies have shown that similar to *Drosophila*, CRY1 and TIM in the monarch co-localize. When exposed to light CRY1 binds to TIM leading to its degradation (Figure 3; Zhu et al, 2008). It is believed that CRY1 along with CRY2 in the brain could be involved in detecting changes in photoperiod during the fall triggering their southern migration (Reppert et al., 2016). In addition, the *cryptochromes* are likely involved in other behaviors such as reproductive diapause, eclosion and metabolic rhythms (Reppert et al., 2016). The coexistence of *Drosophila* and mammalian clock components in the molecular clock mechanism of monarchs suggests that the monarch butterfly does indeed represent the ancestral clock of animals.

Further evidence of an ancestral clock can be found in the ancient bristle worm, *Platynereis dumerilii*. A representative of the subphyla Lophotrochozoa, this marine annelid has all the clock components of the monarch butterfly: *clk*, *Bmal1*, *per*, *cry2*, *tim*, and *cry1* (Zantke et al., 2013; Zantke et al., 2014). Furthermore, *P. dumerilii* contains orthologs to many genes found in vertebrates, which have been lost in invertebrates such as *Drosophila* suggesting that evolutionarily they are ancestral to most species (Zantke et al., 2014). Besides containing all the ancestral clock components, *P. dumerilii* clock genes do not behave similar to those in the monarch. mRNA levels of *clk* and *Bmal1* cycle in phase with each other, *cry1* mRNA is constitutive, and *per* cycles anti-phase to *clk* (Zantke et al., 2013). In *P. dumerillii*, *cry2* and *tim* peak either in the morning or night suggesting that mRNA levels for *cry2* and *tim* are controlled by light (Zantke et a., 2013). In addition, *P. dumerilii cry1* was shown to act as a light receptor and *cry2* was

shown to be a transcriptional repressor (Zantke et al., 2013). Thus, this annelid shows us that the ancestral clock can also be found in much older species besides the monarch butterfly.

Clocks in other Insects (Order Lepidoptera)

Besides the monarch butterfly other insects have provided valuable insight into the diversity of insect clocks. In the Chinese oak silk moth, *Antheraea pernyi*, the clock genes *clk*, *Bmal1*, *tim*, *cry1*, *cry2*, and *per* were cloned (Chang et al., 2003; Yuan et al., 2007). It was discovered that when the C-terminus of *Bmal1* was truncated there was no transcriptional activation in S2 cells (Chang et al., 2003). However, when the C-terminus of *clk* was truncated there was no effect on transcriptional activation if a full-length *Bmal1* was co-transfected (Chang et al., 2003). That the main transactivation domain of vertebrates and most insects is found in the C-terminus of *Bmal1* rather than in the C-terminus of *clk* as is seen in *Drosophila*, suggests that the ancestral *Bmal1* must have also had the transactivation domain (Chang et al., 2003). Further, *Drosophila Clk* must have acquired the transactivation domain in a separate event, most likely with the expansion of its poly-Q region (Chang et al., 2003).

In *Drosophila*, PER acts as the main transcriptional repressor (Chang et al., 2003). Both PER and TIM co-localize in the cytoplasm during the afternoon and in the nucleus at night (Chang et al., 2003). When silk moth PER was co-transfected with CLK and BMAL1 in S2 cells, PER was able to repress CLK and BMAL1 transactivation and PER

localized to the nucleus in S2 cells (Chang et al., 2003). However, in the pacemaker neurons of the silk moth, PER is localized in the cytoplasm all day (Chang et al., 2003; Sauman et al., 1996). Silk moth PER was also expressed in arrhythmic *Drosophila per*^{01} mutants and was able to rescue their locomotor activity but shortened the mutant's period (Levine et al., 1995). It would seem then that silk moth PER behaves as a transcriptional repressor similar to *Drosophila* PER (Chang et al., 2003). In addition, CRY2 from the silk moth acts as a transcriptional repressor in S2 cells and like in other insects, silk moth CRY1 had no repressive activity but was degraded by light (Yuan et al., 2007). When silk moth TIM was co-transfected with CLK and BMAL1 in S2 cells it was not able to repress on its own, however when co-transfected with PER it was able to increase the repressive activity of PER (Chang et al., 2003). Like silk moth PER, TIM localized to the nucleus in S2 cells but was strictly cytoplasmic *in vivo* (Chang et al., 2003).

The Egyptian cotton leafworm, *Spodoptera littoralis*, is a highly destructive pest of many economically important crops in Africa, Asia, and Europe (Khedher et al., 2017). It was discovered that not only does *S. littoralis* have a brain pacemaker, but its antennae have a peripheral clock as well (Merlin et al., 2007). This peripheral clock is responsible for circadian rhythms in female calling behavior and pheromone emission as well as male pheromone responsiveness (Merlin et al., 2007). It was also found that both *cry2* and *per* in *S. littoralis* cycle in phase, peaking at the beginning of the day (Zeitgeber time, ZT0) (Merlin et al., 2007). However, *cry1* cycles out of phase with both *cry2* and

per, peaking at the middle of the day (ZT6) (Merlin et al., 2007). Though *cry1*, *cry2*, and *per* do not all cycle in phase together they do cycle in a circadian manner in both the brain and the antennae of *S. littoralis* (Merlin et al., 2007). The cycling of *cry2* in *S. littoralis* is similar to how *cry2* cycles in the monarch butterfly suggesting that *cry2* in *S. littoralis* is also a transcriptional repressor (Reppert, 2007).

Clocks in other Insects (Order Hymenoptera)

The honeybee, *Apis mellifera*, is an insect species well known for its circadian controlled behaviors like time-compensated sun compass navigation, division of labor, and time memory (Rubin et al., 2006). Interestingly, unlike most insects but similar to humans, honeybees emerge from the pupae essentially "clockless" with neither rhythmic behavior nor clock gene expression occurring until later in life (Eban-Rothschild et al., 2012). Using bioinformatic analysis it was discovered that the adult honeybee has the clock components *clk*, *Bmal1*, *cry2*, and *per* (Rubin et al., 2006). However, in the honeybee, *tim* and *cry1* have been lost, thus it is not known how the honeybee clock is entrained by light (Rubin et al., 2006; Yuan et al., 2007). The main transactivation domain in honeybees is located in the C-terminus of BMAL1 as is seen in the monarch butterfly and the silk moth (Rubin et al., 2006). As in monarchs, honeybee CRY2 acts as the main transcriptional repressor, inhibiting its own transcription by binding to BMAL1 and CLK (Reppert, 2007).

It is well known that the clock plays an important role in the division of labor as well as the social organization of insects like bees and ants (Ingram et al., 2012). The fire ant, *Solenopsis invicta*, is an invasive pest whose clock shares similarities with the honeybee (Ingram et al., 2012). Analysis of its genome revealed that like the honeybee, fire ants have *clk*, *Bmal1*, *period*, and *cry2*, but lack *cry1* and *tim* (Ingram et al., 2012). In fire ants, mRNA cycling of *per* and *cry2* peak during the night and *Bmal1* cycles anti-phase to *per* (Ingram et al., 2012). Although not discussed with honeybees, fire ants do have *cwo*, which has a conserved C-terminal domain found only in other insects but not the mammalian ortholog *dec2* (Ingram et al., 2012). Thus, it would appear that although the honeybee and fire ant clocks are more similar to the ancestral monarch butterfly clock the loss of *cry1* and *tim* are another example of how the clock has evolved in different ways.

Clocks in other Insects (Order Coleoptera)

The red flour beetle, *Tribolium castaneum* is found worldwide and is a major pest of stored-grains (Cato et al., 2017). Bioinformatic analysis has shown that the red flour beetle has *Bmal1*, *clk*, *cry2*, *per*, and *tim* but has lost *cry1* (Yuan et al, 2007; Li et al., 2017). Like the monarch butterfly, when *T. castaneum* CRY2 is expressed in S2 cells it is able to potently repress CLK and BMAL1 and has no light sensitivity (Yuan et al., 2007). It has also been shown that *T. castaneum*, *tim* is necessary for eclosion and embryonic development (Li et al., 2017). The peak of *tim* mRNA in *T. castaneum* occurs during the middle of the day (ZT4), and appears to be less sensitive to light probably due

to the loss of *cry1* (Li et al., 2017). *T. castaneum Bmal1* and *per* cycle with two peaks at ZT4 and ZT12 and *clk* peaks at ZT12 (Li et al., 2017). Thus, the red flour beetle clock provides further evidence for yet another form of the clock mechanism in insects.

Organisms living in caves are forced to live life without sunlight and exhibit adaptions such as decreased eye size and reduced body pigmentation (Friedrich et al., 2011). It would be assumed that cave dwellers would eventually lose their clocks, however there is clear evidence that this is not the case (Friedrich, 2013). The cave dwelling beetle, *Ptomaphagus hirtus*, has a fully functional clock with all the core clock components and it has been postulated that it can be weakly entrained by light despite having a highly reduced visual system (Friedrich et al., 2011). Like the red flour beetle, *P. hirtus* has *Bmal1*, *clk*, *per*, *tim*, and *cry2* but has lost *cry1* (Friedrich et al., 2011). How *Ptomaphagus* may be entrained by light without *cry1* is unknown, but it may be similar to the red flour beetle (Friedrich et al., 2011).

Clocks in other Insects (Order Hemiptera)

The bean bug, *Riptortus pedestris*, is a major pest of soybean and is one of the few hemimetabolous insects to have its circadian clock examined (Ikeno et al., 2008). So far all the insects discussed have been holometabolous, which undergo complete metamorphosis, whereas the bean bug has incomplete metamorphosis. The clock of the bean bug has *clk*, *Bmal1*, *per*, and *cry2* (Ikeno et al., 2008). Due to limited sequence information it is unsure whether *R. pedestris* has *cry1* or *tim* (Ikeno et al., 2008). The

amino acid sequences for PER, BMAL1 and CRY2 in the bean bug appear to be conserved among insects, and domains important for a functional *Drosophila* clock have also remained well conserved in the bean bug (Ikeno et al., 2008). In *R. pedestris*, mRNA levels for *clk*, *per*, *Bmal1*, and *cry2* appear to remain constant or show weak rhythms (Ikeno et al., 2008; Ikeno et al., 2013). However, the presence of *cry2* suggests that the bean bug's clock is more closely related to the ancestral clock than to the *Drosophila* clock (Ikeno et al., 2008). In fact, it has been shown that *R. pedestris* CRY2 along with PER both act as transcriptional repressors of CLK and BMAL1 (Ikeno et al., 2011). As more and more non-drosophilid insect genomes are sequenced it is possible to assume that other variations of the ancestral clock in addition to the ones already described will show up among the insects. And with this information the evolution of the clock in insects will be revealed, making known how the clock evolved to meet the needs of insects in different niches.

Clocks in other Insects (Order Diptera)

Diptera, also known as "true flies" are one of the most diverse orders of Insecta, accounting for 1 in 10 species on earth (Wiegmann et al., 2011; Zhang et al., 2017). The origin of Diptera and their relative emergence has remained elusive, but they were first noted in the fossil records from the Miocene, Oligocene, and Eocene (Krzeminski, 1992). Although many true flies are pests and parasites, a great deal more are important pollinators (Wiegmann et al., 2011; Zhang et al., 2017). In fact, they are considered to be some of the most ancient pollinators of angiosperms and are believed to have "played an

important role in the origin of co-evolutionary relationships with flowering plants and insects" (Zhang et al., 2017). Despite difficulties in determining the evolution of Dipterans due to "conflicting anatomical and genetic evidence" true flies are generally divided into two major groups: lower Diptera (mosquitoes) and Brachycera (*Drosophila* and houseflies) (Wiegmann et al., 2011).

The southern house mosquito, *Culex quinquefasciatus*, is the main vector for West Nile encephalitis, St. Louis encephalitis, and the nematode that causes lymphatic filariasis, better known as elephantiasis (Meireles-Filho et al., 2013). The transmission of the filarial worm to human hosts by *C. quinquefasciatus* is "one of the best examples of coevolution of circadian rhythms of parasitism… the worms reach their highest density in human peripheral blood at the time when the vector shows its biting activity peak, increasing their chances of transmission into the insect" (Meireles-Filho et al., 2013). Its clock is most similar to the monarch butterfly, and has the clock genes *clk*, *Bmal1*, *per*, *cry1*, *tim*, and *cry2* (Meireles-Filho et al., 2013). Like other insects, *C. quinquefasciatus* CRY1 is responsible for photoentrainment of the clock and CRY2 acts as the main transcriptional repressor (Meireles-Filho et al., 2013). In *C. quinquefasciatus*, *clk* mRNA levels peak around ZT22-ZT2, which is similar to the phase of *Drosophila Clk* mRNA (Meireles-Filho et al., 2013). Unlike *Drosophila*, *cqBmal1* is rhythmic and peaks around ZT1-5 (Meireles-Filho et al., 2013). In *C. quinquefasciatus*, *cry1* expression is constitutive which is different from *Drosophila cry1*, which peaks at ZT5 (Meireles-Filho et al., 2013). And *cry2* in *C. quinquefasciatus* is rhythmic and peaks around ZT15-

19 (Meireles-Filho et al., 2013). Although, *C. quinquefasciatus* is a close relative to *Drosophila*, this species has a clock that resembles the ancestral clock. This suggests that somewhere in the dipteran lineage, there was a loss of *cry2* and a truncation of BMAL1 leading to PER taking over as the main transcriptional repressor and CLK as the main transcriptional activator.

Seemingly harmless and found nearly everywhere on the planet, houseflies are actually carriers of over 100 human and animal diseases (Scott et al., 2009). They are passive vectors not only for bacterial disease agents such as typhoid fever, salmonella, tuberculosis, and cholera but also for the protozoan parasites causing amebic dysentery, and parasitic worms such as tapeworms and roundworms (Scott et al., 2009). *M. domestica* is well known for transmiting several ocular diseases, most importantly trachoma, which is responsible for six million annual cases of childhood blindness (Scott et al., 2009). They also infect open wounds with bacteria, causing yaws and leprosy (Scott et al., 2009). However, this shouldn't come as a surprise since houseflies come in contact with garbage, carcasses, sewage, our food supply and our homes on a daily basis (Scott et al., 2009). But houseflies don't only impact the lives of humans, their infestation of dairy, poultry, and hog facilities not only expose these animals to diseases but also lowers egg and milk production causing large economic losses (Scott et al., 2009). Because houseflies spread such dangerous diseases and cause major economic losses each year there was a push to have their genome sequenced. This sequence has in

turn, allowed us to use them as our model to understand the evolution of the clock in Diptera.

The housefly, *Musca domestica*, has a clock, with both similarities and notable differences from *Drosophila*. Like *Drosophila*, the housefly has the core clock components *Clk*, *cyc*, *per*, *tim*, and *cry1* (Codd et al., 2007). The mRNAs for the housefly clock components *(per*, *tim*, and *Clk*) cycle in phase with *Drosophila per*, *tim*, and *Clk* (Codd et al., 2007). However, housefly *cry1* mRNA does not robustly cycle like *cry1* in *Drosophila* (Codd et al., 2007). Locomotor activity rhythms of *M. domestica* in 12:12 hour LD conditions showed rhythmic behavior but their activity was restricted almost exclusively to the day unlike the locomotor behavior of *Drosophila* (Codd et al., 2007). Their average period in LD conditions was around 24 hours similar to *Drosophila* and other organisms (Codd et al., 2007). The cells responsible for controlling locomotor activity in *M. domestica* have been shown to be located outside the optic lobe suggesting that the localization of clock cells in their brains maybe similar to *Drosophila* (Balys et al., 2001).

Western blots done on *M. domestica* heads entrained in LD conditions showed that although mdPER does not appear to cycle, mdTIM does cycle with a peak early in the morning (Codd et al., 2007). This finding is in contrast with *Drosophila* where both TIM and PER cycle robustly (Codd et al., 2007). In addition, mdTIM is degraded in response to light, similar to TIM in *Drosophila*, but the degradation is immediate suggesting that

mdPER is not stabilized by mdTIM (Codd et al., 2007). With the sequencing of the housefly genome in 2014, we were able to determine that the C-terminus of *Clk* has the transactivation domain and its *cyc* is more similar in sequence to *Drosophila cyc* than *Bmal1* in the monarch butterfly (Scott et al., 2014). When *mdPer* was cloned, it was discovered that although there are conserved regions between *dmPer* and *mdPer*, the structure of the *Musca* gene is different due to the increased length and number of introns (Piccin et al., 2000). However, Piccin et al., 2000 also found that circadian locomotor rhythms could be rescued in *Drosophila* per^{01} mutants when these flies expressed a single copy of housefly *per*. This suggested that mdPER could be used interchangeably with dmPER. With the knowledge of the differences and similarities between the *M. domestica* and *Drosophila* clocks, and the fact that they are such close relatives, I investigated their clock mechanism further in the hopes of uncovering clues to the evolution of the clock in dipterans.

Objectives

The clock mechanisms in insects all function as auto-regulatory feedback loops, but the components and inner workings of the clock vary between species. There are at least four known versions of the clock in insects: the *Drosophila* clock, the ancestral (monarch butterfly) clock, the beetle clock, and the honeybee clock (Figure 4). The *Drosophila* clock, which has served as the model for all insect clocks until recently has the clock components *Clk*, *cyc*, *per*, *tim*, and *cry1* (Figure 4). The ancestral monarch clock has the *Drosophila* clock components *clk*, *per*, *tim*, and *cry1* and the mammalian-

like clock genes *Bmal1* and *cry2* (Figure 4). The beetle clock is most similar to the ancestral monarch clock but has lost *cry1* (Figure 4). And the honeybee clock is also similar to the ancestral clock but has lost both *cry1* and *tim* (Figure 4).

Insects within the order Diptera are "divided into two major groups: the lower Diptera… mosquito-like flies with long antennae, and Brachycera, stout and fast-moving flies with short antennae" (Wiegmann et al., 2011). Based on sequence analysis, flies in lower Diptera have an ancestral clock, while flies in Brachycera have a *Drosophila* clock. Somewhere in between the ancestral clock has diverged into the *Drosophila* clock. I propose to determine how the clock evolved within the dipteran lineage by using the housefly. Its clock mechanism is not fully understood, but differs from the *Drosophila* clock, and could provide clues into how the clock diverged in dipterans (Figure 4).

The following hypotheses will be tested:

- Hypothesis 1: The housefly has a hybrid clock, which has both *Drosophila* and ancestral clock components. To test this hypothesis, I first cloned the housefly clock components *Clk*, *cyc*, *per*, and a possible *cry2* ortholog. I then transfected them into S2 cells to test whether they activated or repressed transcription and found that neither mdPER nor mdCRY2 were able to repress mdCLK:mdCYC. Further sequence analysis and testing for (6-4) DNA photolyase activity showed that what we thought was *mdCry2* was actually a (6-4) DNA photolyase.

- Hypothesis 2: Since houseflies lack *cry2* and mdPER has no repressive role, mdCWO acts as the main transcriptional repressor. To test this hypothesis, I first cloned housefly *cwo*. I then transfected *mdCwo* along with *mdClk* and *mdCyc* into S2 cells to test whether mdCWO could repress mdCLK:mdCYC. I found that mdCWO acts as a potent repressor in S2 cells, however when I co-transfected *mdCwo* and *mdPer*, there was no synergistic effect, despite what was been shown previously with *dmCwo* and *dmPer*.

My thesis will be broken down into 3 parts: the materials and methods (Chapter II) which I used to conduct this research, the results (Chapter III) which describes how the clock functions in the housefly, and the discussion (Chapter IV) which discusses the major findings of this research, the significance of those findings, and future directions for this work.

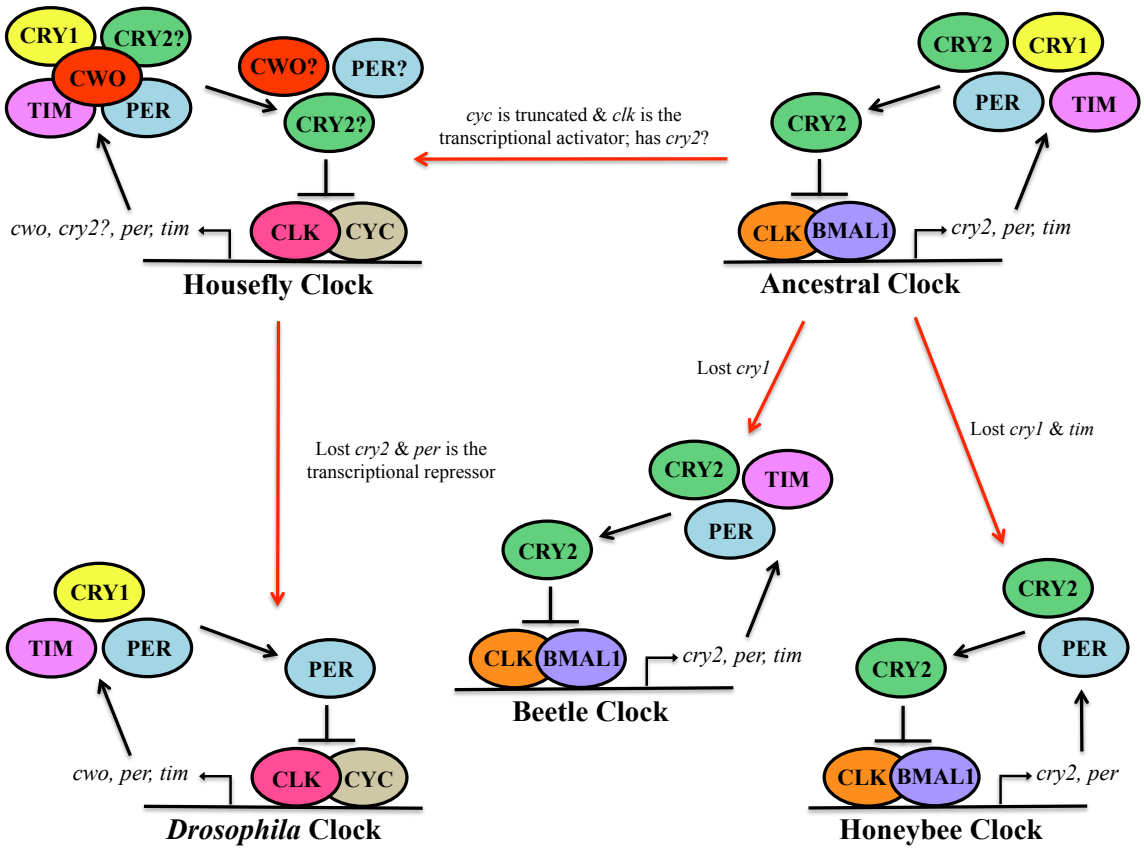

Figure 4: Clock variation in insects. Based off already published data and sequence information there appears to be four known variations of the clock in insects. There is the ancestral clock (mammalian and *Drosophila* clock components), the honeybee clock (lost *cry1* and *tim*), the beetle clock (lost *cry1*), and the *Drosophila* clock. The housefly clock has a truncated *cyc* and *Clk* is the transcriptional activator instead of *Bmal1*, and may have *cry2*, thus perhaps the *Drosophila* clock diverged from the housefly clock.

CHAPTER II

MATERIALS AND METHODS

Fly Maintenance

Houseflies were reared in a BugDorm observation cage (#1452, BioQuip) and were maintained at 25ºC under 12:12 hour LD conditions. We received wild-type *M. domestica* from Jeffery Tomerlin (Texas A&M University) and an inbred *aabys* strain from Rich Meisel (University of Houston), which were both used for cloning. *M. domestica* embryos and larvae were raised on a medium of wheat bran (80 g), alfalfa mill (48 g), and corn meal (32 g) mixed with 240 mL of water. Embryos and larvae were kept in an open container to allow for airflow, which would eventually dry the food out, a condition necessary for them to pupate (Schoof, 1964). They were allowed to grow on the media for around 10 days until they reached adulthood.

Adult houseflies were raised on a mixture of 1:1 sugar and nonfat instant dry milk. Adults were also given honey as this seemed to encourage them to mate. A quarter-size amount of honey was placed in a petri dish and then a kimwipe was folded twice and placed on top of the honey, which allowed the houseflies to eat the honey through the kimwipe without getting stuck. Adults got water from a plastic flask, which had a paper towel stuffed inside that was secured at the rim of the flask with a rubber band.

Once the adults emerged, they were allowed to breed for approximately 5 days then their embryos were collected. To collect the embryos, two kimwipes were wrapped around a small ball of larval food and placed in a small plastic cup in their cage. Every 4 hours, the ball of larval food was taken out and the embryos were rinsed into a beaker with distilled water. Fertilized embryos sink and non-fertilized embryos float, so the fertilized embryos were collected and placed on fresh larval medium.

Cloning and Sequence Analysis

My collaborator, Christine Merlin (Texas A&M University) generously provided the following constructs for our transcriptional assays: *dpPer:Luc*, *Ren:Luc*, *dpClk*, *dpBmal1*, *dpCry2*, *dpPer*, *dmClk*, *dmPer* and the pAc5.1V5/HisA vector. I received *mdPer* as a gift from Charalambos Kyriacou (University of Leicester). I generated S2 cell expression constructs by subcloning *mdClk*, *mdCyc*, *mdPer*, *mdCry2*, and *mdCwo* into the pAc5.1V5/HisA vector. cDNA templates for PCR were prepared from RNA purified from *Musca domestica* heads. To generate the cDNA, I used SuperScript II Reverse Transcriptase (#18064014, Thermo Fisher Scientific) and Oligo$(dT)_{12\text{-}18}$ primer (#18418012, Thermo Fisher Scientific).

Primers used to amplify *mdClk*, *mdCyc*, *mdCry2*, *mdPer*, and *mdCwo* are listed in Table 1. PCR was performed using Q5 Hot Start High-Fidelity DNA Polymerase (#M0493S, NEB) to reduce the number of PCR induced mutations. Clones were verified by restriction digest and sequencing. The Gene Technologies Laboratory at Texas A&M

University sequenced the clones. The *mdClk* and *mdPer* cDNAs proved difficult to clone, perhaps due to their large size. For the first section of my results (Does the housefly have a hybrid clock?), the expression constructs for *mdClk* and *mdPer* were generated from cDNA and gDNA fragments. However, for the second section of my results (What is the main transcriptional repressor in the housefly?), I was able to generate full-length cDNA clones for both *mdClk* and *mdPer.* I did so by transfecting *mdClk* and *mdPer* into S2 cells, and then after ~48 hours I harvested the cells and extracted the RNA. After cDNA synthesis, I was able to amplify the full-length cDNA for *mdClk* and *mdPer* by PCR. I then subcloned *mdClk* and *mdPer* into the pAc5.1V5/HisA vector. The clones were verified as previously described.

Insect Cell Culture and Transfections

S2 cells were maintained at 25ºC in Schneider's *Drosophila* medium (#21720-024, Thermo Fisher Scientific) with 10% heat-inactivated fetal bovine serum (FBS; #16140071, Thermo Fisher Scientific) and 1% Penicillin-Streptomycin (Pen-Strep; #15140122, Thermo Fisher Scientific). All transfections were done with S2 cells that were split no more than two days prior to the experiment. Transfections took place in 12-well tissue culture plates; when we used 24-well plates there was large variation between our replicates so we refrained from using them. Prior to transfection, S2 cells were spun down in 15 mL Falcon tubes (3,500 rpm for 3.5 minutes), the media was removed, the cells were re-suspended in fresh media and added to the 12-well plates. They were allowed to grow until they reached the desired density (8×10^5 cells/well). We did three

biological replicates per experiment and repeated each experiment at least twice. The protocol for transient transfection of plasmids is as follows: add 150 µL of serum free Schneider's *Drosophila* medium (SFM) to 200 ng of DNA. In another tube, SFM (40 µL x the number of wells used) was mixed with Cellfectin II Reagent (5 µL x the number of wells used; #10362100, Thermo Fisher Scientific). Then 135 uL of the SFM/Cellfectin mixture was added to the DNA, mixed well by pipetting, and left at room temperature for 15 minutes. The DNA/SFM/Cellfectin was then added to 1.2 mL of SFM, and pipetted up and down at least five times. The media from the cells was aspirated and ~495 µL of the DNA/SFM/Cellfectin mixture was added to each well. Cells were incubated for 4-6 hours at 25ºC and then 1 mL of Schneider's *Drosophila* medium with 10% FBS and 1% Pen-Strep was added to each well. Transfected cells were incubated at 25ºC for ~48 hours before collection.

Transcription Assays

To test mdCLK:mdCYC, dpCLK:dpBMAL1, and dmCLK:dmCYC transcriptional activation in S2 cells, I co-transfected 5 ng of *Clk* (*mdClk*, *dpClk*, or *dmClk*) and 5 ng of *Bmal1/cyc* (*mdCyc*, *dpBmal1*, or *dmCyc*) with 30 ng of *Ren:Luc* and 10 ng of *dpPer:Luc* (Figure 5). To measure the inhibitory activity of *mdCry2*, *mdPer*, *mdCwo*, *dpCry2*, *dpPer*, *dmPer*, and *dmCwo*, I co-transfected 2 ng-150 ng of *mdCry2*, 2 ng-200 ng of *mdPer*, 100 pg- 100 ng of *mdCwo*, 2 ng-150 ng of *dpCry2*, 2 ng-50 ng of *dpPer*, 2 ng-50 ng of *dmPer*, and/or 100 pg-100 ng of *dmCwo* (Figure 5). The total amount of DNA in each transfection was normalized to 200 ng using the empty vector pAc5.1V5/HisA.

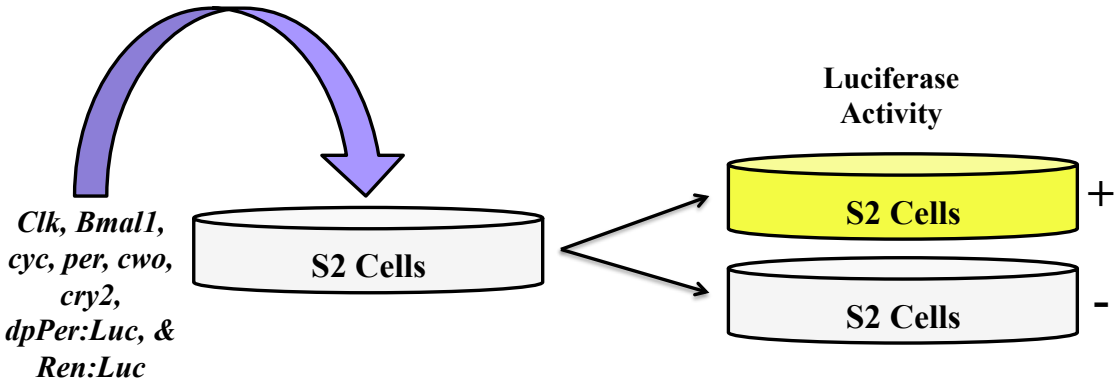

Figure 5: Measuring luciferase activity in S2 cells as an assay for clock function. *Drosophila*, housefly, and monarch butterfly clock genes, the monarch *per*-promoter luciferase vector, and the *Renilla* luciferase vector are co-transfected into S2 cells and the resulting luciferase activity is measured.

After transfection and a 48-hour incubation at 25ºC, the Dual-Luciferase Reporter Assay System (#E1910, Promega) was used to measure luciferase and *Renilla* luciferase activity. The Dual-Luciferase Reporter Assay protocol used is as follows: the media was aspirated from the cells and the cells were washed with 500 µL of phosphate-buffered saline (1X PBS). Cells were then lysed by adding 50 µL of Passive Lysis Buffer and incubated for 15 minutes at room temperature on a shaker set at a low speed. Lysed cells were collected and spun at 4ºC for 5 minutes at 12,000 rpm and 10 µL of each cell extract was added to a 96-well black assay plate (#33-755, Genesee). Luciferase activity was measured on a 96-well plate-reading luminometer with an auto-injector system. The luminometer was programmed to dispense 50 µL of Luciferase Assay Reagent and 50 µL of Stop & Glo into each well. Average ratios of luciferase activity to *Renilla* luciferase activity were calculated and then the ratios were normalized so that the relative activation alone equaled 100%.

Testing for (6-4) DNA photolyase Activity

Photolyase activity was tested according to Daiyasu et al., 2004 and Kobayashi et al., 2000. We first transformed mdCRY2 and two control plasmids: pGEX4T-1-zCRY1B(f) and pGEX4T-3_z6-4phr.prj into SY32 (pRT2) cells (all gifts from Takeshi Todo, Kyoto University). The SY32 (pRT2) cells lack (6-4) DNA photolyase activity. The pGEX4T-1-zCRY1B(f) plasmid was our negative control since it is a known *cryptochrome* in zebrafish (*Danio rerio*). The pGEX4T-3_z6-4phr.prj plasmid was our positive control since it is a known (6-4) DNA photolyase in zebrafish. Transformed cells were exposed to UV light for 30 minutes, and then put under fluorescent light for 1 hour. The cells were grown overnight at 37ºC, and the next day colonies were counted.

Transcript	Primer	Primer Sequence	Product size (bp) (cDNA)
Clock	clk fwd	5'-ATTCTC**GCGGCCGC**ATGTTTTTATTGTTTTTCAGGCGATC	3357
	clk rev	5'-ATTATT**TCTAGA**TGGATTTGTTCCCCCACTTTGC	
cycle	cyc fwd	5'-ATTATA**GCGGCCGC**ATGGAGTATTGCGACAATCTACCGG	1233
	cyc rev	5'- CTTATA**TCTAGA**GAAAAATACCGAATTCTTAGCCATAAT	
period	per fwd	5'-ATTCTC**GCGGCCGC**ATGGAAGGTGAATCTACGGAATCAACA	3144
	per rev	5'-ATTATT**TCTAGA**CAGTCCACCGCCGTGTTGTGTTT	
cryptochrome 2	cry2 fwd	5'-ATTATA**GCGGCCGC**ATGAGTGAGAAGTCAACTCTAATCCAT	1620
	cry2 rev	5'-ATTATT**TCTAGA**TTTCTTCTTTTTCTTTGCTTTCGAGCC	
clockwork orange	cwo fwd	5'-TTATTA**GATATC**ATGGAACCGTCGTATTGGGG	2034
	cwo rev	5'-TTATTT**TCTAGA**CCAGCCATTGGTATTGGAGATAGCT	

Table 1: Primers used in cloning *M. domestica* genes. A list of primers used to clone the housefly clock components *clock*, *cycle*, *period*, *cryptochrome 2*, and *clockwork orange*. Primers include restriction sites used in cloning as well as the addition of ~6 base pairs to allow for efficient cutting by restriction enzymes. Restriction sites are in bold.

CHAPTER III

RESULTS

Does the housefly have a hybrid clock?

Insects from orders outside Diptera, such as Lepidoptera (monarch butterfly), Hymenoptera (honeybee), Coleoptera (red flour beetle), and Hemiptera (bean bug) have clocks with mammalian components, suggesting that the *Drosophila* clock does not represent all insect clocks (Reppert, 2007). The first step was to look at the order Diptera and establish which fly species were most closely related to *Drosophila* and determine which have an ancestral and/or *Drosophila*-like clock in order to uncover where in the dipteran lineage the *Drosophila* clock diverged. I used the dipteran tree published in Wiegmann et al., 2011 as a guide for my analysis.

Based on published dipteran sequences, I selected species for analysis that had the most complete sequencing information on the core clock genes. I used *Drosophila* clock protein sequences from FlyBase, mouse (*Mus musculus*) clock protein sequences from NCBI, and monarch butterfly clock protein sequences from MonarchBase to BLAST against different dipteran species. From this analysis I identified four potential candidates: the malaria mosquito (*Anopheles gambiae*), the melon fly (*Bactrocera cucurbitae*), the housefly (*Musca domestica*), and the tsetse fly (*Glossina pallidipes*). In order to check for full-length protein sequences and similarity between species, I aligned

the sequences from each of the species for each clock protein using the multiple sequence alignment program Clustal Omega.

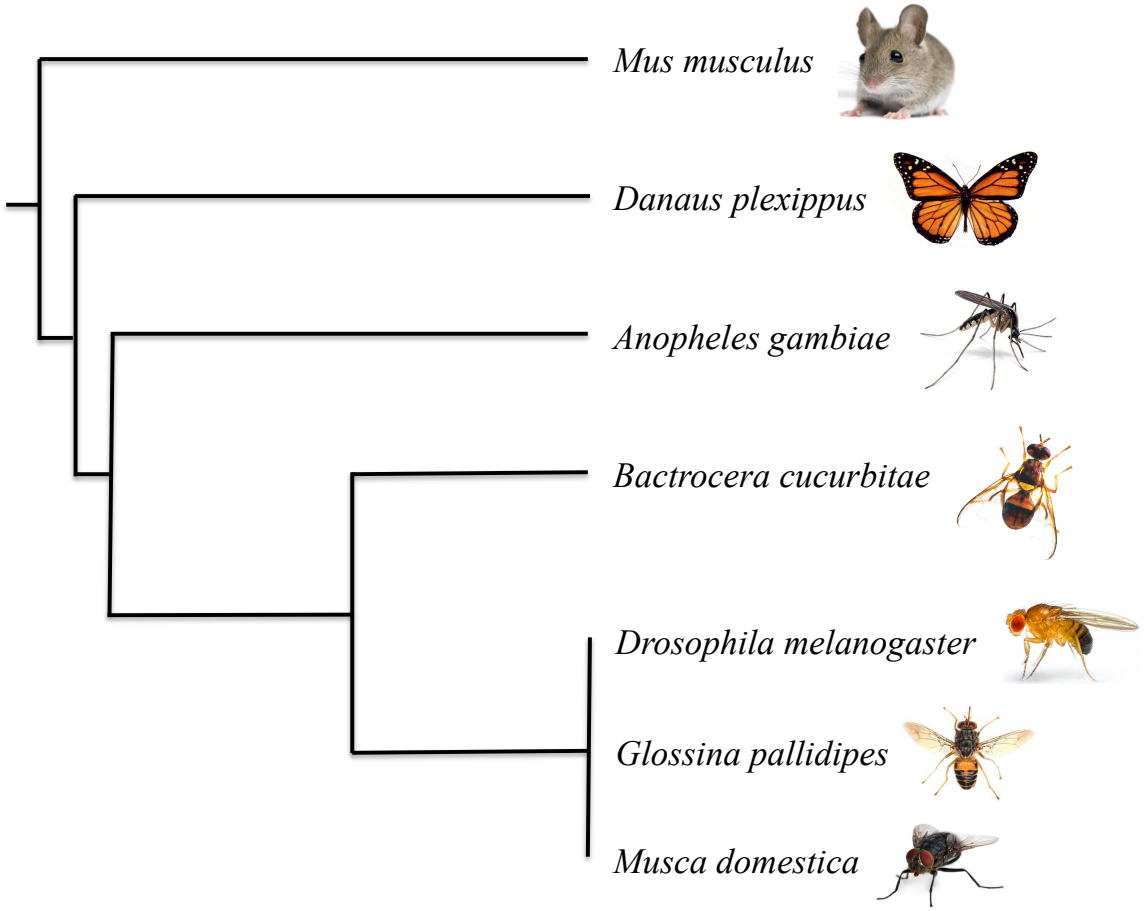

Figure 6: A cladogram of phylogenetic relationships. The mouse (*Mus musculus*), has a mammalian clock and is the outgroup. The monarch butterfly is the most distant insect to *Drosophila* to have an ancestral clock. The dipteran, *Anopheles gambiae* is the closest insect to *Drosophila* to have an ancestral clock. The dipterans, *Bactrocera cucurbitae*, *Musca domestica*, and *Glossina pallidipes* have the clock components *Clk*, *cyc*, *per*, *tim*, *cry1*, and may also have *cry2*.

I found that *Anopheles gambiae* is the most distantly related dipteran to *Drosophila* to have an ancestral clock (Figure 6). Like the monarch butterfly it has a *Drosophila cry1*, but also has a mammalian *Bmal1* and *cry2* (Table 2). The melon fly, the housefly, and

the tsetse fly are the most closely related dipterans to *Drosophila*, and may have an intermediate clock (Figure 6). All three species have a *Drosophila cyc* and *cry1* but what was interesting is they also have what appeared to be a mammalian-like *cry2* (Table 2). For the purpose of this study, however, I am focusing on the housefly due to its availability, fully sequenced genome, and previous clock studies done on this organism.

Species	CLOCK	PERIOD	BMAL1	CYCLE	CRY1	CRY2
Mus musculus	+	+	+	-	+	+
Danaus plexippus	+	+	+	-	+	+
Anopheles gambiae	+	+	+	-	+	+
Drosophila melanogaster	+	+	-	+	+	-
Bactrocera cucurbitae	+	+	-	+	+	?
Musca domestica	+	+	-	+	+	?
Glossina pallidipes	+	+	-	+	+	?

Table 2: Clock components of different species. The mouse (*Mus musculus*) has the clock components *clk*, *Bmal1*, *per*, and mammalian *cry* (*cry2*). *Drosophila* has the clock components *Clk*, *cyc*, *per*, and *Drosophila cry* (*cry1*). The monarch butterfly and *Anopheles gambiae* have the *Drosophila* clock components *clk*, *per*, and *cry1* as well as the mammalian clock components *Bmal* and *cry2*. The melon fly (*Bactrocera cucurbitae*), the housefly (*Musca domestica*), and the tsetse fly (*Glossina pallidipes*) have the *Drosophila* clock components *Clk*, *cyc*, *per*, and *cry1* but may also have *cry2*.

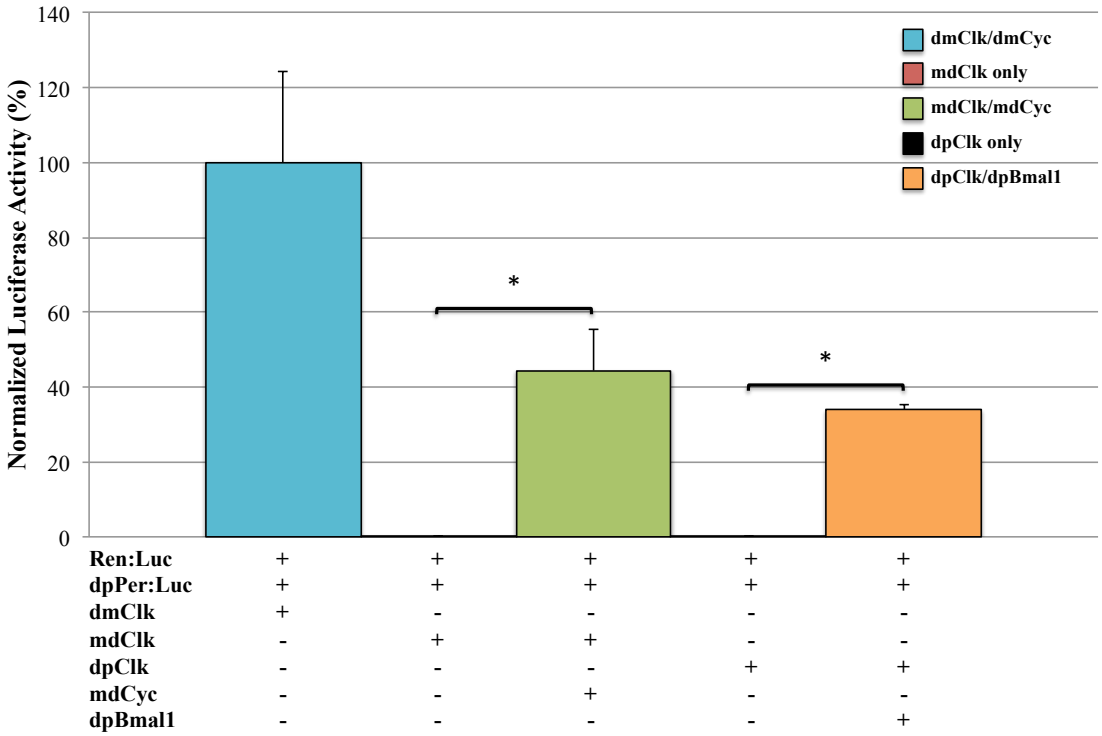

Figure 7: *mdClk* requires *mdCyc* for activation. *Clk* from the monarch and housefly cannot activate with *dmCyc*. They require their counterparts: *dpBmal1* and *mdCyc* for activation. The *mdClk* construct used was generated from cDNA and gDNA fragments. *p<0.05, Student's t-test.

S2 cells are a useful system to test how the housefly clock functions, because they do not express most *Drosophila* clock genes except *dmCyc* (Chang et al., 2003).

Since S2 cells express *cyc*, we first had to test whether mdCLK:dmCYC would mediate transcription and whether it was necessary for us to eliminate or knock down *dmCyc*. Previous work showed that dpCLK:dmCYC does not mediate transcription; only *dpClk* and *dpBmal1* activate transcription upon co-transfection (Yuan et al., 2007). Our *mdClk* construct was generated with both cDNA and gDNA fragments, which could affect its ability to activate if transcripts were not properly spliced. However, mdCLK:dmCYC

does not mediate transcription except when *mdClk* is transfected along with *mdCyc* (Figure 7).

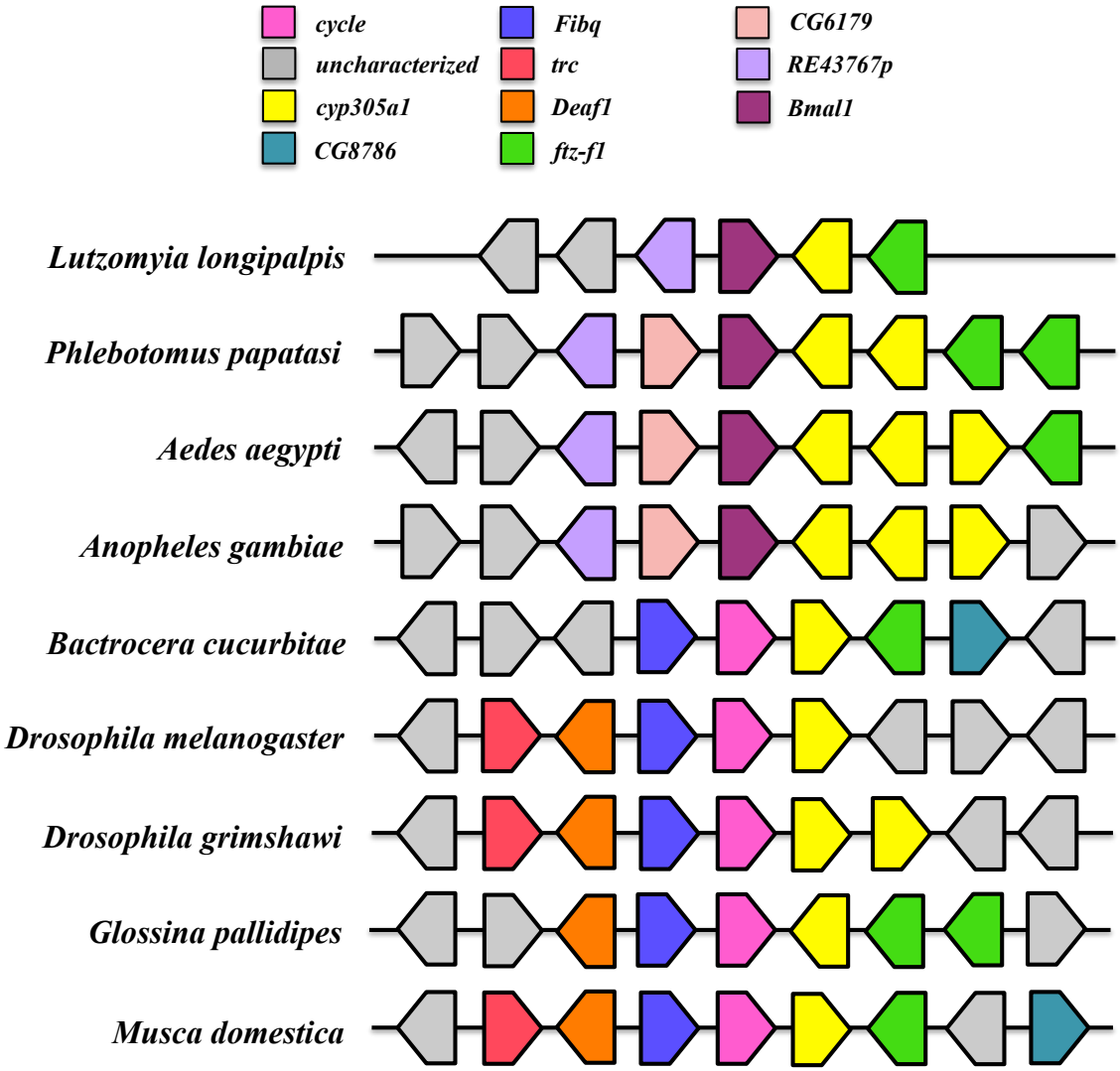

Figure 8: A synteny map of *Bmal1* and *Cycle*. Based on sequence analysis lower dipterans (*Lutzomyia longipalpis*, *Phlebotomus papatasi*, *Aedes aegypti*, and *Anopheles gambiae*) have *Bmal1*, and flies in Brachycera (*Bactrocera cucurbitae*, *Drosophila melanogaster*, *Drosophila grimshawi*, *Glossina pallidipes*, and *Musca domestica*) have *cycle*. Although, there is no information on how the C-terminus of *Bmal1* was lost in Brachycera, it is interesting to note that some synteny is conserved between these species.

The protein sequences for mdCYC and dmCYC are more similar than to dpBMAL1 as they are both truncated forms of *Bmal1*. Also, mdCYC and dmCYC, both contain bHLH-DNA binding domains and two PAS protein-protein interaction domains (Maurer et al., 2009). When transfected with their counterparts (mdCLK or dmCLK), they are able to activate transcription, thus it can be assumed that the bHLH and PAS domains are functional for both. So it is curious as to why mdCLK cannot bind with dmCYC to mediate transcription in two closely related species. This suggests that despite sequence similarity, a shared ability to bind to DNA and to each other, *mdCyc*/*dmCyc* or perhaps *mdClk*/*dmClk* have diverged enough that mdCLK and dmCYC can no longer form heterodimers.

I was also curious to discover where in the dipteran lineage the C-terminus of *Bmal1* was lost, as it is still found in mosquitoes but is absent in *Drosophila*. So I constructed a synteny map (Figure 8) to determine which dipterans had *Bmal1* and which species had *cyc*. A synteny map is useful as it can show us the "preserved order of genes on chromosomes of related species which results from decent from a common ancestor" (Duran et al., 2009). While only a limited number of dipterans have had their genomes sequenced, I was able to determine that sandflies (*Lutzomyia longipalpis* and *Phlebotomus papatasi*) and mosquitoes (*Aedes aegypti* and *Anopheles gambiae*) have *Bmal1*. The melon fly (*Bactrocera cucurbitae*), two *Drosophila* species (*Drosophila melanogaster* and *Drosophila grimshawi*), the tsetse fly (*Glossina pallidipes*), and the housefly all have *cyc*.

There is also some conservation in genes surrounding *cyc* and *Bmal1*. In mosquitoes and sandflies the gene *RE43767p*, is present in those species but has been lost in flies containing *cyc*. The gene *Fibq* is present in all species where *cyc* is present, and is always adjacent and upstream of *cyc*. The gene, *Cyp305a1* is present in all dipteran species and has been duplicated or triplicated in some species, mostly in mosquitoes and sandflies. Because of the close proximity of *Cyp305a1* to both *Bmal1* and *cyc* and its duplication in mosquitoes and sandflies, I investigated whether the C-terminus of BMAL1 could have been lost when the gene duplication was lost. However, upon taking a closer look at the DNA sequences between the genes and when I looked for a sequence similar to the C-terminus of BMAL1 in *Cyp305a1* it was absent. Perhaps with the sequencing of more dipterans we could make a more definitive conclusion as to how the C-terminus of BMAL1 was lost over time.

After establishing that mdCLK does not require dmCYC for activation I began to do transcriptional assays to determine how the housefly clock components compared to clock components of both *Drosophila* and the monarch. My *mdPer* construct was generated with both cDNA and gDNA fragments, which may have affected its ability to repress in unexpected ways. When *mdPer* was co-transfected with *dmClk*, it acted as a potent repressor of dmCLK:dmCYC (Figure 9). In fact, it appeared to be a more effective repressor than *dmPer* (Figure 9). These results support the findings of Piccin et al., 2000, who found that they could rescue locomotor rhythms of *Drosophila* per^{01} mutants using a single copy of *mdPer*. Interestingly, *dpPer*, which does not act as the

main transcription repressor in the monarch, and had repression levels similar to *dmPer* with *Drosophila* activator components (Figure 9).

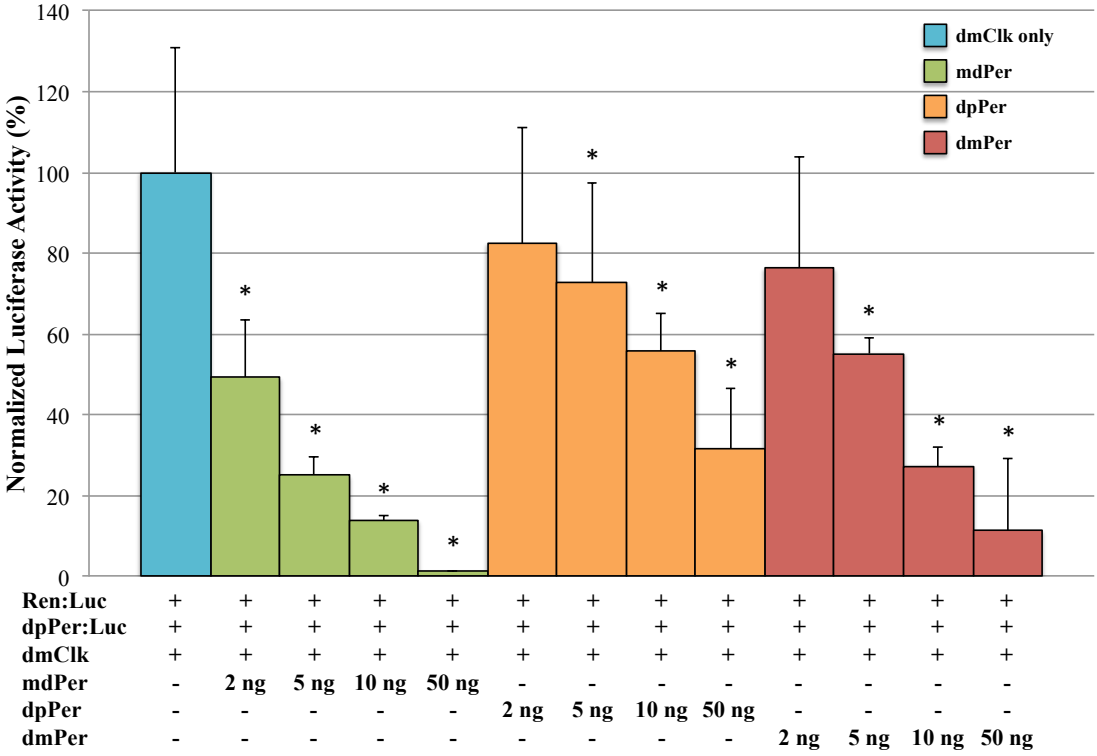

Figure 9: mdPER potently represses dmCLK:dmCYC. In S2 cells, mdPER, dpPER, and dmPER repress dmCLK:dmCYC but mdPER appears to repress more strongly than either dpPER or dmPER. The *mdPer* construct used was generated from cDNA and gDNA fragments. *p<0.05, Student's t-test.

When *mdCry2* was co-transfected with *dpBmal1* and *dpClk* there appeared to be no robust repression of dpBMAL1:dpCLK by mdCRY2, but rather only weak repression at high doses (Figure 10). However, dpCRY2 was able to strongly repress dpBMAL1:dpCLK with as little as 2 ng of dpCRY2 (Figure 10). This suggested to me that since mdPER was able to potently repress dmCLK:dmCYC, perhaps *mdPer* acted as the main transcriptional repressor in the housefly and not *mdCry2*. I thought that this

might provide evidence that PER, rather than CRY2, serves as the main trancriptional repressor in *Drosophila*.

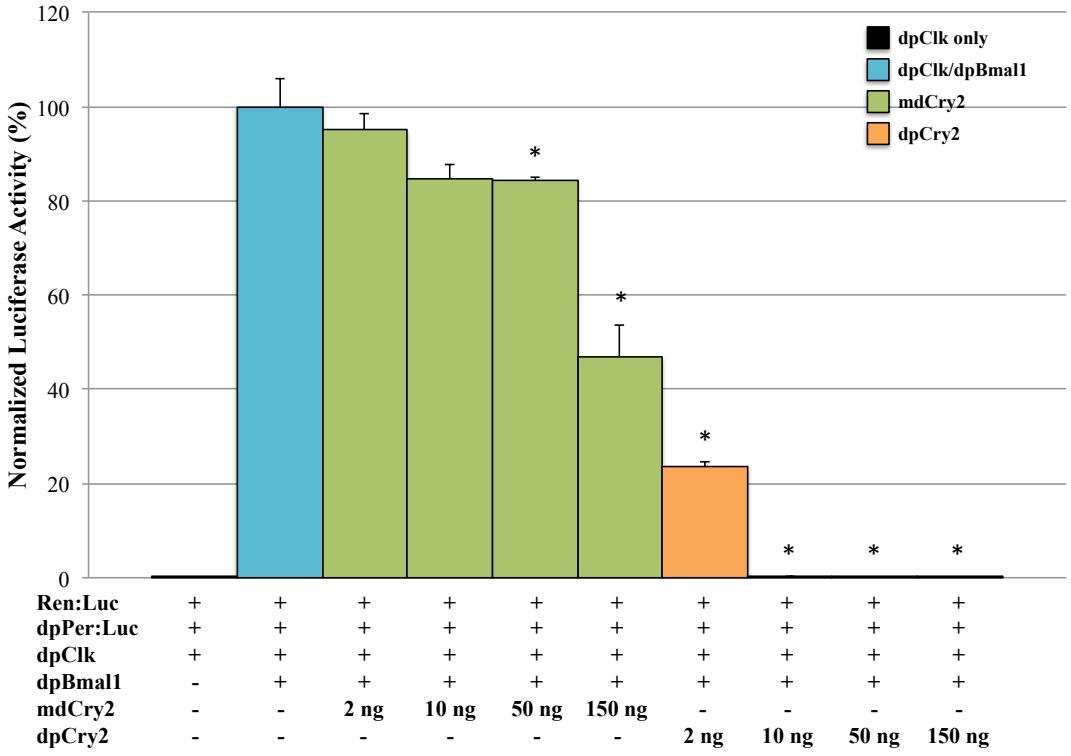

Figure 10: mdCRY2 does not repress dpCLK:dpBMAL1. In S2 cells, mdCRY2 does not repress dpCLK:dpBMAL1 but dpCRY2 is a potent repressor of dpCLK:dpBMAL1. *p<0.05, Student's t-test.

However, this proved not to be the case. When *mdPer* was transfected with *mdClk* and *mdCyc*, mdPER was unable to repress its own activators (Figure 11). Where previously it had potently repressed dmCLK:dmCYC it could not do so with mdCLK:mdCYC. This result suggests that mdPER does not behave like a transcriptional repressor in houseflies. But, it is possible that perhaps mdPER requires another protein such as mdTIM to help stabilize and translocate it back to the nucleus where it can repress. However, this seems unlikely since mdPER alone in S2 cells is able to repress dmCLK:dmCYC.

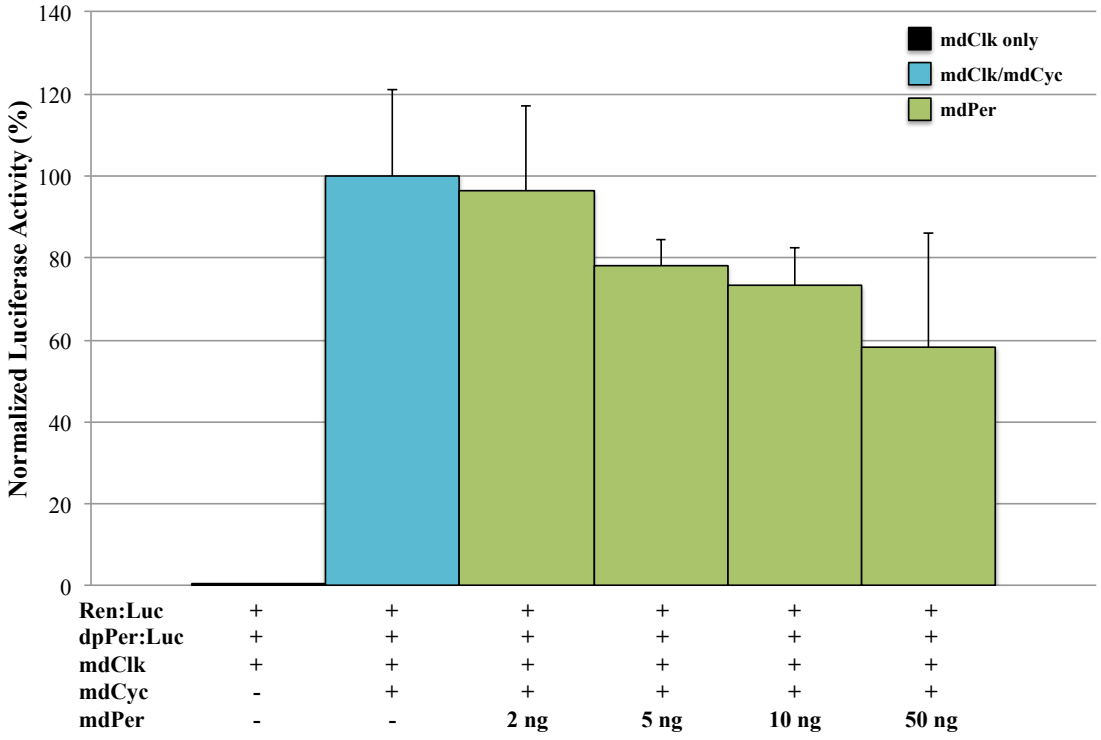

Figure 11: mdPER does not repress mdCLK:mdCYC. In S2 cells, mdPER shows no repression of mdCLK:mdCYC despite being able to potently repress dmCLK:dmCYC.

I then determined whether mdCRY2 could repress mdCLK:mdCYC, and found that like mdPER, mdCRY2 could not repress mdCLK:mdCYC (Figure 12). This suggested that perhaps neither mdCRY2 nor mdPER acted as transcriptional repressors in the housefly. Once again it is possible that perhaps mdCRY2 requires mdTIM to translocate back into the nucleus and repress its transcription. However, this seems unlikely since dpCRY2 in S2 cells is able to repress dpBMAL1:dpCLK without the presence of dpTIM.

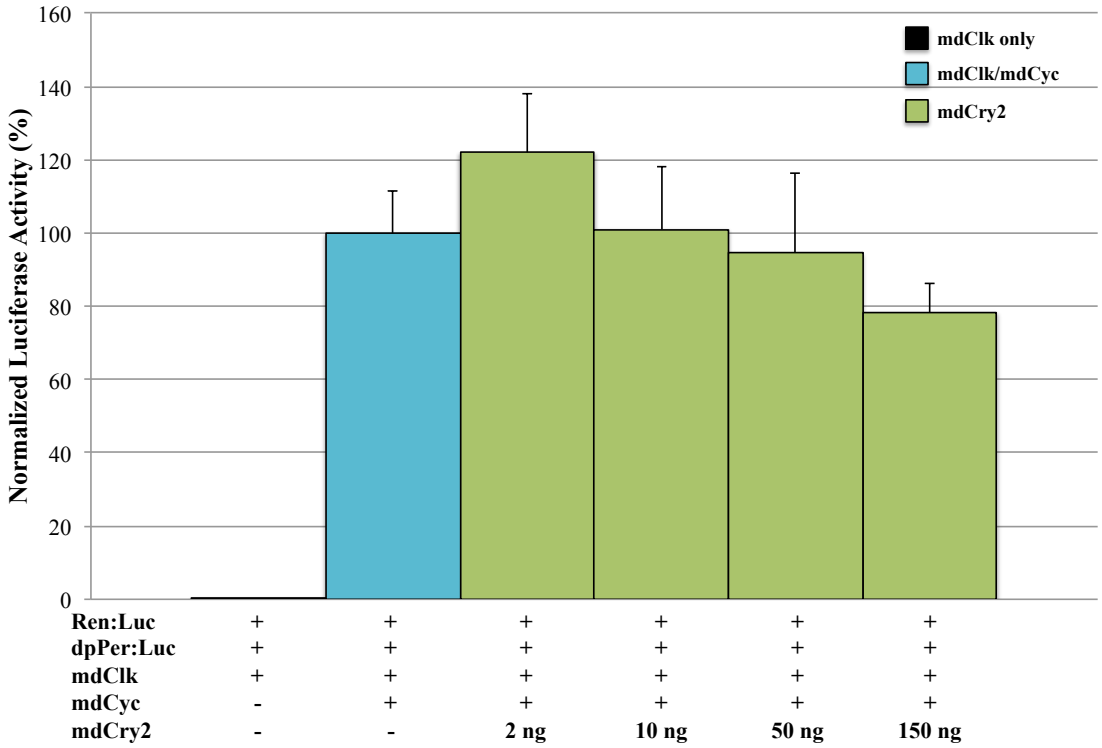

Figure 12: mdCRY2 does not repress mdCLK:mdCYC. In S2 cells, mdCRY2 shows no repression of mdCLK:mdCYC, which is not unexpected since it was also unable to repress dpCLK:dpBMAL1.

Since there was no repression of mdCLK:mdCYC by either mdPER or mdCRY2, I reasoned that perhaps in houseflies the two proteins work together to repress mdCLK:mdCYC. To test this hypothesis, I co-transfected varying amounts of *mdPer* and *mdCry2*. When *dmCwo* and *dmPer* are co-transfected in S2 cells, they were able to repress CLK activation by five to ten fold (Zhou et al., 2016). This suggested that CWO acts in synergy with PER to repress CLK and CYC (Zhou et al., 2016). However, I found that no matter the dose of *mdPer* or *mdCry2*, there was no repression of mdCLK:mdCYC transcription, thus there is no synergistic effect between these two proteins (Figure 13).

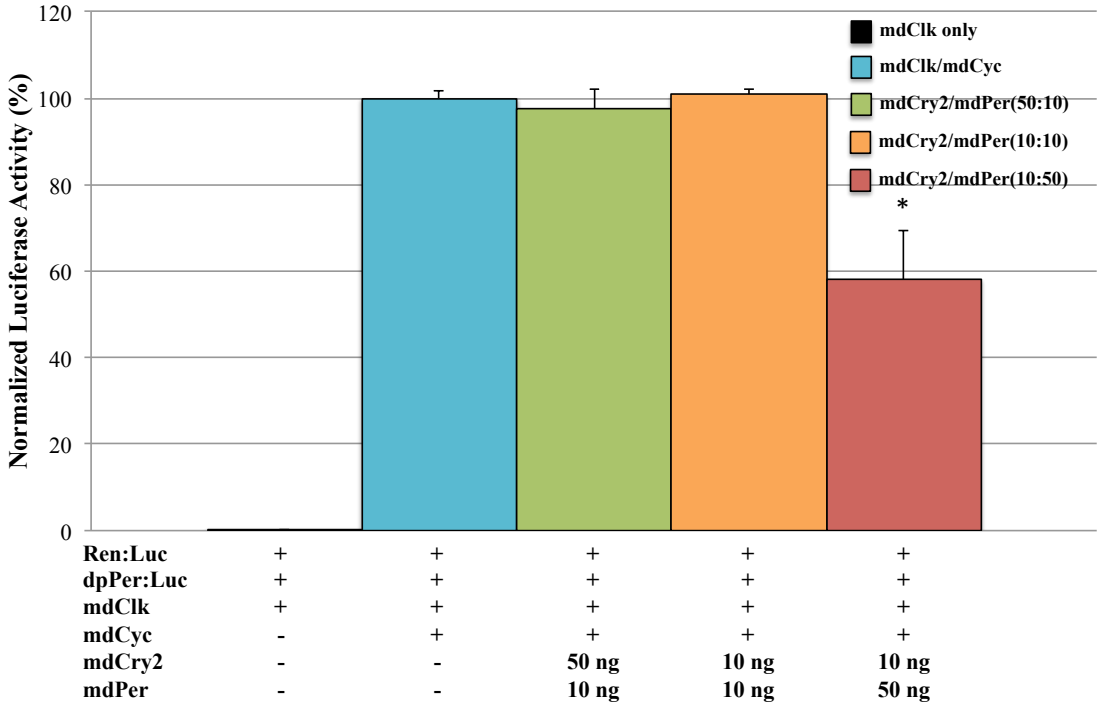

Figure 13: There is no synergistic effect between mdPER and mdCRY2. In S2 cells, there is no repression even when mdPER and mdCRY2 are co-transfected together. Thus, there is no synergistic effect between mdPER and mdCRY2. *p<0.05, Student's t-test.

My only evidence that mdCRY2 is a *cryptochrome* is based on a BLAST search using the dpCRY2 protein sequence to BLAST against the housefly genome. Since, mdCRY2 does not appear to repress either dpBMAL1:dpCLK or mdCLK:mdCYC, it is possible that *mdCry2* was misidentified as a mammalian-like *cry*. To test this possibility, I generated a synteny map for *cry2* to determine if there was any conservation of gene blocks between dipterans known to have *cry2* and those dipterans, which I thought might have a *cry2* (Figure 14).

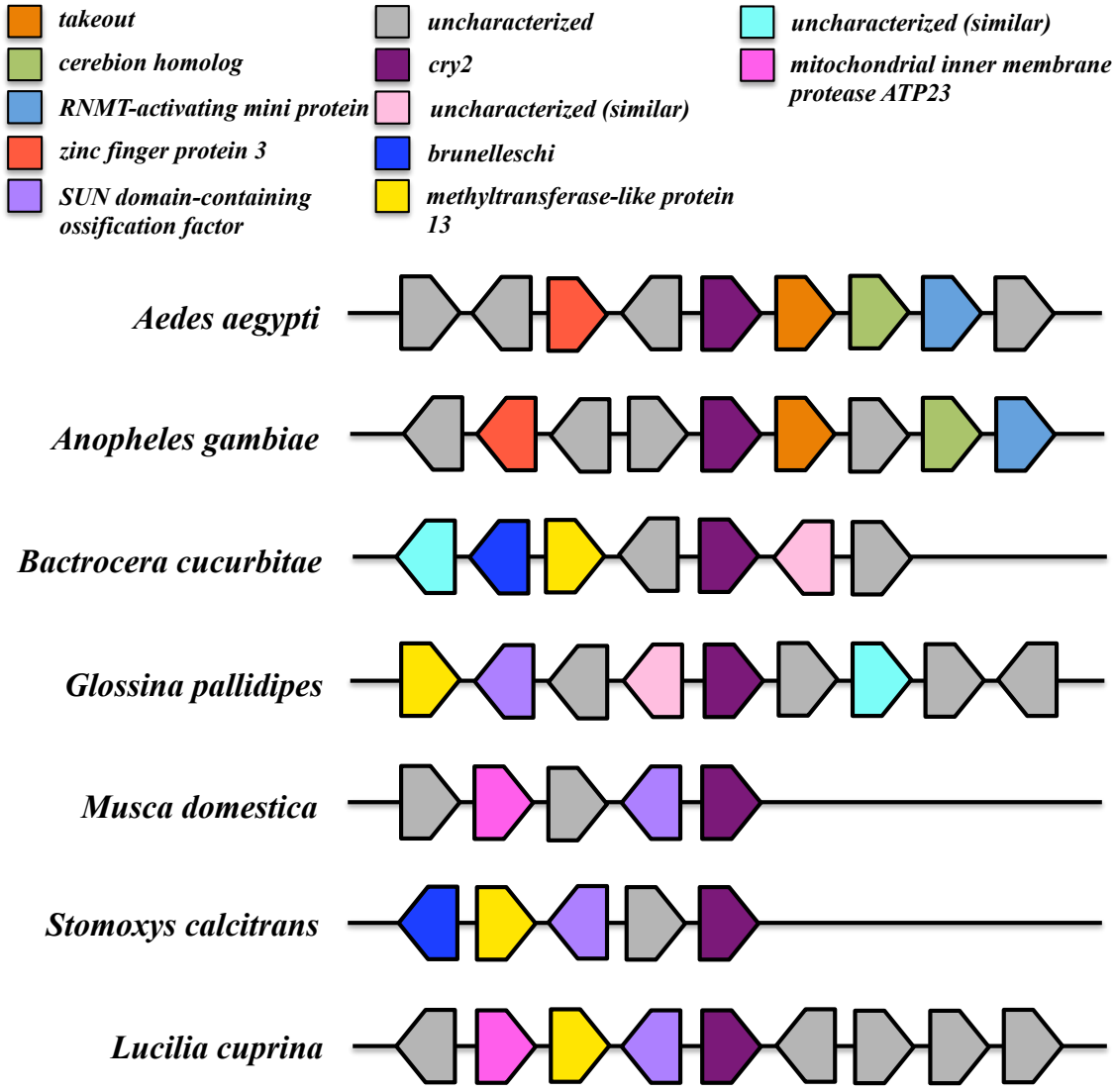

Figure 14: A synteny map of mammalian cryptochrome (*cry2*). There appears to be no synteny between lower dipterans (*Aedes aegypti* and *Anopheles gambiae*), which are known to have *cry2* and brachycerans (*Bactrocera cucurbitae*, *Glossina pallidipes*, *Musca domestica*, *Stomoxys calcitrans*, and *Lucilia cuprina*) suggesting that perhaps flies in Brachycera do not have *cry2*.

When species, known to have CRY2 (*Aedes aegypti* and *Anopheles gambiae*) were compared to species suspected of having CRY2 based on sequence similarity to dpCRY2 (*Bactrocera cucurbitae*, *Glossina pallidipes*, *Musca domestica*, *Stomoxys*

calcitrans, and *Lucilia cuprina*), we found that there was no synteny between those two groups. There appears to be synteny between *Aedes aegypti* and *Anopheles gambiae* as there is conservation of several genes next to *cry2*. These genes are *takeout, zinc finger protein 3, cerebion homolog,* and *RNMT-activating mini protein* (Figure 14). There is also some conservation of genes between species suspected of having *cry2*, but those genes are not found in either *Aedes aegypti* or *Anopheles gambiae*. This result argues that mdCRY2 might not be in fact a mammalian *cryptochrome*.

In a BLAST search, using the *Drosophila* and monarch CRY1 protein sequences as our search query against the housefly genome, we identified housefly CRY1. Thus, I knew with a high degree of certainty that *mdCry2* which I had originally identified was not the blue light photoreceptor, *cry1*. Another possible explanation was that mdCRY2 could be a (6-4) DNA photolyase. It is known that both *cryptochromes* and *photolyases* belong to the same family of flavoproteins and share structural similarities (Mei et al., 2015). There are two conserved domains found in both *cryptochromes* and *photolyases*: a DNA photolyase related domain and a FAD binding domain (Mei et al., 2015). However, *cryptochromes* have a C-terminal domain that varies in length between species and is not found in *photolyases* (Mei et al., 2015). *Photolyases* are also DNA repair enzymes, which are activated by blue light and repair DNA damage caused by ultraviolet light, whereas *cryptochromes* lack this ability to repair UV-damaged DNA (Mei et al., 2015; Zhao et al., 1998).

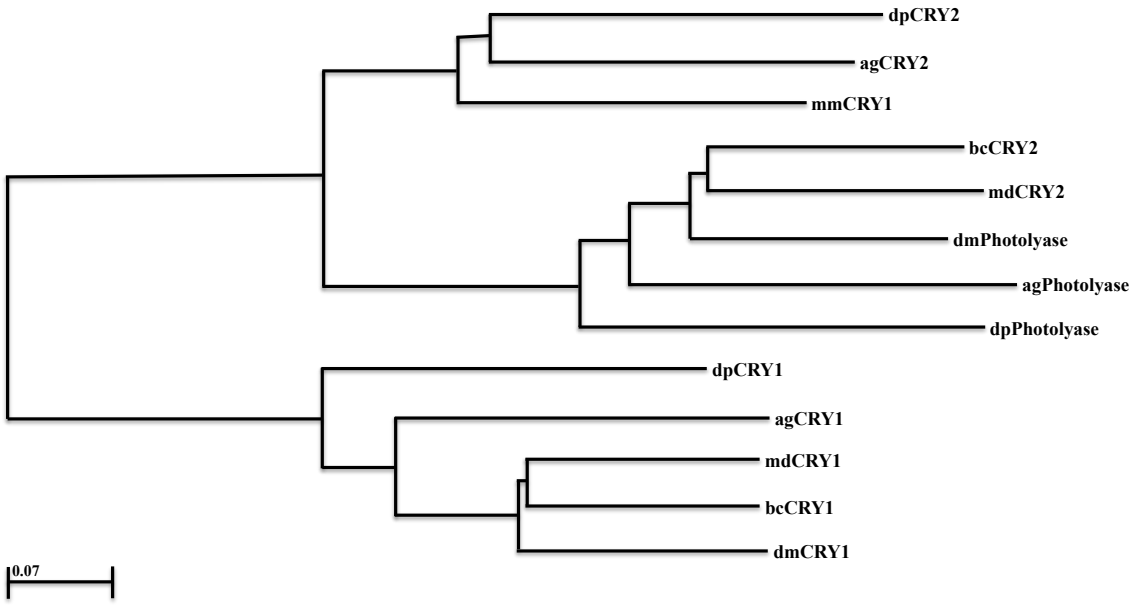

Figure 15: A phylogenetic tree of cryptochromes and (6-4) DNA photolyases. The monarch and *A. gambiae* both have CRY2 and it is grouped with the mouse CRY1. All insects, which have the blue light photoreceptor, CRY1 group together on the phylogenetic tree. However, both B. *cucurbitae* and *M. domestica* CRY2s are grouped with known (6-4) DNA photolyases from *Drosophila*, the monarch, and *A. gambiae*. The phylogentic tree was generated using COBALT and was then modified by the author.

I identified known (6-4) DNA photolyases from *Drosophila*, *A. gambiae*, and the monarch, as well as the mammalian *cryptochrome* from the monarch, mouse, and *A. gambiae*. I also identified *cry1* from the monarch, *A. gambiae*, the housefly, the melon fly, and *Drosophila*. Then using the multiple alignment tool COBALT, the sequences were aligned and compared to the *cry2* sequences we had originally identified for the housefly and the melon fly. A phylogenetic tree was generated to see how the different proteins grouped together. What I discovered from this analysis was that the monarch, mouse (*cry1*), and *A. gambiae* mammalian *cryptochromes* all aligned together and were separated from the *photolyases* and *cry1*s (Figure 15). I also found that *cry1* from the

monarch, *A. gambiae*, the housefly, the melon fly, and *Drosophila* all grouped together and separately on the tree (Figure 15). Interestingly, *cry2* from the housefly and the melon fly grouped together with the (6-4) DNA photolyases from *Drosophila*, *A. gambiae*, and the monarch (Figure 15). This was strong circumstantial evidence that mdCRY2 was a (6-4) DNA photolyase rather than a *cryptochrome*.

However, to determine whether housefly *cry2* was in fact a *photoylase* and not a *cryptochrome*, I tested whether mdCRY2 displayed DNA repair activity. To test this possibility I used a strain of *E. coli* (SY32), which carries the plasmid pRT2 (Daiyasu et al., 2004; Kobayashi et al., 2000). pRT2 is a plasmid which encodes for the *E.coli* phr^{+} gene, which inhibits (6-4) DNA photolyase activity in *E. coli* (Daiyasu et al., 2004). The SY32 cells acted as a host for our expression vectors containing *mdCry2*, *z6-4phr*, or *zcry1b*. *z6-4phr* is a plasmid which has the full length coding sequence of the zebrafish (6-4) DNA photolyase, and *zcry1b* is a plasmid which has the full length coding sequence for a *cryptochrome* in zebrafish (Kobayashi et al., 2000). The *mdCry2*, *z6-4phr*, or *zcry1b* plasmids were transformed into SY32 cells, exposed to UV light, followed by fluorescent light, and then the cells were allowed to grow overnight. Very few colonies grew from the cells transformed with *zcry1b*, which is expected since it is a *cryptochrome* and has no DNA repair activity (Figure 16). However, the number of colonies that grew for *mdCry2* was much more than either the negative control (*zcry1b)* or the positive control (*z6-4phr)*. This result confirmed that *mdCry2* was not a

mammalian-like *cry2*, but was in fact a (6-4) DNA photolyase able to repair UV-induced DNA damage (Figure 16).

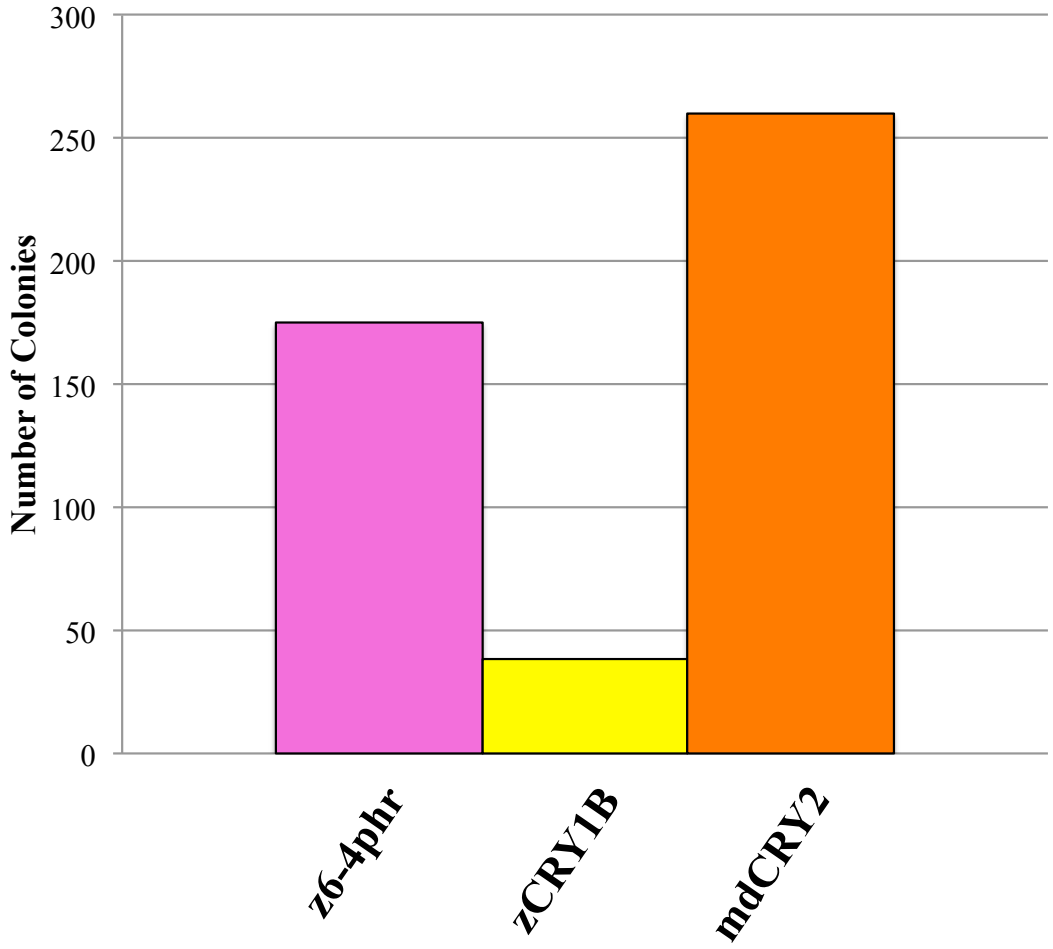

Figure 16: mdCRY2 is not a mammalian-like *cryptochrome*. When tested for (6-4) DNA photolyase activity more colonies grew for mdCRY2 than the zebrafish (6-4) DNA photolyase. This was definitive proof that *mdCry2* was actually a (6-4) DNA photolyase and not a mammalian-like cryptochrome.

What is the main transcriptional repressor in the housefly?

In S2 cells, mdPER could not repress mdCLK:mdCYC and mdCRY2 was discovered to be a (6-4) DNA photolyase rather than a mammalian-like cryptochrome. Thus, how the housefly clock functioned remained a mystery. Specifically, what protein acts as the

main transcriptional repressor in the housefly clock mechanism? One strong candidate was, *clockwork orange*. CWO is a strong transcriptional repressor that acts in synergy with PER to repress CLK:CYC in *Drosophila* (Zhou et al., 2016).

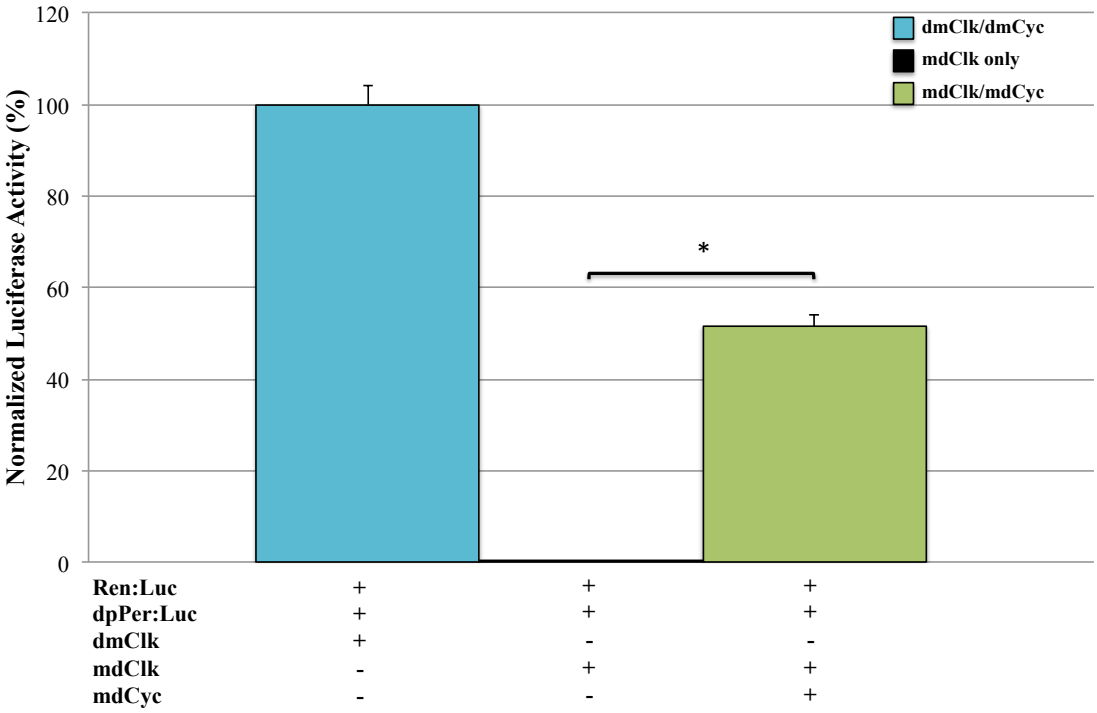

Figure 17: *mdClk* (full-length cDNA) requires *mdCyc* for activation. In S2 cells, *Clk* from the housefly cannot activate with *dmCyc*. It requires *mdCyc* for activation. This *mdClk* construct was generated from full-length cDNA. *p<0.05, Student's t-test.

To test whether CWO functions to repress mdCLK:mdCYC transcription, luciferase assays were carried out using expression plasmids containing full-length cDNAs for *mdClk* and *mdPer* rather than the hybrid cDNA-genomic clones used previously. Expression plasmids containing full-length *mdClk* and *mdPer* cDNAs functioned identically to their hybrid counterparts. As can be seen in Figure 17, mdCLK:dmCYC does not mediate transcription but when *mdClk* is transfected along with *mdCyc* there is

activation. The inability for mdCLK:dmCYC to activate suggests that either *mdCyc*/*dmCyc* or *mdClk*/*dmClk* have diverged enough that mdCLK and dmCYC can no longer bind.

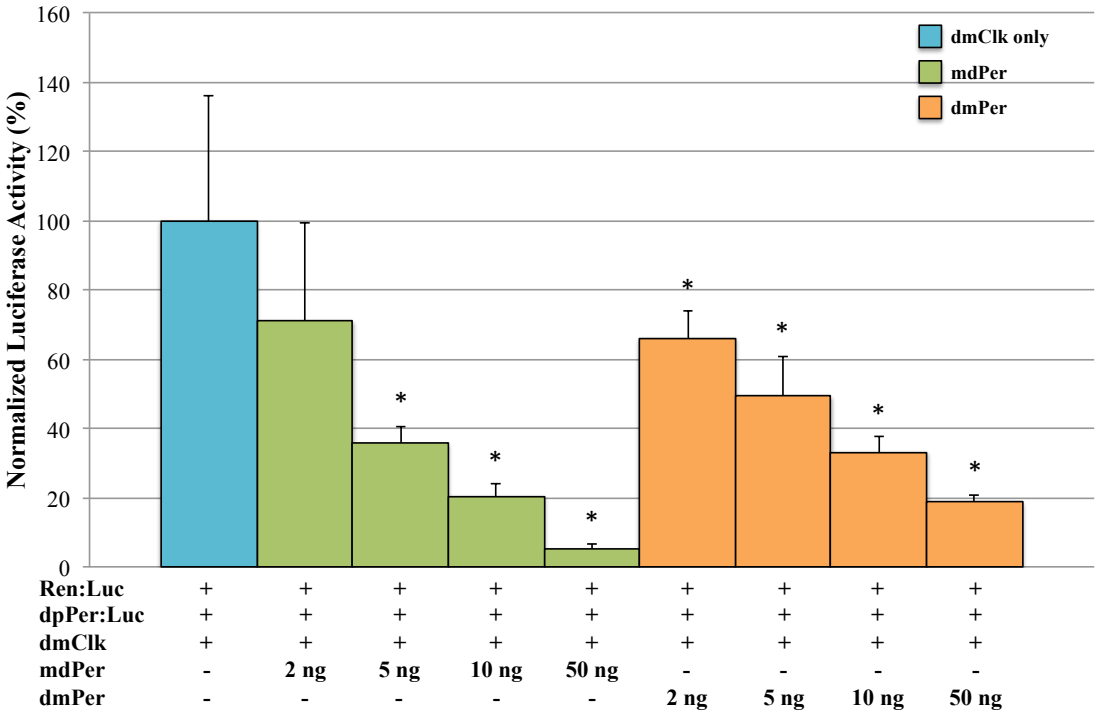

Figure 18: mdPER represses dmCLK:dmCYC. In S2 cells, mdPER and dmPER repress dmCLK:dmCYC but mdPER appears to repress more strongly than dmPER. The *mdPer* construct was generated from full-length cDNA. *p<0.05, Student's t-test.

I also wanted to determine whether mdPER could still potently repress dmCLK:dmCYC as shown previously. When *mdPer* was co-transfected with *dmClk* it was indeed able to potently repress dmCLK:dmCYC (Figure 18). In fact, just like before, it appeared to be a more effective repressor than *dmPer* of dmCLK:dmCYC (Figure 18). I then tested whether mdPER could repress mdCLK:mdCYC and found that like before, mdPER was unable to repress mdCLK:mdCYC (Figure 19). Moreover, dmPER was only able to

weakly repress when 50 ng was transfected. This suggests that perhaps mdCLK has evolved so that PER can no longer bind to it and repress its own gene transcription.

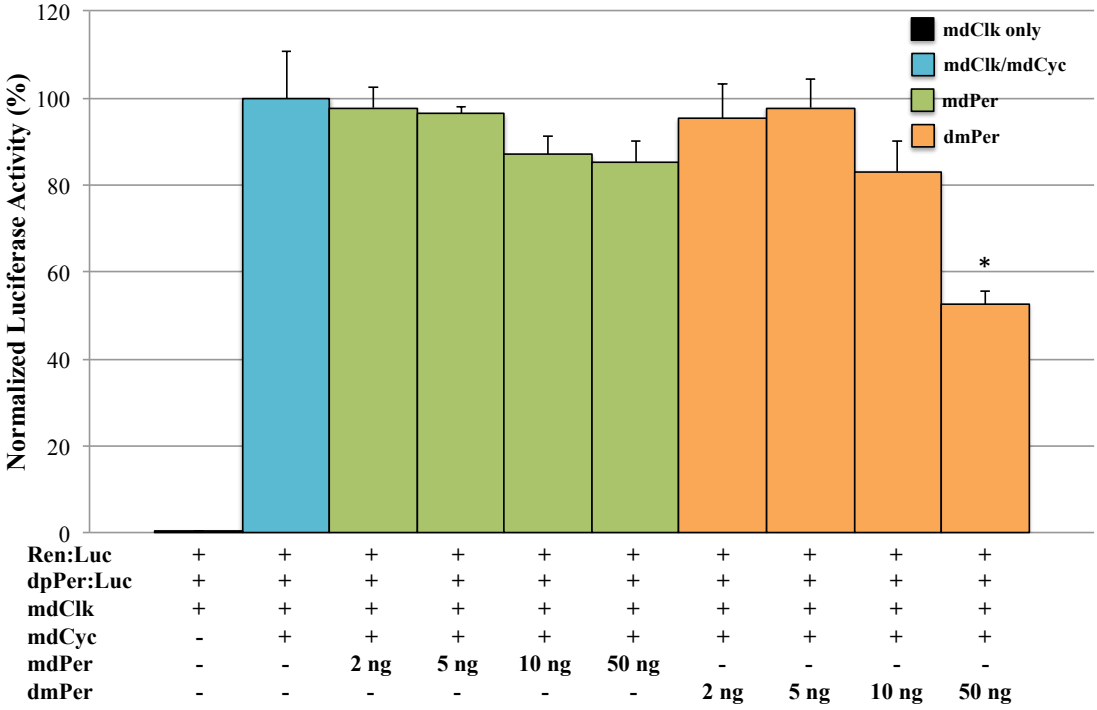

Figure 19: mdPER (full-length cDNA) does not repress mdCLK:mdCYC. In S2 cells, mdPER shows no repression of mdCLK:mdCYC despite being able to potently repress dmCLK:dmCYC. dmPER is able to slightly repress mdCLK:mdCYC when 50 ng is used. *p<0.05, Student's t-test.

After establishing that *mdPer* generated from full-length cDNA still could not repress mdCLK:mdCYC I decided to focus my attention on the candidate repressor, *mdCwo*. I first tried to determine whether mdCWO could repress dmCLK:dmCYC. Thus, I co-transfected *dmClk* along with *mdCwo* in a dose-dependent manner and found that mdCWO was indeed able to repress dmCLK:dmCYC (Figure 20). As a control, *Drosophila* CWO was also able to repress dmCLK:dmCYC with as much potency as mdCWO (Figure 20).

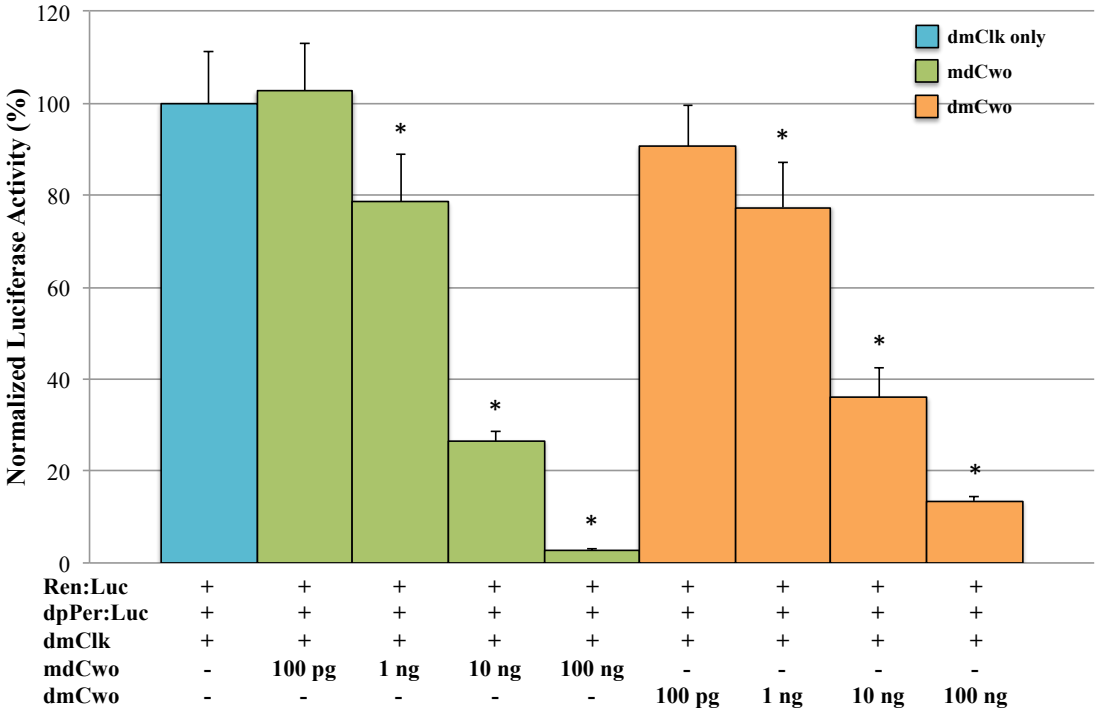

Figure 20: mdCWO represses dmCLK:dmCYC. In S2 cells, both mdCWO and dmCWO can repress dmCLK:dmCYC. *p<0.05, Student's t-test.

I then tested whether mdCWO could repress mdCLK:mdCYC. When 100 pg of *mdCwo* was co-transfected with *mdClk* and *mdCyc*, mdCWO was able to potently repress mdCLK:mdCYC mediated transcription (Figure 21). Whereas, dmCWO also repressed mdCLK:mdCYC although it required 10 ng of dmCWO before repression was observed (Figure 21). These results suggest that mdCWO may act as the main transcriptional repressor in the clock of the housefly. It is curious that in *Drosophila*, previous work by other labs has shown that PER acts as the main transcriptional repressor and CWO acts in synergy with PER to repress transcription, but based on my own work in the closely related species, the housefly, the opposite appears to be true.

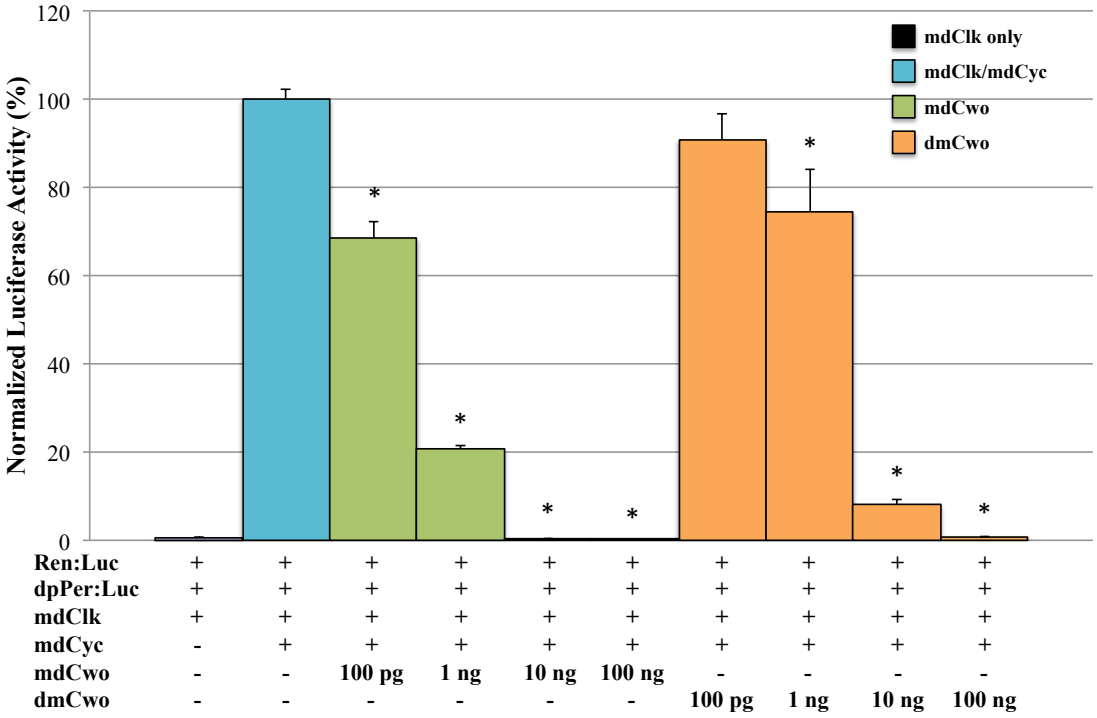

Figure 21: mdCWO potently represses mdCLK:mdCYC. In S2 cells, mdCWO is able to repress mdCLK:mdCYC with as little as 100 pg. dmCWO is also able to repress mdCLK:mdCYC. *p<0.05, Student's t-test.

After demonstrating that mdCWO potently repressed mdCLK:mdCYC, I wanted to test whether mdCWO acted synergistically with mdPER to repress mdCLK:mdCYC since dmCWO and dmPER synergistically repress dmCLK:dmCYC transcription (Kadener et al., 2007). However, when replicating the synergistic action of *Drosophila* CWO and PER I found that co-transfecting varying concentrations of *dmPer* (2 ng-50 ng) and a constant amount of *dmCwo* (2 ng) repressed transcription to the same extent as when *dmPer* was transfected alone (Figure 22). Several different concentrations of *dmCwo* were used with varying concentrations of *dmPer* (data not shown), yet no synergistic repression was observed.

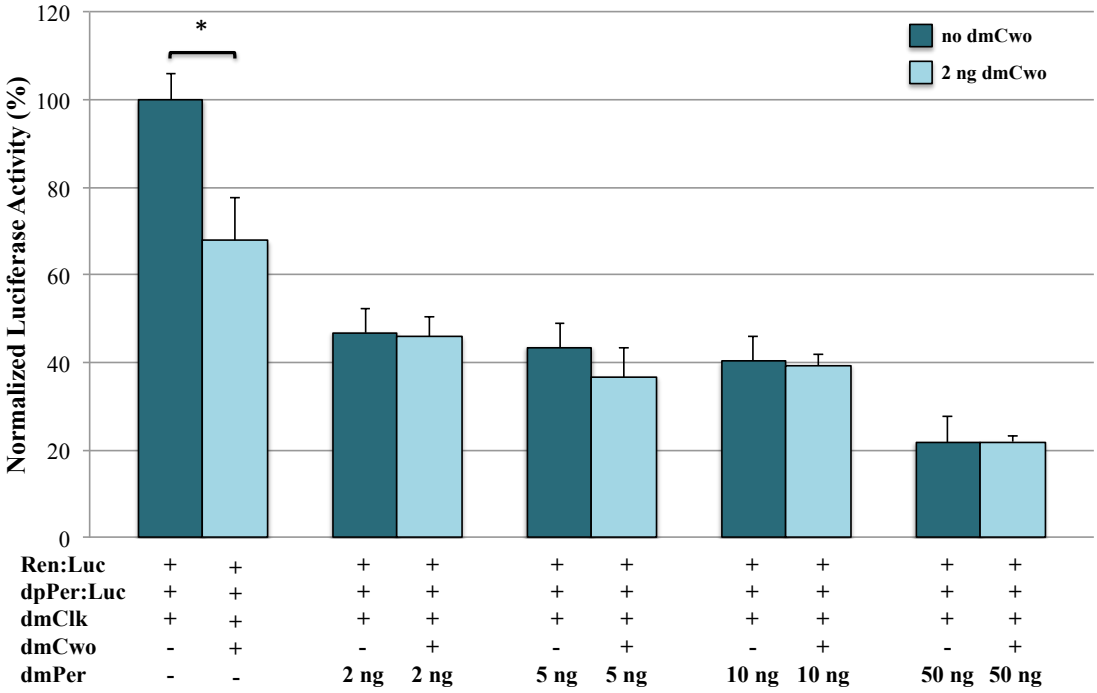

Figure 22: There is no synergy with an increasing dosage of dmPER and constant levels of dmCWO. Although seen previously by other labs, we saw no synergy with an increasing dosage of dmPER and constant levels of dmCWO. *$p<0.05$, Student's t-test.

I then tested synergy by co-transfecting varying concentrations of *dmCwo* (100 pg-100 ng) and a constant amount of *dmPer* (2 ng). Once again, dmCWO did not synergize with dmPER to repress dmCLK:dmCYC (Kadener et al., 2007). When I co-transfected *dmCwo* and *dmPer* the repression was no different than *dmCwo* alone (Figure 23). Several different concentrations of *dmPer* were used with varying concentrations of *dmCwo* (data not shown), yet no synergistic repression was observed. Despite replicating the previous experimental conditions as accurately as possible subtle differences in conditions may account for my different results.

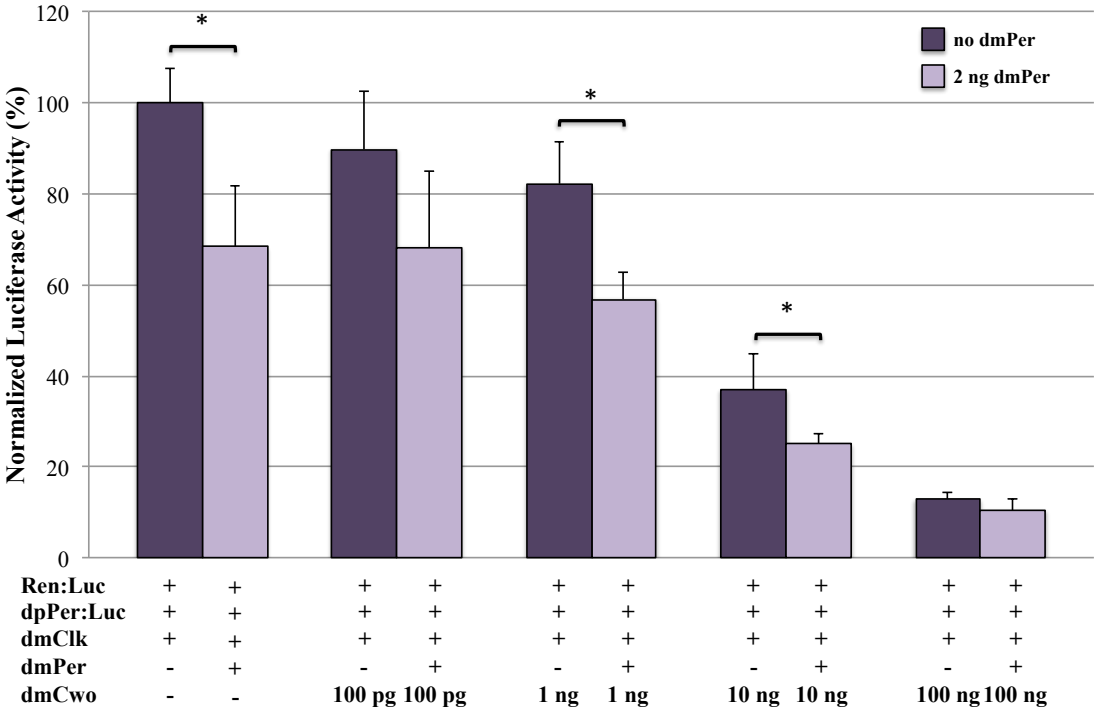

Figure 23: There is no synergy with an increasing dosage of dmCWO and constant levels of dmPER. Although seen previously by other labs, in our hands we saw no synergy with an increasing dosage of dmCWO and constant levels of dmPER. *p<0.05, Student's t-test.

I then tested whether mdPER and mdCWO acted in synergy to repress mdCLK:mdCYC. However, I was unable to detect synergy between mdPER and mdCWO. When varying concentrations of *mdPer* (10 ng-200 ng) were co-transfected with a constant amount of *dmCwo* (100 pg), the low repression levels were no different than when *mdPer* was transfected alone (Figure 24). Several different concentrations of *mdPer* were used with varying concentrations of *mdCwo* (data not shown), but no synergy was detected. Perhaps *mdPer* no longer serves a function in the housefly clock or perhaps its role has become more similar to *dmTim* and acts to stabilize *mdCwo*.

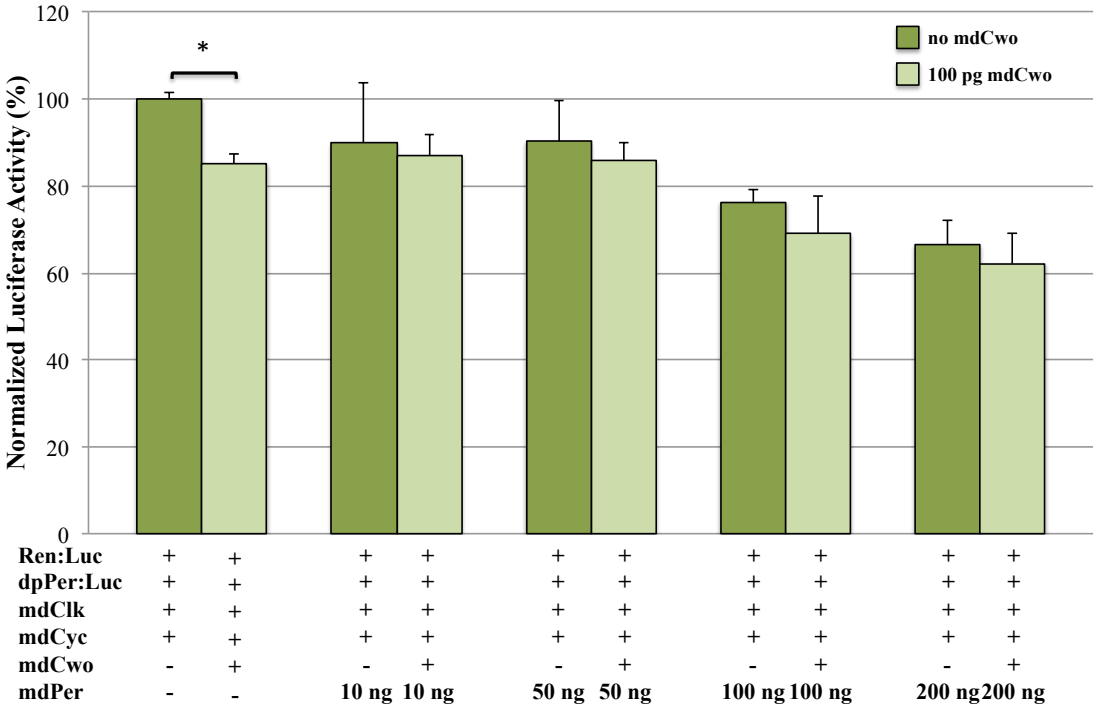

Figure 24: There is no synergy with an increasing dosage of mdPER and constant levels of mdCWO. In S2 cells, we saw no synergy with an increasing dosage of mdPER and constant levels of mdCWO. *p<0.05, Student's t-test.

I then tested synergistic action between mdCWO and mdPER by co-transfecting varying concentrations of *mdCwo* (100 pg-5 ng) with constant amounts of *dmPer* (100 ng). Once again mdCWO and mdPER do not act in synergy to repress mdCLK:mdCYC. When *mdPer* was co-transfected with varying amounts of *mdCwo* repression levels were no different than when *dmCwo* was transfected alone (Figure 25). Since no synergistic effect was detected between dmPER and dmCWO, I cannot determine for sure that there is no synergistic effect between mdPER and mdCWO, but these results suggest that the clock mechanism in the housefly behaves differently from the *Drosophila* clock.

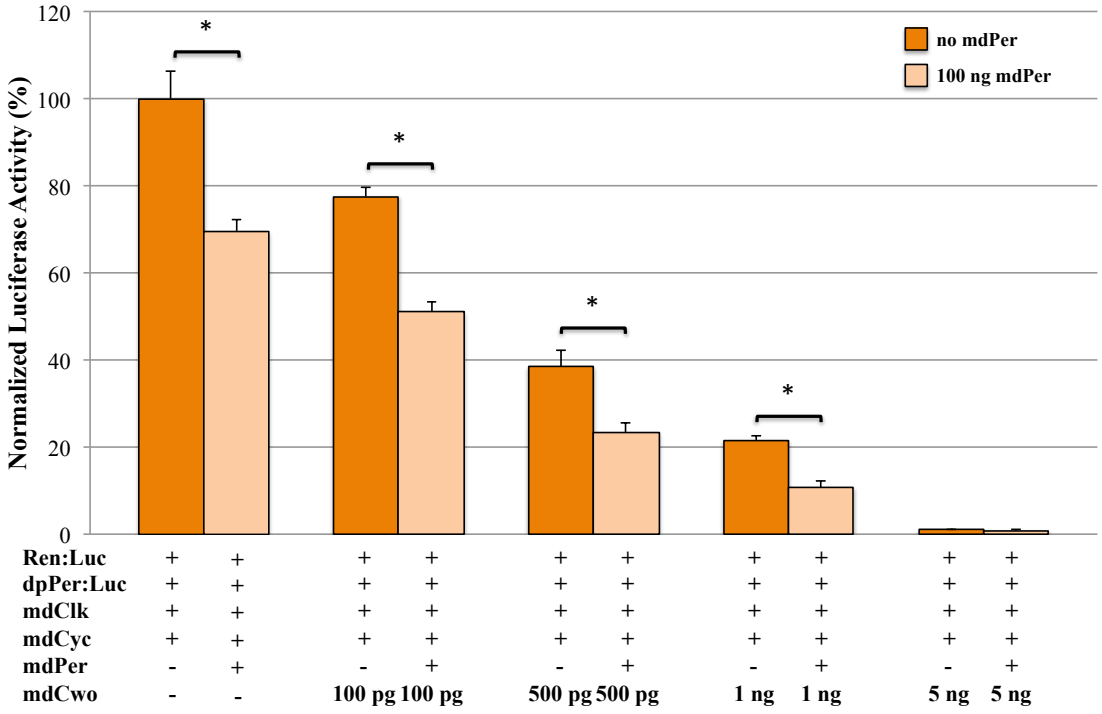

Figure 25: There is no synergy with an increasing dosage of mdCWO and constant levels of mdPER. In S2 cells, we saw no synergy with an increasing dosage of mdCWO and constant levels of mdPER. *p<0.05, Student's t-test.

In addition, I wanted to look at whether there was any synteny among dipterans with *clockwork orange* (Figure 26). I looked at the mosquito (*Anopheles gambiae*), the melon fly (*Bactrocera cucurbitae*), the medfly (*Ceratitis capitata*), two *Drosophila* species (*Drosophila melanogaster* and *Drosophila biarmipes*), the tsetse fly (*Glossina morsitans*), and the housefly. Unfortunately, among this small group of dipterans there was no synteny except in very closely related species such as the *Drosophila* species and the medfly and the melon fly. In the *Drosophila* species there were several conserved genes in close proximity to *cwo*: *jumeau*, *Rfx*, *mthl11*, and *Irp-1B*. For the medfly and the melon fly there were three genes conserved close to *cwo*: *argonaute 3*, *RE26705p*,

and *eclosion hormone*. However, none of these genes appear anywhere close to *cwo* in other dipteran species. This lack of synteny suggests that in Diptera, there may have been translocations or inversions of chromosomes, which led to the loss of conservation of genes around *cwo*.

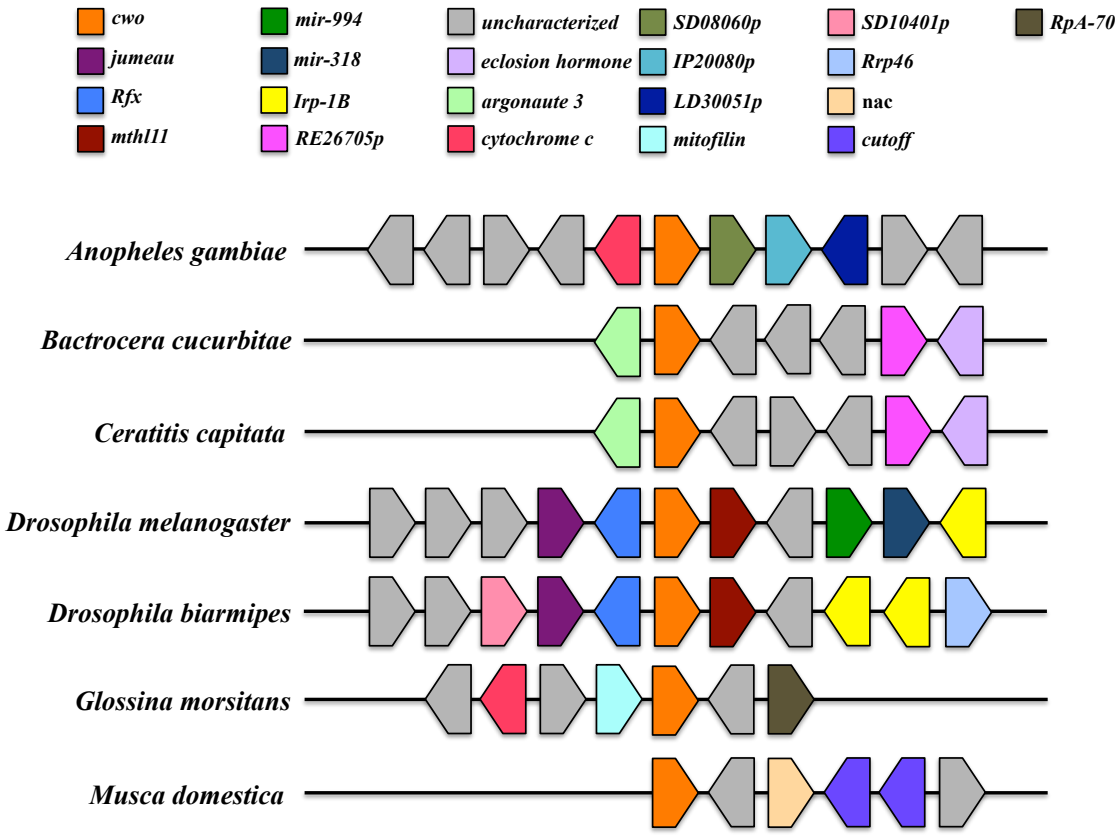

Figure 26: A synteny map of *clockwork orange*. There is no synteny for flies in either lower Diptera (*Anopheles gambiae*) or Brachycera (*Bactrocera cucurbitae*, *Ceratitis capitata*, *Drosophila melanogaster*, *Drosophila biarmipes*, *Glossina morsitans* and *Musca domestica*) suggesting that perhaps *cwo* has evolved various functions in different fly species.

CHAPTER IV

DISCUSSION AND FUTURE DIRECTIONS

M. domestica does not have a mammalian cryptochrome

I initially identified three dipterans, the melon fly (*Bactrocera cucurbitae*), the housefly (*Musca domestica*), and the tsetse fly (*Glossina pallidipes*) as species having a hybrid clock. These species contained the *Drosophila* clock components Clk, *cyc*, *per*, *tim*, *cry1*, and *cwo*, but also the mammalian-like *cryptochrome cry2*. CRY2 in these species was identified from a BLAST search using the monarch CRY2 sequence as our query. The alignment of CRY2 sequences from these dipterans to monarch CRY2 was highly similar, and contained the expected DNA photolyase related domain and FAD binding domain (Mei et al., 2015). I chose the housefly to further our investigation of the evolution of the clock in Diptera because its genome had already been fully sequenced and previous work done on the housefly clock showed that its clock mechanism, though similar to that in *Drosophila* had notable differences.

In S2 cells we found that mdCLK:dmCYC does not mediate transcription except when *mdClk* is transfected with *mdCyc*. However, it was surprising that CLK and CYC from two closely related species could not interact and mediate transcription. Based on sequence alignments they seemed to be more similar to each other than to monarch CLK and BMAL1, an insect with an ancestral clock. When transfected with counterparts (mdCLK or dmCLK), they were able to activate transcription, which suggested that the

E-box and protein-protein binding of these proteins were intact. Despite the sequence similarity among these orthologous proteins, mdCYC/dmCYC or perhaps mdCLK/dmCLK has diverged enough that mdCLK and dmCYC could no longer form functional heterodimers.

We wanted to determine where in the dipteran lineage the C-terminus of BMAL1 was lost. A synteny map of dipterans with *Bmal1* and *cyc* revealed that the sandflies (*Lutzomyia longipalpis* and *Phlebotomus papatasi*) and the mosquitoes (*Aedes aegypti* and *Anopheles gambiae*) have *Bmal1*, while dipterans, like the melon fly (*Bactrocera cucurbitae*), *Drosophila* species (*Drosophila melanogaster* and *Drosophila grimshawi*), the tsetse fly (*Glossina pallidipes*), and the housefly all have *cyc*. The gene, *Cyp305a1*, is in close proximity to dipterans with *Bmal1* and *cyc*. In mosquitoes and sandflies *Cyp305a1* is present in duplicates or triplicates but there is only one copy in the melon fly, *Drosophila* species, the tsetse fly, and the housefly. Because of the close proximity of *Cyp305a1* to both *Bmal1* and *cyc* we wanted to see whether the C-terminus of BMAL1 could perhaps have been lost when the gene duplication of *Cyp305a1* was lost in brachycerans. However, upon taking a closer look at *Cyp305a1* in the melon fly, *Drosophila* species, the tsetse fly, and the housefly as well as at the DNA sequences between *cyc* and *Cyp305a1* we could not locate the C-terminus of BMAL1.

When we co-transfected *mdPer* with *dmClk* we found that mdPER was able to potently repress dmCLK:dmCYC in S2 cells. Although, a potent repressor of *Drosophila*

activators, when we transfected *mdPer* with *mdClk* and *mdCyc*, mdPER was unable to repress mdCLK:mdCYC. In houseflies it would appear that mdPER does not act as the main transcriptional repressor. It is possible that *mdPer* requires *mdTim* or even *mdCry2* for repression. In *Drosophila*, the heterodimerization of TIM and PER is necessary for these proteins ability to translocate into the nucleus and thus repress (Rothenfluh et al., 2000). Heterodimerization of PER and TIM also protects PER from phosphorylation events caused by kinases such as *double-time* (Rothenfluh et al., 2000). Since, we were unable to clone the full-length cDNA of *mdTim* due to the large size of the gene and the cloning of the entire genomic sequence was unfeasible due to multiple large introns, we were unable to test this hypothesis. But it also seems unlikely that mdPER requires mdTIM to repress, since dmPER alone in S2 cells is able to repress dmCLK:dmCYC.

After establishing that mdPER does not act as the main transcriptional repressor in houseflies, I examined whether *mdCry2* fulfilled this function. *mdCry2* was co-transfected with *dpBmal1* and *dpClk* and was unable to repress dpBMAL1:dpCLK transcription, but dpCRY2 was able to strongly repress. Likewise mdCRY2 could not repress mdCLK:mdCYC transcription. Since both mdPER and mdCRY2 could not repress individually we thought perhaps they worked synergistically to repress mdCLK:mdCYC transcription. However, we found that no matter the dose of *mdPer* or *mdCry2*, there was no repression of mdCLK:mdCYC transcription. We concluded that neither mdPER nor mdCRY2 have a repressive capability in the housefly.

Since mdCRY2 is unable to repress either dpCLK:dpBMAL1 or mdCLK:mdCYC, I began to suspect that I had misidentified mdCRY2 as a mammalian-like *cryptochrome*. I constructed a synteny map by comparing species known to have CRY2 (mosquitoes) and species suspected of having CRY2 (melon fly, housefly, and tsetse fly). I found that species known to have CRY2 lacked any synteny with species that were suspected of having CRY2. This evidence made me suspect that perhaps mdCRY2 was in fact a (6-4) DNA photolyase. I came to this idea, because both *cryptochromes* and *photolyases* belong to the same family of flavoproteins that share structural similarities and two conserved domains: a DNA photolyase related domain and a FAD binding domain (Mei et al., 2015).

I then generated a phylogentic tree for *cryptochromes* and *photolyases* to determine which group contained mdCRY. We looked at species known to have a mammalian *cry2* (monarch, mouse, and *A. gambiae*), species known to have a *Drosophila cry1* (monarch, *A. gambiae*, *Drosophila*, housefly, and melon fly), and species known to have (6-4) DNA photolyase (monarch, *A. gambiae*, and *Drosophila*) and compared them to *cry2* from the housefly and the melon fly. I found that all sequences for CRY1 grouped together on the tree and all known sequences for CRY2 grouped together. However, CRY2 from the housefly and the melon fly grouped with the sequences for (6-4) DNA photolyase. Both these tests provided strong circumstantial evidence that mdCRY2 was not a mammalian *cry*. I then tested whether mdCRY2 displayed DNA repair activity like *photolyases*, and found that mdCRY2 is able to repair UV-induced DNA damage.

Therefore, I concluded that *M. domestica* does not have a mammalian *cry2*, and thus the housefly clock mechanism consists only of *Drosophila* clock components.

CWO is the main transcriptional repressor of the housefly

Since mdCRY2 is a (6-4) DNA photolyase and mdPER does not repress mdCLK:mdCYC, we were unsure of how the housefly clock was repressed. Like all other eukaryotic clocks it presumably functioned via feedback loops with positive and negative elements, but which protein acted as the main transcriptional repressor was unclear. We decided to test whether *clockwork orange* could repress mdCLK:mdCYC. It had previously been shown in *Drosophila*, that CWO is a transcriptional repressor that acts in synergy with dmPER to repress dmCLK:dmCYC (Zhou et al., 2016).

I first tested whether mdCWO behaved similarly to dmCWO using *Drosophila* activators. When *dmClk* and *mdCwo* were co-transfected in S2 cells, mdCWO repressed dmCLK:dmCYC with similar repression levels as dmCWO. Likewise, when *mdCwo* was co-transfected with *mdClk* and *mdCyc*, mdCWO was able to potently repress mdCLK:mdCYC. In fact, mdCWO was able to repress mdCLK:mdCYC with as little as 100 pg of mdCWO, whereas it took a thousand times as much dmCWO to repress mdCLK:mdCYC. This data suggests that in the housefly, CWO has taken on the role as the main transcriptional repressor of the clock. It is interesting that in *Drosophila*, PER acts as the main transcriptional repressor and CWO acts in synergy with PER to repress

transcription, but in closely related species like the housefly, the opposite appears to be true.

Once we established that mdCWO represses mdCLK:mdCYC, we decided to test whether we could repeat previous experiments that showed dmPER and dmCWO acted in synergy to repress dmCLK:dmCYC (Kadener et al., 2007). But, when we co-transfected varying concentrations of *dmPer* and a constant amount of *dmCwo* we saw no difference in repression levels from our control (*dmPer* transfected alone). We then tried co-transfecting varying concentrations of *dmCwo* and a constant amount of *dmPer* but we were unable to show a synergistic effect. It is possible that were unable to repeat the results of previous labs because we might not have used the same plasmids or we didn't use the same conditions than they did.

Next, we tested whether mdPER and mdCWO were able to co-repress mdCLK:mdCYC. However, when we co-transfected varying concentrations of *mdPer* and a constant amount of *dmCwo* we saw no difference in repression levels from our control (*mdPer* transfected alone). Then we tried co-transfecting varying concentrations of *mdCwo* and a constant amount of *mdPer*, but we were unable to show a synergistic effect. Because we were unable to show synergy between dmPER and dmCWO we could not determine for sure whether there really wasn't a synergistic effect between mdPER and mdCWO. However, these results suggested that the clock mechanism in the housefly behaves differently from the *Drosophila* clock.

Because mdCWO behaves so differently from dmCWO we looked at whether there was any synteny among dipterans with *cwo*. However, there were very few dipterans with published *cwo* sequences so we could not make any direct conclusions. From the small group of dipterans I compared, there was no synteny except among very closely related species such as the *Drosophila* species (*Drosophila melanogaster* and *Drosophila biarmipes*) and the medfly and the melon fly. Among *Drosophila* species there were several conserved genes in close proximity to *cwo*, indicating that at least among drosophilids genes around *cwo* have been conserved. Like in drosophilids, there was synteny of at least 3 genes near *cwo* from the medfly and the melon fly. However, none of the genes conserved near *cwo* in either the drosophilids or the medfly and melon fly appear in the other dipteran species. This suggests that *cwo* might have been translocated multiple times during the evolution of the dipteran lineage and may have led to its change in function overtime.

Conclusions

The results presented in this thesis have provided additional insights into the clock mechanism of the housefly. Houseflies do not have a mammalian *cryptochrome* and thus do not have an intermediate clock mechanism consisting of both *Drosophila* and ancestral clock components. This was disappointing as we hoped to find clues as to how the ancestral clock found in species from lower Diptera diverged into the *Drosophila* clock in Brachycera species. I found that the housefly has all the core *Drosophila* clock components: *Clk*, *cyc*, *per*, *tim*, *cwo*, and *cry1*.

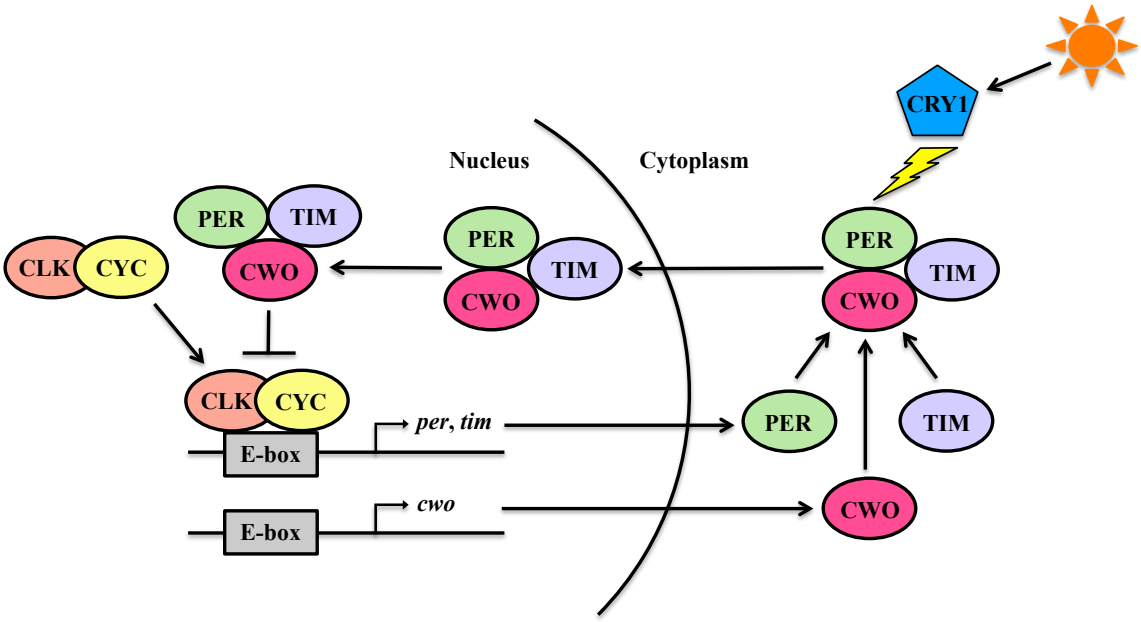

Figure 27: Proposed housefly clock. In the housefly, the clock is composed of the core components *Clock* (*Clk*), *cycle* (*cyc*), *period* (*per*), *timeless* (*tim*), *clockwork orange* (*cwo*), and *cryptochrome* (*cry1)*. However, unlike the *Drosophila* clock *mdCwo* has taken on the main role of transcriptional repressor. Although not known yet, perhaps *mdPer* and *mdTim* help stabilize *mdCwo* or serve some other function in the clock mechanism.

Despite sharing the same clock genes the clock mechanism of the housefly behaves in a very different manner than the *Drosophila* clock. Although the housefly clock still functions via autoregulatory feedback loops, my S2 cell studies show that unlike dmPER, mdCWO has taken over the role as the main transcriptional repressor (Figure 27). It is not yet known whether mdCWO acts in a complex with mdPER, mdTIM, or other proteins yet to be identified, in order to repress mdCLK:mdCYC. However, my experiments in S2 cells suggest that mdCWO does not act in synergy with mdPER to repress mdCLK:mdCYC (Figure 27).

The housefly has provided surprising evidence revealing that even between closely related insects, possessing the same core clock components does not necessarily mean that their clock mechanisms behave in an identical manner. However, most circadian work done on non-model insects (e.g. bean bugs and fire ants) suggest that because an insect has certain core clock components it is either mammalian-like or *Drosophila*-like. The knowledge gained from the housefly experiments described in this thesis argues that clock mechanisms of insects must be evaluated on a species by species basis rather than assuming the clock components in its clock must function like model insects based on the clock components it contains. With the unparalleled diversification of insects came significant diversification of the clock. The purpose of the clock is to enable organisms to adjust to changes in their environment. Thus, with insects inhabiting so many different niches around the world it makes sense that their clock mechanisms have diverged in many different ways too. This suggests that unraveling the evolution of the clock in Diptera and other insects may prove more complicated than initially expected.

Future Experiments

More work should be done to further unravel the housefly clock mechanism before it can truly be compared to those in other insects. First, it would be useful to define *cwo* gene expression in *M. domestica* heads. It would be interesting to see whether *mdcwo* would cycle in phase with either *Drosophila cwo* or *per*. In *Drosophila*, *per* mRNA peaks around ZT12 and *dmcwo* mRNA peaks at ZT14 (Hardin et al., 1990; Kadener et al., 2007). Codd et al., 2007 already showed mRNA levels for *per*, *tim*, *Clk*, *cyc*, and *cry1*

for the housefly. In the housefly *per* and *tim* mRNA peak around ZT12-ZT16 whereas *mdClk* mRNA peaked anti-phase to *mdPer*, and *mdCyc* mRNA was constitutively expressed (Codd et al., 2007). Unlike *Drosophila*, where *dmCry1* cycles rhythmically housefly *cry1* mRNA was constitutively expressed (Codd et al., 2007). Therefore, it would also be interesting to investigate further how *cry1* functions in the housefly clock mechanism. Perhaps neither TIM nor PER play major roles in the negative feedback loop of the housefly. Thus, it may be unnecessary for CRY1 to interact with TIM leading to its degradation and the eventual destabilization and degradation of PER (Rubin et al., 2006).

In addition to examining *cwo* mRNA levels, it would also be important to determine whether relative CWO protein levels in housefly heads cycle or are constitutively expressed. In *Drosophila*, CWO does not cycle (Zhou et al., 2016). Codd et al., 2007 already showed that mdPER protein levels do not cycle in either the head or the body of M. domestica. However, TIM does cycle with a peak at the end of the night phase (Codd et al., 2007). In LD both dmPER and dmTIM cycle robustly and in LL these rhythms in protein expression are dampened (Codd et al., 2007). However, a western blot of mdPER showed that when put in constant light mdPER protein levels are not dampened and only mdTIM is degraded in response to light (Codd et al., 2007). The inability of mdPER to repress, cycle in abundance, and be degraded by light suggests that mdPER may not be stabilized by mdTIM, further suggesting that it no longer serves a function in the housefly clock. Thus, to understand the housefly clock further it would be important to

determine whether other clock proteins in the housefly such as CRY1 and CLK cycle in abundance. In addition, it might also be useful to determine which clock proteins in the housefly interact. This might clear up the mystery of the function of mdPER. However, these experiments may prove difficult since *Drosophila* clock antibodies may not cross-react with housefly clock proteins.

It would be important to examine the subcellular localization of mdCWO throughout the day. Using immunohistochemistry of frozen sections the nuclear translocation of mdCWO could be determined. In some cells mdPER and mdTIM show time-dependent localization to the nucleus or cytoplasm, but in other cells these proteins remains strictly cytoplasmic (Codd et al., 2007). In addition to examining the localization of mdCWO, it would also be important to look at staining of mdCLK. Codd et al., 2007 was only able to show mdPER and mdTIM in a subset of cells, but perhaps by staining for mdCLK we could get a complete spatial pattern of clock cells in the housefly brain. From this we could determine the spatial pattern and number of clock cells in the housefly and compare it to *Drosophila*. However, these experiments may prove difficult on the short term if *Drosophila* CLK and CWO antibodies don't cross-react with housefly CLK and CWO.

REFERENCES

1. Allada, R., B. Y. Chung. *Circadian Organization of Behavior and Physiology in Drosophila*. Annu Rev Physiol, 2010. 72: p. 605-624.
2. Balys, M., E. Pyza. *Localization of the clock controlling circadian rhythms in the first neuropile of the optic lobe in the housefly*. J Exp Biol, 2001. 204: p. 3303-3310.
3. Beckingham, K. M., et al. *Drosophila melanogaster- The Model Organism of Choice for the Complex Biology of Multi-cellular Organisms*. Gravit Space Biol Bull, 2005. 18(2): p. 17-29.
4. Bell-Pedersen, D., et al. *Circadian Rhythms from Multiple Oscillators: Lessons from Diverse Organisms*. Nat Rev Genet, 2005. 6(7): p. 544-556.
5. Cato, A. J., et al. *Geographic Variation in Phosphine Resistance Among North American Populations of the Red Flour Beetle (Coleoptera: Tenebrionidae)*. J Econ Entomol, 2017. p. 1-7.
6. Chang, D. C., et al. *Constructing a Feedback Loop with Circadian Clock Molecules from the Silkmoth, Antheraea pernyi*. JBC, 2003. 278(40): p. 38149-38158.
7. Chang, D. C., S. M. Reppert. *A Novel C-Terminal Domain of Drosophila PERIOD inhibits dCLOCK:CYCLE-Mediated Transcription*. Current Biology, 2003. 13(9): p. 758-762.

8. Clayton, J. D., et al. *Keeping time with the human genome*. Nature, 2001. 409: p. 829-831.
9. Codd, V., et al. *Circadian Rhythm Gene Regulation in the Housefly Musca domestica*. Genetics, 2007. 177: p. 1539-1551.
10. Daiyasu, H. et al. *Identification of cryptochrome DASH from vertebrates*. Genes to Cells, 2004. 9: p. 479-495.
11. Dubowy C., A. Sehgal. *Circadian Rhythms and Sleep in Drosophila melanogaster*. Genetics, 2017. 205(4): p. 1373-1397.
12. Duran, C., et al. *Genetic maps and the use of synteny*. Methods Mol Biol, 2009. 513: p. 41-55.
13. Eban-Rothschild, A., et al. *The colony environment, but not direct contact with conspecifics, influences the development of circadian rhythms in honey bees*. J Biol Rhythms, 2012. 27(3): p. 217-225.
14. Friedrich, M., et al. *Phototransduction and clock gene expression in the troglobiont beetle Ptomaphagus hirtus of Mammoth cave*. Journal of Experimental Biology, 2011. 214: p. 3532-3541.
15. Friedrich, M. *Biological Clocks and Visual Systems in Cave-Adapted Animals at the Dawn of Speleogenomics*. Integr Comp Biol, 2013. 53(1): p. 50-67.
16. Froy, O., et al. *Illuminating the Circadian Clock in Monarch Butterfly Migration*. Science, 2003. 300(5623): p. 1303-1305.

17. Gustafson, C. L., et al. *A slow Conformational Switch in the BMAL1 Transactivation Domain Modulates Circadian Rhythms*. Mol Cell, 2017. 66(4): p. 447-457.

18. Hall, J. C. *Genetics and molecular biology of rhythms in Drosophila and other insects*. Adv Genet, 2003. 48: p. 1-280.

19. Hardin, P. E., et al. *Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels*. Nature, 1990. 343: p. 536-540.

20. Hardin, P. E. *The Circadian Timekeeping System of Drosophila*. Current Biology, 2005. 15: p. 714-722.

21. Hardin, P. E. *Molecular genetic analysis of circadian timekeeping in Drosophila*. Adv Genet, 2011. 74: p. 141-173.

22. Ikeno, T., H. Numata, S. G. Goto. *Molecular characterization of the circadian clock genes in the bean bug, Riptortus pedestris, and their expression patterns under long- and short-day conditions*. Gene, 2008. 419(1-2): p. 56-61.

23. Ikeno, T., et al. *Photoperiodic response requires mammalian-type cryptochrome in the bean bug Riptortus pedestris*. Biochem Biophys Res Commun, 2011. 410(3): p. 394-397.

24. Ikeno, T., et al. *Circadian clock gene Clock is involved in the photoperiodic response of the bean bug Riptortus pededtris*. Physiol Entomol, 2013. 38: p. 157-162.

25. Ingram, K. K., et al. *The Molecular Clockwork of the Fire Ant Solenopsis invicta*. PLoS One, 2012. 7(11): p. 1-11.

26. Jennings, B. H. *Drosophila- a versatile model in biology & medicine*. Materials Today, 2011. 14(5): p. 190-195.

27. Kadener, S., et al. *Clockwork Orange is a transcriptional repressor and a new Drosophila circadian pacemaker component*. Genes Dev, 2007. 21(13): p. 1675-1686.

28. Khedher, S. B., et al. *Combinatorial effect of Bacillus amyloliquefaciens AG1 biosurfactant and Bacillus thuringiensis Vip3Aa16 toxin on Spodoptera littoralis larvae*. J Invertebr Pathol, 2017. 144: p. 11-17.

29. Ko, C. H., J. S. Takahashi. *Molecular components of the mammalian circadian clock*. Human Molecular Genetics, 2006. 15(2): p. 271-277.

30. Kobayashi, Y., et al. *Molecular analysis of zebrafish photolyase/cryptochrome family: two types of cryptochromes present in zebrafish*. Genes to Cells, 2000. 5: p. 725-738.

31. Konopka, R. J., S. Benzer. *Clock Mutants of Drosophila melanogaster*. Proc Nat Acad Sci USA, 1971. 68(9): p. 2112-2116.

32. Krzeminski, W. *Triassic and Lower Jurassic stage of Diptera evolution*. Mitt Schweiz entomol Ges, 1992. 65: p. 39-59.

33. Levi, F., U. Schibler. *Circadian Rhythms: Mechanisms and Therapeutic Implications*. Annu Rev Pharmacol Toxicol, 2007. 47: p. 593-628.

34. Levine, J. D., et al. *Period protein from the giant silkmoth Antheraea pernyi functions as a circadian clock element in Drosophila melanogaster*. Neuron, 1995. 15(1): p. 147-157.

35. Li, C. J., et al. *Functional analysis of the circadian clock gene timeless in Tribolium castaneum*. Insect Sci, 2017. p. 1-11.

36. Maurer, C. et al. *Cytoplasmic interaction with CYCLE promotes the post-translational processing of the circadian CLOCK protein*. FEBS Letters, 2009. 583(10): p. 1561-1566.

37. Maywood, E. S., et al. *Genetic and Molecular Analysis of the Central and Peripheral Circadian Clockwork of Mice*. Cold Spring Harb Symp Quant Biol, 2007. 72: p. 85-94.

38. Mei, Q., V. Dvornyk. *Evolutionary History of the Photolyase/Cryptochrome Superfamily in Eukaryotes*. PLoS One, 2015. 10(9): p. 1-20.

39. Meireles-Filho, A. C., C. P. Kyriacou. Circadian rhythms in insect disease vectors. Mem Inst Oswaldo Cruz, 2013. 108(Suppl. 1): p. 48-58.

40. Merlin, C., et al. *An Antennal Circadian Clock and Circadian Rhythms in Peripheral Pheromone Reception in the Moth Spodoptera littoralis*. J Biol Rhythms, 2007. 22(6): p. 502-524.

41. Meuti, M. E., D. L. Denlinger. *Evolutionary Links Between Circadian Clocks and Photoperiodic Diapause in Insects*. Integrative and Comparative Biol, 2013. 53(1): p. 131-143.

42. Peschel, N., C. Helfrich-Förster. *Setting the clock- by nature: Circadian rhythm in the fruitfly Drosophila melanogaster*. FEBS Lett, 2011. 585(10): p. 1435-1442.

43. Piccin, A., et al. *The Clock Gene period of the Housefly, Musca domestica, Rescues Behavioral Rhythmicity in Drosophila melanogaster: Evidence for Intermolecular Coevolution?*. Genetics, 2000. 154: p. 747-75.

44. Pittendrigh, C. S., S. Daan. *A Functional Analysis of Circadian Pacemakers in Nocturnal Rodents*. J Comp Physiol, 1976. 106: p. 333-355.

45. Reppert, S. M. *The Ancestral Circadian Clock of Monarch Butterflies: Role in Time-compensated Sun Compass Orientation*. Cold Spring Harb Symp Quant Biol, 2007. 72: p. 113-118.

46. Reppert, S. M., et al. *Neurobiology of Monarch Butterfly Migration*. Annu Rev Entomol, 2016. 61: p. 25-42.

47. Roberts, D. B. *Drosophila melanogaster: the model organism*. Entomologia Experimentalis et Applicata, 2006. 121(2): p. 93-103.

48. Rosato, E., et al. *Molecular genetics of the fruit-fly circadian clock*. Eur J Hum Genet, 2006. 14(6): p. 729-738.

49. Rothenfluh, A., et al. *A TIMELESS-independent function for PERIOD Proteins in the Drosophila clock*. Cell, 2000. 26(2): p. 505-514.

50. Rubin, E. B., et al. *Molecular and phylogenetic analyses reveal mammalian-like clockwork in the honey bee (Apis mellifera) and shed new light on the molecular evolution of the circadian clock*. Genome Research, 2006. 16: p. 1352-1365.

51. Sakai, T., N. Ishida. *Circadian rhythms of female mating activity governed by clock genes in Drosophila*. Proc Natl Acad Sci USA, 2001. 98(16): p. 9221-9225.

52. Sandrelli, F., et al. *Comparative analysis of circadian clock genes in insects*. Insect Mol Biol, 2008. 17: p. 447-463.

53. Sauman, I., S. M. Reppert. *Circadian clock neurons in the silkmoth Antheraea pernyi: novel mechanisms of Period protein regulation*. Neuron, 1996. 17(5): p. 889-900.

54. Schoof, H. F. *Laboratory Culture of Musca, Fannia, and Stomoxys*. Bull World Health Organ, 1964. 31(4): p. 539-544.

55. Scott, J. G., et al. *A case for sequencing the genome of Musca domestica (Diptera: Muscidae)*. J Med Entomol, 2009. 46(2): p. 175-182.

56. Scott, J. G., et al. *Genome of the house fly, Musca domestica L., a global vector of diseases with adaptations to a septic environment*. Genome Biol, 2014. 15(10): p. 1-16.

57. Shlizerman, E., et al. *Neural Integration Underlying a Time-Compensated Sun Compass in the Migratory Monarch Butterfly*. Cell Rep, 2016. 15(4): p. 683-691.

58. Tataroglu, O., P. Emery. *Studying circadian rhythms in Drosophila melanogaster*. Methods, 2014. 68(1): p. 140-150.

59. Wiegmann, B. M., et al. *Episodic radiations in the fly tree of life*. PNAS, 2011. 108(14): p. 5690-5695.

60. Yuan, Q., et al. *Insect Cryptochromes: Gene Duplication and Loss Define Diverse Ways to Construct Insect Circadian Clocks*. Mol Biol Evol, 2007. 24(4): p. 948-955.

61. Zantke, J., et al. *Circadian and Circalunar Clock Interactions in a Marine Annelid*. Cell Rep, 2013. 5(1): p. 99-113.

62. Zantke, J., et al. *Genetics and Genomic Tools for the Marine Annelid Platynereis dumerilii*. Genetics, 2014. 197(1): p. 19-31.

63. Zhang, Q., B. Wang. *Evolution of Lower Brachyceran Flies (Diptera) and Their Adaptive Radiation with Angiosperms*. Front Plant Sci, 2017. 8(631): p. 1-6.

64. Zhao, X., D. Mu. *(6-4) Photolyase: Light-dependent repair of DNA damage*. Histol Histopathol, 1998. 13: p. 1179-1182.

65. Zheng, X., A. Sehgal. *Probing the relative importance of molecular oscillations in the circadian clock*. Genetics, 2008. 178(3): p. 1147-1155.

66. Zhou, J., et al. *CLOCKWORK ORANGE Enhances PERIOD Mediated Rhythms in Transcriptional Repression by Antagonizing E-box Binding by CLOCK-CYCLE*. PLoS Genet, 2016. 12(11): p. 1-17.

67. Zhu, H., et al. *The two CRYs of the butterfly*. Current Biology, 2005. 15(23): p. 953-954.

68. Zhu, H., et al. *Cryptochromes Define a Novel Circadian Clock Mechanism in Monarch Butterflies That May Underlie Sun Compass Navigation*. PLoS Biol, 2008. 6(1): p. 138-155.