CIRCADIAN CLOCK REGULATION OF PHOTOPERIODIC RESPONSES IN THE MONARCH BUTTERFLY

A Dissertation

by

SAMANTHA ELIZABETH IIAMS

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

DOCTOR OF PHILOSOPHY

Chair of Committee,	Christine Merlin
Committee Members,	Paul E. Hardin
	Mark J. Zoran
	Zachary N. Adelman
Head of Department,	David W. Threadgill

May 2021

Major Subject: Genetics

ABSTRACT

Since the dawn of life on Earth, organisms have had to overcome tremendous challenges in securing food, territory, and mates in order to survive daily and seasonal environmental fluctuations imposed by the rotation of the Earth around its own axis and its yearly revolution around the sun. Many organisms have adapted responses to the changes in seasons, by modulating their physiology and/or behavior to either withstand the changing conditions or move to more favorable locations. For organisms living at temperate latitudes, daylength (*i.e.* photoperiod) is the most reliable environmental cue of seasonal change. Old classical resonance and night interruption experiments, recently supported by genetic evidence, have suggested the involvement of biological clocks in regulating seasonal photoperiodic sensing/timing and response. However, the molecular mechanisms underlying photoperiodic sensing and responses have still remained poorly understood as most traditional laboratory model organisms lacked robust seasonal responses, as until recently found in flies, and genetic and genomic tools to address the question were lacking in non-traditional, seasonally responsive models. The North American monarch butterfly emerged as a suitable model organism to study the circadian regulation of seasonal responses because it is now possible to maintain healthy lab colonies and genetic and genomic tools that include a draft genome, a monarch specific cell line, and CRISPR/Cas9 to manipulate gene function *in vivo* are

available in this species. Just as importantly, monarch butterflies exhibit a very notable seasonal response: a long-distance migration accompanied by a reproductive arrest (called diapause) that are coincident with declining photoperiods in the fall. The goals of my Ph.D. were to 1) use genetic tools as well as physiological and behavioral assays to reveal and characterize the molecular pathways involved in seasonal photoperiodic responses; and 2) expand the “genetic toolbox” of the monarch butterfly to increase its usability as a model organism for the study of circadian and seasonal rhythms.

ACKNOWLEDGEMENTS

I would like to thank my committee chair and advisor, Dr. Christine Merlin, and the other members of my committee, Dr. Paul Hardin, Dr. Mark Zoran, and Dr. Zachary Adelman for their incredible guidance and rigor in my training over these past several years. The mentorship, leadership, and scientific expertise, especially from my advisor, has helped to shape me into the researcher I am now.

Thank you to the Department of Biology and Interdisciplinary Program of Genetics for providing the support and resources I needed, and as well thank you to my colleagues and friends for making my time here at Texas A&M such a great experience.

I want to also extend my gratitude to my family for their love and encouragement through the years. A final, special thanks to my fiancé and now husband for being so patient, loving, and supportive of me in completing my Ph.D. even though it meant we had to work long distance from each other. It has been worth it.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a committee consisting of Dr. Christine Merlin, Dr. Paul Hardin, and Dr. Mark Zoran of the Department of Biology and of Dr. Zachary Adelman of the Department of Entomology.

For Chapter 2, the research was designed by Dr. Merlin and myself. The research was carried out myself along with Dr. Merlin with her oocyte quantifications while in the lab of Dr. Steven M. Reppert that provided the data for Figs. 2.2 and 2.3 *B* and in generating the RNA-Seq data set, Aldrin B. Lugena and his generation of the heatmaps for Fig. 2.3 *A* and RNA-Seq analyses generating the tables of *Appendix A*, Dr. Ying Zhang in generating the RNA-Seq data set and performing the *ninaB1* eclosion assays for Fig. 2.6 *E*, and Ashley N. Hayden for assistance in generating the *ninaB1* knockout and performing the PER behavioral training assays of Fig. 2.6 *D*. Additionally, Matthew Markert and Justin Vann assisted with the wild-caught migrant collection; undergraduates Sarah Kenny, Kendall Bowen, Catherine Bogdan, and Jason Park assisted with animal husbandry; Dr. Jerome Menet and Ben Greenwell helped with the bioinformatic analyses; the Texas A&M AgriLife Genomics and Bioinformatics facility performed the sequencing services; and Dr. Menet and Dr. Hardin assisted with their commentary and feedback on the manuscript.

For Chapter 3, the research was designed by Dr. Merlin and myself. The work was performed myself with assistance from Dr. Zhang, Aldrin B. Lugena, Ashley N. Hayden, Catherine Bogdan, Anna Subonj, and Alec Judd in animal husbandry.

For Chapter 4, the research was designed by Dr. Merlin and myself. The research was carried out by mostly myself along with Ashley N. Hayden who I trained and worked with in generating the *dpCry1* knockout as shown in Fig. 4.1 and in performing the photoperiodic experiment in Fig. 4.5. Dr. Zhang, Aldrin B. Lugena, Catherine Bogdan, Anna Subonj, Alec Judd, and Jenna Coleman also assisted with the animal husbandry.

For Chapter 5, the research was designed by Dr. Merlin and myself. The generation of the PER::LUC reporter in monarch DpN1 cells was carried out and completed myself. The designing and generation of the reagents for the CRISPR/Cas9-assisted knock was carried out myself along with Catherine Bogdan, Anna Subonj, and briefly by undergraduate research scholar Mandy Eckhardt. The majority of the work to design and generate the transgenesis reagents for *piggybac*- or *tol2*-mediated transposition has been done myself with short term assistance from visiting scholar/masters student Mingqi Cai.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by TAMU Department of Biology Start up funds, NSF Grant IOS-1456985, and an award from the Klingenstein-Simons Foundation all awarded to Dr. Christine Merlin.

NOMENCLATURE

dp	*Danaus plexippus*
dm	*Drosophila melanogaster*
mm	*Mus musculus*
JH	Juvenile Hormone
CCG	Clock-Controlled Genes
Per	*Period* Gene Name or RNA Transcript
PER	*Period* Protein
Tim	*Timeless* Gene Name or RNA Transcript
TIM	*Timeless* Protein
Cry2	*Cryptochrome 2* Gene Name or RNA Transcript
CRY2	*Cryptochrome 2* Protein
Cry1	*Cryptochrome 1* Gene Name or RNA Transcript
CRY1	*Cryptochrome 1* Protein
Bmal1	*Brain and Muscle ARNT-Like 1* Gene Name or RNA Transcript
BMAL1	*Brain and Muscle ARNT-Like 1* Protein
TAD	Transactivation domain
Cyc	*Cycle* Gene Name or RNA Transcript
CYC	*Cycle* Protein
LP	Long Photoperiod

SP	Short Photoperiod
TC	Temperature Cycle
NinaB	*neither inactivation nor afterpotential B* Gene Name or RNA Transcript
NINAB	*neither inactivation nor afterpotential B* Protein
Rdh13	*retinol dehydrogenase 13* Gene Name or RNA Transcript
RDH13	*retinol dehydrogenase 13* Gene Protein
RA	Retinoic Acid
^{o}C	Temperature in Celsius
bp	Base Pair
hr	Hour
min	Minute
mL	Milliliter
µL	Microliter
ng	Nanogram
µg	Microgram
LD	Light:Dark cycle
LL	Constant Light Conditions
DD	Constant Dark Conditions

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1. The Seasonal Migration of the Monarch Butterfly.

The Eastern North American monarch butterfly (*Danaus plexippus, dp*) is a fascinating population of Lepidopterans that undertake each year, in response to decreasing day lengths (*i.e.* photoperiods) in the fall, a massive long-distance migration south to the mountain ranges of central Mexico in order to wait out the harsh northern winters [1] (Fig. 1.1 *A, Left*). Because the photoperiod is the most reliable indicator of seasonal change at temperate latitudes, photoperiodic changes have been proposed to trigger the induction of the migratory state [2] (Fig. 1.1 *B*). Also coincident with declining photoperiods, the monarchs enter a state of reproductive quiescence (*i.e.* diapause) induced by a juvenile hormone (JH) deficiency in order to conserve energy for the long flight ahead [3-6]. Entering this state of reproductive dormancy, alongside showing a marked increase in longevity from their typical ~1 month adult lifespan and exhibiting southward-oriented flight guided by a sun compass that is time-compensated by circadian clocks in the antennae, completes the monarch's induction into the migratory state [3, 7-11]. Upon arriving at the roosting sites in the central Mexican mountains, where they are known to preferably cluster in humid groves of oyamel fir trees to prevent drying out, diapausing migrants hunker down for the winter [1, 6]. When photoperiods and temperature increase in the spring, they become reproductive again, and mate in mass before re-migrating

northwards [6, 12]. Interestingly, the prolonged exposure to overwintering cold temperatures at the roost sites has been shown to effectively reverse flight orientation [13]. Repopulation of their northern ranges occurs over several successive, non-migratory generations which, in contrast to migrants, are reproductive, short lived, and non-oriented in flight. Instead, the monarchs follow the northern emergence of the milkweed plant, the staple of the monarch larvae diet [5, 14] (Fig. 1.1 *A, Right*). The following fall, the progeny of the summer generation cycle starts a new cycle of migration [5]. The fact that migrants are the progeny of non-migrants, and thus share the same core genome, strongly suggests that the seasonal induction of the migratory state is likely triggered by a change in gene expression [15]. If the photoperiod is triggering the induction of the migratory state, measuring RNA levels at a genome-wide scale from the brains of monarchs raised in different photoperiods would allow for the identification of seasonal changes in gene expression and potentially the specific genes and pathways involved in regulating photoperiodic responses

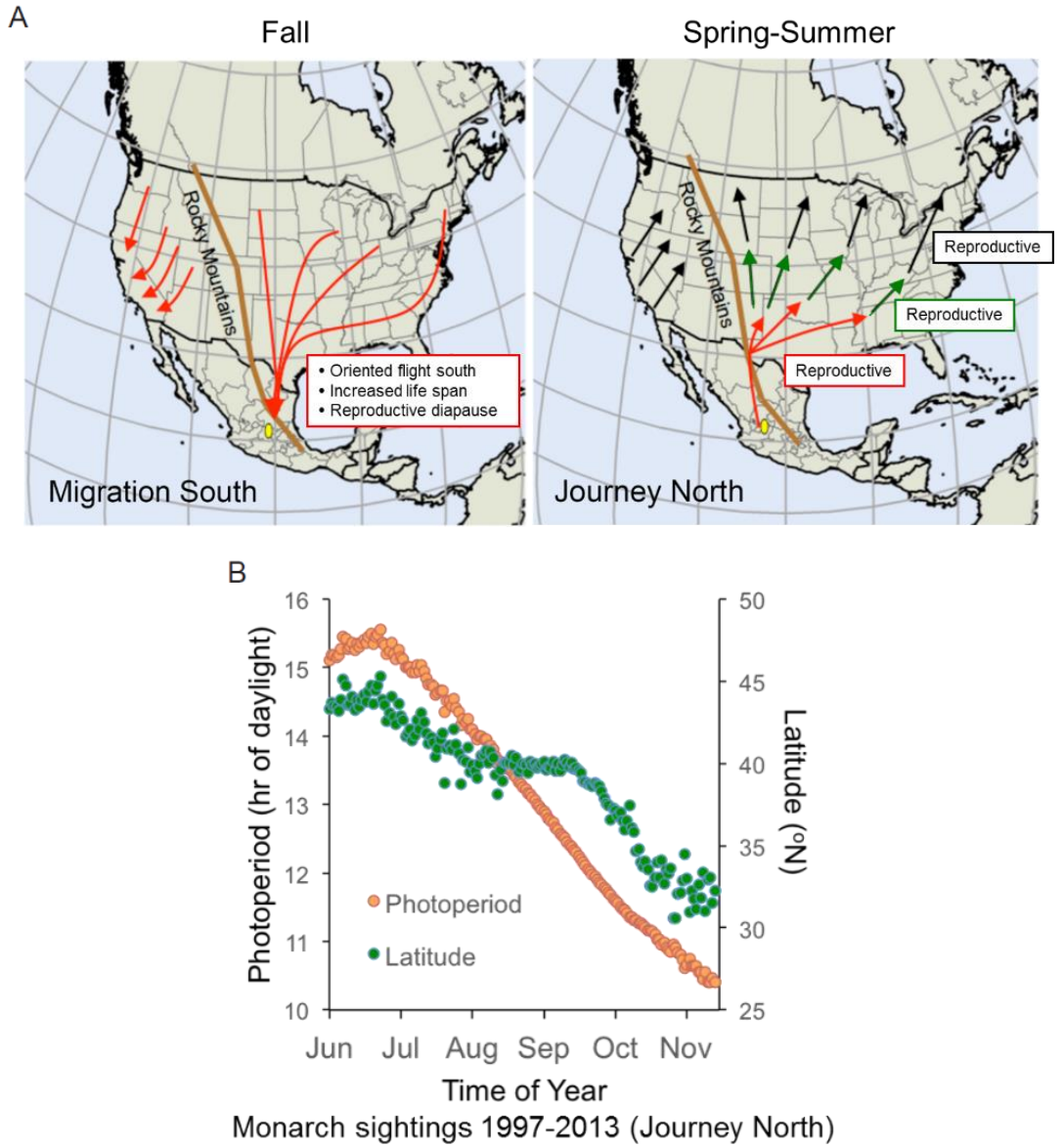


Figure 1.1: The North American monarch migration.
(*A*) (Left panel) Southerly fall migration. Eastern population migrates south to central Mexico. Western population migrates south to southern California. (Right panel) Spring re-migration (red arrows) and successive spring and summer generations (green and black arrows) [9]. (*B*) Daily average latitude of all sighted monarchs from June through November (green dots) from 1997-2013, plotted over the hours of daylight (orange dots) experienced by monarchs at that average latitude each day.

1.2. The Monarch Butterfly Molecular Circadian Clock.

Many organisms have evolved endogenous timekeeping systems in order to synchronize their behavior and physiology with that of the 24 hr rotation of the Earth on its axis. These timekeeping systems, or circadian clocks, oscillate with a stable phase typically entrained by the sun's light:dark (LD) cycles making it possible for organisms anticipate the daily changes in the environment around them. By regulating the 24 hr transcription of clock controlled genes (CCGs), the clock can influence a wide array of behavioral functions such as timing of foraging, migration, and mating, and as well physiological rhythms in body temperature, metabolism, and biosynthesis pathways [16].

The molecular components constituting the core circadian clock were initially characterized in flies and mammals revealing a transcriptional-translational negative feedback loop. [17-23]. In *Drosophila* (*dm*), this loop consists of activators CLOCK (dmCLK) and CYCLE (dmCYC), both members of the basic-helix-loop-helix (bHLH) transcription factor family, which form a heterodimer and bind to enhancer boxes (E-box) to drive the rhythmic transcription of *period* (*dmPer*) and *timeless* (*dmTim*) [19-22, 24, 25]. DmPER and the kinase DOUBLE-TIME (dmDBT) complex together, possibly with or without dmTIM, to repress dmCLK:dmCYC-mediated transcription [18, 26-29]. A blue-light photoreceptor CRYPTOCROME (dmCRY) mediates the degradation of dmTIM in a light-induced manner [30]. Mammalian (*mm*) clocks also utilize the activator mmCLK which instead heterodimerizes to Brain and Muscle ARNT-Like

1 (mBMAL1), a bHLH transcription factor similar to dmCYC but containing an additional trans-activation domain (TAD), to drive the rhythmic transcription of *mmPer1*, *mmPer2*, *mmCry1*, and *mmCry2*. [23, 31]. MmCRYs appear to no longer function in light-dependent roles and instead serve alongside mmPERs to repress mmCLK:mmBMAL1-mediated transcription [31].

The molecular framework of the monarch circadian clock has revealed a hybrid negative feedback loop composed of both mammalian-like and insect-like components [32-36] (Fig. 1.2). Instead of a *Drosophila*-like CYC activator as might be anticipated in an insect model, monarchs use the mammalian-like dpBMAL1 to form a heterodimer with dpCLK and initiate the transcription of several clock genes, including a mammalian-like *dpCry2*, *dpPer,* and *dpTim*. DpPER, dpTIM, and dpCRY2 form complexes in the cytoplasm where they translocate back into the nucleus for dpCRY2 to repress dpCLK:dpBMAL1-mediated transcription. The monarch also possesses a blue-light, sensitive photoreceptor dpCRY1. Like dmCRY [37, 38], dpCRY1 has been shown to be light-sensitive and mediate dpTIM degradation in Dpn1 cells, a monarch cell line containing a light-driven clock [32], however, its role in circadian resetting and synchronization *in vivo* remains to be tested.

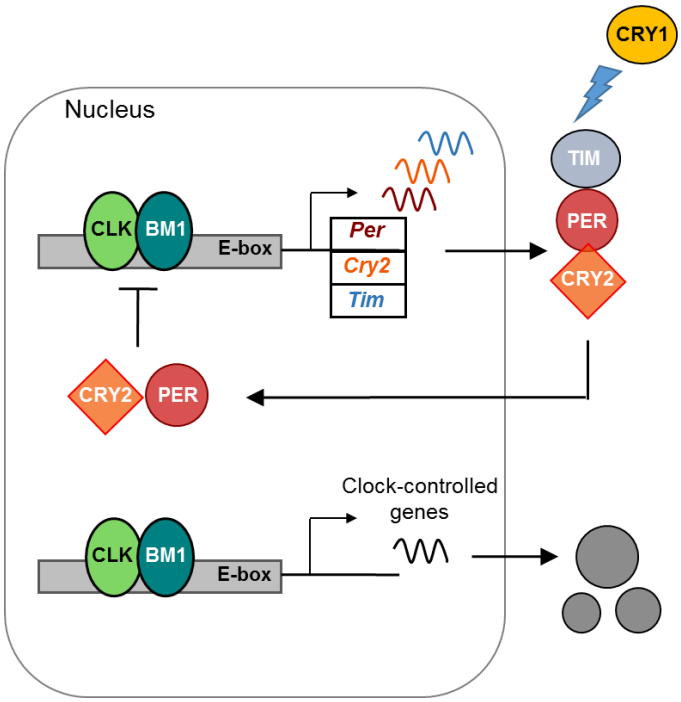


Figure 1.2: The core transcriptional/translational feedback loop of the monarch circadian clock.

Activators CLOCK (CLK) and BMAL1 (BM1) form a heterodimer and bind to enhancer boxes (E-box) to drive the rhythmic transcription of core clock genes *period* (*per*), *cryptochrome 2* (*cry2*) and *timeless* (*tim*). These genes are translated into their functional proteins outside of the nucleus. Upon activation by light, CRYPTOCHROME 1 (CRY1) mediates the degradation of TIM. PER and CRY2 enter the nucleus to rhythmically inhibit CLOCK:BMAL1-mediated transcription. In addition to its own self-regulation, the circadian clock controls many other clock-controlled genes to regulate the biology of the organism. Modified [32].

1.3. Role of the Circadian Clock in Seasonal Photoperiodic Responsiveness.

The photoperiod has been shown to regulate many aspects of physiology and behavior across taxa including dormancy in plants, migration, moulting, gonadal responses in mammals and birds, and developmental or reproductive diapause in insects [39-41]. Classical resonance experiments (denoted Nanda-

Hamner in insects) which expose organisms to a set period of light and varying durations of dark periods revealing peaks and troughs of diapause as well as night-interrupting light pulse experiments have long suggested the role of the circadian clock in regulating photoperiodic responses, and has been genetically tested in insects where ablation of clock neurons or knockdown of clock genes *cry2*, *per*, *tim*, and *cyc* abolished photoperiodic responses [41-48]. Whether the circadian clock functions as a unit to regulate photoperiodic timing or if individual clock genes exert pleiotropic effects is still debated. Some models of photoperiodism have suggested the involvement of the circadian clock, such as the oscillatory external coincidence model in which long day effects are observed when light enters in the dark-phase (scotophil) and short day effects are only seen when light is restricted to the light-phase (photophil) [49, 50]. Evidence supporting this hypothesis has shown that interruption of the scotophil via light pulses disrupted diapausing responses in insects thus suggesting that photoperiodic timing was disrupted [49, 51, 52]. Alternatively, an internal coincidence model involving multiple oscillators has also been suggested where the photoperiod alters the phase-relationship of two or more rhythms, synchronizing them into either a permissive state or inhibitory state, as observed in parasitic wasps which required accumulated exposure to short photoperiods to exhibit a response [50]. Another hypothesis, the hourglass model, instead argues against the involvement of circadian clocks and suggests that photoperiodic timing results from the gradual accumulation of a chemical product

throughout the day where in long day conditions it can reach its threshold to trigger physiological response in contrast to short days where accumulation is cut short [49]. Because genetically tractable models such as mice, and as until recently flies [53], lacked robust seasonal responses, and genetic tools are lacking in most non-traditional model organisms, the role of the circadian clock in regulating photoperiodic responses has remained poorly understood. Testing the role of the circadian clock in another seasonally responsive species, such as the monarch butterfly, could help to further disentangle the mechanism of photoperiodic timing and, if the clock is shown to be involved, reveal the clock-controlled genes and pathways underlying photoperiodic responses.

1.4. The Monarch Butterfly as a Laboratory Model Organism: Life Cycle and Maintenance.

Each generational cycle of the monarch butterfly takes approximately 40 days at 25°C to complete from egg to larvae to pupae to a reproductively active adult. Once a female is mated, which is done by hand pairing the male and female together, the male's sperm is stored in the spermatheca. The egg is fertilized at oviposition through sperm deposition into an orifice at the tip of the egg called the micropyle [54]. The male pronucleus and female pronucleus migrate from this anterior pole to the center of the egg to form the zygote nucleus before dividing into many uncellularized nuclei [54]. Following migration toward the edges of the egg, the nuclei cellularize and continue dividing to form

the monarch embryo. The embryo takes four days to fully develop before hatching as a 1st instar larvae by chewing their way through the egg shell and begin immediately feeding on milkweed leaves, the sole food source of monarch larvae. At the 2nd instar, individual larvae are switched to an artificial diet containing freeze dried milkweed, starch, and other nutrients (https://monarchwatch.org/). This procedure to maintain laboratory stocks composed of large number of butterflies is necessary as a single larva could eat through 3-5 milkweed plants over the course of its development and growing the amount of plants that would be necessary is unpractical. The larvae are maintained on this artificial diet, which is changed every three to four days, through the 3rd, 4th, and 5th instars. At the end of the 5th instar, the frass deposited by the larvae shifts from brown to red indicating it has begun clearing out the gut to prepare for pupation. Soon after the larvae spins a silk pad from a spinneret located at the bottom of the head, attaches itself upside down to the pad, and pupates. During pupation, the larval tissues are lysed but the imaginal discs, which are clusters of embryonic cells, divide and develop to form the adult tissues and structures over the course of the next 9-10 days. The adults then emerge from the chrysalis, a behavior which has been shown to be linked to the early morning light phase and under circadian regulation [35, 55]. The adults are fed daily a 25% honey mixture in water to simulate the nectar they would feed on in the wild. At day 4 post-eclosion, monarchs are reproductively mature and crossed by hand where they remain latched together for just under a day.

Females are then placed in a cage with a milkweed plant to lay their eggs over the course of the day, which are collected and a new cycle begins.

1.5. The Monarch as a Laboratory Model Organism: Molecular, Genetic and Genomic Tools.

The monarch butterfly has emerged as a suitable model organism to study circadian clock regulation of photoperiodic responses given its robust seasonal migratory behavior/diapause and our ability to maintain laboratory colonies. Additionally, we also have several cutting-edge tools including a draft genome [56], a DpN1 cell culture line [57], and the CRISPR/Cas9 technology for genome editing to be able to identify the molecular underpinnings of seasonal responses [36, 58, 59].

The assembly of a monarch draft genome has identified an estimated 16,866 protein-coding genes in the 273 Mb genome [56]. The availability of this genome along with its integration into an easy to search database has provided a foundation for population genetic analyses and comparative analyses between species to begin uncovering the molecular pathways underlying migration [56, 60]. Access to individual monarch gene sequences facilitates the utilization of many essential molecular techniques such as PCR-based genotyping, real-time PCR, cloning, RNAi, high-throughput genome and transcriptome sequencing, and genome editing [36, 58, 59].

Another important tool that has been developed in the monarch, is an embryo-derived DpN1 cell culture line [57]. These cells have the potential to make a significant impact to the chronobiology field as they represent the only known insect cell line to date that possesses a functioning light-driven clock where *dpPer*, *dpTim*, and *dpCry2* cycle under light:dark cycles with a similar phase to that in the monarch brain [32]. As previously mentioned, DpN1 cells have already been used to demonstrate the light sensitivity of dpCRY1 and its mediation of dpTIM degradation, as well as the role of dpCRY2 in transcriptional repression [32]. This cell line has broad appeal for use in the molecular dissection of clockwork mechanisms *in vitro*.

One of the more recently developed tools used in the monarch butterfly is targeted mutagenesis via the CRISPR/Cas9 system, a bacterial-based genome editing system [58]. Bacteria use clustered regularly interspaced short palindromic repeats (CRISPR) integrated into their genome as a part of their immune system to detect and destroy invading bacteriophage DNA with the help of a Cas9 endonuclease [61]. The re-engineering of the bacterial Cas9 endonuclease into a simple two-component system relying on a Cas9 protein to induce a double stranded (ds) DNA break and a designed single guide RNA (sgRNA) that complexes with Cas9 to guide it to a target site of choice has revolutionized the field of genome editing and made this system widely applicable, including in unconventional species [62-67]. This system has been deployed in the monarch to generate gene knockouts by injecting eggs with the

selected sgRNA and Cas9 mRNA within 20 min of fertilization to target nuclei early during their first divisions, as this increases chances that germ line precursors are targeted [58]. Upon dsDNA break induction, cells try to repair the dsDNA cut using one of two mechanisms, the first being homologous recombination (*i.e.* homology directed repair (HDR)) which perfectly repairs the break by using a homologous segment of DNA (ex: the paired chromosome) as template. The other mechanism relies on the non-homologous end-joining pathway (NHEJ), an imperfect repair mechanism that results in the insertion and/or deletion of base pairs at the cut site. Any mutation that is not a multiple of three causes a frameshift mutation, which either alters the amino acid sequence or generates an early stop codon to produce a truncated protein. The CRISPR/Cas9 system has already been used to generate several NHEJ-mediated mutants of core clock genes in the monarch with high efficiency [58]. The ease and effectiveness of CRISPR/Cas9-mediated mutagenesis in the monarch thus allows for the rapid knockout of clock and photoperiodic candidate genes to test their function on seasonal responses, as well as to test the yet unknown role of dpCRY1 in clock entrainment *in vivo.*

As mentioned above, aside from the NHEJ repair pathway, cells often rely on HDR to perfectly repair dsDNA breaks. This property has successfully been leveraged as a method for knocking-in DNA into the genome to insert either transgenes, substitute alleles, or generate gene reporters both *in vitro* and *in vivo* [63, 68-75]. With its status as a newly emerging model species, the

monarch butterfly currently lags behind in the development of knock-in approaches both *in vitro* and *in vivo*. The development of the CRISPR/Cas9 system in this species offers unique opportunities to develop CRISPR-induced HDR to knock-in reporters, including a PERIOD::LUCIFERASE (PER::LUC) reporter in DpN1 cells to generate a real-time reporter of circadian rhythms, similar to what has been done in mammals [72]. Additionally, CRISPR/Cas9 gives us the capability to test knock-in strategies *in vivo* to tag clock neurons with a fluorescent reporter and ultimately map the clock neuronal circuitry in the monarch brain. However, because the use of selective markers for *in vivo* knock-ins is more challenging than in cells, and HDR efficiency differs between cell types and organisms, another potential strategy to consider is transgenesis. In other insect species, transgenesis has been shown to be relatively efficient, steadily expresses the transgene, and has no size limit on the DNA insertion [76-79]. Two of these transposon systems will be tested in the monarch butterfly, piggybac and Tol2, where piggybac inserts at “TTAA” sites and Tol2 inserts at random sites throughout the genome [76-79]. The use of either CRISPR/Cas9-mediated HDR, piggybac- or Tol2-mediated transgenesis will be key to generating an *in vivo* clock neuron reporter to mark clock neurons and an *in vitro* PER::LUC reporter in DpN1 cells, which would further elevate the relevance and usability of the monarch butterfly as a model species for the study of circadian clocks and the regulation of seasonal, photoperiodic responses.

1.6. Specific Aims.

Aim 1: Test the role of monarch circadian clock genes in the regulation of photoperiodically-induced reproductive diapause.

In order to address this goal, I will test the photoperiodically-induced reproductive diapause responses of previously generated clock mutants: a knockout of the circadian repressor *dpCry2*, a knockout of the circadian activator *dpClk*, and a mutant of the activator *dpBmal1* lacking the transactivation domain, similar to *Drosophila dmCyc* (*dpCyc-like)*. Mutants and their control wild-type siblings will be subjected to either long-photoperiod (LP) (15 hr light:9 hr dark) or short-photoperiod (SP) conditions (10 hr light:14 hr dark) at 21^{o}C from egg to adult. Adult females will be dissected on the 14 days post-eclosion and the number of mature oocytes produced counted with averages compared between LP and SP to assess photoperiodic responses.

Aim 2: Identify and characterize clock-controlled, seasonally expressed genes involved in photoperiodically-induced diapause responses.

To look for a seasonal change in gene expression and identify genes potentially involved in photoperiodic responses, we will compare previously generated 24 hr RNA expression profiles generated by high-throughput sequencing from monarch brains raised in different photoperiods: laboratory-raised long-day vs short-day photoperiodic conditions, and laboratory-raised

summer-like vs wild-caught migratory forms. Candidate genes will be identified from an overlap analysis comparing the rhythmicity of genes across all photoperiodic conditions (*i.e.* rhythmic in all, rhythmic in only long day and summer-like, or rhythmic in only short day and wild migrants) and subsequently chosen for functional characterization based on their differential expression between photoperiods and their gene ontology. Chosen candidates will be functionally knocked out *in vivo* using CRISPR/Cas9-mediated mutagenesis and tested for loss of photoperiodic responses.

Aim 3: Test and characterize the role of monarch blue-light, sensitive dpCRY1 in circadian entrainment and in photoperiodically-induced diapause responses.

Given the demonstrated role of dmCRY in entrainment and circadian resetting, we aim to generate a *dpCry1* knockout to test the effect on circadian entrainment in the monarch to determine if there is a conserved or different mode of entrainment between monarch and *Drosophila*. I will test the effect a *dpCry1* knockout on rhythmic eclosion behavior and also measure the expression of *dpPer* and *dpTim* over the course of a circadian day in monarch brains. Additionally, *dpCry1* knockouts will also be tested for their effect on photoperiodically-induced reproductive diapause in order to continue testing the role of the circadian clock in regulating photoperiodic responses.

Aim 4: Establish a PERIOD::LUCIFERASE (PER::LUC) reporter line in monarch DpN1 cells using a CRISPR-based knock-in strategy.

I will use the CRISPR/Cas9 system to create a double stranded break in the C-terminus region of *dpPer*, a core clock gene, and utilize HR at the cut site to knock-in a PER::LUC reporter construct in monarch DpN1 cells. To accomplish this, I will co-transfect an "All-In-One" vector containing Cas9 and the selected sgRNA and a PER::LUC luciferase knock-in construct containing a *hygromycin* selection cassette flanked by Lox p sites. Once a candidate recombinant cell line has been identified, I will "flox" out the *hygromycin* cassette at the knock-in site using Cre recombinase and screen for any unintentional genomic integrations of *Cas9*, *Cre*, *hygromycin*, and *puromycin* (from the "All-in-one" construct but not selected for) by PCR. The PER::LUC DpN1 cell line will represent a useful real-time reporter to rapidly uncover cell autonomous clockwork mechanisms *in vitro*, and will be used to assess the role of dpCRY1 in clock entrainment *in vitro*.

Aim 5: Develop knock-in strategies to tag clock neurons *in vivo* to map the circadian neural circuits in monarch brains.

For this aim I will test several *in vivo* strategies to insert a fluorescent reporter, either *mCherry* or *tdTomato*, under the control of either the *dpCry2* or *dpClk* promoter in the monarch genome using either HDR-mediated knock-in with the CRISPR/Cas9 system, and piggybac-mediated or Tol2-mediated

transgenesis. Monarch eggs will be injected with one of these three systems and tested for the presence of the reporter from the tissue of surviving larvae (G0) and their progeny (G1) in order to determine which system, and what quantities of reagents are most efficient. If a strategy is identified, it will be used to generate a fluorescent reporter monarch line that can be used to map the clock neuronal circuitry, and potentially test for seasonal neuronal remodeling.

1.7. References.

1. Urquhart, F.A., *The monarch butterfly*. 1960, Toronto: University of Toronto Press. xxiv, 361 p.
2. Denlinger, D.L., et al., *Keeping time without a spine: what can the insect clock teach us about seasonal adaptation?* Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1734).
3. Reppert, S.M., R.J. Gegear, and C. Merlin, *Navigational mechanisms of migrating monarch butterflies.* Trends Neurosci, 2010. **33**(9): p. 399-406.
4. Herman, W.S., *Reproductive tract development in monarch butterflies overwintering in California and Mexico.* Journal of the Lepidopterists' Society, 1989. **43**: p. 50-58.
5. Goehring, L. and K.S. Oberhauser, *Effects of photoperiod, temperature, and host plant age on induction of reproductive diapause and development time in Danaus plexippus.* Ecological Entomology, 2002. **27**: p. 674-685.

6. Brower, L., *Monarch butterfly orientation: missing pieces of a magnificent puzzle.* J Exp Biol, 1996. **199**(Pt 1): p. 93-103.

7. 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. 113-8.

8. Merlin, C., R.J. Gegear, and S.M. Reppert, *Antennal circadian clocks coordinate sun compass orientation in migratory monarch butterflies.* Science, 2009. **325**(5948): p. 1700-4.

9. Mouritsen, H. and B.J. Frost, *Virtual migration in tethered flying monarch butterflies reveals their orientation mechanisms.* Proc Natl Acad Sci U S A, 2002. **99**(15): p. 10162-6.

10. Perez, S.M., O.R. Taylor, and R. Jander, *A sun compass in monarch butterflies.* Nature, 1997. **387**.

11. Guerra, P.A., et al., *Discordant timing between antennae disrupts sun compass orientation in migratory monarch butterflies.* Nat Commun, 2012. **3**: p. 958.

12. Reppert, S.M., P.A. Guerra, and C. Merlin, *Neurobiology of Monarch Butterfly Migration.* Annu Rev Entomol, 2016. **61**: p. 25-42.

13. Guerra, P.A. and S.M. Reppert, *Coldness triggers northward flight in remigrant monarch butterflies.* Curr Biol, 2013. **23**(5): p. 419-23.

14. Flockhart, D.T., et al., *Tracking multi-generational colonization of the breeding grounds by monarch butterflies in eastern North America.* Proc Biol Sci, 2013. **280**(1768): p. 20131087.

15. Merlin, C. and M. Liedvogel, *The genetics and epigenetics of animal migration and orientation: birds, butterflies and beyond.* J Exp Biol, 2019. **222**(Pt Suppl 1).

16. Ueda, H.R., et al., *Genome-wide transcriptional orchestration of circadian rhythms in Drosophila.* Journal of Biological Chemistry, 2002. **277**(16): p. 14048-14052.

17. Konopka, R.J. and S. Benzer, *Clock mutants of Drosophila melanogaster.* Proc Natl Acad Sci U S A, 1971. **68**(9): p. 2112-6.

18. Hardin, P.E., J.C. Hall, and M. Rosbash, *Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels.* Nature, 1990. **343**(6258): p. 536-40.

19. Allada, R., et al., *A mutant Drosophila homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless.* Cell, 1998. **93**(5): p. 791-804.

20. Rutila, J.E., et al., *CYCLE is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of Drosophila period and timeless.* Cell, 1998. **93**(5): p. 805-14.

21. Myers, M.P., et al., *Positional cloning and sequence analysis of the Drosophila clock gene, timeless.* Science, 1995. **270**(5237): p. 805-8.

22. Sehgal, A., et al., *Rhythmic expression of timeless: a basis for promoting circadian cycles in period gene autoregulation.* Science, 1995. **270**(5237): p. 808-10.

23. King, D.P., et al., *Positional cloning of the mouse circadian clock gene.* Cell, 1997. **89**(4): p. 641-53.

24. Darlington, T.K., et al., *Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim.* Science, 1998. **280**(5369): p. 1599-603.

25. Hao, H., D.L. Allen, and P.E. Hardin, *A circadian enhancer mediates PER-dependent mRNA cycling in Drosophila melanogaster.* Mol Cell Biol, 1997. **17**(7): p. 3687-93.

26. Bae, K., et al., *dCLOCK is present in limiting amounts and likely mediates daily interactions between the dCLOCK-CYC transcription factor and the PER-TIM complex.* J Neurosci, 2000. **20**(5): p. 1746-53.

27. Lee, C., K. Bae, and I. Edery, *The Drosophila CLOCK protein undergoes daily rhythms in abundance, phosphorylation, and interactions with the PER-TIM complex.* Neuron, 1998. **21**(4): p. 857-67.

28. Menet, J.S., et al., *Dynamic PER repression mechanisms in the Drosophila circadian clock: from on-DNA to off-DNA.* Genes Dev, 2010. **24**(4): p. 358-67.

29. Yu, W., et al., *PER-dependent rhythms in CLK phosphorylation and E-box binding regulate circadian transcription.* Genes Dev, 2006. **20**(6): p. 723-33.

30. Ceriani, M.F., et al., *Light-dependent sequestration of TIMELESS by CRYPTOCHROME.* Science, 1999. **285**(5427): p. 553-6.

31. Takahashi, J.S., *Transcriptional architecture of the mammalian circadian clock.* Nat Rev Genet, 2017. **18**(3): p. 164-179.

32. Zhu, H.S., et al., *Cryptochromes define a novel circadian clock mechanism in monarch butterflies that may underlie sun compass navigation.* Plos Biology, 2008. **6**(1): p. 138-155.

33. Yuan, Q., et al., *Insect cryptochromes: Gene duplication and loss define diverse ways to construct insect circadian clocks.* Molecular Biology and Evolution, 2007. **24**(4): p. 948-955.

34. Zhu, H.S., et al., *The two CRYs of the butterfly.* Current Biology, 2005. **15**(23): p. R953-R954.

35. Merlin, C., et al., *Efficient targeted mutagenesis in the monarch butterfly using zinc-finger nucleases.* Genome Res, 2013. **23**(1): p. 159-68.

36. Zhang, Y., et al., *Vertebrate-like CRYPTOCHROME 2 from monarch regulates circadian transcription via independent repression of CLOCK and BMAL1 activity.* Proc Natl Acad Sci U S A, 2017. **114**(36): p. E7516-E7525.

37. Emery, P., et al., *CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity.* Cell, 1998. **95**(5): p. 669-79.

38. Stanewsky, R., et al., *The cryb mutation identifies cryptochrome as a circadian photoreceptor in Drosophila.* Cell, 1998. **95**(5): p. 681-92.

39. Bunning, E., *The physiological clock*. Longmans Springer-Verlag New York 1967.

40. Hazlerigg, D.G. and G.C. Wagner, *Seasonal photoperiodism in vertebrates: from coincidence to amplitude.* Trends Endocrinol Metab, 2006. **17**(3): p. 83-91.

41. Saunders, D.S., *Insect photoperiodism: measuring the night.* J Insect Physiol, 2013. **59**(1): p. 1-10.

42. Ikeno, T., H. Numata, and S.G. Goto, *Photoperiodic response requires mammalian-type cryptochrome in the bean bug Riptortus pedestris.* Biochem Biophys Res Commun, 2011. **410**(3): p. 394-7.

43. Ikeno, T., H. Numata, and S.G. Goto, *Circadian clock genes period and cycle regulate photoperiodic diapause in the bean bug Riptortus pedestris males.* J Insect Physiol, 2011. **57**(7): p. 935-8.

44. Ikeno, T., et al., *Involvement of the brain region containing pigment-dispersing factor-immunoreactive neurons in the photoperiodic response of the bean bug, Riptortus pedestris.* J Exp Biol, 2014. **217**(Pt 3): p. 453-62.

45. Ikeno, T., et al., *Photoperiodic diapause under the control of circadian clock genes in an insect.* BMC Biol, 2010. **8**: p. 116.

46. Shiga, S. and H. Numata, *Roles of PER immunoreactive neurons in circadian rhythms and photoperiodism in the blow fly, Protophormia terraenovae.* J Exp Biol, 2009. **212**(Pt 6): p. 867-77.

47. Ikegami, K. and T. Yoshimura, *Seasonal time measurement during reproduction.* J Reprod Dev, 2013. **59**(4): p. 327-33.

48. Saunders, D.S. and R.C. Bertossa, *Deciphering time measurement: the role of circadian 'clock' genes and formal experimentation in insect photoperiodism.* J Insect Physiol, 2011. **57**(5): p. 557-66.

49. Saunders, D.S., *Insect Clocks*. 2002: Elsevier. 576.

50. Pittendrigh, C.S., *Circadian surfaces and the diversity of possible roles of circadian organization in photoperiodic induction.* Proc Natl Acad Sci U S A, 1972. **69**(9): p. 2734-7.

51. Beck, S.D., *Photoperiodic Induction of Diapause in an Insect.* Biological Bulletin 1962. **122**(1): p. 1-12.

52. Barker, R.J., A. Mayer, and C.F. Cohen, *Photoperiod effects in Pieris rapae.* Annals of the Entomological Society of America, 1963. **56**: p. 292-294.

53. Nagy, D., et al., *A Semi-natural Approach for Studying Seasonal Diapause in Drosophila melanogaster Reveals Robust Photoperiodicity.* J Biol Rhythms, 2018. **33**(2): p. 117-125.

54. Kobayashi, Y., M. Tanaka, and H. Ando, *Embryology*, in *Lepidoptera, moths and butterflies: Morphology, physiology, and development*, N. Kristensen, Editor. 2003: De Gruyter, Berlin. p. 495–544.

55. Froy, O., et al., *Illuminating the circadian clock in monarch butterfly migration.* Science, 2003. **300**(5623): p. 1303-5.

56. Zhan, S., et al., *The monarch butterfly genome yields insights into long-distance migration.* Cell, 2011. **147**(5): p. 1171-85.

57. Palomares, L.A., et al., *Novel insect cell line capable of complex N-glycosylation and sialylation of recombinant proteins.* Biotechnol Prog, 2003. **19**(1): p. 185-92.

58. Markert, M.J., et al., *Genomic Access to Monarch Migration Using TALEN and CRISPR/Cas9-Mediated Targeted Mutagenesis.* G3 (Bethesda), 2016. **6**(4): p. 905-15.

59. Iiams, S.E., et al., *Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly.* Proc Natl Acad Sci U S A, 2019.

60. Zhan, S. and S.M. Reppert, *MonarchBase: the monarch butterfly genome database.* Nucleic Acids Res, 2013. **41**(Database issue): p. D758-63.

61. Barrangou, R., *The roles of CRISPR-Cas systems in adaptive immunity and beyond.* Curr Opin Immunol, 2015. **32**: p. 36-41.

62. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.

63. Basu, S., et al., *Silencing of end-joining repair for efficient site-specific gene insertion after TALEN/CRISPR mutagenesis in Aedes aegypti.* Proc Natl Acad Sci U S A, 2015. **112**(13): p. 4038-43.

64. Hall, A.B., et al., *SEX DETERMINATION. A male-determining factor in the mosquito Aedes aegypti.* Science, 2015. **348**(6240): p. 1268-70.

65. Kistler, K.E., L.B. Vosshall, and B.J. Matthews, *Genome engineering with CRISPR-Cas9 in the mosquito Aedes aegypti.* Cell Rep, 2015. **11**(1): p. 51-60.

66. Kotwica-Rolinska, J., et al., *CRISPR/Cas9 Genome Editing Introduction and Optimization in the Non-model Insect Pyrrhocoris apterus.* Front Physiol, 2019. **10**: p. 891.

67. Edvardsen, R.B., et al., *Targeted mutagenesis in Atlantic salmon (Salmo salar L.) using the CRISPR/Cas9 system induces complete knockout individuals in the F0 generation.* PLoS One, 2014. **9**(9): p. e108622.

68. Gratz, S.J., et al., *Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease.* Genetics, 2013. **194**(4): p. 1029-35.

69. Hockemeyer, D., et al., *Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases.* Nat Biotechnol, 2009. **27**(9): p. 851-7.

70. Sommer, D., et al., *Efficient genome engineering by targeted homologous recombination in mouse embryos using transcription activator-like effector nucleases.* Nat Commun, 2014. **5**: p. 3045.

71. Zhu, L., et al., *CRISPR/Cas9-mediated knockout of factors in non-homologous end joining pathway enhances gene targeting in silkworm cells.* Sci Rep, 2015. **5**: p. 18103.

72. Yoo, S.H., et al., *PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues.* Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5339-46.

73. Bottcher, R., et al., *Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells.* Nucleic Acids Res, 2014. **42**(11): p. e89.

74. Rong, Y.S., et al., *Targeted mutagenesis by homologous recombination in D. melanogaster.* Genes Dev, 2002. **16**(12): p. 1568-81.

75. Gunawardhana, K.L. and P.E. Hardin, *VRILLE Controls PDF Neuropeptide Accumulation and Arborization Rhythms in Small Ventrolateral Neurons to Drive Rhythmic Behavior in Drosophila.* Curr Biol, 2017. **27**(22): p. 3442-3453 e4.

76. Urasaki, A., K. Asakawa, and K. Kawakami, *Efficient transposition of the Tol2 transposable element from a single-copy donor in zebrafish.* Proc Natl Acad Sci U S A, 2008. **105**(50): p. 19827-32.

77. Urasaki, A., et al., *Transposition of the vertebrate Tol2 transposable element in Drosophila melanogaster.* Gene, 2008. **425**(1-2): p. 64-8.

78. Zhao, S., et al., *PiggyBac transposon vectors: the tools of the human gene encoding.* Transl Lung Cancer Res, 2016. **5**(1): p. 120-5.

79. Gregory, M., et al., *Insect transformation with piggyBac: getting the number of injections just right.* Insect Mol Biol, 2016. **25**(3): p. 259-71.

2. PHOTOPERIODIC AND CLOCK REGULATION OF THE VITAMIN A PATHWAY IN THE BRAIN MEDIATES SEASONAL RESPONSIVENESS IN THE MONARCH BUTTERFLY[1]

2.1. Overview.

Seasonal adaptation to changes in light:dark regimes (i.e., photoperiod) allows organisms living at temperate latitudes to anticipate environmental changes. In nearly all animals studied so far, the circadian system has been implicated in measurement and response to the photoperiod. In insects, genetic evidence further supports the involvement of several clock genes in photoperiodic responses. Yet, the key molecular pathways linking clock genes or the circadian clock to insect photoperiodic responses remain largely unknown. Here, we show that inactivating the clock in the North American monarch butterfly (*dp*) using loss-of-function mutants for the circadian activators CLOCK (dpCLK) and dpBMAL1 and the circadian repressor CRYPTOCHROME 2 (dpCRY2) abolishes photoperiodic responses in reproductive output. Transcriptomic approaches in the brain of monarchs raised in long and short photoperiods, summer monarchs, and fall migrants revealed a molecular signature of seasonal-specific rhythmic gene expression that included several

[1] Reprinted from “Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly” by Iiams SE, Lugena AB, Zhang Y, Hayden AN & Merlin C, 2019. *PNAS*, 116 (50), 25214-25221.

genes belonging to the vitamin A pathway. We found that the rhythmic expression of these genes was abolished in clock-deficient mutants, suggesting that the vitamin A pathway operates downstream of the circadian clock. Importantly, we showed that a CRISPR/Cas9-mediated loss-of-function mutation in the gene encoding the pathway's rate-limiting enzyme, ninaB1, abolished photoperiod responsiveness independently of visual function in the compound eye and without affecting circadian rhythms. Together, these results provide genetic evidence that the clock-controlled vitamin A pathway mediates photoperiod responsiveness in an insect. Given previously reported seasonal changes associated with this pathway in the mammalian brain, our findings suggest an evolutionarily conserved function of vitamin A in animal photoperiodism.

2.2. Introduction.

Understanding how changes in photoperiod are sensed and translated into seasonal changes in animal developmental, physiological, and behavioral processes at the molecular level remains one of the fundamental yet poorly understood questions in the field of biological rhythms. Many classical resonance and night interruption experiments have suggested a circadian basis for photoperiodic responses in nearly all insects, birds, and seasonal-breeding mammals studied so far [1, 2]. However, owing to the fact that most species with strong photoperiodic responses are nontraditional model organisms, a formal

genetic demonstration that the circadian clock, an endogenous timekeeping mechanism that governs 24 hour (h) rhythms, plays a role in photoperiodic responsiveness is generally lacking. A few notable exceptions are in insects in which the use of genetic knockdowns or knockouts has demonstrated that at least some of the core clock genes are necessary for photoperiodic responses [3-7]. Yet, because the impact of a given clock gene disruption on photoperiodic responses can be interpreted as a pleiotropic effect, the notion that the circadian clock acts as a functional system for photoperiod sensing and responsiveness remains debated [8]. Perhaps for this reason, the key molecular pathways linking clock genes or the circadian clock to insect photoperiodic responses have remained virtually unexplored to date.

With a strong seasonal biology, the availability of a draft genome sequence [9] and functional genomic tools [10, 11], the eastern North American migratory monarch butterfly is well-suited not only to assess the effect of clock gene mutants on photoperiodic responses but also to probe the molecular bases underlying photoperiod responsiveness [12, 13]. Each fall, coincident with the decreasing photoperiod, migrating monarchs undergo a switch in physiology and behavior, starting their epic southward migration from the northern United States to their Mexican overwintering sites in a state of reproductive dormancy (i.e., diapause) in which they remain until the spring. Although complete reproductive diapause appears to require more than a single diapause-inducing cue, subjecting laboratory-raised monarchs to a short photoperiod has been shown to

robustly decrease the production of mature oocytes by females [14], making it possible to measure monarch photoperiodic responsiveness in laboratory conditions. Furthermore, the development of gene-editing technologies in this species has also permitted the generation of several clock gene loss-of-function mutants [10, 11, 15]. The monarch circadian clock relies on transcriptional/translational feedback loops initiated by the heterodimeric transcription factor CLOCK (dpCLK):dpBMAL1 that drives the rhythmic transcription of the core clock genes *period* (*dpPer*), *timeless* (*dpTim*), and *cryptochrome* 2 (*dpCry2*). Once translated, dpPER, dpTIM, and dpCRY2 form a repressive complex that in turn rhythmically inhibits dpCLK:dpBMAL1-mediated transcription [16]. Loss-of-function mutants for each of the circadian activators dpCLK and dpBMAL1, and for the main circadian repressor dpCRY2, which are available in our laboratory, can be used to genetically test if both positive and negative elements of the clock are required to elicit monarch photoperiodic responses.

In the work presented here, we demonstrate that both circadian activators and the main circadian repressor are necessary for monarch photoperiodic responses. In addition, RNA-sequencing (RNA-seq) over the course of the day in the brains of laboratory strains of monarchs raised in long and short photoperiods, as well as in the brains of summer monarchs and wild-caught fall migrants, reveals a seasonal signature of photoperiod-dependent rhythmic expression that includes several components of the vitamin A pathway. We

further genetically demonstrate that the vitamin A pathway functions downstream of the brain circadian clock and that it is necessary for photoperiodic responses in the monarch butterfly. Together, these data provide a molecular link between circadian clock genes and insect photoperiodic responsiveness.

2.3. Results.

2.3.1. Circadian Clock Activators and Repressors are Required for Photoperiod Responsiveness.

To genetically determine if circadian clock genes are required for monarch photoperiodic sensing and responses, we tested photoperiodic responsiveness of monarch loss-of-function mutant strains for the 2 circadian activators *dpClk* and *dpBmal1* and for the main circadian repressor *dpCry2* (Fig. 2.1). All of these strains harbor dysfunctional molecular clocks and exhibit arrhythmic adult eclosion behavior [10, 15]. All mutants were raised along with their respective wild-type siblings from eggs to adults under long-photoperiod (LP) conditions (15 hr light:9 hr dark) or short-photoperiod (SP) conditions (10 hr light:14 hr dark) at a constant temperature of 21 °C, and the number of mature oocytes produced by females was quantified 14 days post adult emergence (Fig. 2.1). The temperature of 21 °C was chosen because this temperature not only reflects the average day/night temperature that fall migratory monarchs experience across the eastern United States during their southward migration

but also is permissive to female oocyte development in both LP and SP [14] (Fig. 2.1). Consistent with previous reports [14], we found that wild-type monarch females responded to the photoperiod by producing significantly more mature oocytes in LP than in SP (Fig. 2.1 *A*–*C*). In contrast, females of each of the 3 loss-of-function mutants tested lost the ability to respond to the photoperiod (Fig. 1 *A*–*C*). Loss-of-function mutant females for either *dpClk* or *dpBmal1* [designated *dpCyc-like* as this mutant lacks the C-terminal transactivation domain lost in *Drosophila Cycle* [15] produced, both in LP and SP, significantly more mature oocytes than wild-type females raised in SP and as many as wild-type females raised in LP (Fig. 2.1 *A* and *B*). In contrast to monarchs lacking functional circadian activators, female monarchs lacking a functional dpCRY2 circadian repressor produced, regardless of the photoperiod during which they were raised, significantly fewer oocytes than did wild-type monarchs raised in SP (Fig. 2.1 *C*). We verified that the low level of mature oocyte production in absence of a functional dpCRY2 was not due to sterility but rather to juvenile hormone (JH) deficiency. Topical application of methoprene, a JH analog, restored oocyte maturation in *dpCry2* homozygous mutant females to levels similar to those in treated wild-type females (Fig. 2.2). Therefore, the lack of photoperiodic responses observed in *dpClk*, *dpBmal1*, and *dpCry2* loss-of-function mutants demonstrated that these 3 clock genes are necessary for photoperiodic responses in the monarch.

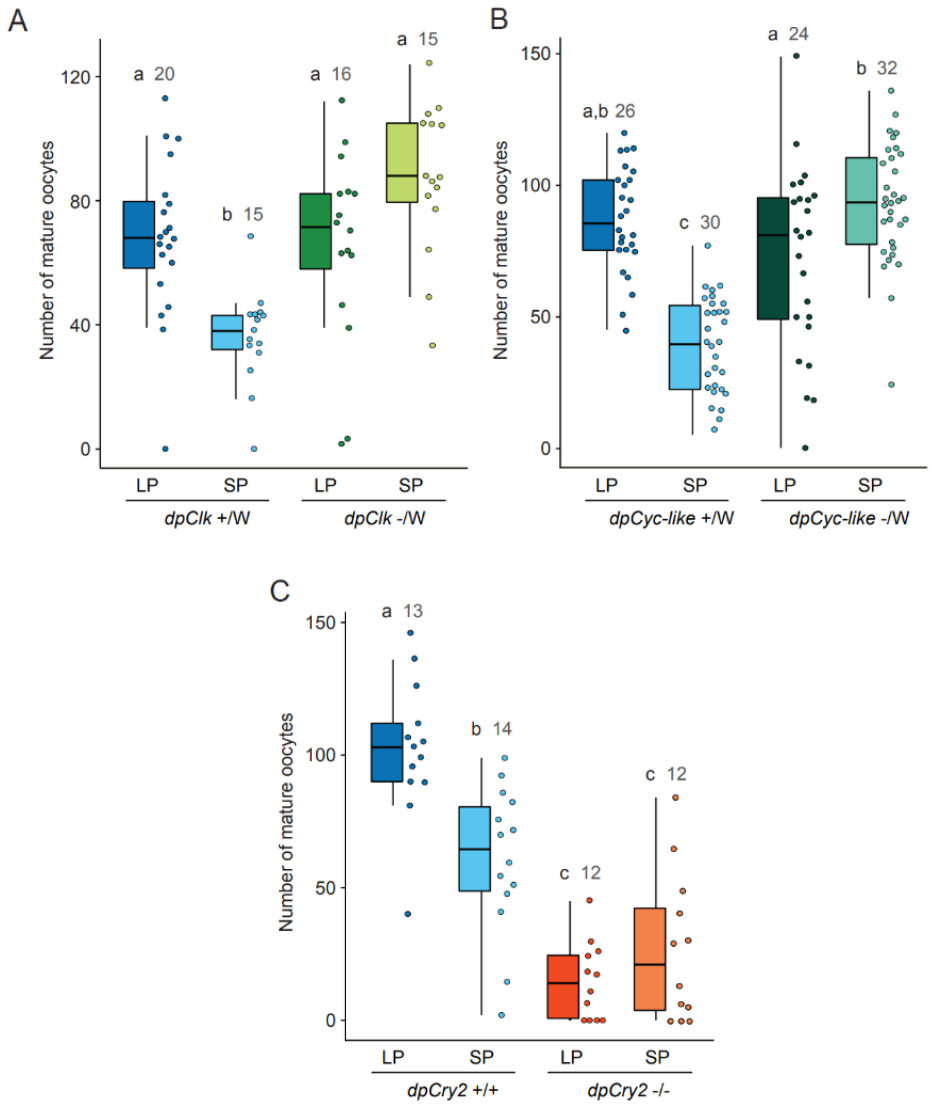


Figure 2.1: Circadian clock genes are required for photoperiodic responses.

Number of mature oocytes produced 14 days post adult emergence in loss-of-function hemizygous *Clock* (*dpClk*; *A*), hemizygous *dpBmal1* (*dpCyc-like*; *B*), and homozygous *Cryptochrome 2* (*dpCry2*; *C*) mutant females, and in their respective wild-type siblings, all raised in LP and SP at 21 °C. Monarchs have a ZW sex-determination system; males are the homogametic sex (ZZ) while females are the heterogametic sex (ZW), and *dpClk* and *dpBmal1* are located on the Z chromosome. For each condition, box plots and raw data points are shown. Error bars on box plots represent 1.5 times the interquartile range. LP: 15 hr light, 9 hr dark; SP: 10 hr light, 14 hr dark. Different letters over bars indicate groups that are statistically significant (interaction genotype × photoperiod, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$, and numbers of dissected females are indicated. Reprinted from [17].

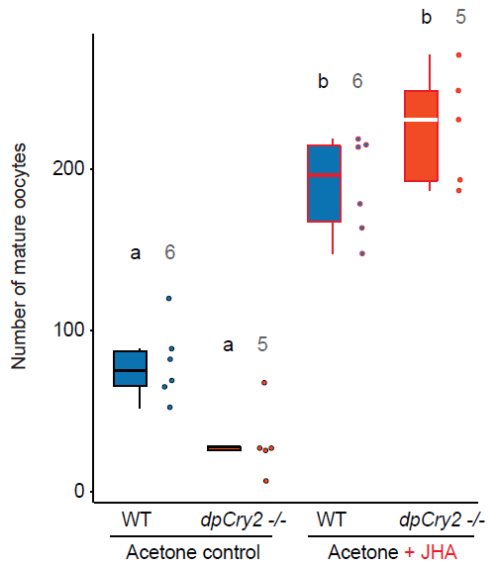


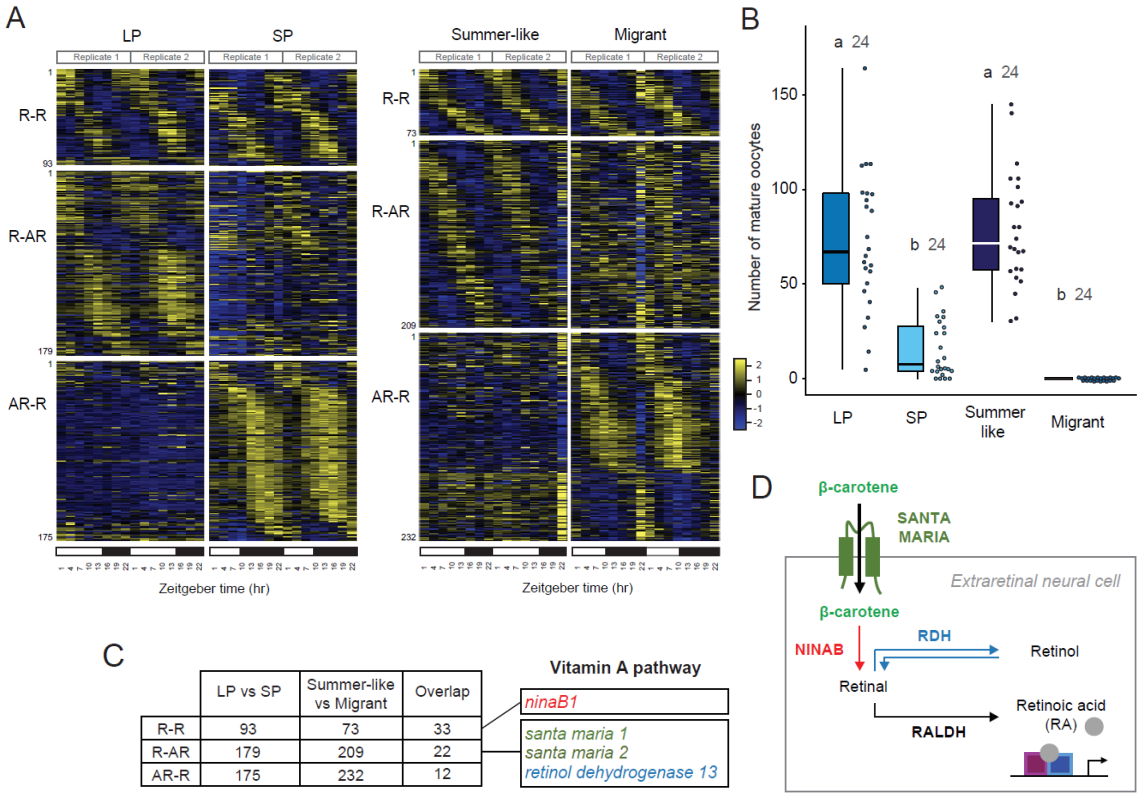
Figure 2.2 (Supplement): Application of a juvenile hormone analogue restores high levels of oocyte production in monarch *dpCry2 -/-*.
Number of mature oocytes produced 10 days post adult emergence in wild-type or *Cryptochrome 2* (*dpCry2*) *-/-* female monarchs treated with either an acetone vehicle or with methoprene, a juvenile hormone analogue (JHA). Monarch were raised in LP at 25˚C. Legends as in Figure 2.1. Interaction genotype x treatment, Two-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$. (Data generated by Dr. Christine Merlin while in the lab of Dr. Steven M. Reppert). Reprinted from [17].

2.3.2. Photoperiodic and Clock Regulation of the Vitamin A Pathway in the Brain.

If circadian clock components are necessary for photoperiodic responses, we reasoned that molecular pathways mediating photoperiod responsiveness may be under rhythmic transcriptional regulation and may show seasonal-specific expression patterns of rhythmic expression in the brain, the organ known to function in photoperiodic reception in lepidopteran and some fly species [18, 19]. To identify photoperiod-specific expression patterns of rhythmic expression in the monarch adult brain, we first carried out RNA-seq at 3 hr

intervals over the 24 hr light:dark (LD) cycle in the brain of monarchs raised in LP and in SP at 21 °C (Fig. 2.3 *A*, *Left*). Expression levels for each gene were examined for rhythmic variation using RAIN [20] and MetaCycle [21], and genes with a maximum/minimum fold-change ≥1.3 and a *P* value corrected for multiple testing ≤0.05 were considered rhythmic. Using these criteria, we identified sets of genes with a seasonal molecular signature of rhythmic gene expression; while 93 genes were expressed rhythmically in both LP and SP conditions (rhythmic-rhythmic; R-R), we found 179 genes expressed rhythmically only in LP (rhythmic-arrhythmic; R-AR), and 175 genes expressed rhythmically only in SP (arrhythmic-rhythmic; AR-R) (Fig. 2.3 *A*, *Left*; *Appendix A*, Table 3; and RNA-Seq dataset, Gene Expression Omnibus (GEO) repository, Accession number: GSE126336). To pinpoint which of these genes may be involved in photoperiodic responsiveness, we next sought to compare the differential rhythmic gene expression found between LP and SP monarchs to the one between different seasonal forms of wild monarchs exposed to summer and fall photoperiods. We predicted that key mediators of photoperiod-regulated responses should show similar patterns of differential expression in the brain of LP and SP monarchs and in the brain of summer monarchs and fall migrants. Unlike fall migrants, which migrate en masse and can therefore be caught in substantial numbers in a single round of collection, wild summer monarchs often move individually, making it challenging to catch the number necessary to perform RNA-seq around the clock at fine temporal resolution. Instead of wild-

caught summer monarchs, we therefore used a laboratory strain of wild-type monarchs raised in a 15 hr light:9 hr dark summer-like long day and summer-like temperatures (25°C). Importantly, summer-like monarchs and fall migrants exhibited similar photoperiodic reproductive responses to LP and SP monarchs raised at 21°C, with summer-like females producing as many mature oocytes as LP females, and fall migrant females being in complete reproductive diapause (Fig. 2.3 *B*). Comparable to results obtained with LP and SP monarchs, 3 classes of genes were identified by RNA-seq in the brains of summer-like monarchs and wild-caught fall migrants, with 73 genes expressed rhythmically in both seasonal forms (R-R), 209 genes expressed rhythmically only in summer-like monarchs (R-AR), and 232 genes expressed rhythmically only in fall migrants (AR-R) (Fig. 2.3 *A*, *Right; Appendix A*, Table 4; and RNA-Seq dataset, GEO repository, Accession number: GSE126336). As predicted, comparing the LP/SP and summer-like/migrant datasets revealed a number of genes expressed in a similar manner in both datasets; specifically, 33 genes were found to be rhythmic in all photoperiodic/seasonal forms (R-R), 22 genes were rhythmic only in LP and summer-like monarchs (R-AR), and 12 genes were rhythmic only in SP and wild-caught migrants (AR-R) (Fig. 2.3 *C*; *Appendix A*, Table 5).



	LP vs SP	Summer-like vs Migrant	Overlap
R-R	93	73	33
R-AR	179	209	22
AR-R	175	232	12

Figure 2.3: Rhythmic gene expression analysis reveals photoperiodic regulation of the vitamin A pathway in the monarch brain.

(*A*) Heatmaps of relative RNA levels for genes rhythmically expressed in brains of monarchs raised in either LP or SP at 21 °C (*Left*) and in brains of either summer-like or wild-caught migrant monarch (*Right*). R-R, genes rhythmic in both conditions; R-AR: genes rhythmic only in summer-like or LP; AR-R, genes rhythmic only in migrants or SP. Columns represent samples collected over a 24 hr LD cycle. Transcripts are arranged by phase, and their order along the vertical axis is conserved in both conditions. White bars: light conditions; black bars: dark conditions. (Generated by Aldrin B. Lugena) (*B*) Number of mature oocytes produced by females 9 to 15 days post adult emergence for summer-like monarchs and 7 to 11 days for LP/SP monarchs. Wild-caught migrants of unknown age were dissected 8 days after capture. Boxplots as in Fig. 2.1. One-way ANOVA, post hoc Tukey test, $P < 0.05$. (Data collected by Dr. Christine Merlin) (*C*) Table showing the number of genes found in any given category and the number of overlapping genes between comparisons. Genes involved in the vitamin A pathway are shown. The complete lists are shown in *Appendix A*, Tables 3, 5, and 7. (Generated and analyzed by Aldrin B. Lugena) (*D*) Proposed model of the vitamin A pathway in insect brain cells. Beta-carotene is transported into extraretinal neural cells of the adult brain via santa maria and converted to retinal by ninaB. Retinal can either be interconverted into retinol by a retinol dehydrogenase (RDH) or converted into RA by a retinaldehyde dehydrogenase (RALDH). RA binds to retinoid receptors to regulate transcription of target genes. Reprinted from [17].

A Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway-enrichment analysis identified circadian rhythm as an enriched term in the R-R category that contained the robustly cycling clock genes of the interlocked core and the stabilizing loops *dpPer*, *dpTim*, *vrille* (*dpVrille*) and *clockwork orange* (*dpCwo*) [9, 22] (Fig. 2.4). Although no other KEGG pathway was found enriched in genes rhythmic in LP and summer-like monarchs (R-AR) or in SP and wild-caught migrants (AR-R), 3 of 22 genes in the R-AR category stood out based on their known function in the vitamin A pathway in both *Drosophila* and mammals. Two of the 3 genes—*scavenger receptor acting in neural tissue 1* and *2* (*santa maria 1* and *2*)—encode homologs of a transmembrane protein responsible for the uptake of beta-carotene into extraretinal neural cells in the *Drosophila* brain [23, 24], and the third gene, *retinol dehydrogenase 13* (*rdh13*), encodes an enzyme belonging to a family known to reversibly convert retinaldehyde (also called retinal) to retinol in mammals [25] (Figs. 2.3 *C* and *D* and 2.5 *A*). Interestingly, a homolog of *neither inactivation nor afterpotential B* (*ninaB1*), the gene encoding the rate-limiting enzyme of this pathway that converts carotenoids into retinaldehyde [26] (Fig. 2.3 *D*), was found expressed rhythmically in all groups (Figs. 2.3 *C* and 2.5 *A*). Together, these data suggested that the vitamin A pathway in the monarch brain could be involved in photoperiodic responsiveness. Because we demonstrated that both functional circadian activators dpCLK and dpBMAL1 and repressor dpCRY2 were necessary for photoperiodic responses, we then used already available RNA-

seq datasets [22] to test whether *ninaB1*, *santa maria 1*, *santa maria 2*, and *rdh13* rhythmic expression in brains of monarchs subjected to 15:9 long day could be *dpClk*- and *dpCry2*-dependent. Consistent with the idea that the vitamin A pathway acts downstream of the circadian clock or clock genes, we found that the rhythmic expression of all 4 genes was abolished in brains of *dpClk* and *dpCry2* loss-of-function monarch mutants (Fig. 2.5 *B*).

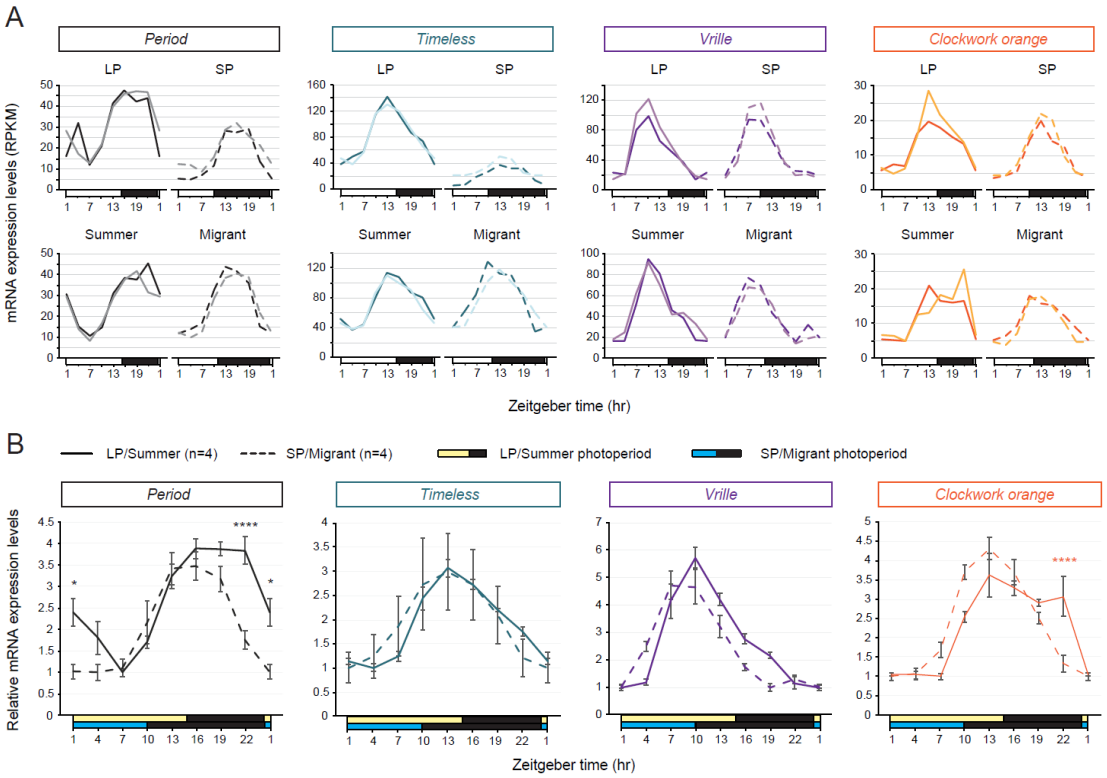


Figure 2.4 (Supplement): Temporal mRNA expression profiles of *period*, *timeless*, *vrille*, and *clockwork orange* in brains of monarchs in different photoperiods.

(*A*) mRNA expression levels in brains of monarchs raised in LP and SP (*Top*) and of summer-like and migrant monarchs (*Bottom*). For each condition, two biological replicates are plotted. White bars: light; black bars: dark. (*B*) Relative mRNA expression levels from pooled biological replicates of LP and summer-like monarchs (solid line),

Figure 2.4 (Supplement) Continued

and SP and migrants (dashed line). Data represent the mean ± s.e.m. Yellow bars: long photoperiod; blue bars: short photoperiod; black bars: dark. Interactions seasonal morph x photoperiod, Two-way ANOVAs: *period* (*dpPer*), $P < 0.0001$; *timeless* (*dpTim*), non significant; *vrille* (*dpVrille*), $P < 0.005$; *clockwork orange* (*dpCwo*), $P < 0.0001$. Tukey's pairwise comparisons, *, $P < 0.05$; ****, $P < 0.0001$. Reprinted from [17].

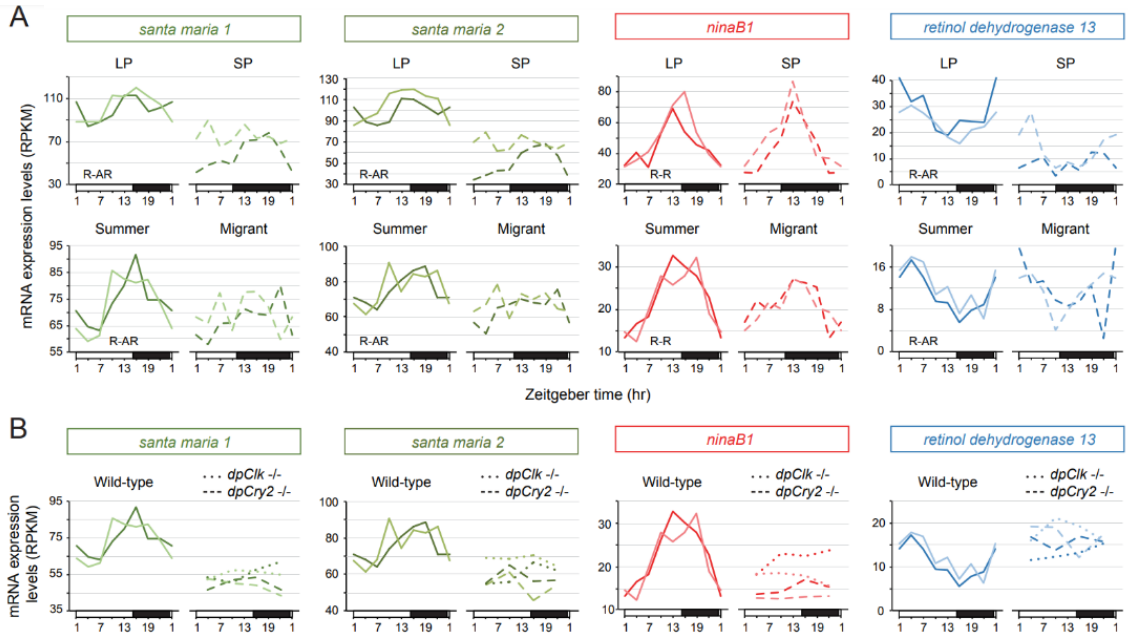


Figure 2.5: The vitamin A pathway is under both photoperiodic and clock regulation.

(*A*) Temporal mRNA expression profiles of *santa maria 1*, *santa maria 2*, *ninaB1*, and *retinol dehydrogenase 13* in brains of monarch raised in LP and SP (*Top*) and summer-like and migrant monarchs (*Bottom*). R-AR: rhythmic in LP and summer-like monarchs, arrhythmic in SP monarchs and migrants; R-R: rhythmic in LP, SP, summer-like monarchs, and migrants. The R-R and R-AR categories were defined based on the analysis of rhythmic gene expression reported in Fig. 2.3. (*B*) Temporal mRNA expression profiles of *santa maria 1*, *santa maria 2*, *ninaB1*, and *retinol dehydrogenase 13* in brains of wild-type summer-like monarchs (plain lines), and *dpClk* (dotted lines) and *dpCry2* (dashed lines) loss-of-function mutants raised in summer-like conditions. For each condition, 2 biological replicates are plotted. Reprinted from [17].

Interestingly, we also identified *collagen type IV subunit alpha 1* (*α-1*)

among the genes expressed rhythmically only in the brain of SP monarchs and

wild-caught migrants (*Appendix A*, Table 5). This gene, which encodes one of 2 subunits of a central component of basement membranes [27], has previously been reported to show strong signatures of divergence in migratory and nonmigratory monarch populations [28]. Based on differential expression in flight muscle of migratory and nonmigratory monarchs that were correlated with flight metabolic rates, *collagen type IV α-1* has been proposed to be associated with the regulation of flight efficiency during long-distance migration [28]. The differential rhythmic expression that we observed in the brain of monarch seasonal forms suggests that *collagen type IV α-1* may serve an additional and/or different function with respect to seasonal migration.

2.3.3. A Functional Vitamin A Pathway is Necessary for Photoperiodic Responses.

Given that several genes belonging to the vitamin A pathway were found differentially expressed in the brains of seasonal forms of monarchs and that seasonal changes associated with this pathway have been reported in the brain of mammals [29-32], we next focused on genetically testing if the vitamin A pathway was necessary for monarch photoperiodic responses. To this end, we generated a CRISPR/Cas9-mediated germline mutation in *ninaB1*, the gene encoding the rate-limiting enzyme responsible for the production of retinal. We selected a 7-bp deletion in the targeted third exon, resulting in a protein truncated by more than 80% to establish a loss-of-function mutant line (Fig. 2.6

A) that we raised in both LP and SP conditions. We found that, similar to clock-deficient monarchs, homozygous mutant *ninaB1* females lost the ability to respond to the photoperiod, producing in both LP and SP as many mature oocytes as their wild-type siblings raised in LP (Fig. 2.6 *B*; Fig. 2.7 *A* and *B*). The LP-like state of oocyte production observed in mutant *ninaB1* females was surprising because *ninaB1* was expressed at constitutively low levels in dpCry2 loss-of-function mutants (Fig. 2.5 *B*), which are themselves locked in a SP-like state of low oocyte production (Fig. 2.1 *C*). The presence of paralogs that have some degree of functional redundancy for the production of retinal and that are expressed in the brain could explain this unexpected result. Blasting ninaB1 against protein sequences encoded in the monarch genome revealed the existence of 2 paralogs, ninaB2 and ninaB3, which showed ~50% sequence conservation at the amino acid level with ninaB1 (Fig. 2.7 *C*). Only *ninaB2* was found expressed at detectable levels by RNA-seq in the monarch brain (Fig. 2.7 *D*), and its product was also the only one of the 2 paralogs to share with ninaB1, a conserved residue known to be essential for the catalytic activity of *Drosophila* ninaB [26] (Fig. 2.7 *C*).

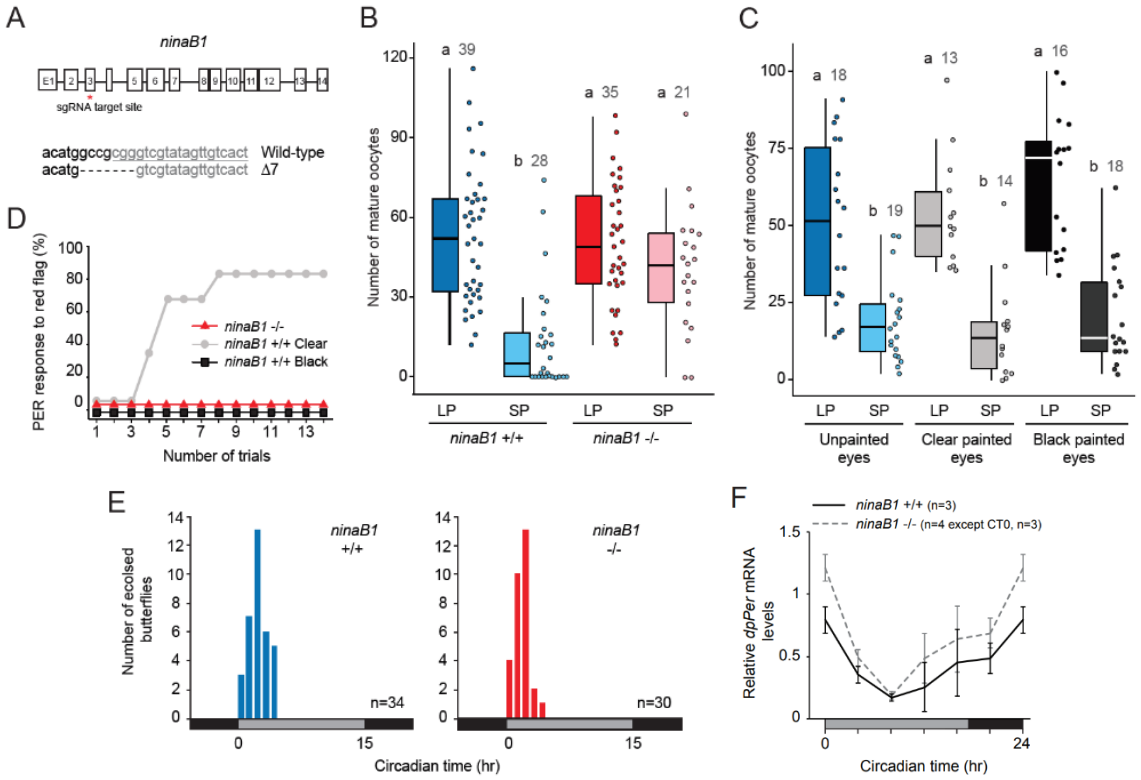


Figure 2.6: Disruption of the vitamin A pathway abolishes photoperiodic responses.

(*A*) Generation of a *ninaB1* loss-of-function mutant using CRISPR/Cas9-mediated targeted mutagenesis. (*Top*) Genomic structure of *ninaB1* in the monarch. Red star, position of the site targeted by sgRNA. (*Bottom*) Nucleotide sequence of the targeted site (underlined) and introduced frame-shifting mutation (7-bp deletion). (*B*) Number of mature oocytes produced 10 days post adult emergence in *ninaB1* homozygous mutant and wild-type sibling females. Boxplots as in Fig. 2.1. Interaction genotype × photoperiod, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$. (*C*) Number of mature oocytes produced 10 days post adult emergence in wild-type female monarchs with unpainted eyes or eyes covered with either clear or black paint at the day of emergence. Interaction genotype × painting condition, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$. (*D*) Associative visual conditioning of the proboscis extension reflex (PER) of *ninaB1* homozygous mutants (red line, $n = 6$) and wild-type monarchs with eyes painted clear (gray line, $n = 6$) or black (black line, $n = 6$) for 14 training trials on day 3 of training. (Performed by Ashley N. Hayden) (*E*) Profiles of adult eclosion in DD of wild-type (blue) and *ninaB1* loss-of-function siblings (red). Gray bars: subjective day; black bars: subjective night. Kolmogorov–Smirnov test, $P > 0.05$. (Performed by Ying Zhang) (*F*) Circadian expression of *period* (*dpPer*) in brains of wild-type (solid black lines) and *ninaB1* loss-of-function (dashed gray lines) siblings. Values are mean ± SEM, and the numbers of animals for each genotype and time point are indicated. Interaction genotype × time, 2-way ANOVA, $P > 0.05$. Reprinted from [17].

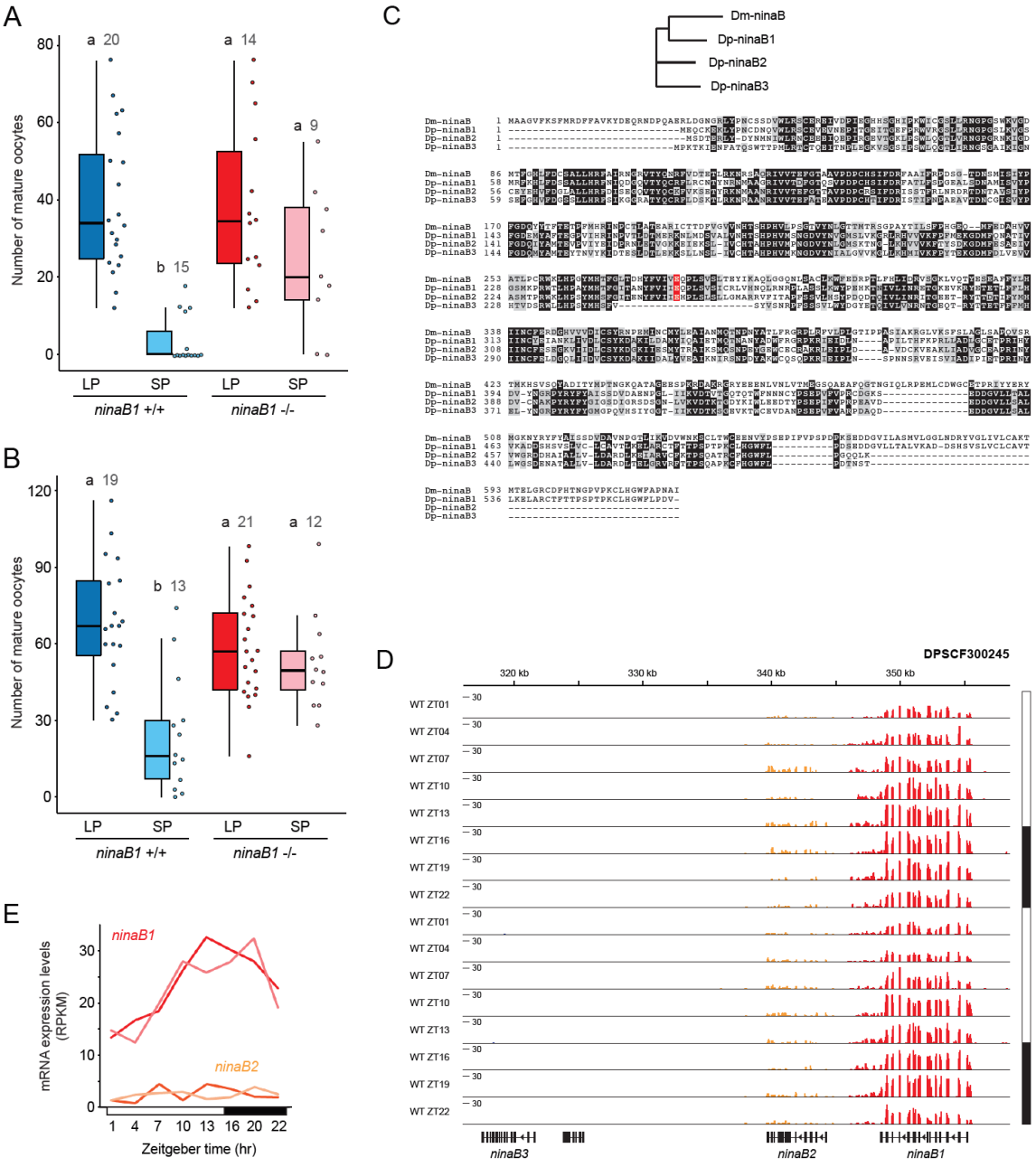


Figure 2.7 (Supplement): The reproductive state exhibited by photoperiod-impaired *ninaB1* loss-of-function could be caused by redundant function of ninaB1 and ninaB2.

(*A, B*) Number of mature oocytes produced 10 days post adult emergence in *ninaB1* homozygous loss-of-function and wild-type sibling females, all raised in LP and SP at 21˚C after either one backcross (A) or three backcrosses (*B*). Legends as in Fig. 2.6 *B*. Interaction genotype x photoperiod, Two-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$. (*C*) Comparisons of *Drosophila* ninaB (Dm; AAF54978) and monarch (Dp) ninaB1 (DPOGS212590), ninaB2 (DPOGS212591), and

Figure 2.7 (Supplement) Continued

ninaB3 (DPOGS212592) sequences. *Top*, phylogenetic relationship. *Bottom*, sequences alignment. Identical and similar residues found in at least three out of four sequences are shaded in black and gray, respectively. The glutamic acid residue highlighted in red, which is conserved between Dm-ninaB, Dp-ninaB1 and Dp-ninaB2 but absent in Dp-ninaB3, corresponds to the residue providing the enzymatic activity of Dm-ninaB [26]. (*D*) Visualization of RNA-seq signal in brains of wild-type monarchs over 24 hr in LP on scaffold DPSCF300245 that contains the three copies of *ninaB* present in the monarch genome. Of the two detectable expressed copies, only *ninaB1* displays rhythmic expression. Two biological replicates are plotted consecutively from top to bottom. White bar: light; dark bar: night. (provided by Aldrin B. Lugena) (*E*) mRNA expression levels of *ninaB1* and *ninaB2* in the brain of wild-type monarchs over the course of the day in LP. For each gene, two biological replicates are plotted. White bar: light; black bar. Reprinted from [17].

Unlike *ninaB1*, *ninaB2* was expressed at low levels and in an arrhythmic manner in the monarch brain (Fig. 2.7 *D* and *E*). If monarch ninaB2 is indeed capable of converting carotenoids into retinal at a high-enough level to promote the production of a mature oocyte in a LP-like state, our data may indicate that the rhythmic production of retinal in the brain could be critical for photoperiodic responsiveness. If this is the case, the phase of ninaB1 expression, which is locked to “lights on” at the messenger RNA (mRNA) level with a peak of expression in the light phase in LP and in the dark phase in SP (Fig. 2.5 *A*), might be important for the induction of a photoperiodic response.

Retinal is most conventionally known to be essential for visual function, and *Drosophila ninaB* loss-of-function mutants are blind [23, 26, 33]. Because our monarch *ninaB1* loss-of-function mutant was not brain-specific, we sought to verify that the observed loss of photoperiodic responses in this mutant did not result from an inability to perceive light through the compound eyes. To this end,

we raised wild-type monarchs in LP and SP conditions and covered the adult compound eyes on the day of adult emergence with either an enamel-based black paint or clear paint (for controls), respectively blocking or allowing light input to monarch tissues [34]. Importantly, we found that the photoperiodic response of monarch females with black painted eyes was indistinguishable from that of control females with either unpainted or clear painted eyes (Fig. 2.6 *C*). We further excluded the possibility that the black paint was not fully blocking eye visual function by showing that, in contrast to wild-type monarchs with clear painted eyes, monarchs with black painted eyes trained to associate a colored flag with a sugar reward were not able to learn the association (Fig. 2.6 *D*), likely because their vision was impaired. Consistent with the defects in vision caused by mutations in *Drosophila ninaB* [26], we also found that monarch homozygous *ninaB1* mutants were unable to associate a color with the reward (Fig. 4*D*). Using similar painting experiments, we also excluded the possibility that the antennae, which harbor the navigational clocks for monarch migration, were necessary for monarch photoperiodic responses (Fig. 2.8). These results strongly suggest that the loss of photoperiodic responsiveness in monarch *ninaB1* loss-of-function mutants results from disruption of the vitamin A pathway in the brain and not in the compound eyes. Finally, we show that the lack of photoperiodic responses in homozygous *ninaB1* monarch mutants did not result from a disrupted circadian clock, as both behavioral rhythms of adult eclosion and molecular rhythms of *dpPer* in the brain were intact in this mutant

(Fig. 2.6 *E* and *F*). Together, our findings provide a functional demonstration that the vitamin A pathway in the brain, which acts downstream of the circadian clock and is seasonally regulated, is necessary to regulate photoperiodic responsiveness in the monarch.

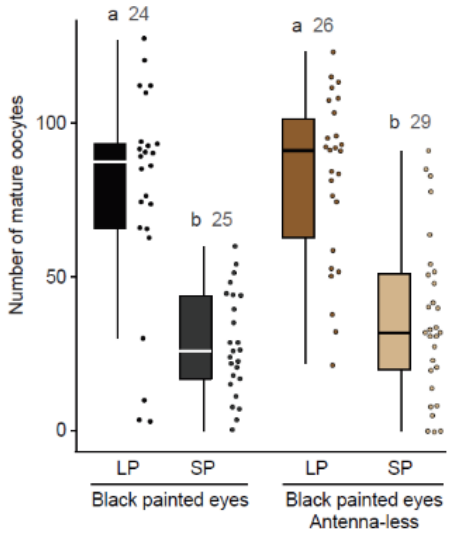


Figure 2.8 (Supplement): The antennae are not necessary for photoperiodic responses.

Number of mature oocytes produced 10 days post adult emergence in wild-type female monarchs with eyes painted black and with or without antennae, all raised in LP and SP at 21˚C. Legends as in Fig. 2.6 *C*. Interaction genotype x photoperiod, Two-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$. Reprinted from [17].

2.4. Discussion.

For decades, progress in understanding the molecular and genetic bases of photoperiodism in insects has been limited by the fact that the key genetically tractable model organism *Drosophila melanogaster* lacks a pronounced diapause, the prominent photoperiodic response exhibited by insects [35]. The

advent of molecular tools and their applicability to nontraditional model insects with a more dramatic photoperiodic response has recently changed this trend. While photoperiodic effects on the oscillation patterns of some circadian clock components provided molecular evidence that clock genes may be involved in sensing and transmitting the photoperiodic information to output cascades [36-39], recent analysis using gene knockdowns and knockouts in the bean bug *Riptortus pedestris* and the mosquito *Culex pipiens* have unambiguously demonstrated the functional involvement of core clock genes in photoperiodism [3-6]. Whether clock gene products act pleiotropically or as part of a functional circadian system remains debatable as the use of individual clock gene knockouts that render the clock dysfunctional cannot formally distinguish between these 2 possibilities. However, our findings that monarchs bearing nonfunctional circadian activators and repressors display opposite egg maturation phenotypes mirror previous findings obtained in the bean bug [3, 5]. Together, these data indicate that circadian activators and repressors may work in a coordinated fashion to sense photoperiod, likely through a feedback loop. As previously demonstrated in other insects [40], photoperiodic induction of reproductive diapause is induced by a reduction of JH synthesis and/or secretion by the corpora allata, a pair of neurohemal organs located behind the brain. Being able to restore the production of mature oocytes to high levels in *dpCry2* knockout females by application of a JH analog suggests not only that JH titers are likely low in this mutant but also that *dpCry2* is not involved in the

cascade of events downstream of JH synthesis and/or secretion. Similarly, while knocking out *dpClk* and *dpBmal1* disrupts photoperiodic responses, this treatment does not affect the levels of mature oocytes produced by mutant females, suggesting that JH titers may be similar between these mutants and wild-type siblings raised in LP. Together, these data indicate that circadian clock genes likely act upstream of JH synthesis and/or secretion, as previously demonstrated in the bean bug [5]. Our discovery that the seasonally regulated vitamin A pathway is also under clock control in the monarch brain further adds support to the idea that a functional circadian system may be involved in photoperiodic responses. It also provides a molecular link in the insect brain between clock genes and the diapause response.

Interestingly, the involvement of vitamin A and carotenoids in insect photoperiodism is not without precedent. Classic dietary-deprivation experiments showed decades ago that many species of insects and mites reared on an artificial diet deficient in carotenoids lose their photoperiodic response [41-44], leading to the hypothesis that a yet-unknown vitamin A-based pigment, retinal-opsin, was the likely photoperiodic receptor. Here, we provide genetic evidence that regulation of the vitamin A pathway in the insect brain mediates photoperiodic responses. The findings that visual function from the compound eyes is dispensable for photoperiodic responses in the monarch is consistent with the major role that extraretinal deep-brain photoreceptors play in photoperiodic reception in other lepidopteran species [45]. By knocking out the

rate-limiting enzyme that converts carotenoids into retinal, we show that retinal conversion is necessary for photoperiodic responses in the monarch. As previously proposed, one possible function of retinal in the brain could be the production of a deep brain photoreceptor for photoperiodic induction that is different from *Cryptochrome 1* (*dpCry1*), the key photoreceptor used for circadian clock entrainment [16]. Determining which of the 5 opsins present in the monarch genome are expressed in monarch brain and knocking them out could help functionally test this hypothesis and identify the potential relevant opsin.

Alternatively, similar to that described in mammals [29-32], retinoic acid (RA), a product of retinal, could be responsible for the seasonal regulation of the transcriptional program and/or the neuronal plasticity in the brain that regulates photoperiodic responses. In rats, concentration and signaling of RA, which regulates gene transcription through the activation of nuclear receptors, are under photoperiod regulation and reduced in SP relative to LP [29, 30, 32]. Until recently, the RA signaling pathway was believed to occur only in vertebrates. However, reports that not only RA but also orthologs of the retinoic acid receptors are expressed in insects [46] suggest that a similar pathway may also operate in these species. While we found that genes involved in beta-carotene uptake (*santa-maria 1 and 2*) and conversion of vitamin A to retinal (*rdh13*) were under photoperiodic regulation and that their rhythms of expression were dampened in SP relative to LP, the gene encoding the enzyme catalyzing the

conversion of retinal to RA (*retinaldehyde dehydrogenase*) was not found under photoperiodic regulation in the monarch brain. However, this does not exclude the possibility of a photoperiodic regulation of concentration and signaling of RA in the monarch brain. Measuring the levels of RA in the brain of monarchs subjected to LP and SP will be necessary to ultimately implicate RA signaling in insect photoperiod responses. In mammals, RA is produced in ependymal tanycytes, a glial cell type lining the third ventricle adjacent to the hypothalamus that controls the season-dependent production of the active form of the thyroid hormone involved in seasonal reproduction [29]. Interestingly, tanycytes have recently been shown to undergo a seasonal change in morphology in hamster and sheep [31, 47], where they enclose the synapses of neurons producing the reproductive hormone gonadotropin-releasing hormone (GRH) during the nonbreeding season, presumably blocking the release of GRH [31]. A similar role of RA-positive cells in brain neuronal plasticity for driving monarch seasonal reproduction is an intriguing possibility. Identifying the cell type in which the vitamin A operates and developing knock-in approaches to mark cells with membrane-tagged fluorescent proteins will be necessary to test whether vitamin A-positive cells also undergo a seasonal remodeling in the monarch brain.

In conclusion, our work in the monarch provides genetic evidence that the vitamin A pathway, which is under seasonal and clock regulation in the brain, mediates photoperiodic responses in an insect. Given the parallel in seasonal changes associated with this pathway in the brain of monarchs and mammals, it

is tantalizing to speculate that the function of vitamin A in animal photoperiodism may be evolutionarily conserved. The continued molecular and genetic dissection of this pathway in the monarch and its involvement in brain photoreception, seasonal regulation of transcription, or seasonal neuronal plasticity should shed light on the mechanisms of action of vitamin A in insect photoperiodic responsiveness and ultimately reveal the degree of conservation with its role in mammalian seasonal biology.

2.5. Materials And Methods.

2.5.1. Animal Husbandry.

For photoperiodic experiments, wild-type and mutant monarchs maintained in 15 hours (h) light: 9 hr dark at 25°C were hand-paired. Fertilized females were placed in a cage to lay eggs on potted tropical milkweed plants (*Asclepias curassavica*). Eggs were transferred to petri dishes onto milkweed leaves and divided in two groups placed in Percival incubators respectively under 15 hr light: 9 hr dark (long photoperiod; LP) and under 10 hr light: 14 hr dark (short photoperiod; SP), both at constant temperature of 21°C and 70% humidity. Starting at the second instar, larvae were raised individually on semi-artificial diet until pupation. Post-eclosion, adult monarchs were housed in glassine envelopes and were manually fed a 25% honey solution daily until dissection.

For RNA-seq experiments, fall migratory monarch butterflies were

captured on October 15, 2014 in College Station, Texas (latitude 30°37'N, longitude 96°20'W). After capture, migrants were housed indoors in glassine envelopes in a Percival incubator under fall-like conditions with light:dark (LD) cycle set to prevailing light conditions (11 h:13 h LD, 0730–1830 Central Standard Time), at a constant temperature of 21 °C with 70% humidity. They were manually fed a 25% honey solution every other day and dissected 8 days after capture. Monarchs used for the LP and SP data sets were raised in the same conditions as for the photoperiodic experiments, and were dissected 7 to 11 days post-eclosion.

2.5.2. Methoprene/Vehicle Treatment.

Monarch wild-type and *Cry2-/-* females were treated topically on their abdomens with 200µg of methoprene in 5µl of acetone or with 5µl of acetone alone on day 1 and day 3 after adult emergence, as previously described [48].

2.5.3. Evaluation of Female Reproductive Status.

Female abdomens were transected dorsally, and mature oocytes were individually counted from each of the ovarioles in both ovaries under a dissecting microscope. Oocytes were considered mature when chorionated, i.e., when the presence of vertical ridges on the chorion was visible.

2.5.4. Compound Eye Painting and Antennae Removal.

Painting of the compound eyes and removal of the antennae were performed on the day of eclosion. For eye painting, the labial palps were first cut at the base to expose the entire compound eyes including the retina and the dorsal rim, which were then covered under a dissecting microscope with enamel based clear paint (Model master clear top coat; Testors no. 2736) or black paint (Glossy black; Testors no. 1147) using a fine paintbrush. Monarchs were harnessed to allow the paint to air-dry. The completeness of painting was verified the next day under a dissecting microscope, and touch ups were performed if needed. The light-permissive and light-blocking properties of the clear and black paints, respectively, had previously been characterized [34]. Unpainted control monarchs were subjected to labial palps removal similar to painted monarchs. Antennae were removed by clipping them with scissors at the base of the flagellum just above the pedicel as described in [34].

2.5.5. RNA-Sequencing Experiments.

Brains of monarchs raised indoors in LP and SP and of wild-caught migrants were dissected in 0.5X RNA later (Invitrogen) to prevent RNA degradation, the retinal pigmented photoreceptor layer was removed, and the brains were stored at -80^{o}C until use. For each seasonal phenotype/photoperiodic condition, three pooled brains were collected in two replicates at Zeitgeber time (ZT)1, ZT4, ZT7, ZT10, ZT13, ZT16, ZT19, and

ZT22. For each sample, total RNA was extracted using an RNeasy Mini kit (Qiagen). For samples from wild-caught migrants, polyA+ RNA was isolated from 2 µg of total RNA with NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs), and multiplexed libraries were prepared using the NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos (New England Biolabs) and amplified with 12 PCR cycles, following the manufacturer's recommendations. RNA-seq datasets from brains of summer-like monarchs and clock-deficient mutants used in this study were already available [22]. Butterflies were raised in 15:9 LD at 25°C, brains were collected at 3 hr intervals for summer-like monarchs starting at ZT1 and at 6 hr intervals for clock-deficient mutants starting at ZT4, and libraries were prepared as for wild-caught migrants. For samples from monarchs raised in LP and SP, multiplexed libraries were prepared by the Texas A&M AgriLife Genomics and Bioinformatics Facility using polyA+ RNA isolated from 1 µg of total RNA, and multiplexed libraries were prepared using the TruSeq Stranded mRNA Library Prep Kit for Illumina, following the manufacturer's recommendations. Libraries quality and size distribution was verified on a Bioanalyzer, libraries were quantified by real-time quantitative PCR, and 16 multiplexed libraries were mixed in equimolar ratios and sequenced on a Hi-seq 2500 (Illumina) using 50bp single end reads.

2.5.6. RNA-seq Data Processing, Mapping and Identification of Cycling Transcripts.

The resulting sequencing files were checked for quality control and demultiplexed by the Texas A&M AgriLife Genomics and Bioinformatics Facility. Reads were mapped to the monarch genome (assembly v3; [49]) using TopHat2 [50] with parameters "--read-realign-edit-dist 2 -g 1 --b2-sensitive". On average across LP, SP, summer-like and migrant monarchs, 91.5% of the reads mapped to the monarch genome (Table S2). After mapping, expression levels of transcripts were quantified in each sample using Cufflinks [51, 52]. To identify cycling transcripts, RAIN [20] and MetaCycle [21] were used with parameters "period=24, deltat=3, period.delta=3, nr.series=2, method='independent', peak.border=c(0.2, 0.8)" and "timepoints=seq(1, 46, by=3), adjustPhase="predictedPer", combinePvalue="fisher")", respectively. Only genes with three or more reads per kilo base per million mapped reads (RPKM) in at least one time point were considered as expressed and used for subsequent analysis. Rhythmically expressed genes were determined based on fold-change and *p*-values adjusted for multiple testing using Benjamini-Hochberg procedure to control for false discovery rate. Transcripts were considered rhythmically expressed when meeting the following criteria: (1) adjusted *p*-values ≤ 0.05 for both RAIN and MetaCycle; and (2) fold-change (maximal/minimal experimental values within a time series) ≥ 1.3. Of the 15,130 genes in the monarch genome, 68.3%, 66.8%, 66.6%, and 68.4% were respectively expressed in the brains of

LP, SP, summer-like, and migrant monarchs, of which 2.70%, 2.66%, 2.80% and 3.04% were respectively determined as rhythmically expressed (Table S7). Comparisons of rhythmic expression were performed between LP and SP monarchs, and between summer-like monarchs and wild-caught migrants. Genes were sorted based on their rhythmic expression in a photoperiod condition or a seasonal form and classified in three groups: (1) rhythmic in both LP and SP or summer-like and migrant monarchs, (2) rhythmic only in LP or summer-like monarchs, and (3) rhythmic only in SP or migrant monarchs. Heatmaps depicting all categories were produced using heatmap.2 in gplots package for R. To identify the genes that may be responsible for photoperiodic responses, genes present in both comparisons were identified. Results of this overlap analysis are summarized in Table S1. Using BLAST [53], homologous proteins from *Drosophila* and mouse that best matched the protein sequences of the overlapped genes were used to annotate them. Enriched gene ontology of biological processes and KEGG pathway were determined using Metascape (metascape.org).

2.5.7. gRNA Design and Construction.

The gRNA site for CRISPR/Cas9-mediated targeted mutagenesis of *ninaB1* was selected within exon 3 of the 14 exons-containing *ninaB1* using CHOPCHOP target site finder (http://chopchop.cbu.uib.no/index.php target site finder; [54, 55]. The gRNA expression vector was constructed by inserting

annealed synthetic oligomers into the DR274 plasmid from Addgene [56] at the *BsaI* cleavage site. Oligomer sequences were as follow (F, forward primer; R, reverse primer): *F*, 5'- TAGGAGTGACAACTATACGACCCG-3' and *R*, 5'-AAACCGGGTCGTATAGTTGTCACT-3'.

2.5.8. Synthesis of Cas9 mRNA and sgRNA.

In vitro transcription of *Streptococcus pyogenes* Cas9 mRNA was performed using the mMessage mMachine T3 transcription kit (Invitrogen) and pCS2-nCas9n expression plasmid from Addgene [57], as previously described [15]. The resulting capped PolyA mRNAs were purified by acid-phenol-chloroform extraction and resuspended in RNase-free water following isopropanol precipitation. The sgRNA was *in vitro* transcribed using T7 RNA polymerase (Promega) from purified PCR products containing the T7 promoter, gRNA and gRNA scaffold amplified from the DR274 vectors using the following primers; F: 5'-ATTGAGCCTCAGGAAACAGC-3' and R: 5'-AAAAGCACCGACTCGGTGCC-3'. The sgRNA was then treated with RQ1-DNase and purified by acid phenol-chloroform extraction and resuspended in RNase-free water after isopropanol precipitation. Cas9 mRNAs and the sgRNA were quantitated by spectrophotometry (NanoDrop 1000) and diluted in RNase-free water to a final concentration of 0.25 µg/µl for the sgRNA and 0.5 µg/µl for Cas9 mRNAs.

2.5.9. Egg Microinjections.

Eggs were collected and microinjected as previously described [58] with a mix of Cas9 mRNA at 0.5 µg/µl, sgRNA at 0.25 µg/µl, and blue food coloring for visual tracking of the injection. After injection, embryos were placed in an incubator at 25°C and 70% humidity. Developing embryos were transferred into individual small petri dishes containing milkweed leaves until larvae hatched. Larvae were fed milkweed leaves until the second larval instar before being transferred onto semi-artificial diet.

2.5.10. Analysis of CRISPR/Cas9-induced Mutations and Generation of a *ninaB1* Monarch Loss-of-Function Line.

PCR fragments flanking the targeted region were amplified from genomic DNA extracted from larval sensors of potential founders at the fifth instar with the following primers: *ninaB1F*, 5'-GTTTCACTTGTACCGTGACTTC-3' and *ninaB1R,* 5'-GGATACTGTTTAGCCAGGTACC-3. PCR products were purified using 2 X modified Sera-Mag Magnetic Speed-beads (GE Healthcare) as previously described [58, 59], and resuspended in 10 µl of RNase-free water. Cas9-based cleavage assays of PCR products (250-350 ng) were performed using a recombinant Cas9 protein (750 ng) and the sgRNA (400-600 ng), as previously described [58]. Digested products were purified using 2 X modified Sera-Mag Magnetic Speed-beads before being resolved with agarose gel electrophoresis and EtBr staining. Larvae presenting a high degree of targeting

in somatic cells, estimated based on the relative abundance of uncleaved fragments, were selected and reared to adulthood. Surviving adults of opposite sexes presenting the highest level of somaticism were hand-paired in individual cages to establish a mutant line. Eggs were collected and the hatched larvae were screened for the presence of mutated alleles as described above. Uncleaved fragments corresponding to mutated alleles were gel purified and sequenced using one of the primers used for PCR amplification. A 7-base pair deletion causing a frameshift and the introduction of a premature stop codon was selected to establish a mutant line.

2.5.11. Proboscis Extension Reflex Assay.

The proboscis extension reflex (PER) is a response in which stimulation of the chemoreceptors on the middle legs with a sugar solution (unconditioned stimulus, US) results in full extension of the proboscis (unconditioned response). Proboscis extension tests were conducted as previously described [34] by touching the middle leg with a cotton-tipped applicator soaked in 50% sucrose solution (wt/wt), with slight modifications. Individual butterflies were fed daily with 150 µl of a 25% honey solution prior to the experiment. The day prior to stimulus conditioning, individuals were harnessed in 15ml polypropylene conical tubes as previously described [34] and starved for 24 h. On the day of testing, individuals were checked for a positive proboscis extension reflex. Wild-type monarchs with eyes covered with either clear or black paint, and *ninaB1* -/- monarchs were

conditioned to a colored stimulus (red flag; conditioned stimulus, CS) by presenting the stimulus for 5 sec (CS only), contacting the middle legs with sucrose solution for 15 sec (CS +US), and removing the stimulus after 5 sec (US only). Individuals were then held for 5 min and this pairing procedure was performed again. US-CS pairing continued in this way for 13 to 14 trials a day for 3 consecutive days until the individual extended its proboscis upon the initial presentation of the CS, in which case a sucrose reward was given. Monarchs were considered to have a positive PER (conditioned response) if they fully extended their proboscis in response to the CS. On the second and third day of testing, butterflies were fed 50 µl of a 25% honey solution to prevent extreme starvation.

2.5.12. Real-time qPCR.

To test for the presence of a functional circadian clock in *ninaB1* loss-of-function monarchs, brains of adult wild-type and *ninaB1* homozygous mutant monarchs entrained to seven 15:9 LD cycles were dissected under red light on the first day of transfer into DD at circadian time (CT) 0, CT4, CT8, CT12, CT16, and CT20. Dissections were performed in 0.5X RNA later (Invitrogen) to avoid RNA degradation, and brains free of eye photoreceptors were stored at -80°C until use. Total RNA was extracted using RNeasy Mini kit (Qiagen), treated with RQ1 Dnase (Promega), and random hexamers (Promega) were used to prime reverse transcription with Superscript II Reverse Transcriptase (Thermo

Scientific), all according to the manufacturers' instructions. Quantifications of gene expression were performed on a QuantStudio™ 6 Flex Real-Time PCR System (Thermo Scientific) using iTaq Universal SYBR Green Supermix (Bio-Rad). PCR reactions were assembled by combining two master mixes: the first mix contained 5 µl of iTaq Universal SYBR® Green Supermix (Bio-Rad) and forward and reverse primers (5 µmol each) per reaction and the second mix contained approximately 7 ng of cDNA template and the water needed to bring each reaction to a final volume of 10 µl. The monarch *per* and control *rp49* primers were as follows (F, forward primer; R, reverse primer): *perF*, 5'-AGTGAAGCGTCCCTCAAAACA-3'; *perR*, 5'-TGGCGACGAGCATCTGTGT-3'; *rp49F*, 5'-TGCGCAGGCGTTTTAAGG-3'; *rp49R*, 5'-TTGTTTGATCCGTAACCAATGC-3'. The near 100% efficiency of each primer set was validated by determining the slope of Ct versus dilution plot on a dilution series. Individual reactions were used to quantify each RNA level in a given cDNA sample, and the average Ct from duplicated reactions within the same run was used for quantification. The data were normalized to *rp49* as an internal control, and normalized to the mean of one sample within a set for statistics.

2.5.13. Eclosion Behavior Assay.

Eclosion behavior was monitored during the first day of constant darkness (DD) as previously described [10] from butterflies raised in LD 15:9 in a Percival

incubator at 25 °C and 70% humidity. Eclosion data were analyzed and plotted as 1 hr bins.

2.5.14. Statistical Analysis.

P values were calculated using Student's t-tests, 1-way and 2-way ANOVAs followed by post hoc analyses, and Kolmogorov–Smirnov tests using online calculators at www.wessa.net/rwasp_Two%20Factor%20ANOVA.wasp [60], http://www.physics.csbsju.edu/stats/KS-test.n.plot_form.html, and https://www.graphpad.com/quickcalcs/posttest1.cfm.

2.5.15. Data Availability.

The RNA-seq datasets have been deposited in the Gene Expression Omnibus repository under accession number GSE126336.

2.6. References.

1. Ikegami, K. and T. Yoshimura, *Seasonal time measurement during reproduction.* J Reprod Dev, 2013. **59**(4): p. 327-33.
2. Saunders, D.S. and R.C. Bertossa, *Deciphering time measurement: the role of circadian 'clock' genes and formal experimentation in insect photoperiodism.* J Insect Physiol, 2011. **57**(5): p. 557-66.

3. Ikeno, T., H. Numata, and S.G. Goto, *Photoperiodic response requires mammalian-type cryptochrome in the bean bug Riptortus pedestris.* Biochem Biophys Res Commun, 2011. **410**(3): p. 394-7.
4. Ikeno, T., H. Numata, and S.G. Goto, *Circadian clock genes period and cycle regulate photoperiodic diapause in the bean bug Riptortus pedestris males.* J Insect Physiol, 2011. **57**(7): p. 935-8.
5. Ikeno, T., et al., *Photoperiodic diapause under the control of circadian clock genes in an insect.* BMC Biol, 2010. **8**: p. 116.
6. Meuti, M.E., et al., *Functional circadian clock genes are essential for the overwintering diapause of the Northern house mosquito, Culex pipiens.* J Exp Biol, 2015. **218**(Pt 3): p. 412-22.
7. Pegoraro, M., et al., *Role for circadian clock genes in seasonal timing: testing the Bunning hypothesis.* PLoS Genet, 2014. **10**(9): p. e1004603.
8. Meuti, M.E. and D.L. Denlinger, *Evolutionary links between circadian clocks and photoperiodic diapause in insects.* Integr Comp Biol, 2013. **53**(1): p. 131-43.
9. Zhan, S., et al., *The monarch butterfly genome yields insights into long-distance migration.* Cell, 2011. **147**(5): p. 1171-85.
10. Markert, M.J., et al., *Genomic Access to Monarch Migration Using TALEN and CRISPR/Cas9-Mediated Targeted Mutagenesis.* G3 (Bethesda), 2016. **6**(4): p. 905-15.

11. Merlin, C., et al., *Efficient targeted mutagenesis in the monarch butterfly using zinc-finger nucleases.* Genome Res, 2013. **23**(1): p. 159-68.

12. Denlinger, D.L., et al., *Keeping time without a spine: what can the insect clock teach us about seasonal adaptation?* Philos Trans R Soc Lond B Biol Sci, 2017. **372**(1734).

13. Reppert, S.M., P.A. Guerra, and C. Merlin, *Neurobiology of Monarch Butterfly Migration.* Annu Rev Entomol, 2016. **61**: p. 25-42.

14. Goehring, L. and K.S. Oberhauser, *Effects of photoperiod, temperature, and host plant age on induction of reproductive diapause and development time in Danaus plexippus.* Ecological Entomology, 2002. **27**: p. 674-685.

15. Zhang, Y., et al., *Vertebrate-like CRYPTOCHROME 2 from monarch regulates circadian transcription via independent repression of CLOCK and BMAL1 activity.* Proc Natl Acad Sci U S A, 2017. **114**(36): p. E7516-E7525.

16. Zhu, H.S., et al., *Cryptochromes define a novel circadian clock mechanism in monarch butterflies that may underlie sun compass navigation.* Plos Biology, 2008. **6**(1): p. 138-155.

17. Iiams, S.E., et al., *Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly.* Proc Natl Acad Sci U S A, 2019.

18. Bowen, M.F., et al., *In vitro reprogramming of the photoperiodic clock in an insect brain-retrocerebral complex.* Proc Natl Acad Sci U S A, 1984. **81**(18): p. 5881-4.

19. Saunders, D.S., *Insect photoperiodism: measuring the night.* J Insect Physiol, 2013. **59**(1): p. 1-10.

20. Thaben, P.F. and P.O. Westermark, *Detecting Rhythms in Time Series with RAIN.* Journal of Biological Rhythms, 2014. **29**(6): p. 391-400.

21. Wu, G., et al., *MetaCycle: an integrated R package to evaluate periodicity in large scale data.* Bioinformatics, 2016. **32**(21): p. 3351-3353.

22. Lugena, A.B., et al., *Genome-wide discovery of the daily transcriptome, DNA regulatory elements and transcription factor occupancy in the monarch butterfly brain.* PLoS Genet, 2019. **15**(7): p. e1008265.

23. Wang, T., Y.C. Jiao, and C. Montell, *Dissection of the pathway required for generation of vitamin A and for Drosophila phototransduction.* Journal of Cell Biology, 2007. **177**(2): p. 305-316.

24. Yang, J. and J.E. O'Tousa, *Cellular sites of Drosophila NinaB and NinaD activity in vitamin A metabolism.* Mol Cell Neurosci, 2007. **35**(1): p. 49-56.

25. Duester, G., *Families of retinoid dehydrogenases regulating vitamin A function: production of visual pigment and retinoic acid.* Eur J Biochem, 2000. **267**(14): p. 4315-24.

26. von Lintig, J., et al., *Analysis of the blind Drosophila mutant ninaB identifies the gene encoding the key enzyme for vitamin A formation invivo.* Proc Natl Acad Sci U S A, 2001. **98**(3): p. 1130-5.

27. Schnorrer, F., et al., *Systematic genetic analysis of muscle morphogenesis and function in Drosophila.* Nature, 2010. **464**(7286): p. 287-91.

28. Zhan, S., et al., *The genetics of monarch butterfly migration and warning colouration.* Nature, 2014. **514**(7522): p. 317-21.

29. Shearer K. D., et al., *Photoperiodic regulation of retinoic acid signalling in the hypothalamus.* Journal of Neurochemistry, 2010. **112**: p. 246-257.

30. Shearer K. D., et al., *Photoperiodic expression of two RALDH enzymes and the regulation of cell proliferation by retinoic acid in the rat hypothalamus* Journal of Neurochemistry, 2012. **122**: p. 789-799.

31. Wood, S.H., et al., *Binary Switching of Calendar Cells in the Pituitary Defines the Phase of the Circannual Cycle in Mammals.* Curr Biol, 2015. **25**(20): p. 2651-62.

32. Helfer G., et al., *Photoperiod regulates vitamin A and Wnt/b-catenin signalling in F344 rats.* Endocrinology, 2012. **153**: p. 815-824.

33. Voolstra, O., et al., *NinaB Is Essential for Drosophila Vision but Induces Retinal Degeneration in Opsin-deficient Photoreceptors.* Journal of Biological Chemistry, 2010. **285**(3): p. 2130-2139.

34. Merlin, C., R.J. Gegear, and S.M. Reppert, *Antennal circadian clocks coordinate sun compass orientation in migratory monarch butterflies.* Science, 2009. **325**(5948): p. 1700-4.

35. Tauber, E. and B.P. Kyriacou, *Insect photoperiodism and circadian clocks: models and mechanisms.* J Biol Rhythms, 2001. **16**(4): p. 381-90.

36. Goto, S.G. and D.L. Denlinger, *Short-day and long-day expression patterns of genes involved in the flesh fly clock mechanism: period, timeless, cycle and cryptochrome.* J Insect Physiol, 2002. **48**(8): p. 803-816.

37. Iwai, S., et al., *Structure and expressions of two circadian clock genes, period and timeless in the commercial silkmoth, Bombyx mori.* J Insect Physiol, 2006. **52**(6): p. 625-37.

38. Majercak, J., et al., *How a circadian clock adapts to seasonal decreases in temperature and day length.* Neuron, 1999. **24**(1): p. 219-30.

39. Qiu, J. and P.E. Hardin, *per mRNA cycling is locked to lights-off under photoperiodic conditions that support circadian feedback loop function*, in *Molecular and Cellular Biology*. 1996. p. 4182–4188.

40. Denlinger, D.L., G.D. Yocum, and J.P. Rinehart, *Hormonal Control of Diapause* in *Insect Endocrinology* 2012.

41. Veerman, A., et al., *Vitamin-a Is Essential for Photoperiodic Induction of Diapause in an Eyeless Mite.* Nature, 1983. **302**(5905): p. 248-249.

42. Veerman, A., et al., *Photoperiodic Induction of Diapause in an Insect Is Vitamin-a Dependent.* Experientia, 1985. **41**(9): p. 1194-1195.

43. Claret, J. and N. Volkoff, *Vitamin A is essential for two processes involved in the photoperiodic reaction in Pieris brassicae.* Journal of Insect Physiology, 1992. **38**: p. 569-574.

44. Saunders, D.S., *Insect Clocks*. 2002: Elsevier. 576.

45. Truman, J.W., *Physiology of Insect Rhythms - .2. Silkmoth Brain as Location of Biological Clock Controlling Eclosion.* Journal of Comparative Physiology, 1972. **81**(1): p. 99-&.

46. Bui-Gobbels, K., et al., *Is retinoic acid a signal for nerve regeneration in insects?* Neural Regeneration Research, 2015. **10**: p. 901-903.

47. Bolborea, M., et al., *Melatonin controls photoperiodic changes in tanycyte vimentin and neural cell adhesion molecule expression in the Djungarian hamster (Phodopus sungorus).* Endocrinology, 2011. **152**(10): p. 3871-83.

48. Zhu, H., et al., *Defining behavioral and molecular differences between summer and migratory monarch butterflies.* BMC Biol, 2009. **7**: p. 14.

49. Zhan, S. and S.M. Reppert, *MonarchBase: the monarch butterfly genome database.* Nucleic Acids Res, 2013. **41**(Database issue): p. D758-63.

50. Kim, D., et al., *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.* Genome Biology, 2013. **14**(4).

51. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nature Protocols, 2012. **7**(3): p. 562-578.

52. Trapnell, C., et al., *Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation.* Nature Biotechnology, 2010. **28**(5): p. 511-U174.

53. Altschul, S.F., et al., *Basic Local Alignment Search Tool.* Journal of Molecular Biology, 1990. **215**(3): p. 403-410.

54. Labun, K., et al., *CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering.* Nucleic Acids Research, 2016. **44**(W1): p. W272-W276.

55. Montague, T.G., et al., *CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing.* Nucleic Acids Research, 2014. **42**(W1): p. W401-W407.

56. Hwang, W.Y., et al., *Efficient genome editing in zebrafish using a CRISPR-Cas system.* Nature Biotechnology, 2013. **31**(3): p. 227-229.

57. Jao, L.E., S.R. Wente, and W.B. Chen, *Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(34): p. 13904-13909.

58. Markert, M.J., et al., *Genomic Access to Monarch Migration Using TALEN and CRISPR/Cas9-Mediated Targeted Mutagenesis.* G3-Genes Genomes Genetics, 2016. **6**(4): p. 905-915.

59. Rohland, N. and D. Reich, *Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture.* Genome Res, 2012. **22**(5): p. 939-46.

60. Wessa, P. *Free Statistics Software*. 2018; version 1.2.1:[Available from: https://www.wessa.net/.

3. FURTHER CHARACTERIZATION OF THE VITAMIN A PATHWAY IN PHOTOPERIODIC RESPONSES: FROM DIETRAY SUPPLEMENTATION TO A POTENTIAL PHOTOPERIODIC SENSOR

3.1. Overview.

In Chapter 2, we have genetically confirmed the role of the vitamin A pathway in regulating photoperiodic responses. This result was consistent with previous findings showing that dietary deprivation of carotenoids led to the loss of photoperiodic responses in some insect and mite species [1-3]. Interestingly, supplemental feeding with various vitamin A derivatives, including β-carotene, 3'-Hydroxyechinenone, and vitamin A acetate, have been shown to restore photoperiodic responses. Since dietary complementation experiments have not yet been tested in the monarch and given a diapausing-like *dpCry2* knockout available in the lab, I could rapidly test the role of individual carotenoids and retinoids in rescuing oocyte production. Unexpectedly, I found that supplemental feeding with β-carotene, retinal, and retinol did not rescue the diapausing-like phenotype of *dpCry2* knockouts, likely due to either insufficient concentration or inefficient uptake of the metabolites by the relevant brain cells.

Despite having identified several candidate genes belonging to the vitamin A pathway, work in the previous Chapter focused solely on the knockout of the rate limiting enzyme, *ninaB1*. Characterizing the roles of the seasonally expressed *santa maria 1* and *2* and *retinol dehydrogenase 13* (*rdh13*) in

photoperiodic responses remained to be tested. Using CRISPR/Cas9-mediated mutagenesis, I generated a knockout of *santa maria 1* and found the mutant was mostly non-viable after pupation. Further study may require dietary supplementation to increase the survival rate through pupation, but a working feeding regime has yet to be found. Additionally, I have generated a knockout of *rdh13*, which will be useful for determining which of the two products of the vitamin A pathway, retinoic acid or vitamin A itself, is they key signaling molecule for photoperiodic responses.

Lastly, in this chapter I searched for a candidate opsin which may be responsible for sensing photoperiodic changes and initiating subsequent signaling pathways to induce a response in the monarch. Studies from reptiles and birds suggest that the photoperiodic sensor must be expressed in the brain. I thus tested if any of the five opsins present in the monarch genome (*blue opsin* (DPOGS214795), *long wavelength opsin* (DPOGS208077), *parapinopsin* (DPOGS200228), *peropsin* (DPOGS211151), and *UV opsin* (DPOGS213510)) would be specifically expressed in the brain. While real-time qPCR did not reveal any opsins with brain specific expression, all were expressed in the brain, making these opsins all interesting candidates for future functional characterization in photoperiodic responses.

3.2. Introduction.

As previously described in greater detail in Chapter 2, the vitamin A

pathway functions in the uptake of carotenoids and their subsequent biosynthesis into retinoids, which are utilized in processes such as vision via formation of the light-sensitive chromophore of the rhodopsin photoreceptor, as well as multiple signaling pathways in the brain [4-10]. While the significance of vitamin A in vision has been established nearly a century ago [10], its role in various physiological and behavioral signaling pathways within the brain are still being brought to light [8, 9].

The impact of the vitamin A pathway on photoperiodic responses was first discovered in spider mites, parasitic wasps, and cabbage moths as rearing them on carotenoid deprived diets completely disrupted their photoperiodic diapause responses [1-3]. Only supplementation with β-carotene, 3'-Hydroxyechinenone, or vitamin A acetate was able to restore these responses [1-3]. Given that wasp larvae have no ocelli, simple eyes or eyespots, and yet still exhibited robust, carotenoid-supported responses to the photoperiod, it was proposed that photoperiodic timing occurs directly in the brain [3]. This was later demonstrated through an array of elegantly designed experiments in several species: (1) administration of light only to the head, which induced photoperiodic responses in aphids and silkworms [11], (2) transplantation of different photoperiod exposed pupal heads into the abdomen of a recipient, which changed the response to that of the donors entrainment in silkworms and cabbage moths [12, 13], (3) and ablation of group I neurosecretory cells within the brains of aphids, which abolished responses, unlike the ablation of the eyes,

optic lobes, or other groups of neurosecretory cells [14]. Despite solid evidence provided by these classical studies that vitamin A and the brain were key to photoperiodic responses, genetic proof of the vitamin A pathway's role in regulating photoperiodic responses and the identification of the photoperiodic sensor was lagging behind.

As genetic tools became more readily available in the model organism *Drosophila melanogaster*, the components of the vitamin A pathway started to be characterized, with *ninaB* being identified as the rate-limiting monooxygenase converting β-carotene into retinal [4, 15], and *santa maria* as class B scavenger receptors for the uptake of carotenoids into extraretinal neural cells [16, 17]. Mutations in *santa maria* rendered flies incapable of producing rhodopsin, and this phenotype was shown to be rescued by dietary supplementation with retinal but not with β-carotene [18]. Contributing further to the characterization of this pathway, our work provided genetic evidence of the vitamin A pathway's brain-based role in regulating photoperiodic responses, via a knockout of monarch *ninaB1* (Fig. 2.6 *A*-*C*) [19].

Taken together, current data prompts two major questions: (1) Which of the vitamin A pathway products is the signaling molecule for photoperiodic responses, retinoic acid (RA) or vitamin A (retinol)?, and (2) What is/are the photoperiodic sensors? One possibility is that RA, also a product of retinal converted by a retinaldehyde dehydrogenase, functions as the photoperiodic response-inducing molecule by binding to retinoic acid receptors to regulate

gene expression [20-23]. RA concentration has been shown to change seasonally in a glial cell type called tanycytes lining the third ventricle of the hypothalamus in seasonal mammals [20-22, 24], and tanycytes have been shown to undergo seasonal remodeling, encasing the terminals of gonadotropin releasing hormone neurons during the non-breeding season [24, 25]. Seasonal changes in RA concentration may thus regulate gene expression and/or neuronal remodeling to induce the behavioral and physiological seasonal switch in the monarch. The other possibility is that it is instead the seasonal production of retinol, converted from retinal by *rdh13*, contributes to the seasonal production of an opsin-based photoperiodic sensor. Opsins are light-sensitive, G protein-coupled receptors (GPCR) found in animals and initiate signal transduction pathways when their conjugated chromophore is activated by light [26]. While many opsins function in vision, several non-visual opsins have been discovered with light sensitive functions in the brains of birds [27-29]. Because invertebrate G_q-coupled opsins, a type of G protein alpha subunit that activates phospholipase C, use the chromophore 11-*cis* retinal, a product of retinol, the production of retinol by the vitamin A pathway may be seasonally regulating the production of a functional deep brain photoperiodic sensor [5-7, 26, 30-33]. To this end, and delving into the second major question, five opsin genes present in the monarch genome appear as candidates, respectively encoding: blue opsin (DPOGS214795), long wavelength opsin (DPOGS208077), parapinopsin (DPOGS200228), peropsin (DPOGS211151), UV opsin (DPOGS213510).

Testing for brain specific expression may help narrow down a candidate opsin responsible for sensing and signaling photoperiodic changes.

Here, I aimed to mimic similar dietary supplementation experiments in the monarch butterfly in order to determine the role of the carotenoid and retinoids products in photoperiodically-induced oocyte production using the diapausing-like *dpCry2* knockout monarch line. Additionally, I sought to further characterize the vitamin A pathway by knocking out the seasonally expressed *santa maria 1* and *rdh13* and test their respective roles in regulating seasonal responses. Although the latter may help to determine whether RA or vitamin A are signaling photoperiodic responses, the two hypotheses are not mutually exclusive. Therefore, regardless of the result of the *rdh13* knockout, the photoperiodic sensor still needs to be identified. To this end, I also began identifying which opsins are expressed in the monarch brain to narrow down a potential candidate for future characterization.

3.3. Results.

3.3.1. Dietary Supplementation of β-carotene, Retinal, or Retinol Does not Rescue the Diapausing Phenotype of *dpCry2* Knockouts.

Our findings in chapter 2 suggested that the rhythmic production of retinal is necessary for photoperiodically-induced reproductive diapause responses. Specifically, differences in *ninaB1* expression levels, constitutively low in *dpCry2* mutants and higher in *dpClk* mutants and, suggested that overall retinal levels

may be key in driving oocyte maturation (Fig. 2.5 *B*). I thus proposed that supplementing *dpCry2* knockouts with retinal would increase the maturation of oocytes, regardless of the photoperiod. Expression of *santa maria* is also low in *dpCry2* mutants compared to wild-type monarchs, and based on experiments from prior studies [18], this is likely to impact the ability of *dpCry2* mutants to efficiently uptake β-carotene, so I included supplementation with β-carotene as a control group. All mutants and their wild-type siblings were raised from eggs to adults in LP (15 hr light:9 hr dark) or SP conditions (10 hr light:14 hr dark) at a constant temperature of 21 °C. Larvae were raised on milkweed leaves for the first 4-5 days after hatching before being switched to artificial diet either containing 0.02mM β-carotene or 0.02mM retinal. Adult females were fed the same concentrations in a 25% honey mixture for 10 days post-eclosion before being dissected to quantify the number of oocytes produced. Curiously, dietary supplementation with 0.02mM of retinal did not rescue the diapausing-like phenotype of *dpCry2* knockouts, where oocyte production from *dpCry2* knockouts in LP and SP was still significantly lower than wild-type monarchs raised in LP (Fig. 3.1 *A*). The β-carotene control group did show an unexpected significant increase in oocyte production in SP *dpCry2* mutants, on par with oocyte production in LP wild-types (Fig. 3.1 *B*). I next decided to test a product further downstream in the pathway, retinol, and also included a vehicle control group to ensure that ethanol, in which each of these derivatives were diluted in, was not causing complications. Mutants and their wild-type siblings were raised

as before but received either 0.08mM β-carotene, 0.08mM retinol, or an equal volume of ethanol as vehicle control in their artificial diet and honey. Oocyte production was not significantly increased in *dpCry2* knockouts in the ethanol control or 0.08mM β-carotene groups as compared to their photoperiodically responsive wild-type siblings in SP (Fig. 3.2, A, B). However, although not statistically significant, an upward trend in oocyte production in females treated with 0.08mM retinol was observed compared to wild-type siblings in SP (Fig. 3.2, C). These results demonstrate that, at least at the concentrations used, neither β-carotene, retinal nor retinol can rescue the diapause-like phenotype of *dpCry2* knockouts.

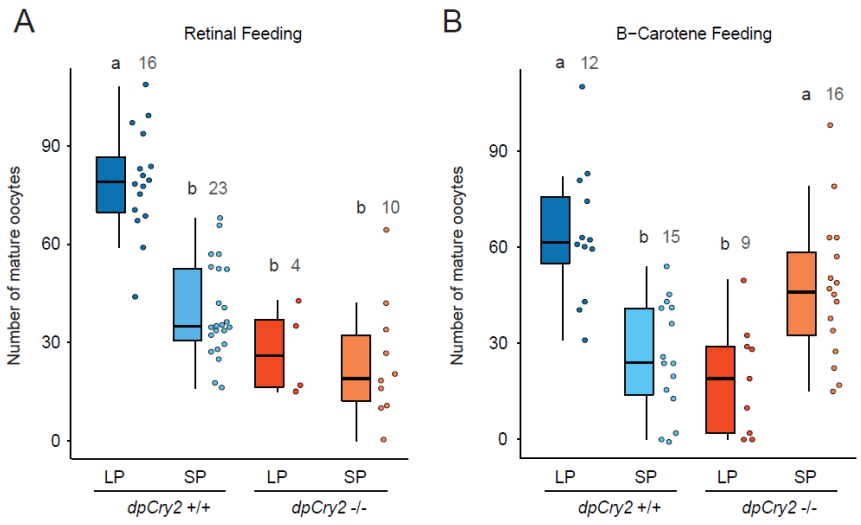


Figure 3.1: Retinal supplementation does not rescue the diapausing phenotype of *dpCry2* knockouts.

(*A*) Number of mature oocytes produced 10 days post adult emergence in *dpCry2* homozygous mutant and wild-type sibling females after feeding larvae with artificial diet and adults with honey supplemented with 0.02mM retinal. Boxplots as in Fig. 2.1. Interaction genotype × photoperiod, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$.(*B*) Number of mature oocytes produced 10 days post adult emergence in *dpCry2* homozygous mutant and wild-type sibling females after feeding larvae with artificial diet and adults with honey supplemented with 0.02mM β-carotene. Boxplots as in Fig. 2.1. Interaction genotype × photoperiod, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$.

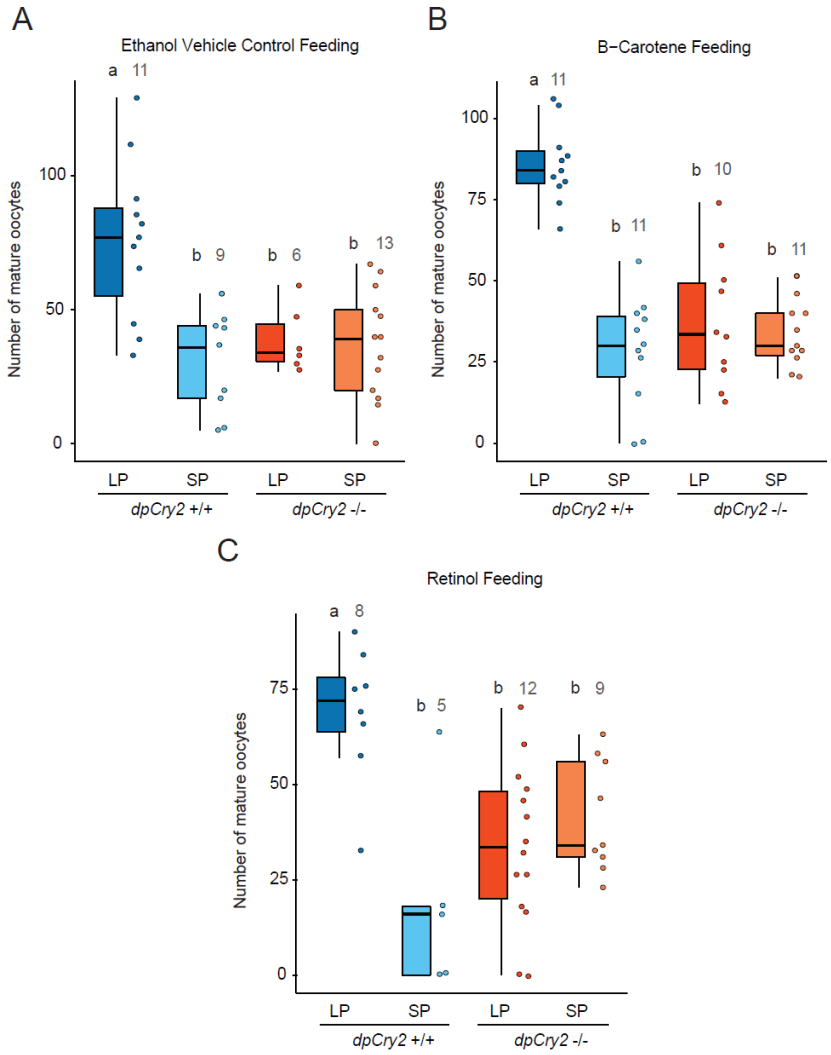


Figure 3.2: Retinol supplementation does not rescue the diapausing phenotype of *dpCry2* knockouts.

(*A*) Number of mature oocytes produced 10 days post adult emergence in *dpCry2* homozygous mutant and wild-type sibling females after feeding larvae with artificial diet and adults with honey supplemented with ethanol, the vehicle control. Boxplots as in Fig. 2.1. Interaction genotype × photoperiod, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$. (*B*) Number of mature oocytes produced 10 days post adult emergence in *dpCry2* homozygous mutant and wild-type sibling females after feeding larvae with artificial diet and adults with honey supplemented with 0.08mM β-carotene. Boxplots as in Fig. 2.1. Interaction genotype × photoperiod, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$. (*C*) Number of mature oocytes produced 10 days post adult emergence in *dpCry2* homozygous mutant and wild-type sibling females after feeding larvae with artificial diet and adults with honey supplemented with 0.08mM retinol. Boxplots as in Fig. 2.1. Interaction genotype × photoperiod, 2-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$.

3.3.2. The Majority of *Santa maria 1* Knockouts are not Viable Past Pupation.

To characterize the roles of the others vitamin A pathway genes in photoperiodic responses, I used CRISPR/Cas9-mediated mutagenesis to target them *in vivo*, starting with the transmembrane protein encoding gene *santa maria 1*. Targeting the second exon resulted in a germline transmitted 80 bp deletion in the nucleotide sequence at the sgRNA targeting site generating an early stop codon that should render the protein non-functional and inhibit β-carotene uptake into extraretinal neural cells (Fig. 3.3 *A*) [18]. After establishing a mutant line, these loss-of-function mutants were raised along with their wild-type and heterozygous carrier siblings in LP and SP regimes as before to observe the female reproductive responses to the photoperiod [19]. Before switching the larvae to the artificial diet, I noticed a distinct difference in coloration among the larvae as they fed on leaves, with some displaying the normal black, white, and yellow striping pattern and others missing this yellow coloration completely (Fig. 3.3 *B*). The larvae were labeled individually for their color phenotype at the day of the dietary switch and within a few days the colorless larvae regained full coloring. Genotyping results confirmed my suspicions that the originally colorless larvae were homozygous for the *santa maria 1* mutation while heterozygous and wild-type siblings had retained their coloration throughout their development. Carotenoids have been shown to contribute to insect pigmentation [4], and the artificial diet (recipe patented with

Monarch Watch at the University of Kansas) may provide carotenoids or vitamin A derivatives downstream of *santa maria* function that are not present in the milkweed our larvae were fed on. In the days shortly before and during the process of pupation, the majority of *santa maria* 1 homozygous mutants exhibited lethal complications leaving only 3 survivors from 50 larvae in the LP conditions and 1 survivor from 51 larvae in SP (Fig. 3.3 *C*). While pupation is an energy-demanding and complex process, the wild-type and heterozygous siblings exhibited a rate of death routinely observed in our colony, typically no more than ~25%, which was significantly lower than that of the full mutants (Fig. 3.3 *C*). Together, these results suggest that the presence of β-Carotene and the ability to it take up plays a significant role in the process of pupation.

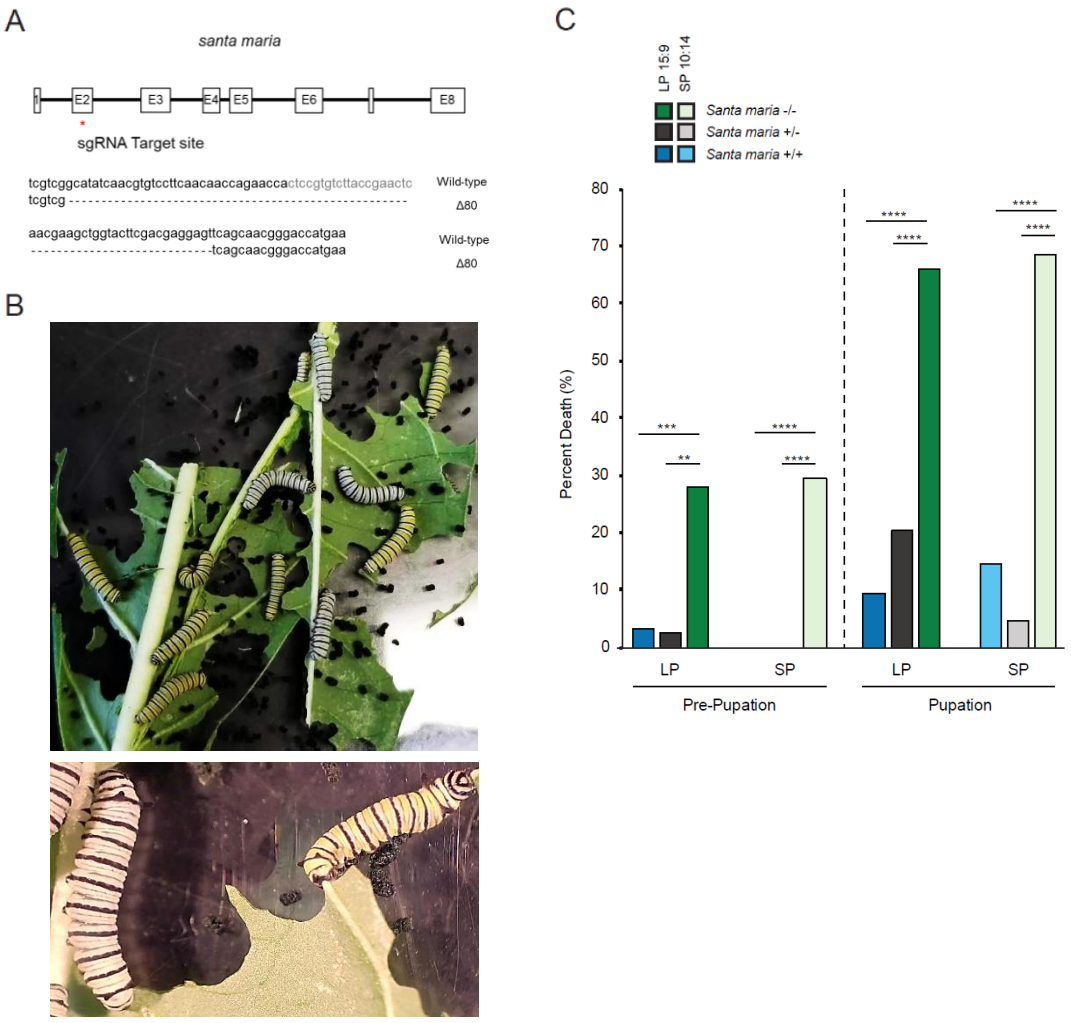


Figure 3.3: *Santa maria 1* knockouts exhibit coloration and pupation deficiencies.

(*A*) Generation of a *santa maria 1* loss-of-function mutant using CRISPR/Cas9-mediated targeted mutagenesis. (*Top*) Genomic structure of *santa maria 1* in the monarch. Red star, position of the site targeted by sgRNA. (*Bottom*) Nucleotide sequence of the targeted site (gray) and introduced frame-shifting mutation (80-bp deletion). (*B*) Two color type patterns, yellow striped and colorless, were observed in the progeny of a *santa maria 1* heterozygous x heterozygous cross. (*C*) Of the larvae that successfully reached the fifth instar stage, chart shows the percentage of death among *santa maria 1* loss-of-function mutants (green) compared to their wild-type (blue) and heterozygous (gray) siblings in LP or SP conditions in the days just before pupation or in the process of pupating. Total larvae numbers are as follows: *santa maria 1* -/-, LP-50 and SP-51; *santa maria 1* +/+, LP-63 and SP-48; and *santa maria 1* +/-, LP-39 and SP-43. Two tailed Z-score test for two population proportions: ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$.

3.3.3. *In vivo* Knockouts of *rdh13* to test if Retinol acts as a Photoperiodic Signaling Molecule of the Vitamin A Pathway.

Following procedures described before, I used CRISPR/Cas9-mediated mutagenesis to target the only exon of the gene encoding the retinal converting enzyme *rdh13 in vivo*. This resulted in a 1 bp deletion in the nucleotide sequence at the sgRNA targeting site generating an early stop codon that should render the truncated protein non-functional (Fig. 3.4, A and B). These loss-of-function mutants, which have not yet been tested due to a colony contamination, are now healthy enough to be tested for their effect on photoperiodically-induced reproductive responses to determine if retinol is the key signaling molecule of these responses.

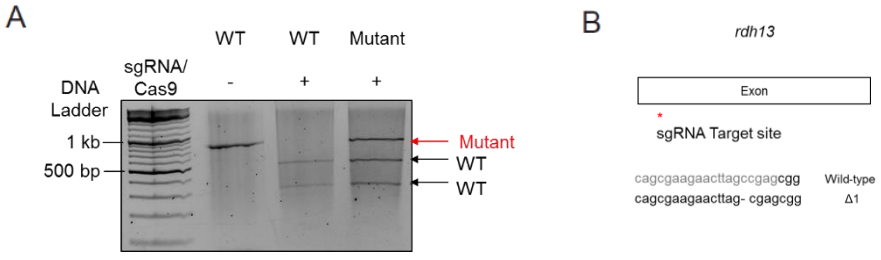


Figure 3.4: Generation of a *rdh13* loss-of-function mutant using CRISPR/Cas9-mediated targeted mutagenesis *in vivo*.
(*A*) Cas9-*in vitro* cleavage assay. An uncut wild-type amplicon and fully cleaved wild-type DNA amplicon subjected to the sgRNA/Cas9 mix are shown as negative and positive controls. The germline mutant used to found the knockout line shows an uncut band when subjected to the sgRNA/Cas9 mix indicating the presence of a mutated allele. (*B*) (*Top*) Genomic structure of *rdh13* in the monarch. Red star, position of the site targeted by sgRNA. (*Bottom*) Nucleotide sequence of the targeted site (gray) and introduced frame-shifting mutation (1-bp deletion).

3.3.4. Identifying Candidates Opsin for a Deep-Brain Photoperiodic Sensing.

While studies have begun to identify putative opsins as deep-brain photoperiodic sensors in birds and reptiles, the seasonal deep-brain receptor in insects is still unknown [29, 34]. Using the available annotated monarch genome, I was able to identify five opsins: *blue opsin* (DPOGS214795*)*, *long wavelength opsin* (DPOGS208077), *UV opsin* (DPOGS213510), *peropsin* (DPOGS211151) [35], and *parapinopsin* (DPOGS200228). I tested the expression of each of these opsins by RT-qPCR in various tissues of both male and female monarchs, including the abdomen, thorax, front pair of legs, antennae, optic lobes, eye photoreceptors, and central brain, to determine if any exhibited brain specific expression. I found that all five opsins were expressed in the eyes of both sexes and also in the brain, with *long wavelength opsin* showing the highest enrichment and *parapinopsin* very little expression (Fig. 3.5 and 3.6). While RT-qPCR of opsin expression did not reveal a brain specific opsin, suggesting that all opsins are deep-brain photoreceptor candidates, it has revealed an intriguing expression pattern for *long wavelength opsin*.

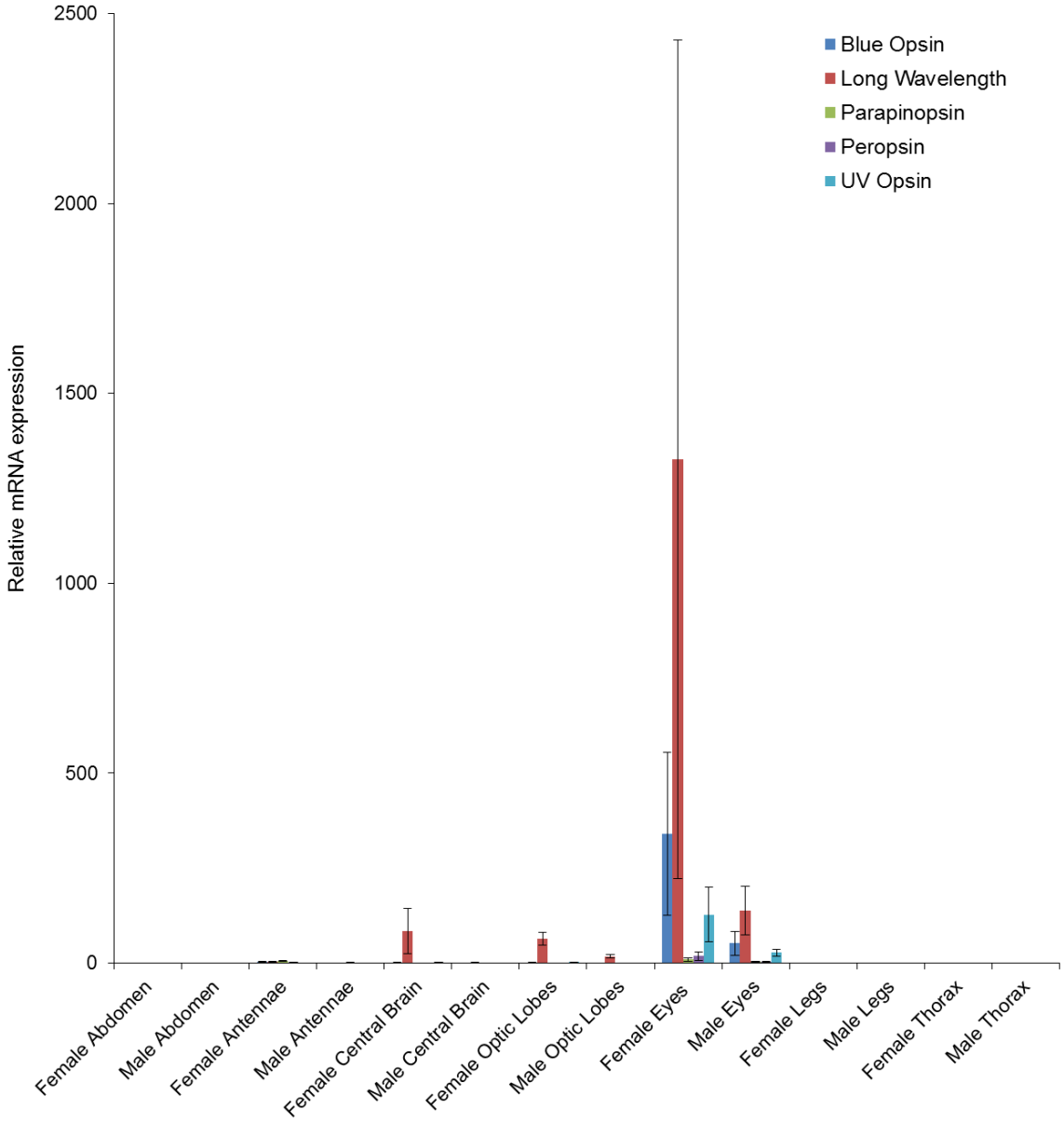


Figure 3.5: Five key monarch opsins exhibit expression mainly in eyes, but are also expressed in brain.

mRNA expression of monarch Blue Opsin, Long Wavelength Opsin, Parapinopsin, Peropsin, and UV Opsin in either male or female tissues from the Abdomen, Antennae, Central Brain, Optic Lobes, Eyes, Legs, and Thorax.

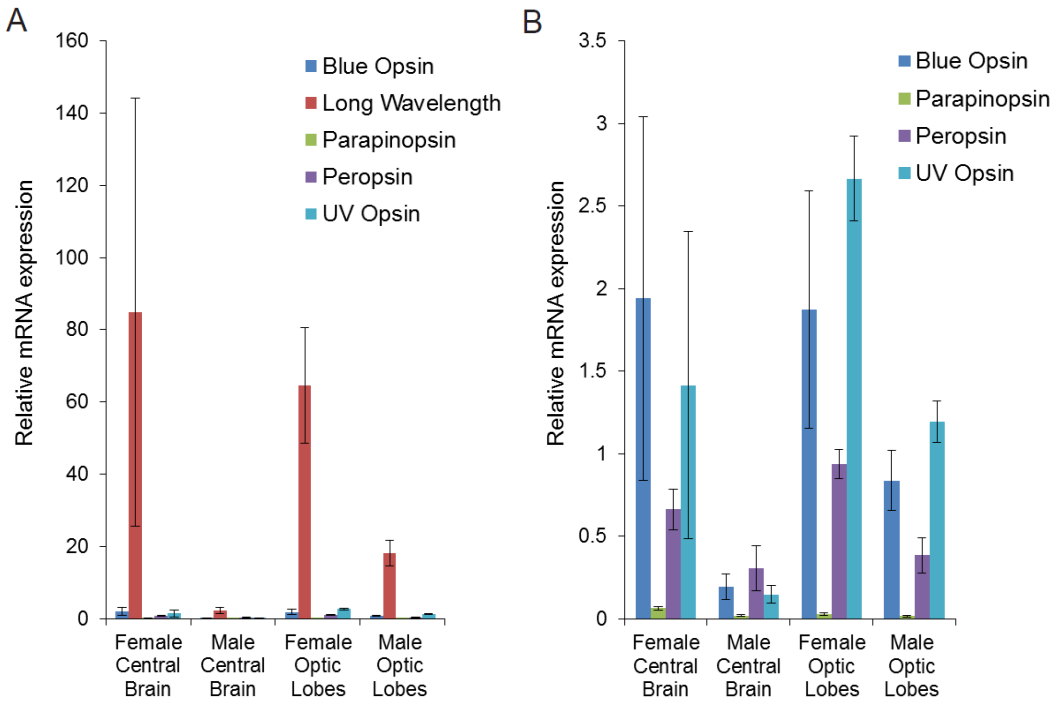


Figure 3.6: Zoomed in mRNA expression of the opsins in the brain. (*A*) Blue Opsin, Long Wavelength Opsin, Parapinopsin, Peropsin, and UV Opsin in male and female central brain and optic lobes from *A*. (*B*) Blue Opsin, Parapinopsin, Peropsin, and UV Opsin in male and female central brain and optic lobes.

3.4. Discussion.

The involvement of vitamin A pathway metabolites in regulating photoperiodic responses has long been observed in several insect species. While dietary supplementation experiments have been shown to restore reproductive photoperiodic responses [1-3], and even rhodopsin production [18], I was not able to restore wild-type levels of oocyte production in a similar manner in diapausing-like monarch *dpCry2* knockouts. The control group fed β-carotene was not expected to restore oocyte production as both *santa maria* and

ninaB1 expression are low in *dpCry2* knockouts, significantly reducing the uptake of β-carotene and its conversion into retinal. Although I saw a significant increase in oocyte production in SP when fed 0.02mM of β-carotene, I did observe the expected result at the 0.08mM concentration. The increase in oocyte production observed in *dpCry2* knockouts raised in SP and fed 0.02mM of β-carotene is difficult to explain and may be an artifact. Repeating the experiment for the 0.02mM concentration of β-carotene supplemental feeding could test this. However, more importantly, feeding *dpCry2* knockouts with retinal or retinol did not significantly increase oocyte production, which was unexpected as these products are downstream of the low expressed *ninaB1* and should in theory be metabolized. Although the data for the retinol fed group appear to show increased oocyte maturation in *dpCry2* knockouts as compared to wild-type SP raised females, the low number of mature oocytes of wild-type monarchs raised in SP does not appear to be an accurate representation as compared to the number observed for this group in other experiments, and this could be due to the fact that I was only able to raise, genotype and phenotype only 5 individuals for this group. Increasing the overall number of dissected females for each group would likely show an oocyte average in wild-type SP raised females similar to those of the β-carotene, retinal, and vehicle control groups. In addition, there could be several reasons why the feeding of retinal or retinol did not show a significant oocyte increase. For one, the concentration of these metabolites may still be too low to induce a response. Increasing their

concentration is however limited by the amount of ethanol that can be used as too much ethanol in the artificial diet causes it to be soupy and not form a solid pad. Another reason could be that the method of administration of the retinoids prevented correct uptake. Retinol is sensitive to degradation by light, and although I prepared the reagent and diet in dark conditions, the exposure to light within the incubators could have caused rapid degradation of the product, making it non-functional by the time it was ingested. In addition, the ethanol-based administration could be limiting the mechanisms of retinoid uptake in the gut, although ethanol itself does not impair overall health as varying volumes were fed to non-diapausing-like *dpCyc-like* mutants and had no effect on oocyte production (data not shown). A third reason could be that the retinoids I used had the wrong form or chirality, as it has been shown that the chirality of carotenoids and retinoids utilized by the vitamin A pathway vary across species and are based on feeding habits rather than phylogeny [4, 10, 36, 37]. Given that only the *all-trans* forms were tested here, switching to a different form of retinal and retinol could restore oocyte maturation in *dpCry2* knockouts. Finally, the *dpCry2* knockout itself may be causing issues. As we cannot eliminate the possibility that dpCRY2 may play a role in regulating the pathways involved in carotenoid/retinoid uptake, continuation of these supplementation experiments are likely not to prove fruitful.

While a knockout of *ninaB1* in the previous chapter was integral to genetically implicating the role of the vitamin A pathway in regulating

photoperiodic responses, knockouts of the two genes of this pathway with seasonal patterns of expression, *santa maria 1* and *rdh13*, had yet to be generated and explored for their effect on photoperiodic responses. Knocking out the transmembrane protein, *santa maria 1*, was of interest as disrupting β-carotene uptake into the extraretinal neural cells should completely disrupt the vitamin A pathway, and thus photoperiodic responses. From the heterozygous by heterozygous crosses, I found that approximately a fourth of the progeny exhibited a colorless phenotype which later genotyping confirmed were homozygous for the *santa maria 1* mutation. This coloration deficiency was not too surprising as carotenoids have been shown to give pigmentation to plants, birds, crustaceans, and insects [4, 38]. Interestingly though, normal coloration was restored in the mutants upon transferring to artificial diet, which could be due to the consumption of additional vitamin A pathway metabolites not found in the milkweed leaf and functioning downstream of *santa maria 1* or possibly the uptake of other carotenoid types through *santa maria 2,* another seasonally expressed scavenger receptor identified in Chapter 2. As the artificial diet recipe is purchased from the Monarch Watch program of the University of Kansas and we do not know the precise composition of the diet, it is difficult to predict which carotenoids may be involved. However, generating a double mutant for both *santa maria* genes and testing for coloration deficiency throughout the larval development could distinguish between these possibilities. The finding that the large majority of *santa maria 1* homozygous mutants failed to pupate and died

suggests that β-carotene transportation into the extraretinal neural cell by SANTA MARIA 1 is essential for the process of pupation, and that *santa maria 2* could have diverged functionally from its paralogue. Although the few homozygous survivors may have been able to utilize *santa maria 2* to uptake carotenoids to some extent, taken together with the robust colorless phenotype, the results support the idea that *santa maria 1* plays a dominant role in carotenoid transportation into the cell and is necessary for the energy-intensive process of pupation. Photoperiodic responses could possibly be tested in these mutants by supplementing the diet with retinoids downstream of *santa maria 1*, however, a successful supplementation regime in the monarch still needs to be developed. The other seasonally expressed vitamin A pathway gene, *rdh13*, has been generated and can now be tested for photoperiodic responses to test if retinol contributes to the signaling of photoperiodic responses. As described in detail in Chapter 2, if these mutants retain the ability to respond to the photoperiod then RA might be responsible for signaling photoperiodic responses and may do so by regulating gene expression and/or neuronal remodeling on a seasonal basis. However, the possibility cannot be excluded that other functionally redundant *rdh* genes, missed by the Chapter 2 RNA-Seq data set, may still be driving photoperiodic responses. Conversely, if the mutants lose photoperiodic responses, it would suggest that retinol is the key signaling molecule and may be regulating these responses by seasonally producing functional photoperiodic sensing opsins.

The cellular location and identity of the photoperiodic sensor has remained poorly understood. Over the years, several behavioral and molecular experiments in insects and birds have provided evidence that the sensor must lie somewhere in the brain [3, 11-14, 27-29]. In quails, a deep-brain Opsin 5 has been identified in the regulation of photoperiodically-induced reproductive responses [29]. However, such an opsin has yet to be identified in any insect. Here, I identified five opsin candidates in the monarch genome and found that all were expressed in the brain and, curiously, higher so in the brains of females than males. The reason for this sex-related difference in opsin expression still remains to be explored, but it does indicate that females may be more sensitive to light, and potentially photoperiodic sensing. Although none of the opsins exhibited brain-specific expression to narrow down a candidate photoperiodic sensor, monarch long wavelength opsin may be may be a contender given in hawkmoths it is co-expressed with melatonin, known to be involved in inducing seasonal responses, in the dorsal and ventral neurons in the optic lobe [25, 39]. Another possible candidate is that of parapinopsin as it has been shown to be expressed in the catfish and lamprey pineal gland, a region of the brain implicated in conveying photoperiodic information via the daily pattern of melatonin secretion, as well as in the parietal eye of reptiles which functions to transmit light signals to the pineal gland [25, 40-42]. While monarchs do not have a pineal gland within their brain, parapinopsin, perhaps located in neurosecretory cells of the pars intercerebralis which have been shown to be

involved in photoperiodically-induced diapause [43], could still serve a photoperiodic function. Monarch parapinopsin could be present in just a few photoperiodic cells, and therefore cannot be excluded because of its low expression level. Knocking either of these candidates out *in vivo* with the CRISPR/Cas9 system would allow for their functional characterization in photoperiodically-induced reproductive diapause. The light-sensitive dpCRY1 is yet another alternative candidate photoperiodic sensor to opsins and its possible involvement should be tested as well.

This work lays out the foundation for a multitude of pharmacological and genetic experiments to continue characterizing the role of the vitamin A pathway in photoperiodic responses, including the identification of the photoperiodic sensor. While the individual components of the vitamin A pathway vary between species, the overall conversion of carotenoids into retinoid derivatives remains the same and as such the work here may be relevant for many seasonally responsive species.

3.5. Materials and Methods.

3.5.1. Animal Husbandry.

For the photoperiodic, dietary supplementation experiments, *dpCry2* mutants and their wild-type siblings were raised in LP and SP photoperiods as described in Chapter 2.5.1 up until the point of the artificial diet switch [19]. The larvae were switched to artificial diet either containing β-carotene, retinal, retinol,

or the vehicle control. Diet was changed every 3 d as we wanted to limit the degradation of β-carotene and the retinoids in the presence of light. At adulthood, monarchs were fed for 10 days with 25% honey supplemented with the β-carotene, the control, or the respective retinoid for that group before females were dissected and the number of mature oocytes quantified.

The progeny of the *santa maria 1* heterozygous x heterozygous cross were raised in 15-hours light: 9-hours dark at 25°C until pupation. The *santa maria 1* line is now maintained at the heterozygous state for potential future rescue experiments.

Rdh13 photoperiodic experiments were done as described in Chapter 2.5.1. [19].

3.5.2. Dietary Complementation with β-carotene, Retinal and Retinol.

β-carotene (Sigma 22040-1G-F), *all-trans*-retinal (Sigma R2500-500MG), and *all-trans*-retinol (Sigma R7632-500MG) were diluted in 100% ethanol to a stock concentration of 10mM and stored at -80^{o}C in a foil covered tube to prevent degradation by light. Based on dietary complementation with retinal in previous studies [18, 44], we initially tested the feeding of 0.4mM β-carotene and *all-trans*-retinal, but a significant amount of death occurred among the larvae and adults so they were not able to be tested. The concentration was then reduced to 0.02mM. Supplemented artificial diet was made in the dark with β-carotene or *all-trans*-retinal for a final concentration of 0.02mM. Spare diet was

stored at 4 ^{o}C in a black trash bag to protect from light degradation until use for the next passage of larvae. Adults were fed 150 µL every day of 0.02mM β-carotene or *all-trans*-retinal containing 25% honey until dissection.

In a second set of experiments, we tried increasing the concentration of the carotenoid, and also included dietary supplementation with *all-trans*-retinol as well as an ethanol vehicle control. Supplemented artificial diet was made in the dark with β-carotene or *all-trans*-retinol to a final concentration of 0.08mM, or an equal volume of 100% ethanol. Diet was stored as mentioned before. Adults were fed 150 µL of 0.08mM β-carotene or *all-trans*-retinol containing 25% honey or 25% honey containing a comparable volume of ethanol.

3.5.3. gRNA Design and Construction.

gRNA sites for CRISPR/Cas9-mediated targeted mutagenesis were identified using CHOPCHOP target site finder as in Chapter 2.5.7. [19]. The gRNA site for mutagenesis of *santa maria 1* was selected within exon 2 of the 8 exons-containing *santa maria 1*, and the one for mutagenesis of *rdh13* was selected closer to the 5' end of the single exon containing *rdh13*. The gRNA expression vector was constructed as referenced in Chapter 2.5.7. using the following oligomer sequences (F, forward primer; R, reverse primer): *santa maria1F*, 5'- TAGGGAGTTCGGTAAGACACGGAG-3' and *santa maria1R*, 5'-AAACCTCCGTGTCTTACCGAACTC-3' and *rdh13F*, 5'-

TAGGCAGCGAAGAACTTAGCCGAG-3’ and *rdh13R*, 5’-AAACCTCGGCTAAGTTCTTCGCTG -3’ [19].

3.5.4. Synthesis of Cas9 mRNA and sgRNA.

In vitro transcription of *Streptococcus pyogenes* Cas9 mRNA was performed using the mMessage mMachine T3 transcription kit (Invitrogen) and pCS2-nCas9n expression plasmid from Addgene [45], as previously described [46]. Purification of the capped PolyA mRNAs, sgRNAs transcription and purification, and quantifications were performed as described in Chapter 2.5.8. [19]. The resulting RNAs were diluted in RNase-free water to a final concentration of 0.25 µg/µl for the *santa maria 1* sgRNA and the *rdh13* sgRNA, and 0.5 µg/µl for Cas9 mRNAs.

3.5.5. Egg Microinjections.

Eggs were collected and microinjected to either target *santa maria 1* or *rdh13* as described in Chapter 2.5.9. [19].

3.5.6. Analysis of CRISPR/Cas9-induced Mutations and Generation of *Santa maria 1* or a *Rdh13* Monarch Loss-of-Function Lines.

PCR fragments flanking either the *santa maria 1* or the *rdh13* targeted region were amplified from genomic DNA extracted from larval sensors of potential founders at the fifth instar with the following primers: *santa maria1F*, 5’-

ACATGAAAACATTTGATGGGAC-3’ and *santa maria1R,* 5’-CACGCCACAGTTATAAACATTC-3’*;* and *rdh13F*, 5’-ACAATCGCTTCAGCTTACGACA-3’ and *rdh13R,* 5’-GTTTGTGCGGCTTCCCAAAC-3’. Purification of the resulting PCR products and subsequent screening for somatic targeting by Cas9-based cleavage assays were done as described in Chapter 2.5.10. [19]. An 80-base pair deletion mutant was selected to establish the *santa maria 1* mutant line and a 1-base pair deletion mutant was selected to establish the *rdh13* mutant line, as each caused a frameshift and the introduction of a premature stop codon. Each of these mutants were backcrossed to wild-type monarchs to found their respective mutant line.

3.5.7. Calculation of *Santa maria 1* Death Rates.

As the *santa maria 1* mutants and their siblings started to reach pupation, I noticed that the majority of the mutants were failing pupation. From this point forward, I counted all larvae from mutants and the wild-type and heterozygous (het) siblings that had made it to fifth instar, the stage preceding pupation. I counted from full knockouts 50 in LP and 51 in SP, in wild-types 63 in LP and 48 in SP, and in the het group 39 in LP and 43 in SP. I then recorded how many larvae died from each genotype, and whether the death occurred pre-pupation or in the process of pupation. Pre-pupation was considered to be any death before the larvae began clearing out all waste from their body which is easily

observable by a change in frass color from brown to bright red. Any failure after the excretion of the red colored frass and during the process of pupation itself was considered death at pupation. Pre-pupation death in LP conditions were as follows: 14 mutant, 2 wild-type, 1 het. Pre-pupation death in SP conditions: 15 mutants. Death in the pupation process in LP: 33 mutant, 6 wild-type, 8 het. Death in the pupation process in SP: 35 mutant, 7 wild-type, 2 het. Deaths were then calculated as percentages.

3.5.8. Evaluation of Female Reproductive Status.

Adult females were dissected and oocyte number quantified as described in Chapter 2.5.3. [19].

3.5.9. Real Time qPCR for Tissue-Specific Expression of Opsin Candidates.

To test for the expression of the five monarch opsins in different tissues, the first pair of legs, abdomen, thorax, and antennae were clipped from 3 males and 3 female monarchs, while the photoreceptor layer of the compound eyes and the brain, itself separated into the optic lobes and central brain, were dissected in 0.5X RNA later (Invitrogen) to avoid RNA degradation during dissection. All samples were stored at -80°C until use. Total RNA from the optic lobes and central brains were extracted using a RNeasy Mini kit (Qiagen) as described in Chapter 2.5.12. [19]. For the thorax, abdomen, legs, eyes, and antennae, the tissue was pulverized with a pestle in liquid nitrogen and RNA

extracted using 350uL of RNA extraction buffer (100 mM Tris pH 7.5, 100 mM LiCl, 20 mM DTT, and 10% SDS) followed by purification with acid-phenol-chloroform as described before [47]. Reverse transcription into cDNA and the following quantifications for gene expression were performed as in Chapter 2.5.12 [19]. The monarch *blue opsin* (*bl)*, *long wavelength opsin* (*lw*), *parapinopsin* (*pp*), *peropsin* (*pe*), *UV opsin* (*uv*) and control *rp49* primers were as follows (F, forward primer; R, reverse primer): *blF*, 5'-CGGTTACCAACTCGGATGTGA-3'; *blR*, 5'- CCGCCGATTCCAGACAGT-3'; *lwF*, 5'-GGGCCTAGCTGCAAAACCA-3'; *lwR*, 5'-ATACCAAGCACTCGCAGCAA-3'; *ppF*, 5'-CGAGTCGTCCGACAGGAAA-3'; *ppR*, 5'-AGCATCCGATCCGGTTTTG-3'; *peF*, 5'-ACCGAAAGCGTAGTTATCATGACA-3'; *peR*, 5'-GGGCGATTGGCAACCA-3'; *uvF*, 5'- CTGCATCGATCCTTGGGTTT-3'; *uvR*, 5'-CGCTGAAGCTCTTGCCTGTA-3'; and *rp49F*, 5'-TGCGCAGGCGTTTTAAGG-3'; *rp49R*, 5'-TTGTTTGATCCGTAACCAATGC-3'. Data normalization and statistics were done as in Chapter 2.5.12 [19].

3.5.10. Statistical Analysis.

P values were calculated using Student's t-tests, 1-way and 2-way ANOVAs followed by post hoc analyses, and two-tailed Z-score tests using online calculators at www.wessa.net/rwasp_Two%20Factor%20ANOVA.wasp [48],

https://www.icalcu.com/stat/anova-tukey-hsd-calculator.html, and https://www.socscistatistics.com/tests/ztest/Default2.aspx.

3.6. References.

1. Claret, J. and N. Volkoff, *Vitamin A is essential for two processes involved in the photoperiodic reaction in Pieris brassicae.* Journal of Insect Physiology, 1992. **38**: p. 569-574.
2. Veerman, A., et al., *Vitamin-a Is Essential for Photoperiodic Induction of Diapause in an Eyeless Mite.* Nature, 1983. **302**(5905): p. 248-249.
3. Veerman, A., et al., *Photoperiodic Induction of Diapause in an Insect Is Vitamin-a Dependent.* Experientia, 1985. **41**(9): p. 1194-1195.
4. Goodwin, T.W., *Metabolism, nutrition, and function of carotenoids.* Annu Rev Nutr, 1986. **6**: p. 273-97.
5. Palczewski, K., et al., *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science, 2000. **289**(5480): p. 739-45.
6. Vogt, K., *Is the fly visual pigment a rhodopsin* Zeitschrift für Naturforschung C, 1983. **38**(3-4): p. 196-197.
7. Wald, G., *The molecular basis of visual excitation.* Nature, 1968. **219**(5156): p. 800-7.
8. Lane, M.A. and S.J. Bailey, *Role of retinoid signalling in the adult brain.* Prog Neurobiol, 2005. **75**(4): p. 275-93.

9. Travis, G.H., et al., *Diseases caused by defects in the visual cycle: retinoids as potential therapeutic agents.* Annu Rev Pharmacol Toxicol, 2007. **47**: p. 469-512.

10. Moore, T., *Vitamin A*. 1957, London: Elsevier. 645.

11. Lees, A.D., *The Location of the Photoperiodic Receptors in the Aphid Megoura Viciae Buckton.* J Exp Biol, 1964. **41**: p. 119-33.

12. Seuge, J. and K. Veith, *Diapause de Pieris brassicae: Role des photorecepteurs cephaliques, etude des carotenoides cerebraux.* Journal of Insect Physiology, 1976. **22**(9): p. 1229-1235.

13. Williams, C.M. and P.L. Adkisson, *Photoperiodic Control of Pupal Diapause in the Silkworm, Antheraea pernyi.* Science, 1964. **144**(3618): p. 569.

14. Steel, C.G. and A.D. Lees, *The role of neurosecretion in the photoperiodic control of polymorphism in the aphid Megoura viciae.* J Exp Biol, 1977. **67**: p. 117-35.

15. von Lintig, J., et al., *Analysis of the blind Drosophila mutant ninaB identifies the gene encoding the key enzyme for vitamin A formation invivo.* Proc Natl Acad Sci U S A, 2001. **98**(3): p. 1130-5.

16. Kiefer, C., et al., *A class B scavenger receptor mediates the cellular uptake of carotenoids in Drosophila.* Proc Natl Acad Sci U S A, 2002. **99**(16): p. 10581-6.

17. von Lintig, J., et al., *Towards a better understanding of carotenoid metabolism in animals.* Biochim Biophys Acta, 2005. **1740**(2): p. 122-31.

18. Wang, T., Y.C. Jiao, and C. Montell, *Dissection of the pathway required for generation of vitamin A and for Drosophila phototransduction.* Journal of Cell Biology, 2007. **177**(2): p. 305-316.

19. Iiams, S.E., et al., *Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly.* Proc Natl Acad Sci U S A, 2019.

20. Helfer G., et al., *Photoperiod regulates vitamin A and Wnt/b-catenin signalling in F344 rats.* Endocrinology, 2012. **153**: p. 815-824.

21. Shearer K. D., et al., *Photoperiodic regulation of retinoic acid signalling in the hypothalamus.* Journal of Neurochemistry, 2010. **112**: p. 246-257.

22. Shearer K. D., et al., *Photoperiodic expression of two RALDH enzymes and the regulation of cell proliferation by retinoic acid in the rat hypothalamus* Journal of Neurochemistry, 2012. **122**: p. 789-799.

23. Bui-Gobbels, K., et al., *Is retinoic acid a signal for nerve regeneration in insects?* Neural Regeneration Research, 2015. **10**: p. 901-903.

24. Wood, S.H., et al., *Binary Switching of Calendar Cells in the Pituitary Defines the Phase of the Circannual Cycle in Mammals.* Curr Biol, 2015. **25**(20): p. 2651-62.

25. Bolborea, M., et al., *Melatonin controls photoperiodic changes in tanycyte vimentin and neural cell adhesion molecule expression in the Djungarian*

hamster (Phodopus sungorus). Endocrinology, 2011. **152**(10): p. 3871-83.

26. Terakita, A., *The opsins.* Genome Biol, 2005. **6**(3): p. 213.
27. Okano, T., T. Yoshizawa, and Y. Fukada, *Pinopsin is a chicken pineal photoreceptive molecule.* Nature, 1994. **372**: p. 94-97.
28. Max, M., et al., *Pineal opsin: a nonvisual opsin expressed in chick pineal.* Science, 1995. **267**(5203): p. 1502-6.
29. Nakane, Y., et al., *A mammalian neural tissue opsin (Opsin 5) is a deep brain photoreceptor in birds.* Proc Natl Acad Sci U S A, 2010. **107**(34): p. 15264-8.
30. Saunders, D.S., *Insect Clocks*. 2002: Elsevier. 576.
31. Smrcka, A.V., et al., *Regulation of polyphosphoinositide-specific phospholipase C activity by purified Gq.* Science, 1991. **251**(4995): p. 804-7.
32. Taylor, S.J., et al., *Activation of the beta 1 isozyme of phospholipase C by alpha subunits of the Gq class of G proteins.* Nature, 1991. **350**(6318): p. 516-8.
33. Waldo, G.L., et al., *Purification of an AlF4- and G-protein beta gamma-subunit-regulated phospholipase C-activating protein.* J Biol Chem, 1991. **266**(22): p. 14217-25.
34. Davies, W.I., et al., *Vertebrate ancient opsin photopigment spectra and the avian photoperiodic response.* Biol Lett, 2012. **8**(2): p. 291-4.

35. Zhan, S. and S.M. Reppert, *MonarchBase: the monarch butterfly genome database.* Nucleic Acids Res, 2013. **41**(Database issue): p. D758-63.

36. Goodwin, T.W., *Biochemistry of the Carotenoids*. 2nd ed. Volume II Animals. Vol. 2. 1984, London: Chapman & Hall. 224.

37. Kayser, H., *Carotenoids in Insects*, in *Carotenoid Chemistry and Biochemistry*, G. Britton and T.W. Goodwin, Editors. 1982, Pergamon: Liverpool. p. 195-210.

38. Edge, R., D.J. McGarvey, and T.G. Truscott, *The carotenoids as anti-oxidants--a review.* J Photochem Photobiol B, 1997. **41**(3): p. 189-200.

39. Lampel, J., A.D. Briscoe, and L.T. Wasserthal, *Expression of UV-, blue-, long-wavelength-sensitive opsins and melatonin in extraretinal photoreceptors of the optic lobes of hawk moths.* Cell Tissue Res, 2005. **321**(3): p. 443-58.

40. Blackshaw, S. and S.H. Snyder, *Parapinopsin, a novel catfish opsin localized to the parapineal organ, defines a new gene family.* J Neurosci, 1997. **17**(21): p. 8083-92.

41. Koyanagi, M., et al., *Bistable UV pigment in the lamprey pineal.* Proc Natl Acad Sci U S A, 2004. **101**(17): p. 6687-91.

42. Wada, S., et al., *Expression of UV-sensitive parapinopsin in the iguana parietal eyes and its implication in UV-sensitivity in vertebrate pineal-related organs.* PLoS One, 2012. **7**(6): p. e39003.

43. Shiga, S., N.T. Davis, and J.G. Hildebrand, *Role of neurosecretory cells in the photoperiodic induction of pupal diapause of the tobacco hornworm Manduca sexta.* J Comp Neurol, 2003. **462**(3): p. 275-85.

44. Malloy, C.A., *PROFILING THE ACTION OF ACETYLCHOLINE IN THE DROSOPHILA MELANOGASTER LARVAL MODEL: HEART, BEHAVIOR, AND THE DEVELOPMENT AND MAINTENANCE OF SENSORIMOTOR CIRCUITS.* Theses and Dissertations--Biology, 2017. **47**.

45. Jao, L.E., S.R. Wente, and W.B. Chen, *Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(34): p. 13904-13909.

46. Zhang, Y., et al., *Vertebrate-like CRYPTOCHROME 2 from monarch regulates circadian transcription via independent repression of CLOCK and BMAL1 activity.* Proc Natl Acad Sci U S A, 2017. **114**(36): p. E7516-E7525.

47. Wan, G., et al., *Cryptochrome 1 mediates light-dependent inclination magnetosensing in monarch butterflies.* Nature Communications, *In Press*.

48. Wessa, P. *Free Statistics Software*. 2018; version 1.2.1:[Available from: https://www.wessa.net/.

4. CRYPTOCHROME 1 IS THE SOLE PHOTORECEPTOR NECESSARY FOR ENTRAINMENT OF RHYTHMIC ECLOSION IN THE MONARCH BUTTERFLY[2]

4.1. Overview.

Light is one of the key Zeitgebers, external or environmental cues, in the entrainment of the circadian clocks as most organisms are subject to the light:dark cycles resulting from the daily Earth's rotation on its axis. Decades of work in the traditional model system *Drosophila* (*dm*) contributed to the discovery of the molecular mechanisms underlying light entrainment of circadian clocks, which rely on multiple rhodopsins, phospholipase-based signal transduction pathways, and the circadian blue-light photoreceptor Cryptochrome (dmCRY). Similar to *dmCry*, the monarch (*dp*) ortholog, *dpCry1*, has also been shown to be photoreceptive and mediate TIMELESS degradation *in vitro*, but its role in monarch clock entrainment *in vivo* has not been tested yet. Here, behavioral and molecular characterization of a CRISPR/Cas9 generated loss-of-function mutant of *dpCry1* has revealed an unexpected complete loss of rhythmicity in eclosion behavior in constant darkness after entrainment to 12-hr light: 12-hr dark cycles, as compared to rhythmic wild-type sibling controls. I further showed that the arrhythmic phenotype of *dpCry1* loss-of-function mutants

[2] Part of the data reported in this chapter is reprinted from "Cryptochrome 1 mediates light-dependent inclination magnetosensing in monarch butterflies" by Wan G, Hayden AN, Iiams SE, and Merlin C, 2021. *Nat Commun*, 12 (1), 771.

was due to a deficit in light-entrainment and not to a dysfunctional circadian clock, as entraining the mutants to 15°C:25°C temperature cycles in constant darkness during the pupal stage rescued rhythmicity of eclosion behavior and the molecular rhythms of core clock genes. Together, these results demonstrated that, in contrast to *Drosophila*, dpCRY1 is likely the sole molecule involved in light-entrainment of the monarch circadian clock. Importantly, we also showed that *dpCry1* loss-of-function mutants lack photoperiodic responses, further supporting the idea that a functional circadian clock is required for photoperiodic responses in monarchs or suggesting that dpCRY1 plays a role in photoperiodic sensing.

4.2. Introduction.

CRYPTOCHROMES (CRYs) belong to a class of blue-light and UV sensitive flavoproteins widely distributed across species in both Eukarya and Bacteria [1]. CRYs have evolved from ancestral DNA repair photolyases, and have retained an amino-terminal photolyase-related (PHR) domain as well as the chromophore access cavity for flavin adenine dinucleotide (FAD) binding [1]. However, CRYs have lost the DNA repair capability of a photolyase and instead function in vision, photoperiodism, phototropism, phototaxis, and notably in circadian clock entrainment and function in plants and animals [1-3]. While some of these CRYs have diverged to function in a light-independent manner as transcriptional regulators, such as those found in mammals and the monarch

(dp) dpCRY2 repressor, others like CRYs of *Arabidopsis* and insect species (dpCRY1 in monarch and dmCRY in *Drosophila* (*dm*)) retain photoreceptive function [4-8].

Drosophila dmCRY functions as a light-sensitive photoreceptor that interacts with core components of the circadian clock by mediating the degradation of TIMELESS (dmTIM) in a light-induced manner, and is also capable of dimerizing with PERIOD (dmPER) [9, 10]**.** The null mutant *crybaby* (cry^b), generated by a point mutation in the region believed to be involved in FAD binding, was key in revealing dmCRY's role in circadian resetting, as it not only disrupted dmPER and dmTIM rhythms in whole heads but also rendered the mutants irresponsive to phase-shifting light pulses as previously shown for *Arabidopsis* CRY mutants light-pulsed with blue light [4, 5, 8]. Curiously, cry^b mutants and a fly cry^0 full knockout retained behavioral rhythms in pupal eclosion and adult locomotor activity and were even entrainable to new light:dark (LD) regimes [5, 11]. This paradox was later explained by the discovery that dmPER and dmTIM are still cycling in a subset of clock neurons within the brain in these mutants [5], suggesting that other dmCRY-independent photoreceptors and pathways must be involved in *Drosophila* clock entrainment. Complete abolishment of molecular and behavioral rhythms was later obtained in a double mutant for cry^b and *glass*, a transcription factor necessary for the development of photoreceptors in the eyes and Hofbauer-Buchner eyelets [12]. In addition, phospholipase, no receptor potential A (norpA), a novel signal transduction

pathway utilizing phospholipase C at 21C (Plc21C) and several rhodopsin photoreceptors, rhodopsins (Rh) Rh1, Rh5, Rh6, Rh7 acting either in the eyes or the brain, have been found to play a role in *Drosophila* clock entrainment [13-16]. Together, these results suggested that *Drosophila* clocks utilize a complex network of photoreceptors and signal transduction pathways to sense and synchronize to their environment.

To date, the role of light-sensitive CRY in clock entrainment in other insects has been restricted to *in vitro* studies in the monarch. Using a monarch embryo-derived DpN1 cell line that expressed all core clock components and exhibits light-driven rhythms, dpCRY1 has been shown to function, as in *Drosophila*, as a blue-light photoreceptor necessary for entrainment to light and mediation of dpTIM degradation [6, 7, 9]. Yet, the extent of its contribution to eclosion and activity rhythms remains unknown as its role in circadian entrainment has never been tested *in vivo*. A monarch knockout of *dpCry1*, harboring a 2-bp deletion in the 4th exon that led a 90% reduction in *dpCry1* mRNA likely caused by nonsense-mediated RNA decay and thus likely no protein production, was generated [17] (Fig. 4.1). Even if the protein is produced, the 2-bp deletion introduces an early stop codon that would remove the C-terminal region containing a group of tryptophan (trp) residues, known as the trp tetrad, that are necessary for CRY photosensing by reducing the FAD chromophore into its active state [18-22]. This FAD reduction is associated with the conformational change in the CRY protein moiety and is proposed to induce

its activity to initiate signaling downstream signaling pathways [23-26]. Using this knockout, I aimed to 1) test and characterize the role of dpCRY1 in circadian entrainment to expand our understanding of insect clock evolution, and 2) test whether the clock functions as a unit to regulate seasonal responses by testing the effect of this mutant on photoperiodic responses.

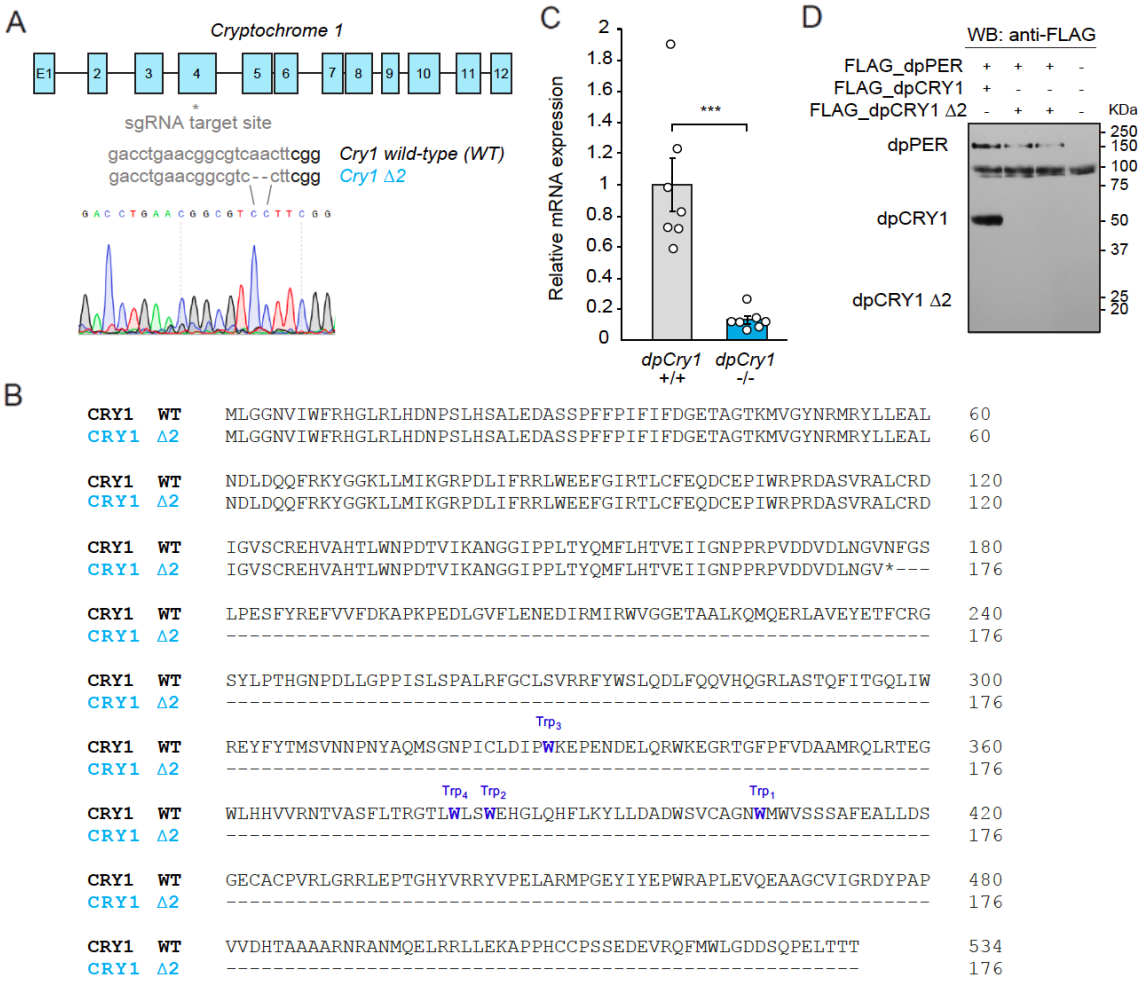


Figure 4.1 (Supplement): *Cryptochrome 1* loss-of-function monarch mutant generated by CRISPR/Cas9-mediated targeted mutagenesis.

(*A*) Schematic representation of monarch *Cryptochrome 1* (*dpCry1*) genomic locus. The gray star indicates the position of the single guide RNA (sgRNA) used to induce indel mutations. The sgRNA sequence at the target site is shown in gray and is followed by a protospacer adjacent motif (PAM; shown in black). A 2-bp deletion (Δ2) in G_1 progeny was used to establish a *dpCry1* loss-of-function mutant line. (*B*) Amino acid sequence alignment of wild-type (WT) and truncated loss-of-function (Δ2) dpCRY1 proteins. The four tryptophan (Trp) residues involved in the Trp tetrad are highlighted in blue. (*C*) Relative *dpCry1* mRNA expression in brains of wild-type and homozygous *dpCry1Δ2* mutant monarchs (n = 7 for each genotype). Student's t-test: ***, $P < 0.0005$. (*D*) Western blot of monarch DpN1 cells transfected with FLAG-tagged dpPER and either FLAG-tagged full length dpCRY1 or FLAG-tagged dpCRY1Δ2 mutant, probed with an anti-FLAG antibody. Reprinted from [17].

4.3. Results.

4.3.1. Loss of dpCRY1 Abolishes Rhythmic Eclosion Behavior.

Given that *cry^{b}* flies still exhibited rhythmic pupal eclosion and adult locomotor activity and were also entrainable to new LD cycles, I wanted to first test the effect of the *dpCry1* loss-of-function mutation on behavioral rhythms. One easily tractable behavior in monarchs is the emergence of the fully mature adult from its chrysalis, also known as eclosion. This behavior is timed to the early morning phase of LD cycles by the circadian clock, and can be easily recorded with infrared cameras [5, 27]. Monarchs were raised from eggs to day 8 of the pupal stage in 12 h:12 hr LD at 25^{o}C and then released into constant dark (DD) conditions two days prior to their expected eclosion. Interestingly, I found that, in contrast to their wild-type siblings that linked their eclosion around the expected timing of lights on (CT0), *dpCry1* knockouts emerged completely arrhythmically with eclosions happening nearly over the entire 24 hr circadian day (Fig. 4.2 *A*). Molecular rhythms of *dpPer* and *dpTim* from monarch brains were also abolished in the mutant (Fig. 4.2 *B*).

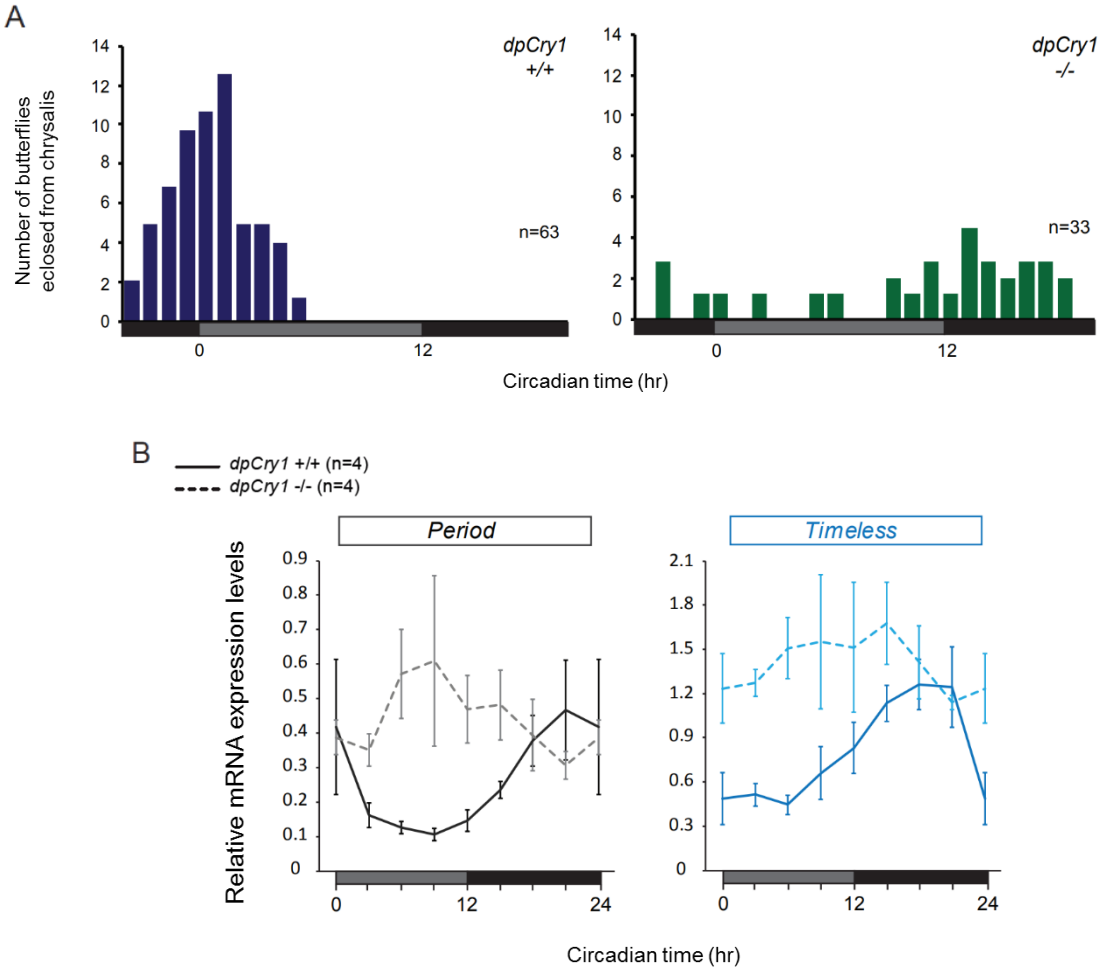


Figure 4.2: *DpCry1* knockouts exhibit arrhythmic eclosion behavior.
(*A*) Profiles of adult eclosion pooled together from eclosion on the first and second day in constant darkness (DD) of wild-type (dark blue) and *dpCry1* homozygous mutant siblings (green) after entrainment in 12 h:12 hr LD at 25ºC. Gray bars: subjective day; black bars: subjective night. Kolmogorov–Smirnov (KS) test maximum difference between data sets: *P*= 0.00. KS normal distribution test wild-type: *P*=0.23. KS normal distribution test *dpCry1* mutant: *P*=0.04. One-way ANOVA between data sets: *P*=0.00 (*B*) Circadian expression of *period* (*dpPer*) and *timeless* (*dpTim*) in the brain of adult monarchs eclosed from A during the first day of release in DD after at least 7 days of re-entrainment to 12 h:12 hr LD cycles at 25ºC. Wild-types: solid lines. *DpCry1* homozygous mutants: dashed lines. Values are mean ± SEM of four brains for each timepoint. *dpPer* in wild-type, 1-way ANOVA: *P*= 0.048. *dpPer* in *dpCry1* mutant, 1-way ANOVA: *P*=0.613. *dpTim* in wild-type, 1-way ANOVA: *P*=0.003, |CT6-CT18|=*P*=0.035; |CT6-CT21|=*P*=0.041. *dpTim* in *dpCry1* mutant, 1-way ANOVA: *P*=0.886. Interaction genotype × time, 2-way ANOVAs: *dpPer*: *P*=0.036, *dpTim*: *P*=0.235.

4.3.2. Monarch dpCRY1-less Mutants Can be Re-Entrained to Temperature Cycles.

To determine whether arrhythmic eclosion behavior in *dpCry1* homozygous mutants was the result of an un-entrained clock or an oscillator-less mutant, I sought to bypass the light Zeitgeber and instead re-entrain the mutants to temperature cycles (TC), as TC has been shown to also act as a potent Zeitgeber in flies [5]. If the behavioral and molecular rhythms can be re-entrained by TC, this would strongly suggest that the arrhythmic phenotypes observed in *dpCry1* homozygous mutants result from a deficiency in entrainment to light rather than to a dysfunctional core clock mechanism. Monarchs were raised in 12 h:12 hr LD at 25^{o}C from eggs to the end of the 4th larval instar stage. For the duration of the 5th instar, the larvae were exposed to constant light (LL) conditions to disrupt core clock oscillations and ensure that no residual rhythms from entrainment to light could persist and impact the entrainment to TC that followed [27]. Pupae were moved to a 12 h-20°C / 12 h-25°C TC, in anti-phase to the LD cycle (with low temperature occurring in previous light phase and high temperatures in previous dark phase), in DD conditions and adult eclosion behavior was observed under the same conditions (Fig. 4.3 *A*). My expectation was that the monarchs would link their emergence close to the switch from 20°C to 25°C, as the rise in temperature should signal the onset of the subjective day, but curiously, eclosions in both the wild-type and *dpCry1* homozygous mutant occurred only during the 20°C phase and were spread

throughout the 12-h period (Fig. 4.3 *B*). These puzzling results may be explained by the possibility that the amplitude of the TC wasn't strong enough for proper entrainment.

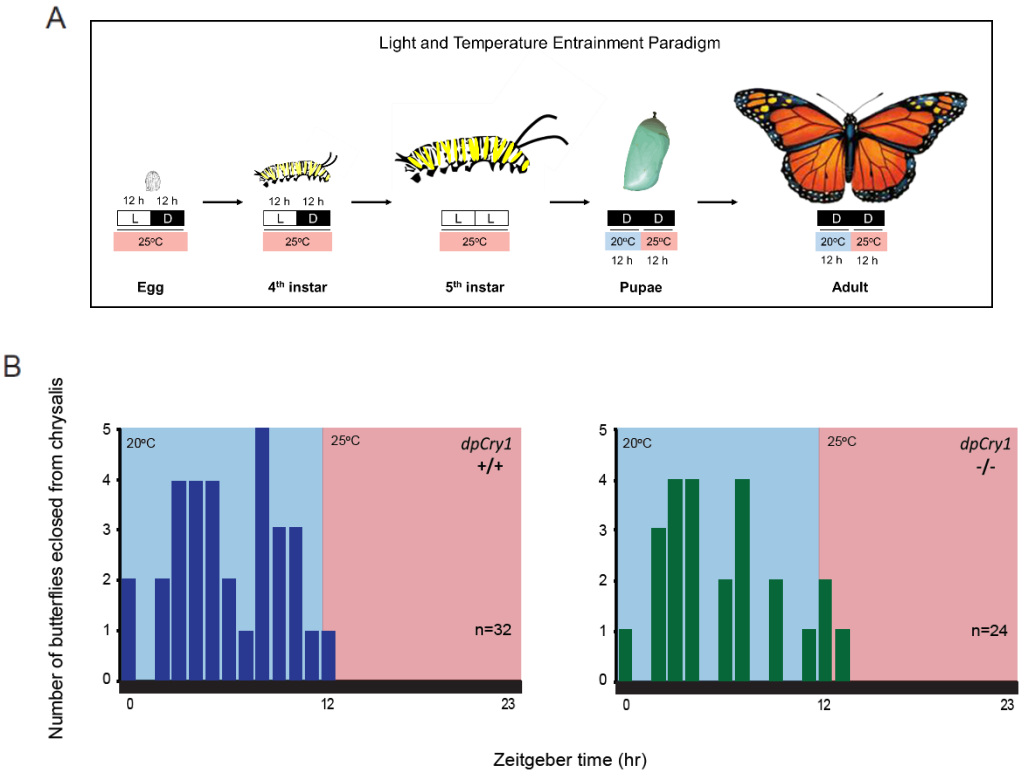


Figure 4.3: A 5°C amplitude temperature cycle (TC) is not strong enough to recapitulate monarch rhythmic eclosion behavior.

(*A*) Light and temperature entrainment paradigm. Monarchs were raised from egg to 4th instar in 12 h:12 hr LD conditions at 25ºC. At fifth instar, larvae were released into LL at 25ºC. At pupation, pupae were shifted to DD in TC of 12 h-20°C/12 h-25°C and maintained for another 7 d as adults in 12 h-20°C/12 h-25°C DD. (*B*) Profiles of adult eclosion after pupal entrainment to 12 h-20°C/ 12 h-25°C DD cycles. Black bars: constant darkness. Blue background: 20ºC. Red background: 25ºC. Wild-type (dark blue) and *dpCry1* homozygous mutant siblings (green). KS test maximum difference between data sets: $P = 0.862$. KS normal distribution test wild-type: $P = 0.58$. KS normal distribution test *dpCry1* mutant: $P = 0.19$. One-way ANOVA between data sets: $P = 0.883$.

To test if a TC of greater amplitude could properly entrain eclosion behavior, I then raised the monarchs using a similar paradigm, but with a 10°C amplitude of change cycling at 12 h-15°C / 12 h-25°C in DD (Fig. 4.4 *A*). In these conditions, both wild-type and *dpCry1* homozygous mutants re-entrained their behavior to the TC and eclosed as expected during the subjective morning after the temperature rose from 15°C to 25°C at CT12 (Fig. 4.4 *B, Left*). To rule out a possible masking effect by temperature, I also raised mutants of a previously generated core clock gene activator, *clock* (*dpClk*) [28], along with their wild-type siblings under the same paradigm. Unlike their wild-type siblings, *dpClk* mutants were unable to entrain to the 10°C amplitude TC, thus demonstrating that the synchronized eclosion behavior of *dpCry1* loss-of-function mutants is not due to an acute response to the increase in temperature, but is an entrained response timed by a functional core clock loop (Fig. 4.4 *B, Right*). To test if re-entrainment of behavioral rhythms by TC would be mirrored at the molecular level, I quantified using RT-qPCR the expression of *dpPer* and *dpTim* in the brain of both adult wild-type and *dpCry1* homozygous mutant subjected to the same TC conditions in DD after eclosion. I found that molecular rhythms of *dpPer* and *dpTim* were also entrained to the anti-phasic TC in both genotypes, where expression peaked at the high to low temperature transition (CT0-CT4) and was at its trough in the early high temperature “daytime” phase (CT16), matching that of *dpPer* and *dpTim* in brain of wild-type monarchs entrained to LD cycles [6] (Fig. 4.4 *C*). Together, these data demonstrate that

dpCRY1 not only functions in the light-entrainment pathway to the circadian clock but, unlike in *Drosophila*, appears to be the sole photoreceptor involved, The results here also provide a new clock gene knockout which can be tested for the effect of abolished clock entrainment on photoperiodic responses, in comparison to the previous findings from knocking out circadian activation or repression [29].

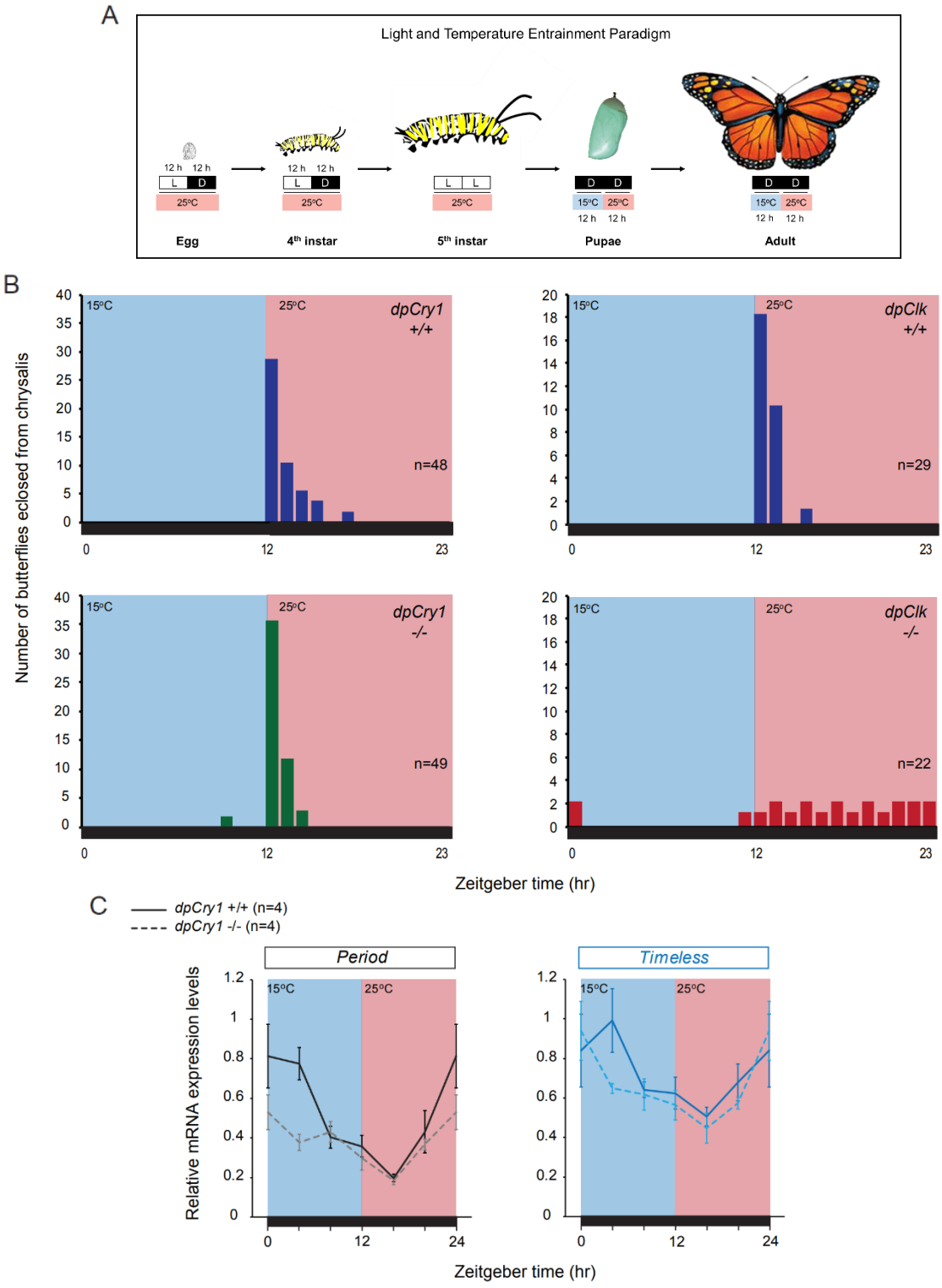


Figure 4.4: 10°C temperature cycles re-entrain eclosion behavior and brain molecular rhythms in wild-type and *dpCry1* knockouts.

Figure 4.4 Continued

(*A*) Light and temperature entrainment paradigm. Monarchs were raised from egg to 4th instar in 12 h:12 hr LD conditions at 25^{o}C. At fifth instar, larvae were released into LL at 25^{o}C. At pupation, pupae were shifted to DD in TC of 12 h-15°C/12 h-25°C and maintained for another 7 d as adults in 12 h-15°C/12 h-25°C DD. (*B*) Profiles of adult eclosion after pupal entrainment to 12 h-15°C/ 12 h-25°C DD cycles. Black bars: constant darkness. Blue background: 15^{o}C. Red background: 25^{o}C. (Left) Wild-type (dark blue) and *dpCry1* homozygous mutant siblings (green). KS test maximum difference between data sets: *P*= 0.639. KS normal distribution test wild-type: *P*=0.00. KS normal distribution test *dpCry1* mutant: *P*=0.00. One-way ANOVA between data sets: *P*=0.016. (Right) Profiles of wild-type (blue) and *dpClk* homozygous mutant siblings (red). KS test maximum difference between data sets: *P*= 0.00. KS normal distribution test wild-type: *P*=0.00. KS normal distribution test *dpCry1* mutant: *P*=0.24. One-way ANOVA between data sets: *P*=0.005. (*C*) Expression of *period* (*dpPer*) and *timeless* (*dpTim*) in the brain of adult wild-type and *dpCry1* mutant monarchs eclosed from C after at least 7 days of adult entrainment to 12 h-15°C/ 12 h-25°C in DD. Black bars: constant darkness. Blue background: 15^{o}C. Red background: 25^{o}C. Wild-types: solid lines. *DpCry1* mutants: dashed lines. Values are mean ± SEM of four brains for each timepoint. *dpPer* in wild-type, 1-way ANOVA: *P*= 0.0007, |ZT0-ZT12|=*P*=0.025; |ZT0-ZT16|=*P*=0.002; |ZT4-ZT12|=*P*=0.046; |ZT4-ZT16|=*P*=0.003. *dpPer* in *dpCry1*, 1-way ANOVA: *P*=0.006, |ZT0-ZT16|=*P*=0.003; |ZT8-ZT16|=*P*=0.043. *dpTim* in wild-type, 1-way ANOVA: *P*=0.09. *dpTim* in *dpCry1*, 1-way ANOVA: *P*=0.012, |ZT0-ZT12|=*P*=0.047; |ZT0-ZT16|=*P*=0.006. Interaction genotype × time, 2-way ANOVAs: *dpPer*: *P*=0.045, *dpTim*: *P*=0.394.

4.3.3. DpCRY1 is Necessary for Photoperiodically-Induced Reproductive Diapause Responses.

In Chapter 2, I demonstrated that disruption of the circadian clock activators or the repressor inhibits photoperiodically-induced reproductive diapause responses, locking females into either a reproductive or non-reproductive state. While these results suggested that the clock may function as a unit to regulate photoperiodic responses [29], the possibility that individual clock genes could have pleiotropic effects could not be ruled out. Testing the effect of *dpCry1* deficiency on photoperiodic responses could provide further evidence to disentangle these two hypotheses. If dpCRY1 is essential for

entraining the monarch circadian clock, and the circadian clock functions in photoperiodic sensing as a unit rather than through the pleiotropic effect of individual clock genes, then I predicted that photoperiodic responses would also be disrupted in *dpCry1* homozygous mutants. Here, we raised *dpCry1* mutants and their wild-type siblings under long-photoperiod (LP) conditions (15 hr light:9 hr dark) or short-photoperiod (SP) conditions (10 hr light:14 hr dark) at a constant temperature of 21°C, and counted the number of mature oocytes as mentioned before [29]. This experiment was performed prior to our discovery that the compound eyes are not necessary for photoperiodic responses, and we thus included two groups for each genotype and lighting condition, one with the compound eye painted black and another with the eyes painted clear (control) [29]. As expected, wild-type monarchs responded to the photoperiod, regardless of paint group, by producing significantly more eggs in LP than SP (Fig. 4.5). However, *dpCry1* mutants were locked in a non-diapausing like phenotype regardless of photoperiod and painting status (Fig. 4.5).

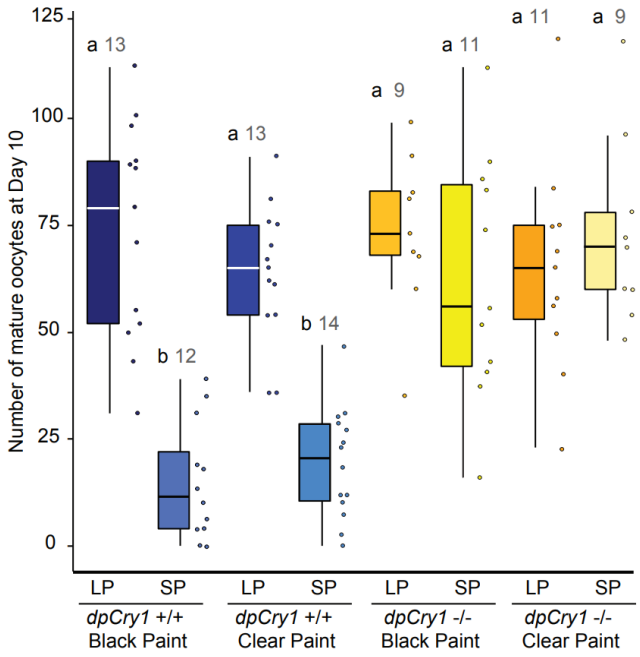


Figure 4.5: *DpCry1* knockouts lose their ability to respond to the photoperiod.

Number of mature oocytes produced 10 days post-eclosion in wild-type or *dpCry1* homozygous mutant female monarchs with eyes painted either with clear or black paint. For each condition, box plots and dot plots are shown. Error bars on box plots represent 1.5 times the interquartile range. LP: 15 hours light, 9 hours dark at 21°C; SP: 10 hours light, 14 hours dark at 21°C. Different letters over bars indicate groups that are statistically significant. Interaction genotype x painting condition, Two-way ANOVA, Tukey's pairwise comparisons, $P < 0.05$.

4.4. Discussion.

Understanding the role of light-sensitive animal CRYs in entrainment of the circadian clock has steadily gained ground over the past few decades. While work in the genetically tractable insect model *Drosophila* has provided evidence for its role as a circadian photoreceptor in clock resetting and the mechanisms behind CRY sensing/signaling, our current knowledge has remained limited to this single species [4, 5, 11-16].

With a conserved dpCRY1 blue-light photoreceptor and suite of genetic tools available, the monarch butterfly provided an opportunity to characterize the ortholog's role of dmCRY in the monarch clock to expand our knowledge of clock entrainment to another insect species. Surprisingly, I found that knocking out the *dpCry1* photoreceptor abolished rhythmic eclosion behavior in the monarch, unlike in *Drosophila,* which remains rhythmic as it utilizes additional dmCRY-independent rhodopsin photoreceptors and signal transduction pathways for circadian entrainment [13, 14, 16]. This finding suggests that dpCRY1 is the sole photoreceptor for entrainment of the monarch clock to light *in vivo*, highlighting striking differences between insect species that could be due to CRY-independent pathways in monarch being either lost in monarchs or gained in *Drosophila*. One could argue that the reason why we see a loss of rhythmic eclosion behavior in *dpCry1* mutants is because dpCRY1-independent pathways are not yet formed and/or have not developed the proper neuronal connections to clock cells in the monarch pupal brain. An argument against this are the results that cry^0 knockout flies not only maintained rhythms in adult activity, but also eclosed in a periodic manner as well, suggesting that CRY-independent pathways are functional by the time of adult eclosion, at least in *Drosophila* [11]. In fact, developmental studies of the fly visual system has found that the optic nerve connects to the circadian clock early in the embryonic stage and that *Rh5*, *Rh6*, and *norpA* are already expressed within larval retina cells [30]. It is likely that such pathways would also already be formed in the larval

stage of monarch butterflies. If this was the case, it would confirm that the loss of rhythmic eclosion behavior in *dpCry1* mutants was due to the fact that dpCRY1 is the sole photoreceptor necessary for entrainment of the clock *in vivo*. Alternatively, assessing flight activity rhythms in adult monarchs would bypass potential developmental issues. One caveat is that not only an assay would have to be developed, but we would only be able to test monarchs in LD conditions, as they do not show flight activity in the dark and testing their activity in DD conditions is not feasible (G. Wan and C. Merlin, personal communication).

The unexpected loss of eclosion behavior in *dpCry1* mutants suggested that either dpCRY1 was essential for the development of a functional circadian clock or indispensable for clock entrainment to LD. To distinguish between these two possibilities, I tested the ability of *dpCry1* mutants to entrain to TCs, as previously done in cry^b mutants [5]. Using a 10^{o}C amplitude TC entrainment paradigm, I was able to show that the circadian clock of *dpCry1* knockouts is intact as eclosion was linked tightly to the increase in temperature, whereas *dpClk* KOs remained arrhythmic, eliminating the possibility of a masking effect. Furthermore, even molecular rhythms of *dpPer* and *dpTim* were restored and exhibited phases of decreasing expression in the cold and increasing in the warm. Together, these results showed not only that the clock could be re-entrained but that it was also anticipating changes in temperature, although through yet unknown mechanisms in the monarch. The ability to entrain circadian clocks to TCs has previously been documented in both insects and

mammals [5, 31-34]. In *Drosophila*, nocte and the Ionotropic Receptor 25a (IR25a) have been shown to mediate clock rhythmicity to TC [35, 36]. Nocte-like (DPOGS213113) and IR25a (DPOGS213837) orthologues exist in the monarch (data not shown). Knocking them out would allow to test if they serve a similar role in temperature entrainment in the monarch. Overall, our results reveal a possible second mode of circadian entrainment in insects, highlighting the importance of expanding the repertoire of model organisms in the field to understand clock evolution.

Finally, we also showed here that *dpCry1* knockouts lose the ability to respond to the photoperiod. Circadian clock genes have previously been shown to be necessary for the regulation of photoperiodic responses where loss of circadian activators or repressors exhibit opposite levels of oocyte maturation (*i.e.* low versus high) [29, 37-40]. Whether this was due to the clock functioning as a unit or to possible pleiotropic effects of individual clock genes had remained a debated topic [29]. Showing that *dpCry1* knockouts are photoperiodically unresponsive adds evidence for the involvement of yet another clock gene in photoperiodic sensing/responsiveness, and strongly suggests that a functional clock is necessary for photoperiodism. One possibility that cannot however be excluded is that of dpCRY1 serving as the photoperiodic sensor. Painting the compound eyes black, as previously shown [29], does not inhibit photoperiodic responses in wild-type females, indicating that the photoperiodic sensor lies within the brain. Whether dpCRY1, or other brain-expressed monarch opsins

function to encode the photoperiod remains to be tested.

Mapping the genes and molecular pathways underlying circadian and photoperiodic entrainment will not only increase our understanding on how the circadian clock functions to regulate the biology of an organism, but also will allow us to start putting together the clues on how clocks evolved over time.

4.5. Materials and Methods.

4.5.1. Generation of a Monarch *dpCry1* Knockout Line.

Monarch *dpCry1* knockouts were generated via CRISPR/Cas9-mediated targeted mutagenesis, as described in [17].

4.5.2. Animal Husbandry.

For all experiments listed, monarch larvae were reared on a semi-artificial diet and adults were manually fed a 25% honey solution every day.

For testing rhythmic eclosion behavior, wild-type and dp*Cry1* homozygous mutant monarchs were raised at 25°C with 70% humidity under the following light:dark cycle (LD) conditions: 15 h: 9 hr LD from eggs to fifth instar larvae, 12 h: 12 hr LD from the day 1-8 of pupation, and constant dark (DD) thereafter. After eclosion, adults were re-entrained to 12 h:12 hr LD cycles for at least 7 days before brains were dissected in DD, as previously described in Chapter 2.5.12. [29].

For testing eclosion behavior under temperature cycles (TC), wild-type,

dpCry1 and *dpClk* homozygous mutant monarchs were raised from eggs to fourth instar in 12 h:12 hr LD at 25°C with 70% humidity, moved to constant light conditions (LL) at same temperature and humidity levels at the fifth instar larvae, and then moved at pupation to DD in either a 12 h-20°C / 12 h-25°C or a 12 h-15°C / 12 h-25°C TC in anti-phase to the previous larval LD cycle. 12 hr of exposure to 25°C occurs in the larvae's previous dark phase and 12 hr of exposure to 20°C or 15°C occurs in the larvae's previous light phase. After eclosion, adults were maintained in the same 12 h-15°C / 12 h-25°C cycles in DD for at least 7 days before dissection.

Rearing of the monarchs for photoperiodic experiments was done as described in Chapter 2.5.1 [29]. In brief, upon eclosion, adult females were painted with either a clear paint or black paint over the compound eye and dissected 10 days post-eclosion.

4.5.3. Compound Eye Painting.

Compound eyes were covered on the day of eclosion with either clear or black paint as described in Chapter 2.5.4. [29].

4.5.4. Evaluation of Female Reproductive Status.

Female abdomens were dissected and oocytes individually counted as previously described in Chapter 2.5.3. [29].

4.5.5. Real-Time qPCR.

For the circadian time course experiment, the brains of adult wild-type and *dpCry1* homozygous mutants were entrained to at least 7 days of 12 h: 12 hr LD cycles and dissected during the first day of DD after entrainment under red light in 0.5X RNA later (Invitrogen). Brains were dissected at circadian time (CT) CT0, CT3, CT6, CT9, CT12, CT15, CT18, and CT21.

For the temperature entrainment time course experiment, brains of adult wild-type and *dpCry1* homozygous mutants were entrained to at least 7 days of 12 h-15°C / 12 h-25°C in DD before being dissected at ZT0, ZT4, ZT8, ZT12, ZT16, and ZT20.

For both experiments, RNA extractions and gene expression quantifications were performed as previously described in 2.5.12. [29], using validated monarch *per*, *tim*, and control *rp49* primers [28]. The primers for *tim* were as follows (F, forward primer; R, reverse primer): *timF*, 5'-GGACGGCAGCGGATACG -3'; *timR*, 5'-CGCCGTTTCGCACACA-3'. The data were normalized to *rp49* as an internal control and normalized to the mean of one sample within a set for statistics.

4.5.6. Eclosion Behavior Assays.

For regular eclosion assays, pupae were raised for 7 days in 12 h: 12 hr LD cycles at 25°C before being released into DD. Eclosion was observed on the second day of DD. For temperature entrained eclosion assays, eclosion was

delayed and observed between 10-15 d after pupation while being continuously entrained to 12 h-20°C / 12 h-25°C or 12 h-15°C / 12 h-25°C in DD. Eclosion data were analyzed and plotted as 1-h bins, as in [28, 29].

4.5.7. Statistical Analysis.

P-values were calculated using one-way ANOVAs or two-way ANOVAs followed by post-hoc analyses, and Kolmogorov–Smirnov tests using online calculators at https://www.icalcu.com/stat/anova-tukey-hsd-calculator.html, https://www.wessa.net/rwasp_Two%20Factor%20ANOVA.wasp [41], and http://www.physics.csbsju.edu/stats/KS-test.n.plot_form.html.

4.6. References.

1. Lin, C. and T. Todo, *The cryptochromes.* Genome Biol, 2005. **6**(5): p. 220.
2. Todo, T., *Functional diversity of the DNA photolyase/blue light receptor family.* Mutat Res, 1999. **434**(2): p. 89-97.
3. Mei, Q. and V. Dvornyk, *Evolutionary History of the Photolyase/Cryptochrome Superfamily in Eukaryotes.* PLoS One, 2015. **10**(9): p. e0135940.
4. Emery, P., et al., *CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity.* Cell, 1998. **95**(5): p. 669-79.

5. Stanewsky, R., et al., *The cryb mutation identifies cryptochrome as a circadian photoreceptor in Drosophila.* Cell, 1998. **95**(5): p. 681-92.
6. Zhu, H.S., et al., *Cryptochromes define a novel circadian clock mechanism in monarch butterflies that may underlie sun compass navigation.* Plos Biology, 2008. **6**(1): p. 138-155.
7. Zhu, H.S., et al., *The two CRYs of the butterfly.* Current Biology, 2005. **15**(23): p. R953-R954.
8. Yanovsky, M.J., et al., *Resetting of the circadian clock by phytochromes and cryptochromes in Arabidopsis.* J Biol Rhythms, 2001. **16**(6): p. 523-30.
9. Ceriani, M.F., et al., *Light-dependent sequestration of TIMELESS by CRYPTOCHROME.* Science, 1999. **285**(5427): p. 553-6.
10. Rosato, E., et al., *Light-dependent interaction between Drosophila CRY and the clock protein PER mediated by the carboxy terminus of CRY.* Curr Biol, 2001. **11**(12): p. 909-17.
11. Dolezelova, E., D. Dolezel, and J.C. Hall, *Rhythm defects caused by newly engineered null mutations in Drosophila's cryptochrome gene.* Genetics, 2007. **177**(1): p. 329-45.
12. Helfrich-Forster, C., et al., *The circadian clock of fruit flies is blind after elimination of all known photoreceptors.* Neuron, 2001. **30**(1): p. 249-61.
13. Ogueta, M., R.C. Hardie, and R. Stanewsky, *Non-canonical Phototransduction Mediates Synchronization of the Drosophila*

melanogaster Circadian Clock and Retinal Light Responses. Curr Biol, 2018. **28**(11): p. 1725-1735 e3.

14. Senthilan, P.R., et al., *Role of Rhodopsins as Circadian Photoreceptors in the Drosophila melanogaster.* Biology (Basel), 2019. **8**(1).
15. Szular, J., et al., *Rhodopsin 5- and Rhodopsin 6-mediated clock synchronization in Drosophila melanogaster is independent of retinal phospholipase C-beta signaling.* J Biol Rhythms, 2012. **27**(1): p. 25-36.
16. Ni, J.D., et al., *A rhodopsin in the brain functions in circadian photoentrainment in Drosophila.* Nature, 2017. **545**(7654): p. 340-344.
17. Wan, G., et al., *Cryptochrome 1 mediates light-dependent inclination magnetosensing in monarch butterflies.* Nat Commun, **12**(1): p. 771.
18. Chaves, I., et al., *The cryptochromes: blue light photoreceptors in plants and animals.* Annu Rev Plant Biol, 2011. **62**: p. 335-64.
19. Kao, Y.T., et al., *Ultrafast dynamics and anionic active states of the flavin cofactor in cryptochrome and photolyase.* J Am Chem Soc, 2008. **130**(24): p. 7695-701.
20. Lin, C., et al., *Circadian clock activity of cryptochrome relies on tryptophan-mediated photoreduction.* Proc Natl Acad Sci U S A, 2018. **115**(15): p. 3822-3827.
21. Martin, R., et al., *Ultrafast flavin photoreduction in an oxidized animal (6-4) photolyase through an unconventional tryptophan tetrad.* Phys Chem Chem Phys, 2017. **19**(36): p. 24493-24504.

22. Nohr, D., et al., *Extended Electron-Transfer in Animal Cryptochromes Mediated by a Tetrad of Aromatic Amino Acids.* Biophys J, 2016. **111**(2): p. 301-311.

23. Berndt, A., et al., *A novel photoreaction mechanism for the circadian blue light photoreceptor Drosophila cryptochrome.* J Biol Chem, 2007. **282**(17): p. 13011-21.

24. Ganguly, A., et al., *Changes in active site histidine hydrogen bonding trigger cryptochrome activation.* Proc Natl Acad Sci U S A, 2016. **113**(36): p. 10073-8.

25. Hoang, N., et al., *Human and Drosophila cryptochromes are light activated by flavin photoreduction in living cells.* PLoS Biol, 2008. **6**(7): p. e160.

26. Vaidya, A.T., et al., *Flavin reduction activates Drosophila cryptochrome.* Proc Natl Acad Sci U S A, 2013. **110**(51): p. 20455-60.

27. Froy, O., et al., *Illuminating the circadian clock in monarch butterfly migration.* Science, 2003. **300**(5623): p. 1303-5.

28. Markert, M.J., et al., *Genomic Access to Monarch Migration Using TALEN and CRISPR/Cas9-Mediated Targeted Mutagenesis.* G3 (Bethesda), 2016. **6**(4): p. 905-15.

29. Iiams, S.E., et al., *Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly.* Proc Natl Acad Sci U S A, 2019.

30. Malpel, S., A. Klarsfeld, and F. Rouyer, *Larval optic nerve and adult extra-retinal photoreceptors sequentially associate with clock neurons during Drosophila brain development.* Development, 2002. **129**(6): p. 1443-53.

31. Brown, S.A., et al., *Rhythms of mammalian body temperature can sustain peripheral circadian clocks.* Curr Biol, 2002. **12**(18): p. 1574-83.

32. Buhr, E.D., S.H. Yoo, and J.S. Takahashi, *Temperature as a universal resetting cue for mammalian circadian oscillators.* Science, 2010. **330**(6002): p. 379-85.

33. Currie, J., T. Goda, and H. Wijnen, *Selective entrainment of the Drosophila circadian clock to daily gradients in environmental temperature.* BMC Biol, 2009. **7**: p. 49.

34. Wheeler, D.A., et al., *Behavior in light-dark cycles of Drosophila mutants that are arrhythmic, blind, or both.* J Biol Rhythms, 1993. **8**(1): p. 67-94.

35. Chen, C., et al., *Drosophila Ionotropic Receptor 25a mediates circadian clock resetting by temperature.* Nature, 2015. **527**(7579): p. 516-20.

36. Chen, C., et al., *nocte Is Required for Integrating Light and Temperature Inputs in Circadian Clock Neurons of Drosophila.* Curr Biol, 2018. **28**(10): p. 1595-1605 e3.

37. Ikeno, T., H. Numata, and S.G. Goto, *Photoperiodic response requires mammalian-type cryptochrome in the bean bug Riptortus pedestris.* Biochem Biophys Res Commun, 2011. **410**(3): p. 394-7.

38. Ikeno, T., H. Numata, and S.G. Goto, *Circadian clock genes period and cycle regulate photoperiodic diapause in the bean bug Riptortus pedestris males.* J Insect Physiol, 2011. **57**(7): p. 935-8.

39. Ikeno, T., et al., *Photoperiodic diapause under the control of circadian clock genes in an insect.* BMC Biol, 2010. **8**: p. 116.

40. Pegoraro, M., et al., *Role for circadian clock genes in seasonal timing: testing the Bunning hypothesis.* PLoS Genet, 2014. **10**(9): p. e1004603.

41. Holliday, I.E. *Two-Way ANOVA (v1.0.6) in Free Statistics Software (v1.2.1)*. 2019 [cited 2017 November 21]; Available from: https://www.wessa.net/rwasp_Two%20Factor%20ANOVA.wasp/.

5. DEVELOPMENT OF ADDITIONAL GENETIC TOOLS IN THE MONARCH BUTTERFLY: FROM A LUCIFERASE REPORTER DPN1 CELL LINE TO IN VIVO HOMOLOGY-DIRECTED REPAIR AND TRANSGENESIS

5.1. Overview.

With the development of contemporary tools such a draft genome [1], DpN1 cell culture line [2], and the availability of CRISPR/Cas9 technology for genome editing [3-5], the monarch has emerged as an additional model organism to Drosophila for mechanistic studies of the circadian clock at both molecular and cellular levels. Despite these developments, a number of key tools are still lacking to further enhance the usability of the monarch as a powerful model system. Here, I used CRISPR-assisted homologous recombination to generate a stable, rhythmic PERIOD::LUCIFERASE reporter in the light-driven monarch DpN1 cells. The establishment of this reporter cell line provides an essential tool, similar to that of the PER2::LUCIFERASE reporter in mammals [6], for high-throughput studies of insect clockwork mechanisms *in vitro.* Additionally, I also established plasmid constructs and mRNAs for testing various strategies for tagging clock genes *in vivo* with the fluorescent reporter *tdTomato* to ultimately gain insights into the clock neuronal network in the monarch brain. One strategy relies on CRISPR-assisted homologous recombination and the other on transgenesis using either a *piggybac* or a *tol2* transposase. Both CRISPR-assisted and transgenesis-based strategies can be

tested to determine the most efficient method for tagging the monarch clock gene *cryptochrome 2* with *tdTomato*.

5.2. Introduction.

The monarch butterfly has steadily been gaining ground over the past decade as a compelling model organism for studying the evolution of the core circadian clock mechanism due to its Drosophila/mammalian hybrid composition, and for dissecting the molecular and cellular mechanisms underlying the circadian regulation of seasonal responses, such as that of its migration and associated reproductive diapause. However, as expected for a non-conventional model, the monarch currently lags behind conventional model systems in the development of cutting-edge tools such as *in vitro* and *in vivo* reporters to rapidly gain molecular mechanistic insights in cell cultures or the living organism itself, as well as identify their cellular basis in the brain.

Firefly luciferase, the enzyme that gives fireflies their bioluminescence [7], is a commonly used circadian reporter that produces light upon the catalyzation of luciferin, ATP, and oxygen. The luciferase's luminescence production has been leveraged as a simple tool for both measuring gene expression at the transcriptional level and reporting circadian activity at the protein level [6, 8, 9]. Work in the mammalian circadian field has been greatly facilitated by the generation of a mouse PERIOD2::LUCIFERASE (PER2::LUC) fusion protein that provides real-time reporting of circadian dynamics [6]. This reporter has

been used for testing core-peripheral clock signaling, identifying modulators of the clock, and has allowed the establishment of PER2::LUC cell cultures to assess tissue-specific clock function [6, 10-13]. Such a tool is however currently lacking in insect models as *Drosophila* Schneider 2 (S2) cells are clock-less [14, 15]. The only insect cell line known to date to harbor a clock, albeit driven by light/dark cycles, is a monarch cell line called DpN1 in which the clock genes *period* (*dpPer*), *timeless* (*dpTim*), and *cryptochrome 2* (*dpCry2*) cycle at both the RNA and protein levels with similar phases than in monarch brains [2, 16]. CRISPR/Cas9-mediated homology directed repair could be used to knock-in LUCIFERASE (LUC) in frame with PERIOD (dpPER) to ultimately generate a DpN1 cell line containing a PER::LUC reporter as a tool to rapidly dissect insect clockwork mechanisms *in vitro*.

In addition to facilitating *in vitro* studies, reporters could also be used to tag clock neurons *in vivo* for mapping the brain clock circuitry and testing for its possible seasonal plasticity. Fluorescent proteins are ideal reporters to report cellular processes *in vivo* such as in cell-lineage tracing, transplantation studies, cell morphology analysis, and fusion protein applications [17, 18]. In the context of temporal studies over the 24-hr day, the fluorescent protein tdTOMATO, a genetic fusion of two dTOMATO proteins, is one of the most suitable as its excitation/emission spectra is above the blue-light sensitive photoreceptor dpCRY1 and thus excitation of this molecule would not disturb the clock at night [19, 20]. Focusing first on tagging clock neuron with *tdTomato* under the control

of the promoter of the circadian repressor *dpCry2*, two main strategies can be tested: (1) insertion of *tdTomato* at the endogenous locus via CRISPR-assisted homologous recombination or (2) insertion of *dpCry2-tdTomato* somewhere in the genome via transgenesis. Many CRISPR-based HDR strategies have been developed [21], and of those methods relying on either a single sgRNA to create the DSB or two sgRNAs to remove a DNA region of equivalent size to *tdTomato* should both be tested. Although the insertion of *tdTomato* at the *dpCry2* endogenous locus would be ideal, it is anticipated to be a low efficiency process as recombination events are rare and unlike *in vitro* it is not possible to select for knock-in events [22]. Another, presumably more efficient strategy for tagging *dpCry2* positive neurons, would be to use transgenesis. Transgenesis utilizes a donor transposon plasmid containing the desired transgene flanked by the transposase recognition sequences. The transposase recognizes these sequences and “cuts” it from the donor plasmid to “paste” it into the genome. Given their availability and possible compatibility in the monarch model, we have selected two transposase systems to test from amongst those that have been demonstrated to work in insect models, the cabbage looper moth *piggybac* or the medaka fish *tol2*, where *piggybac* “pastes” into “TTAA” sites and *tol2* at random sites [23-25]. With more possible locations of transposition into the genome, and the recent demonstration that *piggybac* display reduced efficiency in lepidopterans compared to many other insect classes [24], *tol2*-mediated transgenesis might be the most efficient option in monarchs. Here, both

CRISPR-assisted knock-in and transgenesis-mediated strategies will be tested to find the most efficient strategy for tagging clock genes *in vivo*.

5.3. Results.

5.3.1. Generation of a PERIOD::LUCIFERASE Reporter Rhythmic DpN1 Stable Cell Line Using CRISPR-Assisted Homologous Recombination.

In order to generate a PER::LUC reporter line in monarch DpN1 cells using CRISPR/Cas9-assisted homologous recombination, I first identified a potential knock-in site in the last intron of the *period* (*dpPer*) locus (Fig. 5.1 *A*) and subcloned the sequence into an “All-in-One” sgRNA and Cas9 expressing plasmid (Kusakabe lab; [26]). I then validated the efficiency of the sgRNA for inducing double stranded breaks (DSB) using a Cas9-based *in vitro* cleavage assay (not shown). I then designed and constructed a knock-in plasmid containing the left and right homology arms each with ~1000 bp of complementary sequence upstream (3’ end of *dpPer*) and downstream (3’ UTR) of the selected target site, respectively (Fig. 5.1 *B*). Between these homology arms, I subcloned the firefly *luciferase* (*luc*) sequence (lacking the start codon) in frame with the last exon of *dpPer* (lacking the stop codon) so that if the knock-in construct was successfully recombined into the genome after a CRISPR-generated DSB, it would generate a protein fusion reporter (Fig. 5.1 *B*). Additionally, I also included a selectable *hygromycin* resistance cassette under the control of an actin promotor to select for DpN1 cells with recombination

events, and flanked the cassette with Cre recombinase recognition sites, lox p, for later removal (Fig. 5.1 *B*). Transfection of DpN1 cells with the linearized knock-in construct and the "All-in-One" plasmid, followed by several passages in hygromycin containing media led to the formation of isolated colonies that were picked and cultured individually. Once confluent in a 25-cm^2 flask, cells were harvested from each colony and screened by PCR to confirm recombination using a primer within the construct and a primer hybridizing to the adjacent genomic region. Several colonies were identified to have the correct sized amplicon and sequence for the PER::LUC knock-in, but they also contained other sized amplicon variants which when sequenced revealed deletions/additions of DNA sequence, likely due to mis-recombination events (Fig. 5.1 *C*). One colony (#10) however produced a single amplicon of the correct size and showed the proper *per::luc* fusion upon sequencing (Fig. 5.1 *C* and *D*).

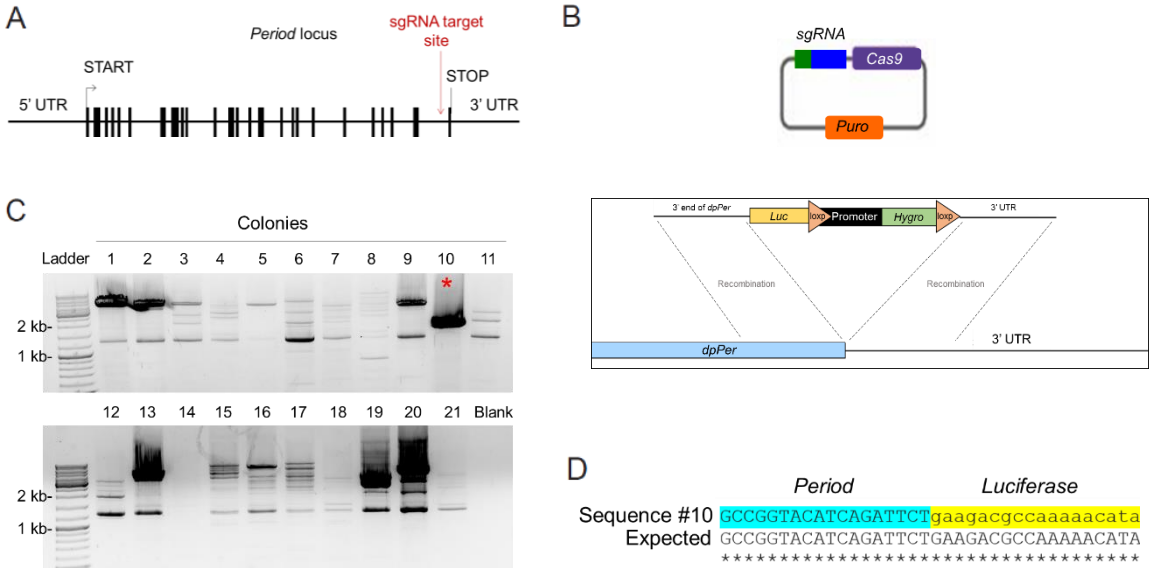


Figure 5.1: Generating a PER::LUC knock-in at the *period* locus in monarch DpN1 cells.

(*A*) Gene model of protein coding region for *period* (*dpPer*) and location of the sgRNA target site tested. Vertical boxes: exons; Horizontal lines: introns. (*B*) Schematic of knock-in strategy and constructs. (*Top*) The selected 20bp guide site was cloned into a plasmid upstream of a guide scaffold sequence and driven by a U6 promoter. The plasmid also contains a *Cas9* sequence under the control of an OpIE2 promoter, along with a *puromycin* (*puro*) cassette under the control of an actin A3 mini promoter. (*Bottom*) Insertion of the construct relies on homology directed repair (HDR) of a DSB where the gene of interest can be knocked-in via recombination with left and right flanking homology arms. Construct designed with left homology arm containing the 3' end sequence of *dpPer* lacking a stop codon and fused to *luciferase* (*luc*) lacking a start codon. Knock-in construct includes a *hygromycin* (*hygro*) selection cassette which is flanked by Cre recombinase recognition sites (loxp) for later removal. Right homology arm consists of the 3' UTR of *dpPer*. (*C*) PCR screen for knock-in events using a forward primer located in the genome upstream of the left homology arm and a reverse primer located in *luc*. Recombinant cells should yield a 2,111bp fragment. Red star denotes a colony with a singular correct sized PCR fragment. (*D*) Sequencing of PER::LUC fusion region in colony 10. *dpPer* lacks the stop codon and *luc* lacks a start codon.

Because recombinant cells contained the selectable *hygromycin* resistance cassette displacing *luc* from the 3' UTR of *dpPer,* which could disrupt rhythmic expression of LUC, I next transfected this colony with a pBA plasmid

driving the expression of *Cre recombinase* to “flox” out the *hygromycin* resistance cassette and its actin promoter. After transfection, the colony was diluted across a 96-well plate in order to separate cells with floxed alleles from those with unfloxed ones. Once individual colonies were confluent in a 25-cm^2 flask, they were each screened by PCR for the presence or absence of the *hygromycin* resistance cassette and promoter using flanking primers. Three colonies, #10, #15, and #34, were found amongst the first batch of screening to exhibit the shortened PCR amplicon corresponding to the floxed allele (Fig. 5.2 *A*). To test for any unintended genomic integrations of the plasmids used to generate this reporter, I also screened for the presence of *Cas9*, *Cre recombinase*, *hygromycin*, and *puromycin* in these three colonies by PCR (Fig. 5.2 *B*). While colony #10 did appear to have some *hygromycin* integration elsewhere in the genome, all three colonies were devoid of the more detrimental integrations for *Cas9* and *Cre recombinase* which could have potentially been driving mutations across the genome (Fig. 5.2 *B*). An initial screen for rhythmic LUC expression across the three colonies was promising, but colonies #10 and #15 died soon after and before being able to stock them in liquid nitrogen. Colony #34 remained healthy and exhibited a stable, rhythmic expression of LUC under light:dark (LD) cycles with a phase similar to that of dpPER in monarch brains [16] (Fig. 5.2 *C*). Because *dpPer* could be present in several copies in DpN1 cells, we could not ascertain whether other wild-type *dpPer* copies would be absent. To test for this possibility, I used PCR primers flanking

luc. I found that there are indeed other copies of *dpPer* in this cell line as well as another wild-type-like *dpPer* variant; however, the amplicon size indicating the presence of the *luc* knock-in did not amplify from the PER::LUC sample (Fig. 5.2 *D, Left*). To ensure that the reporter was not lost, I then did two follow up PCRs using primers within and outside of *luc*. With these primers, I found that the PER::LUC was still present in the cell line, and may not have been amplified in the initial PCR maybe due to primers being used up by the smaller, wild-type sized amplicons (Fig. 5.2 *D*, *Middle, Right*).

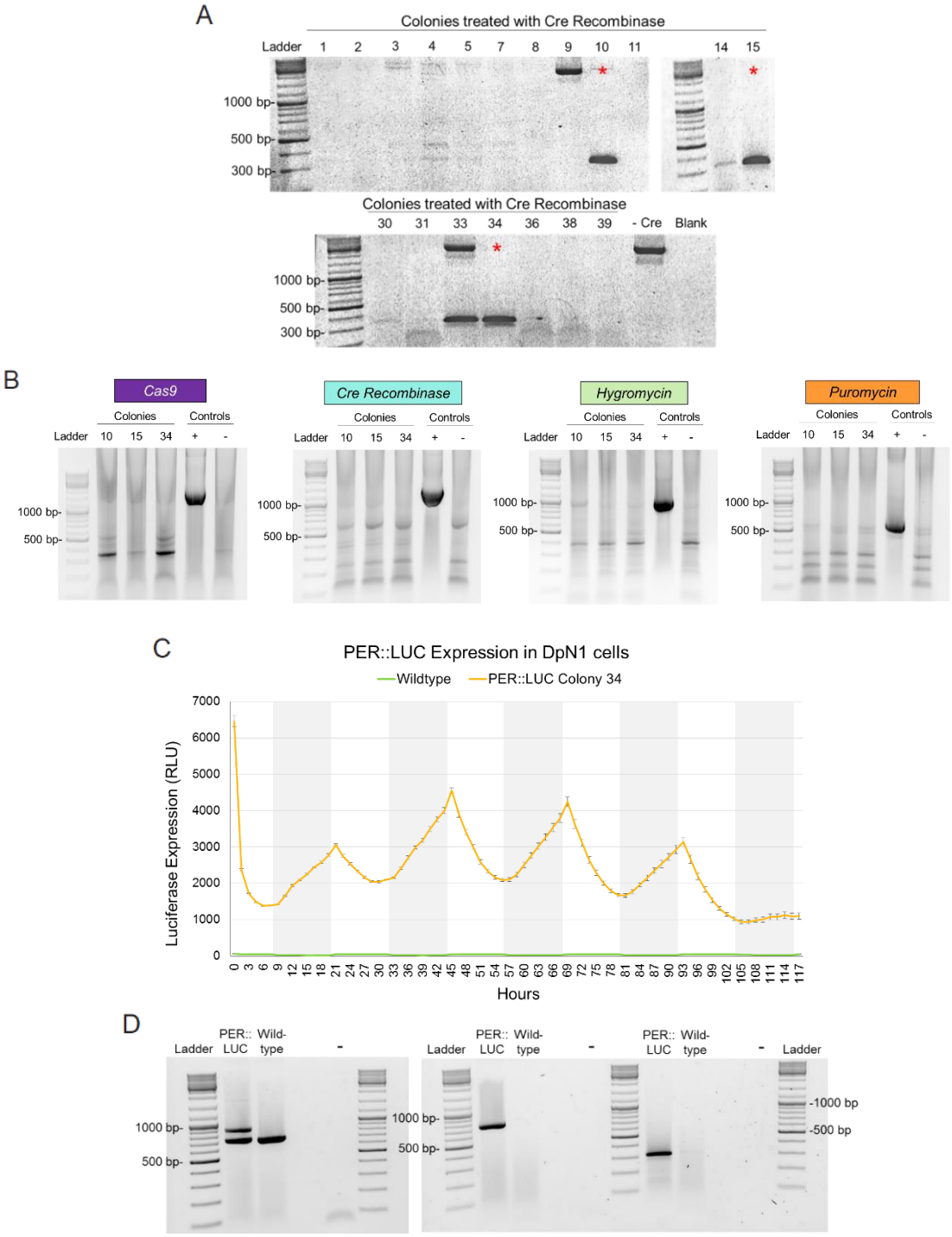


Figure 5.2: Identification of a PER::LUC DpN1 colony and establishment of a stable cell line with rhythmic LUC expression.

(*A*) PCR screen of surviving DpN1 colonies using primers flanking the lox p sites. Full length, unfloxed amplicon: 2,587bp. Floxed amplicon: 309bp. Red stars: fully floxed

Figure 5.2 Continued

colonies. (*B*) The three colonies were screened for unintended integrations of: (Far left) the *Cas9* gene, 1,272bp; (Left middle) *Cre recombinase*, 1,348bp; (Right middle) *Hygromycin*, 1,026bp; (Far right) *Puromycin*, 600bp. (*C*) Luciferase expression in relative light units (RLU) of colony 34. Under 12:12 light (white):dark (grey) cycles, colony 34 (yellow) shows rhythmic LUC expression in phase with dpPER expression in the monarch brain [16]. Wild-type control DpN1 cells are shown in green; Timepoints are taken every 1.5 h. n=9 (*D*) PCR screen for additional copies of *dpPer*. (*Left*) PCR screen by size using primers only in the *dpPer* sequence. *dpPer* fused to *luc* amplifies 2,425 bp, wild-type *dpPer* amplifies 711 bp. (*Middle*) PCR using a forward primer in *dpPer* and a reverse in *luc*, *dpPer* fused to *luc* amplifies 736 bp, wild-type *dpPer* amplifies nothing. (*Right*) PCR using a forward primer in *luc* and a reverse primer in *dpPer*, *dpPer* fused to *luc* amplifies 309 bp, wild-type *dpPer* amplifies nothing. For each panel, samples were loaded as follows: ladder, PER::LUC DpN1 cell DNA, wild-type DpN1 DNA, empty well, and negative control.

5.3.2. Testing the Usability of the PER::LUC Reporter for Rapid Assessment of Clockwork Mechanisms.

As described in Chapter 4, *Cryptochrome 1* (*dpCry1*) appears to be the sole photoreceptor necessary for entrainment of the monarch circadian clock. To provide evidence that it does do in a cell autonomous manner as well as demonstrate the usability of the newly made PER::LUC line in testing clockwork mechanisms *in vitro*, I tested the effect of *dpCry1* dsRNA knockdown on rhythmic LUC expression in LD by targeting the 5’ UTR, cDNA, and 3’ UTR of *dpCry1*. If *dpCry1* acts in the light entrainment pathway in a cell autonomous manner, I would expect LUC rhythm to be disrupted in its absence. As compared to untransfected PER::LUC cells and PER::LUC cells transfected with a control dsRNA targeting *eGFP*, dsRNA-mediated knockdown of *dpCry1* disrupted rhythmic LUC expression (Fig. 5.3). Some LUC rhythm was still present in the dsRNA-*dpCry1* knockdown cells, but with a dramatically reduced amplitude.

Given the major disruption in LUC rhythm amplitude, these results support the role of dpCRY1 as a cell autonomous photoreceptor necessary for monarch circadian clock entrainment. It also provides a proof-of-principle for the usability of this reporter line for future mechanistic clockwork studies.

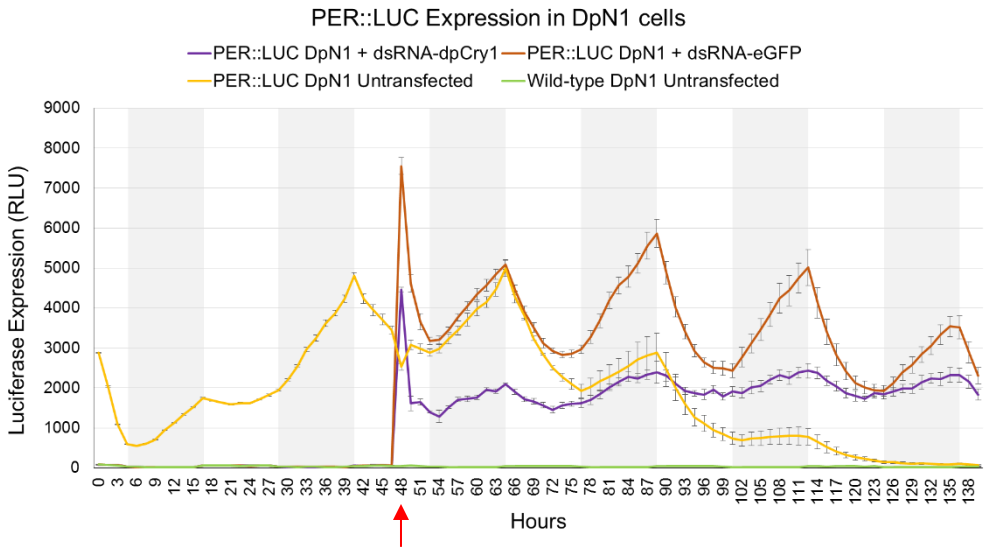


Figure 5.3: dsRNA knock-down of *dpCry1* disrupts rhythmic LUC expression in the PER::LUC reporter.

Wild-type DpN1 cells and PER::LUC reporter DpN1 cells were entrained for two days in 12:12 light (white):dark (grey) cycles where the untransfected wild-type DpN1 cells (green) and untransfected PER::LUC DpN1 cells (yellow) served as controls and were grown in luciferin containing media from the day of entrainment. PER::LUC cells for knock-down were transfected 2 days later with dsRNA either targeting *dpCry1* (purple) or the control *eGFP* (Dark orange). Luciferin containing media was then added to both the PER::LUC cell groups at the end of the transfection (red arrow). n=5 each timepoint. Timepoints taken every 1.5 h. Red arrow denotes when transfection occurred.

5.3.3. Generation of the Single or Two sgRNA-Based CRISPR/Cas9 Constructs for the *in vivo* Knock-In of *tdTomato* Under the Control of the *dpCry2* Promoter.

As previously shown for the generation of a DpN1 PER::LUC stable cell line, one possible strategy for tagging clock genes with a *tdTomato* fluorescent reporter *in vivo* is via CRISPR-assisted homology repair. Here, the aim is to knock-in a membrane (Mb)-tagged *tdTomato* just downstream of the *dpCry2* 5' UTR so that the reporter will be under the regulation of the endogenous clock gene promoter. To this end, I started by designing two constructs to test two possible strategies: (1) using a single sgRNA to create a DSB at the end of the *dpCry2* 5' UTR, and (2) using two sgRNAs to remove a portion of the *dpCry2* 5' genomic sequence of size roughly equivalent to the transgene to be inserted (Fig. 5.4 *A*). While both will rely on HDR for to knock-in Mb-*tdTomato* under the control of the endogenous *dpCry2* promoter, the two sgRNA-based method may fare better for topology reasons. Both methods are awaiting injection with the sgRNA(s), Cas9 mRNA, and corresponding knock-in construct into monarch eggs and must be done within 20 min of fertilization to target the nuclei during the first divisions (Fig. 5.4 *A*). Surviving larvae should be screened by PCR for knock-in events, candidates backcrossed or intercrossed, and the resulting progeny also screened for inherited germline knock-in events. If a founder is identified amongst the progeny, it can be used to establish the *dpCry2:tdTomato* monarch line and allow for verification of clock-cell specific expression, mapping

of the neuronal circuitry, and testing for seasonal rewiring between long (LP) and short photoperiods (SP) (Fig. 5.4 *B*).

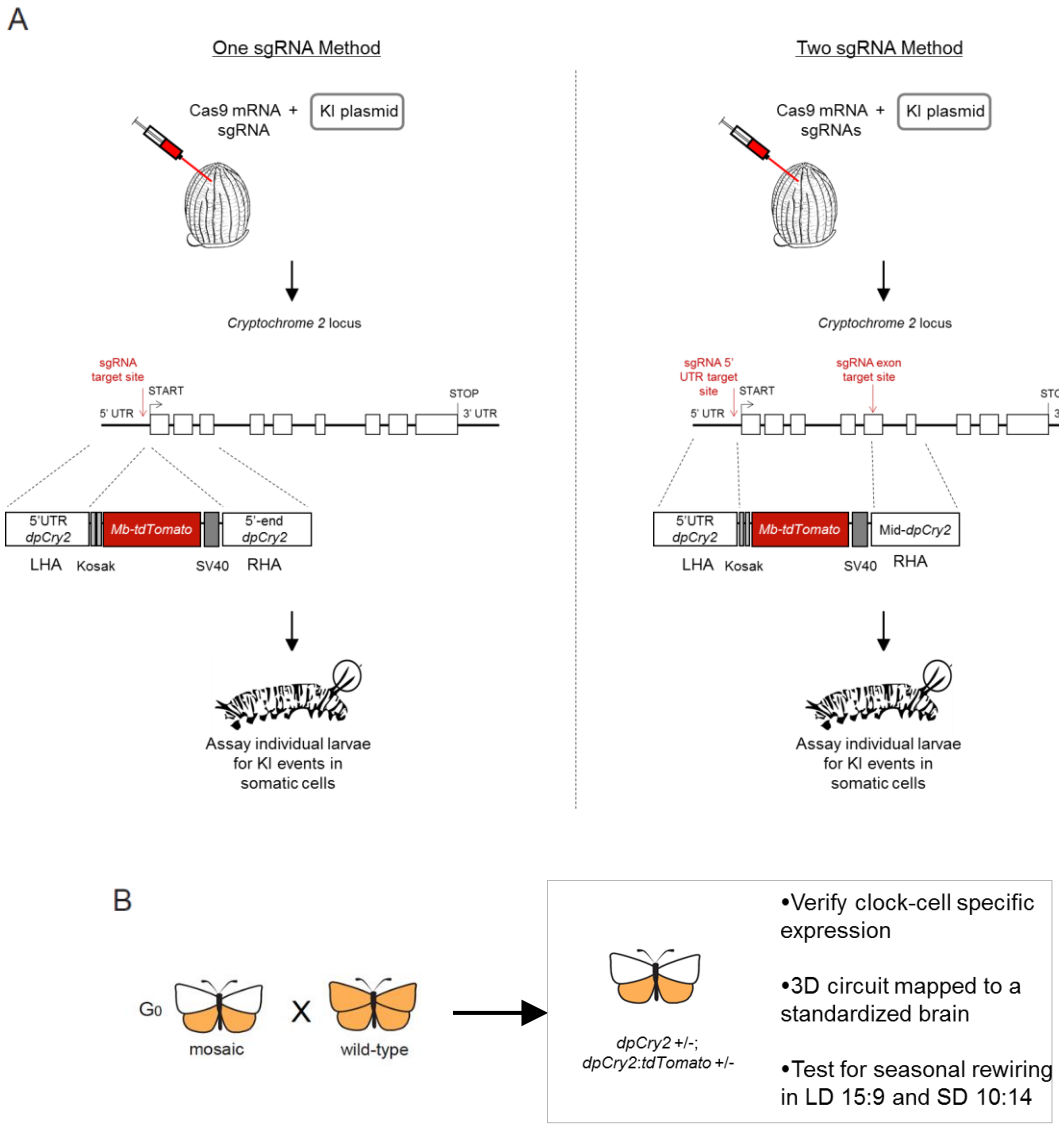


Figure 5.4: Strategy using CRISPR-assisted homology repair to knock-in *tdTomato* under the control of the endogenous *dpCry2* promoter.

(*A*) Strategy for injecting either a single sgRNA (*Left*) or two sgRNAs (*Right*) to respectively generate one or two DSBs in *dpCry2* to allow for HDR and knock-in (KI) of the *tdTomato* KI contruct. The single sgRNA strategy will use a construct with a left

Figure 5.4 Continued

homology arm (LHA) in the *dpCry2* 5' UTR and a right homology arm (RHA) in the 5' end of the *dpCry2* gene. The two sgRNA method will use a construct with the same LHA, but a RHA in the mid region of the *dpCry2* genomic locus. Dotted lines represent areas of recombination. Any larvae surviving the injections will be screened for KI events. (*B*) Somatic mutants will be intercrossed or backcrossed to wild-type and the progeny will be screened for inherited germline knock-in events. KI positive larvae can be used to found the *dpCry2:tdTomato* line which can be used to verify clock-cell specific expression, map the clock neuronal circuitry, and to test for seasonal neuronal remodeling between long photoperiods (LP) and short photoperiods (SP)

5.3.4. Generation of the Constructs for the Transposition of *mCD8::tdTomato* Under the Control of the *dpCry2* Promoter via Piggybac and Tol2 Transgenesis.

Given that, as shown in other species [21], CRISPR-assisted homologous directed repair is anticipated to be a low frequency event in the monarch, transgenesis could be an alternative, more efficient, strategy. Transgenesis instead relies on a transposase that "cuts and pastes" the contents of a donor construct randomly into in the genome. In this case, the expression of *tdTomato* needs to be driven by the *dpCry2* promoter such that no matter where *tdTomato* is transposed, it is under clock gene regulation. I generated transposon plasmids for piggyback and tol2 transgenesis by fusing a membrane tag, the mouse lymphocyte marker mCD8 [27], to *tdTomato* to drive its expression to the cell membrane, and subcloned the *mCD8::tdTomato* sequence just downstream to ~1kb of the *dpCry2* promoter sequence within either flanking piggybac or tol2 recognition sequences into appropriate plasmids (Fig. 5.5). These donor constructs along with their corresponding mRNA encoding transposase,

piggybac or *tol2*, are now awaiting injection into monarch eggs as in the CRISPR knock-in strategy (Fig. 5.5 and 5.4). Surviving larvae may be screened for transposition events by PCR, but there is a chance the larvae may be harboring the unintegrated plasmid or that transposition events may only occur in a small number of cells and be undetectable by PCR. To overcome this, all surviving butterflies should be intercrossed and their resulting progeny screened for a more confirmatory presence of transposition. Alternatively, one could perform a southern blot to probe for genomic integrations of *mCD8::tdTomato* by using a single restriction enzyme to cleave the genomic DNA into many pieces from the parents and the progeny and compare the resulting band sizes to that of the same, but once cleaved transposon plasmid control. Any bands running at a size other than that of the linearized plasmid should indicate genomic integration. The tol2 transgenesis strategy may be more advantageous as tol2 can insert anywhere in the genome while piggybac transposes only into “TTAA” sites which comprise only about 0.78% of the monarch genome [23]. Regardless, these are two viable methods of transgenesis that may be key to tagging clock genes *in vivo* with a fluorescent reporter.

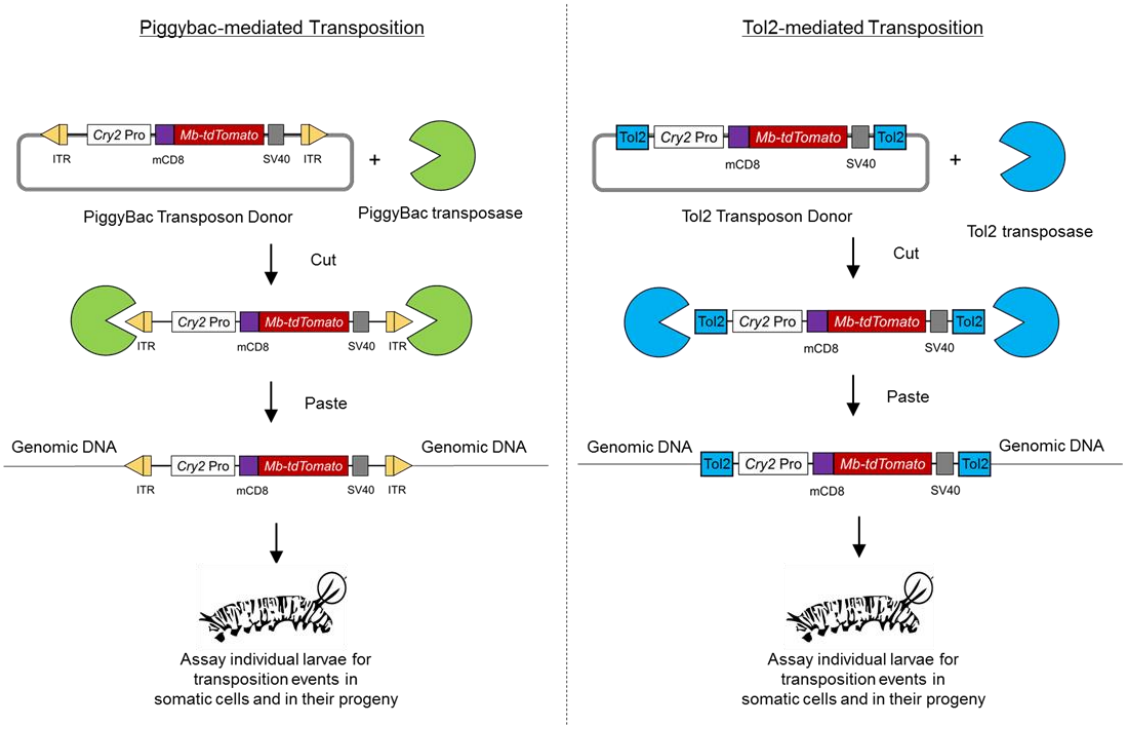


Figure 5.5: Strategy using either piggybac or tol2-mediated transgenesis to transpose *mCD8::tdTomato* under the control of the endogenous *dpCry2* promoter into the genome.

Schematic of the two transgenesis methods. The *dpCry2* promoter (*Cry2* Pro) was cloned upstream of the *mCD8::tdTomato* fusion sequence followed by a SV40 polyA tail. This was then cloned into either a piggybac or tol2 donor plasmid. The piggybac plasmid is flanked by inverted terminal repeats (ITRs), the recognition sequences of the *piggybac* transposase. The tol2 plasmid is flanked by *tol2* recognition sequences. The donor construct and the transposase can be injected into monarch eggs where the transposase will "cut" out the donor sequence and "paste" it into the genome. *Piggybac* will insert at "TTAA" sites and *tol2* will insert at random sites. Surviving larvae and their progeny will be screened for transposition events. Figure modified and expanded from [23].

5.4. Discussion.

In this chapter, using monarch DpN1 cells, I successfully generated to my knowledge the first rhythmic PER::LUC reporter in insect cells. The rhythm of LUC expression was in a phase similar to that of dpPER expression in the

monarch brain, peaking at the dark to light transition [16], suggesting that it can be used to dissect clockwork mechanisms *in vitro*. While screening for other copies of *dpPer* in this reporter line, I did identify at least one additional wild-type copy of *dpPer*, as well as a larger sized variant. This variant may be the result of a mis-recombination event where the *dpPer* target site was cleaved by the sgRNA but did not get the *luc* knock-in, and instead was poorly repaired by the repair machinery adding additional sequences. Sequencing of this larger wild-type-like amplicon would answer this question. Although we assume that the PER::LUC fusion protein is functional, we cannot exclude the possibility that the PER::LUC rhythms in our DpN1 stable cell line are actually driven by the wild-type copie(s). Importantly, by knocking down *dpCry1* mRNA in this cell line, I provided evidence that, just as in *Drosophila*, dpCRY1 acts cell-autonomously in the light entrainment pathway, adding to its role as the sole photoreceptor necessary for circadian entrainment *in vivo*. The residual, low amplitude rhythm observed in the dsRNA-mediated *dpCry1* knockdown cells is likely due to the fact that *dpCry1* expression was not fully eliminated, as is usually the case with knockdown approaches [28, 29].

Using similar approaches than those used *in vitro*, I also started to test several strategies for tagging the clock gene *dpCry2* with the fluorescent reporter *tdTomato* by generating the relevant constructs for both CRISPR-assisted homologous recombination and piggybac/tol2 transgenesis. While CRISPR-assisted homologous recombination is a viable strategy for knocking in

membrane-tagged *tdTomato* under the control of the endogenous *dpCry2* promoter, the chances of this occurring *in vivo* without a selective marker are very low, and would likely take thousands of monarch egg injections to identify a single candidate. Thus deploying *piggybac* or *tol2* transgenesis-based strategies to randomly integrate the *dpCry2-mCD8::tdTomato* sequence into the genome are valuable strategies to develop in parallel to CRISPR-assisted HDR. The generation of this *in vivo* reporter would give us the ability to verify clock-cell expression, map the clock neuronal circuitry in the monarch brain, and test for its possible seasonal, photoperiodic-driven, rewiring.

5.5. Material and Methods.

5.5.1. DpN1 Cell Maintenance.

DpN1 cells were maintained as previously described in Grace's media complete with 10% FBS (VWR 97068-085) and 0.5% Penicillin-Streptomycin (Sigma P0781) in a 28^{o}C incubator [16].

5.5.2. Generation of the PERIOD::LUCIFERASE Reporter in Monarch DpN1 Cells.

Please refer to Table 5.1 for sequences of primers used to generate the PERIOD::LUCIFERASE construct and screen for the reporter cell line.

5.5.2.1. Designing and Validating the sgRNA.

For the CRISPR/Cas9-assisted knock-in of *luciferase* (*luc*) fused to the 3' end of *period* (*dpPer*) in DpN1 cells, the guide RNA (gRNA), 5'-CGCCATCTATCAAAGGCATA-3', was selected in the last intron of *dpPer* using CHOPCHOP [30, 31].

Targeting efficiency of the sgRNA in DpN1 cells was tested by subcloning the guide directly upstream of a guide scaffold in an "All in One" plasmid (Kusakabe lab; [26]) using inverse PCR with a KOD Xtreme Hot Start DNA Polymerase (Novagen 71975, Toyobo), accompanying reagents, and the following primers at 10 μM in concentration: SI266/SI290. The thermocycling steps were done as follows: 94^{o}C 2 min, 5X(98^{o}C 10 sec, 51^{o}C 30 sec, 72^{o}C 10 min), 5X(98^{o}C 10 sec, 51^{o}C 30 sec, 70^{o}C 10 min), 25X(98^{o}C 10 sec, 51^{o}C 30 sec, 68^{o}C 10 min), 68^{o}C for a final elongation of 10 min, and cool to 4^{o}C. In addition to the gRNA and scaffold sequence, the plasmid also contains a *Cas9* sequence and a selectable *puromycin* (*puro*) cassette, each under the control of their own promoter. Once the DpN1 cells reached 90% confluency in a 25-cm^2 flask, the media was removed and the cells washed with 5 mL 1X Phosphate Buffer Saline (PBS) (Gibco 10010-023) to remove any residual serum. The adherent cells were dissociated from the bottom of the flask with 1mL of 1X Trypsin/EDTA solution (Gibco R-001-100) and further suspended in an additional 7 mL of complete Grace's media. The 8 mL mixture was then diluted into 20 mL of complete Grace's media and seeded across 6-well plates at 2

mL/well. Two days later and at 50% confluency, the cells were incubated with a 600 µL mixture of 500ng/well of plasmid, 15 µL/well of Cellfectin (Gibco 10362-100), and the remaining serum-free Grace's media for 5hr at 28^{o}C similar to the previously described protocol [16]. 1,400 µL of complete Grace's media was added at the end of the transfection. Two days post-transfection, the media was replaced with 5 µg/mL Puromycin (VWR 97064-280) containing complete media and passed with fresh selection media every 2 days. After 3 passages, cells were harvested, DNA extracted, and the target region amplified via PCR with the following primers (F, forward; R, reverse): F SI025 and R SI026. Cutting efficiency was tested by *in vitro* Cas9 cleavage assays [32]. To generate the gRNA necessary for this assay, annealed forward and reverse oligonucleotides corresponding to the 20bp target sequence (F SI088 and R SI089) were subcloned into the gRNA expression vector pDR274 (AddGene) at the BsaI cleavage site. Note that the oligonucleotides are designed to contain sequences generating BsaI overhangs flanking the target sequence to facilitate subcloning [33]. The single guide RNA (sgRNA) was then transcribed *in vitro* using T7 RNA polymerase with PCR amplified DNA template from the DR274 expression vector as previously described [4]. The sgRNA quality was then tested using *in vitro* Cas9 cleavage assays, where purified PCR product of the target region (250 to 350 ng) amplified with primers F SI025 and R SI026 was incubated with recombinant Cas9 protein (750 ng) and the sgRNA (400 to 600 ng), as previously described [32]. Complete cleavage of the full length PCR amplicon

into the two expected sized fragments indicated the presence of a wild-type allele. Additional, uncut, mutant amplicons were sequenced to verify the presence of mutations.

5.5.2.2. Designing the Knock-in Construct.

The PER::LUC knock-in construct was designed to contain ~1000bp of left and right homology arms flanking the CRISPR/Cas9 target site, *dpPer* lacking a stop codon fused in frame to *luc* lacking a start codon, and a selectable *hygromycin* resistance cassette under the control of an actin promoter and flanked by loxp recognition sites for later removal.

The SV40 polyA tail of a *luc* containing *dpPer4Ep-Luc* reporter plasmid [16] was removed via inverse PCR with KOD Xtreme Hot Start DNA Polymerase using primers SI342/SI343 which also added new restriction enzyme sites PacI and AgeI, respectively.

The left homology arm (LHA) was amplified from genomic DNA and constructed in two separate pieces in order to mutate the protospacer adjacent motif (PAM) of the gRNA sequence and prevent re-cutting within the knock-in construct. The upstream region of the left homology arm was amplified with primers F SI395 containing a KpnI site on the 5’ end and R SI366 containing a new 3 bp to mutate the PAM sequence. The downstream sequence was amplified with F SI365 containing the complementary mutated PAM sequence and 28 bp of overlap with SI366 on the 5’ and R SI396 with a NcoI site on the 5’.

These fragments were fused via overlapping PCR with Phusion High Fidelity DNA Polymerase (NEB M0530S) by annealing the 28 bp of overlap for 10 cycles and then adding distal primers F SI395 and R SI396 for another 25 cycles of amplification. The newly fused amplicon was then digested and ligated into *dpPer4Ep-Luc* using KpnI-HF (NEB) and NcoI (NEB). The stop codon of *dpPer* and the start codon of *luc* were then removed from the plasmid using inverse PCR with KOD Xtreme Hot Start DNA Polymerase and primers SI397/SI398.

To insert a *hygromycin* selection cassette, an actin A3 mini promoter was amplified from the “All in One” plasmid [26] using primers F SI345 and R SI346 both containing BamHI sites on their 5’ ends. The amplicon was digested and ligated into the pCoHygro plasmid (Thermo Fisher (Invitrogen; [34]) at the BamHI site, replacing the copia promoter. The left flanking loxp site was inserted into this new pA3Hygro plasmid by annealing primers F SI385 and R SI386 to create KasI and SalI 5’ overhangs (restriction sites will not be reformed once ligated, to prevent recutting) and also insert a PacI site, similar to a previously used method [33]. The right flanking loxp was generated by annealing primers F SI387 and R SI388 that would create KpnI and SacI 3’ overhangs (restriction sites will not be reformed once ligated, to prevent recutting) and insert another PacI site [33]. The loxp-A3pro-*hygro*-loxp region was excised by digestion with PacI and ligated into the respectively digested pLHA-*Luc* plasmid.

The right homology arm (RHA) was amplified from genomic DNA with F SI351 and R SI420 both with AgeI sites on the 5’ end. Both the RHA amplicon

and pLHA-*Luc*-loxp-A3pro-*hygro*-loxp plasmid were digested and ligated at the AgeI sites to complete the PER::LUC knock-in construct.

5.5.2.3. Generating the DpN1 Reporter Line.

Cells were seeded onto a 6-well plate following the procedures described above. The PER::LUC knock-in construct was linearized with SalI-HF and purified by a Zymoclean Gel DNA Recovery Kit (Zymo Research D4001) while fresh “All in One” plasmid was purified from bacterial cultures with a QIAprep Spin Miniprep Kit (Qiagen 27104). On day two, DpN1 cells were transfected with 200 ng/well of the linearized construct and 500ng/well of the “All in One” construct, as described above. Two days post-transfection, the media was replaced with 400 µg/mL Hygromycin B (GoldBio H-271-500) containing complete media and refreshed every 3 days. After 6 passages, isolated colonies growing in the wells were picked and cultured individually in 25-cm^2 flasks. Knock-in events were screened via PCR with a primer outside the knock-in construct, F SI434, and a primer within *luc*, R SI074, from DNA extracted from the cells in 0.01 × proteinase K (Sigma) in lysis buffer as described before in [3]. A positive colony was identified and verified by sequencing.

Cre recombinase was amplified from a pPGK-Cre-bpA plasmid (Adggene #11543) using primers F SI497/R SI496 containing NotI and XhoI sites in the 5’ tail, respectively, and ligated into a pBA (Reppert Lab[16]) plasmid at those restriction sites. The positive PER::LUC colony was then transfected with a

1.5mL mixture containing 1,250 ng of pBA-Cre Recombinase, 37.5 µL of Cellfectin, and the remaining volume serum-free Grace's media. 3.5 mL of complete media was added to the flask at the end of transfection. Two days post-transfection, cells were washed with 1X PBS and trypsinized. 5 µL of the trypsin/cell mix was diluted into 20 mL of complete media and distributed across a 96-well plate at 200 µL/well. Colonies were grown and moved to larger well plates as they reached confluency until maintainable in a 25-cm^2 flask. Colonies were screened for rhythmic LUC expression by trypsinizing a flask, diluting 500 µL of the trypsin/cell mixture into 1500 µL of complete media for a 1:3 dilution and distributing 150 µL/well across 10 wells of an OptiPlate-96 microplate (PerkinElmer 6005290). The remaining 500 µL was distributed as 50 µL/well across another 10 wells with 100 µL of additional complete media for a 1:11 dilution. The microplate was placed in a Percival incubator at 28^{o}C and the cells entrained to two 12:12 light:dark (LD) cycles. Media was replaced with media containing 0.025 mM luciferin (GoldBio LUCNA-100) and the plate placed in an EnVision Xcite Multilabel reader (PerkinElmer 2101-0020) located inside the incubator. Luminescence readings were taken every 1.5 hours for 3-4 days. Two candidates were identified with strong LUC rhythms and PCR screened for complete "floxing" of *hygromycin* from the construct (F SI072/R SI026) and any unintended genomic integrations of *hygromycin*, *puromycin*, *Cas9*, and *Cre recombinase* (respectively using the following primers: F SI516/R SI517, F SI509/R SI510, F SI525/R SI526, and F SI497/R SI496). One candidate

contained integration of *Cas9* was discarded while the second candidate was free of integrations but exhibited partial floxing of *hygromycin*, probably due to a mix of floxed and unfloxed cells. To isolate the floxed ones, the colony was trypsinized, 5 µL of trypsin/cell mix diluted into 20mL of complete media, and distributed across a 96-well as previously described. Colonies were screened for rhythmic LUC expression and then PCR screened for *hygromycin* floxing and all possible integrations as tested in the previous generation. A rhythmic PER::LUC colony fully floxed and free of plasmid integrations was identified, and stocked in serum-free Grace's media containing 10% DMSO (Sigma-Aldrich D2438) in liquid nitrogen.

Table 1: List of oligonucleotides used to generate the PER::LUC reporter in DpN1 cells.

SI025 5′-GTGGCCCTGTGTCCTAATCC-3′
SI026 5′-GATGCGAAATTTGTATCGCCCG-3′
SI037 5′-GCCGGTACATCAGATTCTGAAGACGCCAAAAACATA-3′
SI038 5′-GTGTAGTAAACATTCCAAAACC-3′
SI072 5′-TGTTTGTGGACGAAGTACCG -3′
SI074 5′-AGTACTCAGCGTAAGTGATGTC-3′
SI088 5′-TAGGCGCCATCTATCAAAGGCATA-3′
SI089 5′-AAACTATGCCTTTGATAGATGGCG-3′
SI266 5′-CACTTGTAGAGCACGATATT-3′
SI290 5′-CGCCATCTATCAAAGGCATAGTTTTAGAGCTAGAAATAGCAAGTT-3′
SI342 5′-TTTGTGTTAATTAATTACACGGCGATCTTTCCGCCCTTCTTGGCC-3′
SI343 5′- TGGGGGACCGGTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAG-3′
SI345 5′-TGTTGGGGATCCGCGCGTTACCATATATGGTG-3′
SI346 5′-TTTGTTGGATCCTCGATATCAAGCTTATCGATACCG-3′
SI351 5′-GTGTGGACCGGTTATTACGATTATGGGCGAGAGC-3′
SI365 5′-CAATATTATTTACTGAACAAAATAAACTAAAGC-3′
SI366 5′-AGTTTATTTTGTTCAGTAAATAATATTGTATGCCTTTGATAGATGGCG-3′
SI385 5′-GCGCTTAATTAAATAACTTCGTATAGCATACATTATACGAAGTTAT-3′

Table 1 Continued

SI386 5′-TCGAATAACTTCGTATAATGTATGCTATACGAAGTTATTTAATTAA-3′
SI387 5′-ATAACTTCGTATAGCATACATTATACGAAGTTATTTAATTAAAGCT-3′
SI388 5′-CTTAATTAAATAACTTCGTATAATGTATGCTATACGAAGTTATGTAC-3′
SI395 5′-TGTGTGGGTACCGGGTCAATTGTACTTGGATCTG-3′
SI396 5′-AAATAACCATGGCTAAGAATCTGATGTACCGG-3′
SI397 5′- GAAGACGCCAAAAACATAAAGAAAGGCCCG-3′
SI398 5′-AGAATCTGATGTACCGGCAACGTTATCTGG-3′
SI434 5′-CGGACGAAGAGTTTAGAGGAAG-3′
SI496 5′-TTTAAACTCGAGCCATAGAGCCCACCGC-3′
SI497 5′-TTTAAAGCGGCCGCCCCAAGAAGAAGAGGAAGGTGT-3′
SI509 5′-ATGACCGAGTACAAGCCCACGG-3′
SI510 5′-TCAGGCACCGGGCTTGCG-3′
SI516 5′-ATGGGAAAGCCTGAGCTGAC-3′
SI517 5′-CTATTCCTTTGCCCTCGGACG-3′
SI525 5′-TGGTGAATCAGTGTCGCGTC-3′
SI526 5′- CCTGTGGAAAATACCCAATTGC-3′

5.5.3. DsRNA Knockdown of *dpcry1* in PER::LUC DpN1 Cells.

dsRNA targeting the 5’ UTR, the cDNA, or the 3’ UTR of *dpcry1* was amplified from genomic DNA using primers with T7 promoter containing 5’ tails. 5’ UTR primers: *5’utrF*, 5’-TAATACGACTCACTATAGGAACTTAAAAATACAGGTT-3’; *5’utrR*, 5’-TAATACGACTCACTATAGGCTTGACTTCAATATATCT-3’. cDNA primers: *cDNAF*, 5’- TAATACGACTCACTATAGGTATACGTACGCTATGCTTCG-3’; *cDNAR*, 5’- TAATACGACTCACTATAGGCTGCAGAATGTCTCGTACTCC-3’. 3’ UTR primers: *3’utrF*, 5’-TAATACGACTCACTATAGGATATAACCTTACTGCCAACTAG-3’; *3’utrR*, 5’-TAATACGACTCACTATAGGTTTGGTTTCTTAGACAAATTGGG-3’. The resulting amplicons were purified using a NucleoSpin Gel and PCR Clean up kit

(Machery-Nagel 740609), *in vitro* transcribed using the T7 RNA polymerase and the resulting RNA was purified with acid-phenol, as previously described [4]. The RNA was then heated to 65^{o}C for 10 min before being allowed to cool to room temperature to anneal the RNA. For the transfection, 152 ng of 5' UTR targeting, 9225 ng of cDNA targeting, and 1214 ng of 3' UTR targeting dsRNA were combined together, and transfected in five replicates of PER::LUC DpN1 cells seeded in 96-wells plate (~2.12 µg per well/replicate). Five control replicates of PER::LUC DpN1 cells were also transfected with a 2 µg per well of dsRNA targeting *eGFP* [4]. Transfections were done as described in Chapter 5.5.2.3. Five additional replicates of PER::LUC DpN1 cells and five replicates of wild-type DpN1 cells were also seeded as control but not transfected. Both wild-type and PER::LUC cells were entrained for two days in 12 light: 12 dark cycles. Wild-type cells and untransfected PER::LUC cells were maintained in 0.025mM luciferin containing media from the day of entrainment while the PER::LUC cells transfected with either dsRNA-*dpcry1* or dsRNA-*eGFP* were administered 0.025mM luciferin media at the end of the 5 hr transfection on day 2.

5.5.4. Generation of a *tdTomato* Knock-in Construct to Tag Clock Genes *in vivo* via CRISPR/Cas9-assisted Knock-in.

5.5.4.1. Designing the sgRNAs.

For a single sgRNA assisted knock-in strategy, two potential knock-in sites were identified near the 3' of the *dpCry2* 5' UTR using CHOPCHOP target

site finder (http://chopchop.cbu.uib.no/index.php target site finder; [30, 31]. Each was subcloned into gRNA expression DR274 plasmid from Addgene [33] by inserting annealed synthetic oligomers at the *BsaI* cleavage site. F JM001/R JM002 were annealed for 5' UTR Target site 1 and F JM003 /R JM004 for 5' UTR Target site 2.

The the two sgRNA insertion strategy, two other guide sites about 1000 bp downstream of the 5' UTR Target 2 site within the exons of *dpCry2* were identified and subcloned into the DR274 expression vector as described above using oligonucleotides F SI460/R SI461 for Exon Target 1 and F SI466/R SI467 for Exon Target 2.

5.5.4.2. Synthesis of Cas9 mRNA and sgRNAs.

The synthesis of the Cas9 mRNA and each sgRNA was performed as described previously in Chapter 2.5.8. [5].

5.5.4.3. Egg Microinjections to Test sgRNAs Targeting the 5' UTR.

Eggs were collected and microinjected to either target either the 5' UTR Target 1 or 5' UTR Target 2 potential knock-in sites, as described in Chapter 2.5.9. [5].

5.5.4.4. Analysis of CRISPR/Cas9-induced Mutations and Identification of a Knock-in site in the 5' UTR.

PCR fragments flanking the targeted regions were amplified from genomic DNA extracted from larval sensors of potential founders at the fifth instar with the following primers F JM005/R JM006. PCR products were purified and subjected to a Cas9-based cleavage assay as described in Chapter 2.5.10. [5].

Based on the relative abundance of uncleaved fragments relative to cleaved ones, a higher degree of targeting in somatic cells and a greater number of larvae exhibiting targeting were observed in the group injected with the sgRNA targeting 5' UTR Target 2. This site was thus chosen as the knock-in site for the 5' UTR.

5.5.4.5. Testing the *dpCry2* Exon Targeting sgRNAs Target 1 and Target 2 *in vitro.*

The sgRNAs targeting in the exon at either Target 1 or Target 2 were tested *in vitro* using Cas9-based cleavage assays on PCR product amplified from genomic DNA with primers F SI462/R SI463. Exon Target 2 exhibited complete cleavage of the PCR product and was chosen as the downstream target site for a two sgRNA knock-in strategy.

5.5.4.6. Generation of the *tdTomato* Knock-in Construct to Tag *dpCry2.*

Initially, *mCherry* was the selected fluorescent reporter. A plasmid containing membrane-localized *mCherry* and accompanying SV40 polyA tail was ordered from AddGene (#53750, pCS-memb-mCherry) and modified with left and right homology arms containing ~1000 bp of complementary DNA upstream and downstream of the 5' UTR Target 2 site for the single sgRNA knock-in strategy, respectively. The LHA was amplified from genomic DNA using primers F SI382 containing a HindIII site in the 5' tail and R SI383 containing a EcoRI site in the 5' tail, and ligated into the pCS-memb-mCherry plasmid using those restriction sites. The RHA was amplified from genomic DNA using primers F SI361, containing a NotI site in the 5' tail and designed to skip the "ATG" of the *dpCry2* start codon, and R SI362 containing a KpnI in the 5' tail, and ligated into the pCS-LHA-memb-mCherry plasmid at those restriction sites, just downstream of the SV40 polyA tail. It was then found that the product of *mCherry* may not be stable in solvents necessary for brain clearing, a technique that would allow to rapidly image tagged neurons using light sheet microscopy, and we decided to use *tdTomato* instead as its product is known to be stable. The restriction sites flanking *mCherry* were replaced with BamHI sites using inverse PCR and the KOD Xtreme Hot Start DNA Polymerase (Novagen 71975, Toyobo) with primers SI491/SI492 to insert a BamHI on the 5' end and SI470/SI471 to insert BamHI on the 3' end. *tdTomato* was amplified from the plasmid pCSCMV:tdTomato

(Addgene, #30530) using primers F SI480/R SI472 containing BamHI sites in the 5’ tails, and subcloned into the knock-in construct at BamHI sites.

For the two sgRNA knock-in strategy, the RHA of the pCS-LHA-memb-tdTomato-RHA was removed with NotI and KpnI restriction sites and replaced with a new RHA containing ~1000 bp of sequence downstream of the Exon Target 2 site. The RHA was amplified with primers F SI474/R SI476 containing NotI and KpnI sites in the 5’ tails, respectively, and ligated in place of the former RHA in the knock-in construct just downstream of the SV40 polyA tail.

5.5.5. Generating the Piggybac Construct for the Transposition of *mCD8::tdTomato* Under the Control of the *dpCry2* Promoter.

tdTomato was amplified from the pCSCMV:tdTomato plasmid using the KOD Xtreme Hot Start DNA Polymerase and primers F SI500/R SI499 containing BmtI and HpaI sites in the 5’ tails, respectively. This amplicon was digested and ligated into a plasmid containing a donor sequence flanked by inverted terminal repeats for piggybac transposase recognition, pim148-hr5-ie1p-mcherry-p10-3utr-mcs-pbac (Joseph Parker, CalTech), to replace *mCherry* and the *p10* 3’ UTR region. The *dpCry2* promoter was amplified from genomic DNA using primers F SI501/R SI502 containing SpeI and BmtI 5’ tails, respectively. The amplicon was digested and ligated into the newly made pim148-hr5-ie1p-tdTomato-mcs-pbac plasmid to replace the *hr5-ie1p* promoter sequence. To target tdTOMATO to the membrane of monarch *dpCry2* positive

cells, a gene encoding the mouse lymphocyte marker CD8 (*mCD8*) was amplified from genomic DNA of UAS-mCD8-tagged GFP *Drosophila* (Paul Hardin, Texas A&M University) using primers F SI540/R SI541 containing BmtI tails on each of their 5' tails. The amplicon was digested and ligated into the pim148-dpCry2-tdTomato-mcs-pbac plasmid to fuse mCD8 in frame with tdTOMATO.

5.5.6. Synthesis of *Piggybac* Transposase mRNA.

The *piggybac* transposase mRNA was transcribed *in vitro* from a hyperactive-pbac-transposase plasmid (Joseph Parker, CalTech) using a T7 RNA polymerase (Thermo Fischer). The plasmid was linearized using a NotI-HF (NEB) and purified by a NucleoSpin Gel and PCR Clean up kit (Machery-Nagel). 1-3 µg of the linearized plasmid was combined with 10 µL of 5X Transcription Buffer (Thermo Fischer), 10 µL of 0.1 M DTT (Promega), 4 µL 25 mM each mix rNTPs (Promega), 50 µL of 2X T3 NTP/CAP (mMessage mMachine T3 transcription kit (Invitrogen)), 2 µL of the T7 RNA Polymerase, and HPLC water up to 100 uL. The mixture was incubated overnight at 37^{o}C. The mRNA was then treated with an RQ1 Dnase for 15 min at 37^{o}C and purified by acid-phenol extraction.

5.5.7. Generating the Tol2 Construct for the Transposition of *mCD8::tdTomato* Under the Control of the *dpCry2* Promoter.

The *dpCry2-mCD8::tdTomato* sequence was amplified from the transgenesis donor plasmid for piggybac using the primers F SI554/R SI555 containing ApaI and XhoI restriction sites, respectively. The amplicon was digested and ligated into a tol2-hsp-mcs plasmid (Bruce Riley, Texas A&M University) using the selected restriction sites to replace the *hsp* gene.

5.5.8. Synthesis of *Tol2* Transposase mRNA.

The tol2 mRNA was transcribed as described for piggybac but using a tpase-tol2 plasmid (Bruce Riley, Texas A&M University) linearized with NotI-HF and a SP6 RNA polymerase (Promega).

Table 2: List of oligonucleotides used to generate the CRISPR-based knock-in strategy construct and piggybac-mediated and tol2-mediated transgenesis constructs.

JM001 5′-TAGGTGCATGATCACCGATGATTT-3′
JM002 5′-AAACAAATCATCGGTGATCATGCA-3′
JM003 5′-TAGGATAATTGTAGGCTGTGATGA-3′
JM004 5′-AAACTCATCACAGCCTACAATTAT-3′
SI460 5′-TAGGGTTGAGACTAGAACGACATT-3′
SI461 5′-AAACAATGTCGTTCTAGTCTCAAC-3′
SI466 5′-TAGGCACTCCACGGTCAGATACTT-3′
SI467 5′-AAACAAGTATCTGACCGTGGAGTG-3′
JM005 5′-AATGCATAAAATAGTGACAAACCAGC-3′
JM006 5′-ATTTATTAATGCCGACGTTAGACG-3′
SI462 5′-CGATTTGGAGTGCCGACAC-3′
SI463 5′-GCCCATTTCTTGAGAGCCTC-3′
SI382 5′-TGTGGGAAGCTTAGTAGGTGAGTAACATAATCCC-3′

Table 2 Continued

SI383 5′-GTTGGGGAATTCTCACAGCCTACAATTATACTCTC-3′
SI361 5′- GGGTGGGCGGCCGCTGACGGTAACGTAATTTCGGTTGCCGAGACGCTG-3′
SI362 5′-GTGCGCGGTACCCGCGTAAATTCCCTTCGGGGTG-3′
SI491 5′-CGCCACCATGGTGAGCAAGGGCGAG-3′
SI492 5′-GGATCCGCCTCGTCGTCGTTCAGGTTGTCCTTGCG-3′
SI470 5′-GGATCCGATCCGGACATGATAAGATACATTGATGAGTTTGGAC-3′
SI471 5′-TTACTTGTACAGCTCGTCCATGCCGC-3′
SI480 5′-TTTGCAGGATCCCGCCACCATG-3′
SI472 5′-TATTAAGGATCCTTACTTGTACAGCTCGTCCATGCCG-3′
SI474 5′-TTTTTTGCGGCCGCCTTTGGAGGGAATTTTTCTATTGC-3′
SI476 5′-TTTTTTGGTACCCCAGCGCTGCTTAGAGAAAG-3′
SI500 5′-TTTAAAGCTAGCATGGTGAGCAAGGGCGAG-3′
SI499 5′-GCAATTGTTGTTGTTAACTTGTTTATTGCAGC-3′
SI501 5′- TTTAAAACTAGTGGTGACTTTTGTAAGACGTGAAG-3′
SI502 5′- TTTAAAGCTAGCAATTACGTTACCGTCATCACAG-3′
SI540 5′- TTTAAAGCTAGCATGGCCTCACCGTTGACC-3′
SI541 5′- TTTAAAGCTAGCGCGGCTGTGGTAGCAGATG-3′
SI554 5′- TATATAGGGCCCGGTGACTTTTGTAAGACGTGAAGT-3′
SI555 5′- TATATACTCGAGTTACTTGTACAGCTCGTCCATGC-3′

5.6. References.

1. Zhan, S., et al., *The monarch butterfly genome yields insights into long-distance migration.* Cell, 2011. **147**(5): p. 1171-85.
2. Palomares, L.A., et al., *Novel insect cell line capable of complex N-glycosylation and sialylation of recombinant proteins.* Biotechnol Prog, 2003. **19**(1): p. 185-92.
3. Markert, M.J., et al., *Genomic Access to Monarch Migration Using TALEN and CRISPR/Cas9-Mediated Targeted Mutagenesis.* G3 (Bethesda), 2016. **6**(4): p. 905-15.

4. Zhang, Y., et al., *Vertebrate-like CRYPTOCHROME 2 from monarch regulates circadian transcription via independent repression of CLOCK and BMAL1 activity.* Proc Natl Acad Sci U S A, 2017. **114**(36): p. E7516-E7525.
5. Iiams, S.E., et al., *Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly.* Proc Natl Acad Sci U S A, 2019.
6. Yoo, S.H., et al., *PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues.* Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5339-46.
7. Deluca, M., *Firefly luciferase.* Adv Enzymol Relat Areas Mol Biol, 1976. **44**: p. 37-68.
8. Gelmini, S., P. Pinzani, and M. Pazzagli, *Luciferase gene as reporter: comparison with the CAT gene and use in transfection and microinjection of mammalian cells.* Methods Enzymol, 2000. **305**: p. 557-76.
9. Ueda, H.R., et al., *A transcription factor response element for gene expression during circadian night.* Nature, 2002. **418**(6897): p. 534-9.
10. Baba, K., et al., *Circadian regulation of the PERIOD 2::LUCIFERASE bioluminescence rhythm in the mouse retinal pigment epithelium-choroid.* Mol Vis, 2010. **16**: p. 2605-11.
11. van der Veen, D.R., et al., *Cardiac atrial circadian rhythms in PERIOD2::LUCIFERASE and per1:luc mice: amplitude and phase*

responses to glucocorticoid signaling and medium treatment. PLoS One, 2012. **7**(10): p. e47692.

12. Baba, K., A.J. Davidson, and G. Tosini, *Melatonin Entrains PER2::LUC Bioluminescence Circadian Rhythm in the Mouse Cornea.* Invest Ophthalmol Vis Sci, 2015. **56**(8): p. 4753-8.
13. Moore, S.R., et al., *Robust circadian rhythms in organoid cultures from PERIOD2::LUCIFERASE mouse small intestine.* Dis Model Mech, 2014. **7**(9): p. 1123-30.
14. Darlington, T.K., et al., *Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim.* Science, 1998. **280**(5369): p. 1599-603.
15. Saez, L. and M.W. Young, *Regulation of nuclear entry of the Drosophila clock proteins period and timeless.* Neuron, 1996. **17**(5): p. 911-20.
16. Zhu, H.S., et al., *Cryptochromes define a novel circadian clock mechanism in monarch butterflies that may underlie sun compass navigation.* Plos Biology, 2008. **6**(1): p. 138-155.
17. Bjorkoy, G., et al., *p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death.* J Cell Biol, 2005. **171**(4): p. 603-14.
18. Muzumdar, M.D., et al., *A global double-fluorescent Cre reporter mouse.* Genesis, 2007. **45**(9): p. 593-605.

19. Drobizhev, M., et al., *Two-photon absorption properties of fluorescent proteins.* Nat Methods, 2011. **8**(5): p. 393-9.

20. Shaner, N.C., et al., *Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein.* Nat Biotechnol, 2004. **22**(12): p. 1567-72.

21. Yang, H., et al., *Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks.* Int J Mol Sci, 2020. **21**(18).

22. Andersen, S.L. and J. Sekelsky, *Meiotic versus mitotic recombination: two different routes for double-strand break repair: the different functions of meiotic versus mitotic DSB repair are reflected in different pathway usage and different outcomes.* Bioessays, 2010. **32**(12): p. 1058-66.

23. Zhao, S., et al., *PiggyBac transposon vectors: the tools of the human gene encoding.* Transl Lung Cancer Res, 2016. **5**(1): p. 120-5.

24. Gregory, M., et al., *Insect transformation with piggyBac: getting the number of injections just right.* Insect Mol Biol, 2016. **25**(3): p. 259-71.

25. Urasaki, A., et al., *Transposition of the vertebrate Tol2 transposable element in Drosophila melanogaster.* Gene, 2008. **425**(1-2): p. 64-8.

26. Zhu, L., et al., *CRISPR/Cas9-mediated knockout of factors in non-homologous end joining pathway enhances gene targeting in silkworm cells.* Sci Rep, 2015. **5**: p. 18103.

27. Burr, A.A., et al., *Using membrane-targeted green fluorescent protein to monitor neurotoxic protein-dependent degeneration of Drosophila eyes.* J Neurosci Res, 2014. **92**(9): p. 1100-9.

28. Mocellin, S. and M. Provenzano, *RNA interference: learning gene knock-down from cell physiology.* J Transl Med, 2004. **2**(1): p. 39.

29. Terenius, O., et al., *RNA interference in Lepidoptera: an overview of successful and unsuccessful studies and implications for experimental design.* J Insect Physiol, 2011. **57**(2): p. 231-45.

30. Labun, K., et al., *CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering.* Nucleic Acids Research, 2016. **44**(W1): p. W272-W276.

31. Montague, T.G., et al., *CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing.* Nucleic Acids Research, 2014. **42**(W1): p. W401-W407.

32. Markert, M.J., et al., *Genomic Access to Monarch Migration Using TALEN and CRISPR/Cas9-Mediated Targeted Mutagenesis.* G3-Genes Genomes Genetics, 2016. **6**(4): p. 905-915.

33. Hwang, W.Y., et al., *Efficient genome editing in zebrafish using a CRISPR-Cas system.* Nature Biotechnology, 2013. **31**(3): p. 227-229.

34. Johansen, H., et al., *Regulated expression at high copy number allows production of a growth-inhibitory oncogene product in Drosophila Schneider cells.* Genes Dev, 1989. **3**(6): p. 882-9.

6. GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

6.1. Monarch Photoperiodic Responses are Regulated via the Clock-Controlled Vitamin A Pathway.

The idea that circadian clock genes and/or the circadian clock play a role in regulating insect photoperiodic responses is not an entirely new concept, as it had previously been demonstrated that RNA interference (RNAi) knockdown of clock gene activators and repressors, as well ablation of clock containing neurons disrupts photoperiodically-induced diapause responses in insects [1-8]. However, whether the clock is functioning as a unit to regulate these responses or whether it is though the action of individual clock genes acting outside of their core clock function has remained an open and debated question. In Chapter 2, using mutants of the monarch circadian activators, *Clock* (*dpClk*) and *Bmal1* (*dpCyc-like*), and the repressor, *Cryptochrome 2* (*dpCry2*), we demonstrated that disruption of the negative or positive elements of the clock locked reproductive responses into either a diapausing-like or non-diapausing-like response, regardless of photoperiodic conditions [9]. These results have provided evidence in another insect model than the bean bug that circadian clock genes are essential for the regulation of photoperiodic responses [9]. As disruption of the monarch circadian activators or the repressor resulted in reproductive responses inverse to those observed in the bean bug studies, this also more strongly supports the role that the circadian clock functions as a unit

to regulate photoperiodic responses. Additionally, in Chapter 4 I also found abolished photoperiodic responses in the newly generated clock gene knockout of the monarch blue-light photoreceptor *dpCry1*. While *dpCry1* knockouts have been shown to eliminate circadian entrainment, and thus render the clock non-functional, the possibility cannot be excluded that dpCRY1 may instead be functioning to sense the photoperiod to cause the loss of reproductive responses.

In addition to the negative feedback regulation of the core circadian loop, the clock is also responsible for the rhythmic expression of many clock-controlled genes which control many behavioral and physiological inductive pathways [9, 10], and could regulate photoperiodic responses through the expression of some of these genes. Based on the RNA-Seq data in Chapter 2, several clock-controlled genes belonging to the vitamin A pathway exhibited a seasonally-dependent expression pattern as would be expected for photoperiodic candidate genes [9]. This pathway was of interest to test functionally as it has been shown over the past few decades that carotenoids and retinoids of the vitamin A pathway are capable of restoring photoperiodic responses in insects [11-13]. Generating a monarch knockout of the rate-limiting enzyme, *neither inactivation nor afterpotential B* (*ninaB1*), and testing reproductive diapause responses confirmed the role of the vitamin A pathway in regulating photoperiodic responses [9]. In Chapter 3, knocking out another seasonally expressed gene, *santa maria 1*, which encodes a transmembrane

transporter acting in the same pathway, did not yield any conclusive results. However, the recently generated *retinol dehydrogenase 13* (*rdh13*) knockout is expected to yield exciting data as to which one of the pathway's products, retinol or retinoic acid (RA), is inducing photoperiodic responses. If the knockouts are still capable of responding to the change in photoperiod then this may suggest that RA activation of the retinoic acid receptor (RAR) and RAR's dimerization to retinoid X receptor (RXR) is controlling seasonal gene expression. While it is unknown whether a RAR homolog exists in the monarch, developing a chromatin immunoprecipitation (ChIP) grade antibody targeting monarch RXR (DPOGS214735) or ultraspiracle (DPOGS214137 and DPOGS214138) [14], a *Drosophila* identified homolog of RXR [15], for use in ChIP-Sequencing would allow us to identify the target genes of the RA-activated transcription factor complex. Another possible explanation if the knockouts are still responsive to the photoperiod, is that other functionally redundant *rdh* genes are compensating to produce retinol so there is no disruption to the vitamin A pathway. Determining which of the two reasons, regulation of seasonal gene expression by RA or lack of perturbance to the pathway due to other compensatory *rdh* genes, will be difficult to disentangle. However, if the knockout renders the mutants incapable of responding to the photoperiod, then this would suggest that retinol is the key molecule via a seasonal production of photoperiodic sensing opsins. To identify which photoreceptor(s) may be involved, the five candidate opsins identified in Chapter 3 should each be knocked out and tested for the effect on photoperiodic

responses. Long wavelength opsin appears as a possible candidate given the high expression level observed in monarch brains and its co-expression with the hormone melatonin, known to be involved in inducing seasonal responses [16], in a subset of neurons in hawkmoths [17]. Parapinopsin is another particularly attractive candidate, as its orthologues are expressed in the pineal gland of fish, lampreys, and reptiles [16, 18-20] a region of the brain known to relay photoperiodic information via the secretion of melatonin. While the possibility of dpCRY1 serving as the photoperiodic sensor cannot be completely excluded, if the *rdh13* mutant results indicate retinol as the key signaling molecule this would strongly suggest that the photoperiodic sensor utilizes a retinal chromophore, and argue against the flavin-based dpCRY1 as the photoperiodic photoreceptor [21-25].

Together, work from Chapters 2 and 3 demonstrates that the monarch butterfly, with robust seasonal responses and applicable set of cutting-edge genetic tools, is a powerful system to study the genetics underlying photoperiodic responses. Given the discovery of the vitamin A pathway's role in regulating photoperiodic responses and its overall conservation in flies and mammals to produce vitamin A and RA products, the results discussed in this dissertation and revelations from continued exploration of this pathway are likely to pave the way for understanding at a molecular level how photoperiod is sensed and transduced into an appropriate physiological or behavioral response in a range of other seasonally responsive species.

6.2. Circadian Entrainment Relies Solely on the dpCRY1 Photoreceptor in the Monarch.

Two decades of research on the *Drosophila* circadian clock have revealed that while the blue-light photoreceptor dmCRY functions in clock entrainment, it acts in concert with a network of rhodopsins and signal transduction pathways to entrain circadian rhythms *in vivo* [26-32]. In Chapter 4 I sought to explore the role of dpCRY1 in the entrainment of the monarch circadian clock, as no studies have been conducted outside of *Drosophila*. I uncovered what appears to be a simpler mechanism of circadian entrainment. Unlike in *Drosophila*, in which removal of dmCRY and all other photoreceptors is required to abolish behavioral and molecular rhythms, a single knockout of *dpCry1* abolished rhythmic pupal eclosion and rhythms of *period* (*dpPer*) and *timeless* (*dpTim*) in the adult monarch brain [28]. Additionally, data from Chapter 5 showed dramatically reduced amplitudes of LUCIFERASE (LUC) rhythm in PERIOD::LUCIFERASE (PER::LUC) DpN1 cells in which *dpCry1* was knocked down, demonstrating that the effect is similar even in cell culture. The arrhythmicity observed in the whole organism was in fact not due to a defective core clock as both the molecular clock and eclosion rhythms of *dpCry1* knockouts could be entrained by temperature cycles. So far, we cannot exclude the possibility that the lack of eclosion rhythms in *dpCry1* knockouts results from not yet developed dpCRY1-independent entrainment pathways at the pupal stage. However, the fact that pupal eclosion is still rhythmic in *Drosophila* cry^0

knockouts and that several of the rhodopsins used in dmCRY-independent entrainment pathways are already expressed at the larval stage [33, 34] argues against this possibility. Future experiments focused on developing a rhythmic behavior assay in the adult monarch and testing the effect of *dpCry1* knockout on adult locomotor rhythms will address this question. Additionally, eclosion timing of wild-type monarchs under different wavelengths of light within and outside of dpCRY1's functional UV-A/blue light range of ~380 - 420 nm could also be tested. If monarch pupae were only able to eclose rhythmically under UV-A/blue light then this would eliminate the participation of any dpCRY1-independent opsins functioning outside of this light range.

Altogether, results of Chapter 4 suggest the potential role of dpCRY1 as the sole photoreceptor necessary for entrainment of circadian rhythms in the monarch butterfly. This apparent, striking difference in clock entrainment mechanisms between the *Drosophila* clock and the monarch clock *in vivo* highlights the importance of expanding clockwork studies to more model insect species in the field to understand the evolution of the circadian clock.

6.3. Developing New Genetic Tools will Enhance the Status of the Monarch Butterfly as a *Bona Fide* Model Organism.

Chapter 5 focused on the development of new genetic tools in the monarch butterfly in order to elevate its status as a powerful model organism and provide more insect-based reagents that are likely to be of broad interest to

the field. Specifically, I generated a rhythmic PER::LUC fusion protein reporter in monarch DpN1 cells to be able to test and study clockwork dynamics *in vitro* in an insect model, which has been lacking due to the fact that the key insect cell line Drosophila S2 is clock-less [35-39]. The dramatic disruption of LUC rhythms observed when knocking down *dpCry1* via dsRNA provided a proof of concept for the usability of this reporter line to study clockwork mechanisms. One important drawback is that the rhythms exhibited in this cell line under LD cycles do not persist in constant conditions [35]. Future work will need to develop a method for maintaining rhythms under constant conditions to be able to assess the effects on circadian period. Some possibilities that should be tested include entrainment via temperature under constant dark (DD) conditions and release into constant temperature, forskolin administration which is known to chemically reset rhythms by elevating levels of cyclic adenosine monophosphate (cAMP), or serum shock which has been shown to be capable of inducing circadian expression of various genes in mammalian cell cultures, including core clock genes [40, 41]. Once a method is developed, this line can be used to test the effect of modulators of the circadian clock on rhythmic phase, amplitude, and period of the PER::LUC reporter.

In Chapter 5 I also worked on developing several strategies and reagents for tagging clock genes and visualize the clock neuronal network *in vivo* with the fluorescent reporter *tdTomato*. Two constructs for CRISPR-assisted knock-in have been generated to either work with a single sgRNA to induce a single

double stranded DNA break (DSB) at the knock-in site or with two sgRNAs to cut the DNA in two spots and remove a region of similar size to the construct to be knocked-in. The efficiency of both of these methods should be tested and compared by injecting several hundred eggs with the knock-in construct, Cas9 mRNA, corresponding sgRNA(s) and screened for inherited germline knock-in events in the resulting progeny to establish the reporter line. Reagents were also developed for alternative transgenesis strategies. The piggybac or tol2 transposon, each carrying *tdTomato* under the control of the *dpCry2* promoter, along with the corresponding *piggybac* or *tol2* transposase mRNA should be injected into monarch eggs and the progeny of the surviving larvae screened for transposition events. Because injected plasmids can be maintained in the progeny of the injected larvae (SE Iiams and ZN Adelman, personal communication), the best way to screen for genomic transposition events in surviving injected larvae or their progeny would be through southern blots using a DNA probe. Three probes can be designed to target *tdTomato*, *mCD8*, or the fusion region of *mCD8::tdTomato* and tested for their hybridization efficiency on the transposon plasmid. Once a probe is selected, we will next need to choose an enzyme that cleaves the genome at many places but only once in the transposon plasmid outside of the probe itself as it would allow us to discriminate based on size between hybridization of the probe to the plasmid itself and hybridization of the probe to genomic regions of size that should be in theory different from that of the plasmid. When a candidate(s) with genomic

integration in the germline is identified, it can be backcrossed to found the reporter line.

Whichever of the two strategies, CRISPR-assisted knock-in or transgenesis, is more efficient to tag *dpCry2* with *tdTomato* can then be used to tag the activator *dpClk*. Overlapping the expression of *tdTomato* tagged to either the repressor or activator would allow the identification of clock expressing cells and the mapping of the clock circuitry of the monarch brain. Additionally, these reporters could also be tested for changes in the clock brain circuitry between long and short photoperiods. If changes in the clock neuronal projections are seen between photoperiods, it may suggest that the clock is rewired on a seasonal basis. This may be particularly relevant to understand whether the vitamin A pathway involved in regulating photoperiodic responses, as shown in Chapter 2 and 3, do so by seasonally regulating the neuronal circuitry in the brain as shown in mammals [42].

Generation of the PER::LUC reporter and the future generation of the *in vivo* reporters will significantly expand the potential of the monarch as a model organism and increase our ability to dissect clockwork mechanisms in a seasonally responsive species.

6.4. References.

1. Ikeno, T., H. Numata, and S.G. Goto, *Photoperiodic response requires mammalian-type cryptochrome in the bean bug Riptortus pedestris.* Biochem Biophys Res Commun, 2011. **410**(3): p. 394-7.
2. Ikeno, T., H. Numata, and S.G. Goto, *Circadian clock genes period and cycle regulate photoperiodic diapause in the bean bug Riptortus pedestris males.* J Insect Physiol, 2011. **57**(7): p. 935-8.
3. Ikeno, T., et al., *Involvement of the brain region containing pigment-dispersing factor-immunoreactive neurons in the photoperiodic response of the bean bug, Riptortus pedestris.* J Exp Biol, 2014. **217**(Pt 3): p. 453-62.
4. Ikeno, T., et al., *Photoperiodic diapause under the control of circadian clock genes in an insect.* BMC Biol, 2010. **8**: p. 116.
5. Shiga, S. and H. Numata, *Roles of PER immunoreactive neurons in circadian rhythms and photoperiodism in the blow fly, Protophormia terraenovae.* J Exp Biol, 2009. **212**(Pt 6): p. 867-77.
6. Ikegami, K. and T. Yoshimura, *Seasonal time measurement during reproduction.* J Reprod Dev, 2013. **59**(4): p. 327-33.
7. Saunders, D.S., *Insect photoperiodism: measuring the night.* J Insect Physiol, 2013. **59**(1): p. 1-10.

8. Saunders, D.S. and R.C. Bertossa, *Deciphering time measurement: the role of circadian 'clock' genes and formal experimentation in insect photoperiodism.* J Insect Physiol, 2011. **57**(5): p. 557-66.
9. Iiams, S.E., et al., *Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly.* Proc Natl Acad Sci U S A, 2019.
10. Ueda, H.R., et al., *Genome-wide transcriptional orchestration of circadian rhythms in Drosophila.* Journal of Biological Chemistry, 2002. **277**(16): p. 14048-14052.
11. Claret, J. and N. Volkoff, *Vitamin A is essential for two processes involved in the photoperiodic reaction in Pieris brassicae.* Journal of Insect Physiology, 1992. **38**: p. 569-574.
12. Veerman, A., et al., *Vitamin-a Is Essential for Photoperiodic Induction of Diapause in an Eyeless Mite.* Nature, 1983. **302**(5905): p. 248-249.
13. Veerman, A., et al., *Photoperiodic Induction of Diapause in an Insect Is Vitamin-a Dependent.* Experientia, 1985. **41**(9): p. 1194-1195.
14. Zhan, S. and S.M. Reppert, *MonarchBase: the monarch butterfly genome database.* Nucleic Acids Res, 2013. **41**(Database issue): p. D758-63.
15. Oro, A.E., M. McKeown, and R.M. Evans, *Relationship between the product of the Drosophila ultraspiracle locus and the vertebrate retinoid X receptor.* Nature, 1990. **347**(6290): p. 298-301.

16. Bolborea, M., et al., *Melatonin controls photoperiodic changes in tanycyte vimentin and neural cell adhesion molecule expression in the Djungarian hamster (Phodopus sungorus).* Endocrinology, 2011. **152**(10): p. 3871-83.

17. Lampel, J., A.D. Briscoe, and L.T. Wasserthal, *Expression of UV-, blue-, long-wavelength-sensitive opsins and melatonin in extraretinal photoreceptors of the optic lobes of hawk moths.* Cell Tissue Res, 2005. **321**(3): p. 443-58.

18. Blackshaw, S. and S.H. Snyder, *Parapinopsin, a novel catfish opsin localized to the parapineal organ, defines a new gene family.* J Neurosci, 1997. **17**(21): p. 8083-92.

19. Koyanagi, M., et al., *Bistable UV pigment in the lamprey pineal.* Proc Natl Acad Sci U S A, 2004. **101**(17): p. 6687-91.

20. Wada, S., et al., *Expression of UV-sensitive parapinopsin in the iguana parietal eyes and its implication in UV-sensitivity in vertebrate pineal-related organs.* PLoS One, 2012. **7**(6): p. e39003.

21. Palczewski, K., et al., *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science, 2000. **289**(5480): p. 739-45.

22. Wald, G., *The molecular basis of visual excitation.* Nature, 1968. **219**(5156): p. 800-7.

23. Saunders, D.S., *Insect Clocks*. 2002: Elsevier. 576.

24. Terakita, A., *The opsins.* Genome Biol, 2005. **6**(3): p. 213.

25. Lin, C. and T. Todo, *The cryptochromes.* Genome Biol, 2005. **6**(5): p. 220.

26. Emery, P., et al., *CRY, a Drosophila clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity.* Cell, 1998. **95**(5): p. 669-79.

27. Stanewsky, R., et al., *The cryb mutation identifies cryptochrome as a circadian photoreceptor in Drosophila.* Cell, 1998. **95**(5): p. 681-92.

28. Helfrich-Forster, C., et al., *The circadian clock of fruit flies is blind after elimination of all known photoreceptors.* Neuron, 2001. **30**(1): p. 249-61.

29. Ni, J.D., et al., *A rhodopsin in the brain functions in circadian photoentrainment in Drosophila.* Nature, 2017. **545**(7654): p. 340-344.

30. Ogueta, M., R.C. Hardie, and R. Stanewsky, *Non-canonical Phototransduction Mediates Synchronization of the Drosophila melanogaster Circadian Clock and Retinal Light Responses.* Curr Biol, 2018. **28**(11): p. 1725-1735 e3.

31. Senthilan, P.R., et al., *Role of Rhodopsins as Circadian Photoreceptors in the Drosophila melanogaster.* Biology (Basel), 2019. **8**(1).

32. Szular, J., et al., *Rhodopsin 5- and Rhodopsin 6-mediated clock synchronization in Drosophila melanogaster is independent of retinal phospholipase C-beta signaling.* J Biol Rhythms, 2012. **27**(1): p. 25-36.

33. Dolezelova, E., D. Dolezel, and J.C. Hall, *Rhythm defects caused by newly engineered null mutations in Drosophila's cryptochrome gene.* Genetics, 2007. **177**(1): p. 329-45.

34. Malpel, S., A. Klarsfeld, and F. Rouyer, *Larval optic nerve and adult extra-retinal photoreceptors sequentially associate with clock neurons during Drosophila brain development.* Development, 2002. **129**(6): p. 1443-53.

35. Zhu, H.S., et al., *Cryptochromes define a novel circadian clock mechanism in monarch butterflies that may underlie sun compass navigation.* Plos Biology, 2008. **6**(1): p. 138-155.

36. Yoo, S.H., et al., *PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues.* Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5339-46.

37. Schneider, I., *Cell lines derived from late embryonic stages of Drosophila melanogaster.* J Embryol Exp Morphol, 1972. **27**(2): p. 353-65.

38. Darlington, T.K., et al., *Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim.* Science, 1998. **280**(5369): p. 1599-603.

39. Saez, L. and M.W. Young, *Regulation of nuclear entry of the Drosophila clock proteins period and timeless.* Neuron, 1996. **17**(5): p. 911-20.

40. Yagita, K. and H. Okamura, *Forskolin induces circadian gene expression of rPer1, rPer2 and dbp in mammalian rat-1 fibroblasts.* FEBS Lett, 2000. **465**(1): p. 79-82.

41. Balsalobre, A., F. Damiola, and U. Schibler, *A serum shock induces circadian gene expression in mammalian tissue culture cells.* Cell, 1998. **93**(6): p. 929-37.

42. Wood, S.H., et al., *Binary Switching of Calendar Cells in the Pituitary Defines the Phase of the Circannual Cycle in Mammals.* Curr Biol, 2015. **25**(20): p. 2651-62.

APPENDIX A - TABLES

Table 3: List of genes with similar or differential temporal expression patterns in the brain of monarchs raised in LP and SP. R-R: Genes rhythmic in both conditions; R-AR: Genes rhythmic in LP monarchs and arrhythmic in SP monarchs; AR-R: Genes arrhythmic in LP monarchs and rhythmic in SP monarchs. The annotation is based on *Drosophila* (no shading) or mouse (colored shading) orthologues. NA, monarch genes without orthologues in *Drosophila* or the mouse.

Gene ID	R-R
DPOGS211878	Cyclin D
DPOGS200190	similar
DPOGS214070	stumps
DPOGS205548	Ubiquitin conjugating enzyme 7
DPOGS207780	multiple edematous wings
DPOGS209874	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS207730	CG10082
DPOGS205355	Ucp4A
DPOGS213007	Cytochrome P450-18a1
DPOGS207942	synaptotagmin XI
DPOGS206908	Glucose dehydrogenase
DPOGS205549	G protein-coupled receptor 158
DPOGS205086	Multidrug-Resistance like Protein 1
DPOGS209797	Vacuolar H[+] ATPase 100kD subunit 2
DPOGS215479	monooxygenase, DBH-like 1
DPOGS210745	transmembrane protein 135
DPOGS201345	tocopherol (alpha) transfer protein-like
DPOGS215439	lethal (2) essential for life
DPOGS211121	ubiquitin-conjugating enzyme E2Q family-like 1
DPOGS204644	phospholipid phosphatase 1
DPOGS209511	solute carrier family 17 (anion/sugar transporter), member 5
DPOGS215865	THUMP domain containing 1
DPOGS213900	Heat shock protein 68
DPOGS208883	Glycerol 3 phosphate dehydrogenase
DPOGS215738	eiger
DPOGS213901	Heat shock protein 68
DPOGS215489	6-phosphofructo-2-kinase
DPOGS205929	sprite

Table 3 Continued

DPOGS202078	Cyp9f2
DPOGS204253	calsyntenin 2
DPOGS202178	Glutamate oxaloacetate transaminase 1
DPOGS207058	choline dehydrogenase
DPOGS215494	1,4-Alpha-Glucan Branching Enzyme
DPOGS210128	PAPS synthetase
DPOGS201488	Mid1 interacting protein 1 (gastrulation specific G12-like (zebrafish))
DPOGS213925	Heat shock protein 68
DPOGS204642	GTPase regulator associated with FAK
DPOGS205928	sprite
DPOGS213114	Choline transporter-like 2
DPOGS205386	phosphoglucose mutase
DPOGS214409	solute carrier family 22 (organic cation transporter), member 1
DPOGS201746	Carbonic anhydrase 1
DPOGS202237	DnaJ heat shock protein family (Hsp40) member B4
DPOGS200817	alkaline phosphatase, liver/bone/kidney
DPOGS214695	meteorin, glial cell differentiation regulator-like
DPOGS200691	Enhancer of split mbeta, helix-loop-helix
DPOGS213804	Inositol 1,4,5-triphosphate kinase 1
DPOGS213896	Mitf
DPOGS200684	Enhancer of split mbeta, helix-loop-helix
DPOGS200409	Integrin betanu subunit
DPOGS203128	argos
DPOGS210886	rau
DPOGS208606	vrille
DPOGS201013	high mobility group box transcription factor 1
DPOGS205643	solute carrier family 36 (proton/amino acid symporter), member 4
DPOGS215160	Trehalose transporter 1-2
DPOGS214408	solute carrier family 22 (organic cation transporter), member 21
DPOGS211073	cysteine-rich secretory protein 2
DPOGS207391	clavesin 2
DPOGS204091	Salt-inducible kinase 3
DPOGS215159	Trehalose transporter 1-1
DPOGS215378	chondroitin sulfate synthase 1

Table 3 Continued

DPOGS205105	ATP binding cassette subfamily G member 4
DPOGS201012	Leucine-rich pentatricopeptide repeat containing 2
DPOGS214162	Na[+]/H[+] hydrogen antiporter 1
DPOGS211050	Insulin-like receptor
DPOGS209785	Aminolevulinate synthase
DPOGS207504	Farnesyl pyrophosphate synthase
DPOGS214179	timeless
DPOGS209175	phosphatidylinositol-specific phospholipase C, X domain containing 3
DPOGS203974	Calcium-dependent secretion activator
DPOGS209926	hattifattener
DPOGS214215	Na[+]/H[+] hydrogen antiporter 1
DPOGS205855	solute carrier family 46, member 3
DPOGS200290	penguin
DPOGS212590	neither inactivation nor afterpotential B
DPOGS207444	phospholipase A2, group III
DPOGS201101	solute carrier family 2 (facilitated glucose transporter), member 6
DPOGS209146	solute carrier family 22 (organic cation transporter), member 3
DPOGS209925	clockwork orange
DPOGS211474	pastrel
DPOGS204463	Matrix metalloproteinase 1
DPOGS213552	Ecdysone-induced protein 28/29kD
DPOGS216102	lethal (2) k09913
DPOGS201925	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS215419	dumpy
DPOGS203908	period
DPOGS208959	Vacuolar H[+] ATPase 100kD subunit 2
DPOGS203797	Heat shock factor
DPOGS209585	FCH domain only 2
DPOGS212685	sarcoplasmic calcium-binding protein
DPOGS208406	CG30069
DPOGS204552	solute carrier family 2, (facilitated glucose transporter), member 8
Gene ID	R-AR
DPOGS215700	primo-2

Table 3 Continued

DPOGS207634	tribbles
DPOGS214920	Multicopper oxidase-1
DPOGS201668	short gastrulation
DPOGS205889	Rab40
DPOGS202087	Cytochrome P450-9b2
DPOGS215481	no mechanoreceptor potential C
DPOGS213661	stearoyl-Coenzyme A desaturase 1
DPOGS208470	CG13272
DPOGS215992	PDGF- and VEGF-receptor related
DPOGS210839	solute carrier family 16 (monocarboxylic acid transporters), member 14
DPOGS209237	Slowpoke binding protein
DPOGS206903	SIK family kinase 3
DPOGS206136	LON peptidase N-terminal domain and ring finger 2
DPOGS207671	branchless
DPOGS209909	carboxylesterase 2F
DPOGS214094	myospheroid
DPOGS202126	very low density lipoprotein receptor
DPOGS208999	retinol dehydrogenase 13 (all-trans and 9-cis)
DPOGS213594	Glycogen binding subunit 76A
DPOGS203211	Cuticular protein 97Ea
DPOGS201245	CAP
DPOGS215826	solute carrier family 46, member 3
DPOGS203379	periaxin
DPOGS213639	xin actin-binding repeat containing 2
DPOGS200414	twist
DPOGS201544	Organic anion transporting polypeptide 74D
DPOGS203582	CAP
DPOGS214692	phosphoribosyl pyrophosphate synthetase-associated protein 2
DPOGS207676	lethal (3) 73Ah
DPOGS200195	transmembrane protein 135
DPOGS212327	F-box protein 32
DPOGS201486	windpipe
DPOGS201894	Maltase A4
DPOGS212552	Cyp9f2
DPOGS205180	lipase, member H

Table 3 Continued

DPOGS210627	carbonic anhydrase 3
DPOGS200316	globin 1
DPOGS214402	Glycogen synthase
DPOGS205764	U2A
DPOGS206008	microtubule associated monooxygenase, calponin and LIM domain containing 3
DPOGS214921	Multicopper oxidase-1
DPOGS210848	starvin
DPOGS204087	ribbon
DPOGS210850	leucine rich repeat containing G protein coupled receptor 5
DPOGS207651	UDP glucuronosyltransferase 2 family, polypeptide B1
DPOGS212108	straw
DPOGS201241	CAP
DPOGS204250	phenazine biosynthesis-like protein domain containing 2
DPOGS207000	Myosin heavy chain-like
DPOGS212996	Ecdysone-inducible gene E1
DPOGS215099	transmembrane protease, serine 11f
DPOGS208825	Protostome-specific GEF
DPOGS202077	Cyp9f2
DPOGS209025	LIM domain only 7
DPOGS207287	selenium binding protein 1
DPOGS207994	family with sequence similarity 193, member A
DPOGS202356	mutagen-sensitive 312
DPOGS205888	Serine hydroxymethyltransferase 2
DPOGS213632	pallbearer
DPOGS215384	RIO kinase 2
DPOGS208881	DnaJ-like-2
DPOGS200771	stall
DPOGS215969	Transaldolase
DPOGS202029	membrane steroid binding protein
DPOGS205833	amylo-1,6-glucosidase, 4-alpha-glucanotransferase
DPOGS200067	Trehalose transporter 1-2
DPOGS208971	leucine rich melanocyte differentiation associated
DPOGS213899	Heat shock protein 83
DPOGS209508	phosphoglucose mutase
DPOGS206382	lethal (2) essential for life
DPOGS202781	Hsp70/Hsp90 organizing protein

Table 3 Continued

DPOGS203810	Phosphofructokinase
DPOGS210890	Autophagy-related 3
DPOGS214304	spenito
DPOGS215007	Brahma associated protein 170kD
DPOGS215703	firelighter
DPOGS203355	string
DPOGS210704	COP9 signalosome subunit 7B
DPOGS201784	MEP-1
DPOGS201296	Cyclic nucleotide-gated ion channel subunit B
DPOGS207433	dumpy
DPOGS210130	folliculin interacting protein 2
DPOGS201291	Ecdysone-induced protein 93F
DPOGS209745	Reversion-inducing-cysteine-rich protein with kazal motifs
DPOGS204036	crowded by cid
DPOGS200075	Hexokinase A
DPOGS205799	Desaturase 1
DPOGS205647	ATP binding cassette subfamily G member 4
DPOGS202409	dally-like
DPOGS208054	ENL/AF9-related
DPOGS211554	Partner of paired
DPOGS202573	Phosphoenolpyruvate carboxykinase
DPOGS214934	distal antenna
DPOGS210087	unc-5
DPOGS207462	Beadex
DPOGS212184	lysyl oxidase-like 2
DPOGS209609	MLF1-adaptor molecule
DPOGS203228	Cln7
DPOGS203167	solute carrier family 22 (organic cation transporter), member 3
DPOGS207999	MYC-associated zinc finger protein (purine-binding transcription factor)
DPOGS201447	Chitin deacetylase-like 5
DPOGS202150	Chondrocyte-derived ezrin-like domain containing protein
DPOGS202152	Chondrocyte-derived ezrin-like domain containing protein
DPOGS211888	otopetrin 2

Table 3 Continued

DPOGS209166	Bre1
DPOGS208905	kuzbanian
DPOGS201136	Ecdysone-induced protein 74EF
DPOGS202153	Chondrocyte-derived ezrin-like domain containing protein
DPOGS200402	Prosap
DPOGS205510	CG42304
DPOGS208589	Adherens junction protein p120
DPOGS208870	cadherin 15
DPOGS203088	solute carrier family 5 (sodium/glucose cotransporter), member 12
DPOGS204939	fibronectin type III domain containing 5
DPOGS210490	pyroglutamyl-peptidase I
DPOGS203127	Pyruvate kinase
DPOGS214140	Fatty acid (long chain) transport protein
DPOGS206803	Ubiquitin conjugating enzyme 4
DPOGS207157	Formin homology 2 domain containing
DPOGS202825	fuseless
DPOGS205514	happyhour
DPOGS200399	schlank
DPOGS205915	folded gastrulation
DPOGS201347	tocopherol (alpha) transfer protein
DPOGS212492	CG31324
DPOGS215237	Mid1
DPOGS207272	ADP ribosylation factor-like 4
DPOGS211190	minidiscs
DPOGS208314	Mitochondrial pyruvate carrier
DPOGS213044	mediator complex subunit 14
DPOGS202609	Acetylcholine esterase
DPOGS208868	Membrin
DPOGS209828	jaguar
DPOGS214045	transmembrane and tetratricopeptide repeat containing 2
DPOGS200370	discs overgrown
DPOGS214228	Protein kinase, cAMP-dependent, catalytic subunit 3
DPOGS216090	ral guanine nucleotide dissociation stimulator-like 3
DPOGS210086	Prp18
DPOGS205077	neurotrimin

Table 3 Continued

DPOGS209183	sevenless
DPOGS208166	Fimbrin
DPOGS203456	Fasciclin 3
DPOGS208790	proline rich 22
DPOGS203229	solute carrier family 22 (organic cation transporter), member 3
DPOGS204749	solute carrier family 22 (organic cation transporter), member 1
DPOGS214622	BIR repeat containing ubiquitin-conjugating enzyme
DPOGS203436	RIKEN cDNA 3425401B19 gene
DPOGS212948	BCL2/adenovirus E1B interacting protein 3-like
DPOGS208674	GLI pathogenesis-related 2
DPOGS204941	hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase)
DPOGS210475	Adipokinetic hormone receptor
DPOGS201412	CG11370
DPOGS201957	carboxylesterase 2A
DPOGS209128	carboxylesterase 4A
DPOGS202827	scavenger receptor acting in neural tissue and majority of rhodopsin is absent
DPOGS203890	scavenger receptor acting in neural tissue and majority of rhodopsin is absent
DPOGS213662	Pyrokinin 1 receptor
DPOGS214271	Sodium/solute co-transporter-like 5A11
DPOGS213185	missing-in-metastasis
DPOGS201183	Peptidoglycan recognition protein LC
DPOGS210599	Cytochrome b5
DPOGS206439	glutaryl-Coenzyme A dehydrogenase
DPOGS201885	nervana 2
DPOGS201896	NA
DPOGS208079	cryptochrome 2
DPOGS205727	Glycoprotein hormone alpha 2
DPOGS205014	red Malpighian tubules
DPOGS202323	CG12947
DPOGS206847	I'm not dead yet
DPOGS215688	solute carrier family 34 (sodium phosphate), member 1
DPOGS214174	RIKEN cDNA 1110008P14 gene
DPOGS202921	puffyeye

Table 3 Continued

DPOGS213305	Trehalose transporter 1-2
DPOGS203808	arc
DPOGS200805	Chondrocyte-derived ezrin-like domain containing protein
DPOGS208132	CG8854
DPOGS207346	NA
DPOGS208658	Multiple C2 domain and transmembrane region protein
Gene ID	**AR-R**
DPOGS215693	Kinesin-like protein at 31E
DPOGS205193	pita
DPOGS213910	dawdle
DPOGS213990	Valyl-tRNA synthetase, mitochondrial
DPOGS209800	absent, small, or homeotic discs 2
DPOGS204630	Cytochrome b5
DPOGS213168	thickveins
DPOGS209055	breast carcinoma amplified sequence 3
DPOGS201195	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2
DPOGS208537	fritz
DPOGS213902	Unc-89
DPOGS213038	RNA-binding protein 6
DPOGS211622	purine-nucleoside phosphorylase 2
DPOGS213604	family with sequence similarity 205, member A2
DPOGS215784	narrow
DPOGS215002	Tetraspanin 97E
DPOGS208998	ethanolamine phosphate phospholyase
DPOGS205152	Ecdysone-inducible gene L2
DPOGS205265	sulfotransferase family 2A, dehydroepiandrosterone (DHEA)-preferring, member 6
DPOGS208044	sodium potassium chloride cotransporter
DPOGS210295	Phosphoglucose isomerase
DPOGS210772	ubiquitin protein ligase E3 component n-recognin 5
DPOGS215272	SET and MYND domain containing, arthropod-specific, member 4
DPOGS213903	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS200764	wunen

Table 3 Continued

DPOGS200490	Malate dehydrogenase 1
DPOGS211592	Mucin 14A
DPOGS206048	mannan-binding lectin serine peptidase 1
DPOGS209724	transmembrane protease, serine 9
DPOGS211461	family with sequence similarity 43, member A
DPOGS201483	Rrp42
DPOGS213101	Collagen type IV alpha 1
DPOGS203208	Novel nucleolar protein 3
DPOGS200182	Dual oxidase
DPOGS203734	Nckx30C
DPOGS208305	lethal (1) G0020
DPOGS213282	zelda
DPOGS210429	Rab23
DPOGS209220	takeout
DPOGS209293	ACAT-related protein required for viability 1
DPOGS203267	ELF1 homolog, elongation factor 1
DPOGS209941	glutamate receptor, ionotropic, kainate 2 (beta 2)
DPOGS203459	Sphingosine kinase 1
DPOGS215663	DIM1 dimethyladenosine transferase 1-like (S. cerevisiae)
DPOGS213088	Rrp4
DPOGS215556	PSEA-binding protein 95kD
DPOGS210537	ATP synthase mitochondrial F1 complex assembly factor 1
DPOGS208596	geminin
DPOGS211313	JNK1/MAPK8-associated membrane protein
DPOGS207276	karyopherin alpha1
DPOGS207451	Glutaminyl-tRNA synthetase
DPOGS212326	tRNA isopentenyltransferase 1
DPOGS211640	dehydrogenase/reductase (SDR family) member 4
DPOGS202064	mitochondrial ribosomal protein L32
DPOGS201773	Regulatory particle non-ATPase 12
DPOGS203161	merry-go-round
DPOGS211728	jitterbug
DPOGS202638	PWP2 periodic tryptophan protein homolog (yeast)
DPOGS210122	Phospholipase A2 activator protein
DPOGS213171	Mitochondrial trifunctional protein alpha subunit
DPOGS202243	ribosomal RNA processing 15 homolog (S. cerevisiae)
DPOGS210477	Glutathione S transferase D3

Table 3 Continued

DPOGS204283	transmembrane protein 256
DPOGS207617	noisette
DPOGS204316	Isocitrate dehydrogenase
DPOGS213570	methyltransferase like 22
DPOGS201067	Vacuolar protein sorting 33B
DPOGS208233	Phenylalanyl-tRNA synthetase, beta-subunit
DPOGS214668	arsA arsenite transporter, ATP-binding, homolog 1 (bacterial)
DPOGS205156	NAD synthetase
DPOGS203689	ariadne 2
DPOGS203457	Fasciclin 3
DPOGS206624	mitochondrial ribosomal protein S2
DPOGS204467	rhomboid-7
DPOGS209768	Signal recognition particle receptor beta
DPOGS200790	motile sperm domain containing 2
DPOGS205949	GTP-binding protein 10 (putative)
DPOGS208966	SPOUT domain containing methyltransferase 1
DPOGS206283	beta-4-galactosyltransferase 7
DPOGS213207	transmembrane protein 173
DPOGS208468	coiled-coil domain containing 97
DPOGS209268	thioredoxin domain containing 9
DPOGS215730	kurtz
DPOGS212615	SLC22A family member
DPOGS210816	nth (endonuclease III)-like 1 (E.coli)
DPOGS203217	Chaperonin containing TCP1 subunit 8
DPOGS207298	thioredoxin 2
DPOGS214837	dynein cytoplasmic 2 heavy chain 1
DPOGS200566	RIKEN cDNA 0610037L13 gene
DPOGS213209	URB1 ribosome biogenesis 1 homolog (S. cerevisiae)
DPOGS211595	Yip1 domain family, member 7
DPOGS200382	transducin (beta)-like 3
DPOGS214128	acid phosphatase 2, lysosomal
DPOGS201455	mitochondrial ribosomal protein L42
DPOGS206673	DnaJ heat shock protein family (Hsp40) member C17
DPOGS213506	Gdap1
DPOGS213027	Obg-like ATPase 1
DPOGS205508	CCAAT/enhancer binding protein zeta

Table 3 Continued

DPOGS200102	epidermal growth factor receptor pathway substrate 8
DPOGS207615	pitchoune
DPOGS213587	Myb/SANT-like DNA-binding domain containing 3
DPOGS209577	Receptor component protein
DPOGS213219	RIO kinase 3
DPOGS215914	Esa1-associated factor 6
DPOGS210215	Protoporphyrinogen oxidase
DPOGS209063	taxilin alpha
DPOGS207888	smg-9 homolog, nonsense mediated mRNA decay factor (C. elegans)
DPOGS207088	heterogeneous nuclear ribonucleoprotein H3
DPOGS203418	Nucleoporin 107kD
DPOGS215076	syntaxin 18
DPOGS212345	small nuclear ribonucleoprotein E
DPOGS213025	Sp1
DPOGS204880	COP9 signalosome subunit 1b
DPOGS212049	short coiled-coil protein
DPOGS207464	Mediator complex subunit 7
DPOGS212376	telomerase RNA component interacting RNase
DPOGS205969	larsen
DPOGS216155	PSEA-binding protein 45kD
DPOGS211976	coiled-coil domain containing 22
DPOGS213675	Bekka
DPOGS205848	NAD(P)HX epimerase
DPOGS202592	DnaJ heat shock protein family (Hsp40) member B2
DPOGS204859	mitochondrial ribosomal protein S34
DPOGS200836	Ribosomal RNA processing 40
DPOGS214256	Zizimin
DPOGS216147	solute carrier family 2 (facilitated glucose transporter), member 9
DPOGS201420	acyl-Coenzyme A oxidase 1, palmitoyl
DPOGS213908	ATP-binding cassette, sub-family D (ALD), member 2
DPOGS215178	solute carrier organic anion transporter family, member 6c1
DPOGS202770	ubiquitin protein ligase E3 component n-recognin 7 (putative)
DPOGS213674	Chaperonin containing TCP1 subunit 4
DPOGS205058	Ciao1
DPOGS203048	FtsJ RNA methyltransferase homolog 3 (E. coli)

Table 3 Continued

DPOGS205401	Chaperonin containing TCP1 subunit 3
DPOGS212825	transmembrane protein 179
DPOGS216039	dumpy
DPOGS213135	will die slowly
DPOGS212137	pescadillo ribosomal biogenesis factor 1
DPOGS203037	malectin
DPOGS213133	mitochondrial ribosomal protein L30
DPOGS208977	Dynactin 6, p27 subunit
DPOGS202415	Glutamyl-tRNA synthetase, mitochondrial
DPOGS207758	RNA binding motif protein 19
DPOGS212995	Isocitrate dehydrogenase
DPOGS205462	NA
DPOGS203662	Regulator of cullins 2
DPOGS207772	signal sequence receptor, gamma
DPOGS201718	THO complex 2
DPOGS212484	NA
DPOGS209419	zinc finger, MYND-type containing 8
DPOGS204681	Autophagy-related 8a
DPOGS209186	Cadherin 86C
DPOGS207380	leucine-rich repeats and guanylate kinase domain containing
DPOGS210994	Cellular Repressor of E1A-stimulated Genes
DPOGS215977	Dpr-interacting protein eta
DPOGS210673	Eclosion hormone
DPOGS207305	ATP binding cassette subfamily G member 4
DPOGS214673	Hepatocyte nuclear factor 4
DPOGS207246	Rho guanine nucleotide exchange factor at 64C
DPOGS202067	RAS-like, estrogen-regulated, growth-inhibitor
DPOGS212324	tRNA isopentenyltransferase 1
DPOGS210097	Limpet
DPOGS204377	beta-1,4-glucuronyltransferase 1
DPOGS215770	SH2 ankyrin repeat kinase
DPOGS212331	lethal (1) G0469
DPOGS211865	branched chain aminotransferase 1, cytosolic
DPOGS206255	Chromatin assembly factor 1, p180 subunit
DPOGS212259	JmjC domain-containing histone demethylase 2
DPOGS203004	CNMamide Receptor

Table 3 Continued

DPOGS213560	nuclear fallout
DPOGS208273	erb-b2 receptor tyrosine kinase 3
DPOGS206025	Peroxidasin
DPOGS204181	Fermitin 1
DPOGS211900	anterior open
DPOGS206135	Multi drug resistance 65

Table 4: List of genes with similar or differential temporal expression patterns in the brain of summer-like monarchs and wild-caught migrants. R-R: Genes rhythmic in both conditions; R-AR: Genes rhythmic in summer-like monarchs and arrhythmic in migrants; AR-R: Genes arrhythmic in summer-like monarchs and rhythmic in migrants. The annotation is based on *Drosophila* (no shading) or mouse (colored shading) orthologues. NA, monarch genes without orthologues in *Drosophila* or the mouse.

Gene ID	R-R
DPOGS212327	F-box protein 32
DPOGS211121	ubiquitin-conjugating enzyme E2Q family-like 1
DPOGS204552	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS209874	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS207730	CG10082
DPOGS209797	Vacuolar H[+] ATPase 100kD subunit 2
DPOGS206136	LON peptidase N-terminal domain and ring finger 2
DPOGS208537	fritz
DPOGS204644	phospholipid phosphatase 1
DPOGS212996	Ecdysone-inducible gene E1
DPOGS214070	stumps
DPOGS205823	polypeptide N-acetylgalactosaminyltransferase 10
DPOGS215384	RIO kinase 2
DPOGS200190	similar
DPOGS205549	G protein-coupled receptor 158
DPOGS212022	klarsicht
DPOGS210186	Pyruvate dehydrogenase kinase
DPOGS204253	calsyntenin 2
DPOGS203791	Semaphorin 1b
DPOGS215738	eiger

Table 4 Continued

DPOGS202608	maternal gene required for meiosis
DPOGS207058	choline dehydrogenase
DPOGS207000	Myosin heavy chain-like
DPOGS209025	LIM domain only 7
DPOGS213925	Heat shock protein 68
DPOGS202993	solute carrier family 36 (proton/amino acid symporter), member 1
DPOGS213900	Heat shock protein 68
DPOGS202607	maternal gene required for meiosis
DPOGS208881	DnaJ-like-2
DPOGS213901	Heat shock protein 68
DPOGS214402	Glycogen synthase
DPOGS213114	Choline transporter-like 2
DPOGS213326	CG7110
DPOGS204626	transmembrane protein 86B
DPOGS210128	PAPS synthetase
DPOGS213327	predicted gene 973
DPOGS210295	Phosphoglucose isomerase
DPOGS212829	carboxyesterase 2B
DPOGS207764	Enolase
DPOGS202781	Hsp70/Hsp90 organizing protein
DPOGS215969	Transaldolase
DPOGS203810	Phosphofructokinase
DPOGS205152	Ecdysone-inducible gene L2
DPOGS206692	pudgy
DPOGS202237	DnaJ heat shock protein family (Hsp40) member B4
DPOGS208606	vrille
DPOGS200811	hairy
DPOGS215159	Trehalose transporter 1-1
DPOGS209166	Bre1
DPOGS201013	high mobility group box transcription factor 1
DPOGS215160	Trehalose transporter 1-2
DPOGS201012	Leucine-rich pentatricopeptide repeat containing 2
DPOGS203088	solute carrier family 5 (sodium/glucose cotransporter), member 12
DPOGS212605	ATP binding cassette subfamily G member 4
DPOGS209175	phosphatidylinositol-specific phospholipase C, X domain containing 3

Table 4 Continued

DPOGS214215	Na[+]/H[+] hydrogen antiporter 1
DPOGS212590	neither inactivation nor afterpotential B
DPOGS215394	NAD-dependent methylenetetrahydrofolate dehydrogenase
DPOGS214162	Na[+]/H[+] hydrogen antiporter 1
DPOGS214179	timeless
DPOGS205105	ATP binding cassette subfamily G member 4
DPOGS205647	ATP binding cassette subfamily G member 4
DPOGS215419	dumpy
DPOGS209925	clockwork orange
DPOGS212884	sushi, nidogen and EGF-like domains 1
DPOGS203908	period
DPOGS212885	Lysine (K)-specific demethylase 2
DPOGS207088	heterogeneous nuclear ribonucleoprotein H3
DPOGS201668	short gastrulation
DPOGS213282	zelda
DPOGS208079	cryptochrome 2
DPOGS203797	Heat shock factor
DPOGS206718	ArfGAP with SH3 domain, ankyrin repeat and PH domain1
Gene ID	**R-AR**
DPOGS206156	hyaluronoglucosaminidase 2
DPOGS201294	crinkled
DPOGS208747	cacophony
DPOGS203552	acyl-CoA synthetase family member 3
DPOGS211223	transcriptional Adaptor 1-1
DPOGS212916	ferric-chelate reductase 1
DPOGS210716	ariadne 2
DPOGS212858	atlastin
DPOGS215318	short spindle 7
DPOGS212482	gemini
DPOGS211034	Optic atrophy 1
DPOGS211138	Ral guanine nucleotide dissociation stimulator-like
DPOGS209507	S-adenosylmethionine decarboxylase
DPOGS210627	carbonic anhydrase 3
DPOGS207651	UDP glucuronosyltransferase 2 family, polypeptide B1
DPOGS212685	sarcoplasmic calcium-binding protein

Table 4 Continued

DPOGS208406	CG30069
DPOGS201345	tocopherol (alpha) transfer protein-like
DPOGS201488	Mid1 interacting protein 1 (gastrulation specific G12-like (zebrafish))
DPOGS215783	hook
DPOGS216126	Gcn5 acetyltransferase
DPOGS207942	synaptotagmin XI
DPOGS215488	fatty acid desaturase 1
DPOGS203813	sphingomyelin phosphodiesterase 1, acid lysosomal
DPOGS202434	ebony
DPOGS211900	anterior open
DPOGS201894	Maltase A4
DPOGS209585	FCH domain only 2
DPOGS208044	sodium potassium chloride cotransporter
DPOGS204181	Fermitin 1
DPOGS210257	ebony
DPOGS201241	CAP
DPOGS205824	pickled eggs
DPOGS208999	retinol dehydrogenase 13 (all-trans and 9-cis)
DPOGS201544	Organic anion transporting polypeptide 74D
DPOGS203103	epithelial membrane protein
DPOGS201445	integrator complex subunit 13
DPOGS202439	NA
DPOGS201195	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2
DPOGS213560	nuclear fallout
DPOGS201376	Neuralized E3 ubiquitin protein ligase 4
DPOGS215489	6-phosphofructo-2-kinase
DPOGS211148	unzipped
DPOGS213546	dual specificity phosphatase 22
DPOGS215479	monooxygenase, DBH-like 1
DPOGS213594	Glycogen binding subunit 76A
DPOGS215439	lethal (2) essential for life
DPOGS201194	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2
DPOGS212594	Trehalose-6-phosphate synthase 1
DPOGS201446	CG10527
DPOGS214041	beta Spectrin

Table 4 Continued

DPOGS200407	midline uncoordinated
DPOGS202602	Minichromosome maintenance 2
DPOGS200764	wunen
DPOGS205265	sulfotransferase family 2A, dehydroepiandrosterone (DHEA)-preferring, member 6
DPOGS215494	1,4-Alpha-Glucan Branching Enzyme
DPOGS201573	cabut
DPOGS202178	Glutamate oxaloacetate transaminase 1
DPOGS214921	Multicopper oxidase-1
DPOGS212113	straw
DPOGS209035	transketolase-like 2
DPOGS205929	sprite
DPOGS209948	regucalcin
DPOGS215720	Arginine kinase
DPOGS214139	Fatty acid (long chain) transport protein
DPOGS206348	Picot
DPOGS213878	synaptic vesicle glycoprotein 2 b
DPOGS205027	Glycogen phosphorylase
DPOGS212608	androgen dependent TFPI regulating protein
DPOGS202126	very low density lipoprotein receptor
DPOGS211169	mitochondrial ribosomal protein L3
DPOGS212844	Copper transporter 1A
DPOGS200490	Malate dehydrogenase 1
DPOGS200089	Triose phosphate isomerase
DPOGS202687	ATP-binding cassette, sub-family C (CFTR/MRP), member 4
DPOGS208998	ethanolamine phosphate phospholyase
DPOGS213690	beat-IIIc
DPOGS205927	split ends
DPOGS206646	Salt-inducible kinase 2
DPOGS204296	Poly(ADP-ribose) glycohydrolase
DPOGS206596	Megalin
DPOGS205557	Centromeric protein-C
DPOGS207241	5-hydroxytryptamine (serotonin) receptor 1A
DPOGS211921	Myb-interacting protein 130
DPOGS211191	Zinc/iron regulated transporter-related protein 71B
DPOGS213064	Phosphoglycerate kinase

Table 4 Continued

DPOGS207671	branchless
DPOGS207964	CG13024
DPOGS212595	Trehalose-6-phosphate synthase 1
DPOGS202111	Tob
DPOGS215458	glutathione S-transferase, C-terminal domain containing
DPOGS215868	family with sequence similarity 107, member B
DPOGS200356	HEAT repeat containing 2
DPOGS208542	meiotic 9
DPOGS215460	Glyceraldehyde 3 phosphate dehydrogenase 2
DPOGS201746	Carbonic anhydrase 1
DPOGS208579	tektin 4
DPOGS214349	wunen
DPOGS207621	Leucine-rich-repeats and calponin homology domain protein
DPOGS210992	Mig-2-like
DPOGS208578	Vesicular glutamate transporter
DPOGS215853	NMDA receptor 1
DPOGS209910	Sox box protein 14
DPOGS215947	methyltransferase like 4
DPOGS203786	male fertility factor kl3
DPOGS215377	chondroitin sulfate synthase 1
DPOGS213804	Inositol 1,4,5-triphosphate kinase 1
DPOGS204494	metabotropic GABA-B receptor subtype 1
DPOGS205706	CG42588
DPOGS200691	Enhancer of split mbeta, helix-loop-helix
DPOGS203238	non-stop
DPOGS212889	histone PARylation factor 1
DPOGS203920	solute carrier family 36 (proton/amino acid symporter), member 4
DPOGS207222	roadblock
DPOGS205079	serine hydrolase-like
DPOGS214099	mannose-P-dolichol utilization defect 1
DPOGS215591	THAP domain containing, apoptosis associated protein 2
DPOGS204071	lactase
DPOGS200817	alkaline phosphatase, liver/bone/kidney
DPOGS204091	Salt-inducible kinase 3
DPOGS204153	Neurochondrin

Table 4 Continued

DPOGS214218	glycerophosphocholine phosphodiesterase 1
DPOGS211367	CoRest
DPOGS202815	Excitatory amino acid transporter 2
DPOGS212557	hippocalcin-like 1
DPOGS202145	Dpr-interacting protein gamma
DPOGS202879	Nucleoporin at 44A
DPOGS203017	Ataxin-2
DPOGS213275	chondroadherin
DPOGS215237	Mid1
DPOGS207224	pudgy
DPOGS202306	kon-tiki
DPOGS205911	Vacuolar protein sorting 20
DPOGS204363	kekkon-2
DPOGS212492	CG31324
DPOGS214212	c12.1
DPOGS204463	Matrix metalloproteinase 1
DPOGS212989	POC1 centriolar protein A
DPOGS203710	major facilitator superfamily domain containing 12
DPOGS206978	nessun dorma
DPOGS214045	transmembrane and tetratricopeptide repeat containing 2
DPOGS202609	Acetylcholine esterase
DPOGS205855	solute carrier family 46, member 3
DPOGS207157	Formin homology 2 domain containing
DPOGS200409	Integrin betanu subunit
DPOGS214408	solute carrier family 22 (organic cation transporter), member 21
DPOGS214262	zinc finger protein 27
DPOGS205964	Ubiquitin specific protease 20/33
DPOGS203155	Cullin 5
DPOGS213576	Hexosaminidase 2
DPOGS211174	Cwc25
DPOGS200295	dachshund family transcription factor 2
DPOGS205353	Mystery 45A
DPOGS213527	Zinc/iron regulated transporter-related protein 48C
DPOGS203890	scavenger receptor acting in neural tissue and majority of rhodopsin is absent

Table 4 Continued

DPOGS203868	solute carrier family 36 (proton/amino acid symporter), member 4
DPOGS202827	scavenger receptor acting in neural tissue and majority of rhodopsin is absent
DPOGS200883	CG16791
DPOGS203228	Cln7
DPOGS201881	nervana 1
DPOGS201625	3-hydroxybutyrate dehydrogenase, type 2
DPOGS213552	Ecdysone-induced protein 28/29kD
DPOGS209521	Fumarylacetoacetase
DPOGS215624	Glutaminase
DPOGS200617	Runt related B
DPOGS200157	Rab5
DPOGS215437	Pyrroline-5-carboxylate reductase-like 2
DPOGS214155	DEAH (Asp-Glu-Ala-His) box polypeptide 33
DPOGS207405	Protein 1 of cleavage and polyadenylation factor 1
DPOGS205442	king tubby
DPOGS209063	taxilin alpha
DPOGS205077	neurotrimin
DPOGS207714	astray
DPOGS208868	Membrin
DPOGS200168	Tetraspanin 29Fa
DPOGS206850	pantothenate kinase 3
DPOGS209776	Ugt86Di
DPOGS209371	choline dehydrogenase
DPOGS206851	CG17162
DPOGS205076	kirre like nephrin family adhesion molecule 1
DPOGS212748	scheggia
DPOGS200091	rotatin
DPOGS204998	Ccp84Ae
DPOGS207585	tropomyosin 1, alpha
DPOGS202068	Ras-related protein interacting with calmodulin
DPOGS203923	lethal (1) G0289
DPOGS208959	Vacuolar H[+] ATPase 100kD subunit 2
DPOGS202283	pre-mRNA processing factor 39
DPOGS206412	endoplasmic reticulum metallopeptidase 1
DPOGS201651	Negative Cofactor 2alpha

Table 4 Continued

DPOGS209186	Cadherin 86C
DPOGS201463	defective proventriculus
DPOGS209864	jim
DPOGS207448	Bestrophin 1
DPOGS214570	Melanization Protease 1
DPOGS213793	vermilion
DPOGS208231	kallikrein related-peptidase 13
DPOGS208188	Semaphorin 5c
DPOGS201845	wing blister
DPOGS207077	sperm antigen with calponin homology and coiled-coil domains 1-like
DPOGS205727	Glycoprotein hormone alpha 2
DPOGS209262	peroxidasin
DPOGS212060	Fanconi anemia, complementation group M
DPOGS214879	prune
DPOGS200824	zinc finger CCCH type containing 14
DPOGS200985	Leucine-rich repeat-containing G protein-coupled receptor 1
DPOGS203808	arc
DPOGS213355	lipase, member I
DPOGS213007	Cytochrome P450-18a1
Gene ID	**AR-R**
DPOGS210620	miranda
DPOGS210559	NA
DPOGS214770	testis expressed gene 2
DPOGS205355	Ucp4A
DPOGS206716	mind-meld
DPOGS215438	cell division cycle 14
DPOGS200105	Purine-rich binding protein-alpha
DPOGS213926	lethal (2) essential for life
DPOGS215242	chico
DPOGS207729	centromere protein I
DPOGS205032	predicted gene 5141
DPOGS212156	lola like
DPOGS207552	Mitogen-activated protein kinase phosphatase 3
DPOGS206443	gamma-glutamyltransferase 1
DPOGS212331	lethal (1) G0469

Table 4 Continued

DPOGS203967	Glucose transporter 1
DPOGS213347	rudimentary-like
DPOGS204413	F-box and leucine-rich repeat protein 6
DPOGS214030	mulet
DPOGS210848	starvin
DPOGS205386	phosphoglucose mutase
DPOGS206473	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS210324	weary
DPOGS209591	outspread
DPOGS202650	jing
DPOGS211108	BRWD3
DPOGS207955	Tissue inhibitor of metalloproteases
DPOGS211622	purine-nucleoside phosphorylase 2
DPOGS215723	Zinc finger CCHC-type containing 7
DPOGS208796	transmembrane protease, serine 2
DPOGS205218	eukaryotic translation release factor 1
DPOGS200508	zinc finger protein 85
DPOGS209575	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 2
DPOGS216018	RIKEN cDNA 5031439G07 gene
DPOGS202658	laminin, alpha 3
DPOGS213604	family with sequence similarity 205, member A2
DPOGS213632	pallbearer
DPOGS203702	naked cuticle
DPOGS205398	rha
DPOGS207508	forked
DPOGS215797	nicotinate phosphoribosyltransferase
DPOGS215758	Transportin
DPOGS205469	G patch domain containing 2 like
DPOGS206880	aldehyde dehydrogenase 1 family, member L1
DPOGS205541	CD68 antigen
DPOGS211133	phosphorylase kinase beta
DPOGS213985	XRCC1
DPOGS204403	solute carrier family 7 (cationic amino acid transporter, y+ system), member 9
DPOGS209508	phosphoglucose mutase
DPOGS206382	lethal (2) essential for life

Table 4 Continued

DPOGS211598	Peroxidase
DPOGS205122	pollux
DPOGS212285	PNN interacting serine/arginine-rich
DPOGS207495	real-time
DPOGS205231	protease, serine 30
DPOGS201368	ER degradation enhancer, mannosidase alpha-like 2
DPOGS205510	CG42304
DPOGS211989	family with sequence similarity 98, member A
DPOGS203197	pasiflora 2
DPOGS212041	heartless
DPOGS210450	glycine-N-acyltransferase-like 3
DPOGS213101	Collagen type IV alpha 1
DPOGS209745	Reversion-inducing-cysteine-rich protein with kazal motifs
DPOGS202664	membrane steroid binding protein
DPOGS215931	Aldehyde dehydrogenase type III
DPOGS207783	Dual oxidase
DPOGS213914	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS200925	CG13643
DPOGS210130	folliculin interacting protein 2
DPOGS208609	Adh transcription factor 1
DPOGS214409	solute carrier family 22 (organic cation transporter), member 1
DPOGS215544	zinc finger, FYVE domain containing 28
DPOGS206919	lethal (1) G0196
DPOGS201015	Sterol carrier protein X-related thiolase
DPOGS206369	Mediator complex subunit 27
DPOGS214271	Sodium/solute co-transporter-like 5A11
DPOGS207312	Src oncogene at 64B
DPOGS208891	Relish
DPOGS213513	lethal (3) 72Dn
DPOGS204840	ecto-NOX disulfide-thiol exchanger 1
DPOGS212145	Adenosine receptor
DPOGS210680	Kallmann syndrome 1
DPOGS214160	Na[+]/H[+] hydrogen antiporter 1
DPOGS211339	N(alpha)-acetyltransferase 20 A
DPOGS206184	mitochondrial ribosomal protein L35

Table 4 Continued

DPOGS203649	Integrator 6
DPOGS210429	Rab23
DPOGS202082	lectin, mannose-binding 2-like
DPOGS200944	Required for cell differentiation 1
DPOGS207391	clavesin 2
DPOGS207599	mitochondrial ribosomal protein L55
DPOGS205264	aldehyde dehydrogenase 18 family, member A1
DPOGS209146	solute carrier family 22 (organic cation transporter), member 3
DPOGS207504	Farnesyl pyrophosphate synthase
DPOGS208519	Usf
DPOGS203115	Tyrosyl-tRNA synthetase
DPOGS205880	RIKEN cDNA 1110038F14 gene
DPOGS207273	Histone deacetylase 1
DPOGS207041	BUD23, rRNA methyltransferase and ribosome maturation factor
DPOGS215524	Glutathione Synthetase
DPOGS202243	ribosomal RNA processing 15 homolog (S. cerevisiae)
DPOGS213269	amidohydrolase domain containing 2
DPOGS209232	Replication factor C 38kD subunit
DPOGS201279	clavesin 1
DPOGS205906	gem nuclear organelle associated protein 7
DPOGS206662	N(alpha)-acetyltransferase 38, NatC auxiliary subunit
DPOGS203217	Chaperonin containing TCP1 subunit 8
DPOGS200290	penguin
DPOGS210994	Cellular Repressor of E1A-stimulated Genes
DPOGS209768	Signal recognition particle receptor beta
DPOGS206466	PTC7 protein phosphatase homolog
DPOGS214393	Glia maturation factor
DPOGS204523	eukaryotic translation initiation factor 2 subunit alpha
DPOGS202974	MAP kinase activated protein-kinase-2
DPOGS213401	Origin recognition complex subunit 5
DPOGS215555	Minichromosome maintenance 6
DPOGS201992	x16 splicing factor
DPOGS205993	Malignant T cell amplified sequence 1
DPOGS202915	Signal peptide peptidase
DPOGS206597	Fasciclin 2

Table 4 Continued

DPOGS203773	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)
DPOGS212221	NOP2-Sun domain family, member 2
DPOGS201325	mitochondrial translation elongation factor Ts
DPOGS207501	Actin-related protein 3
DPOGS206108	Dodeca-satellite-binding protein 1
DPOGS207351	PIN2/TERF1 interacting, telomerase inhibitor 1
DPOGS207279	NA
DPOGS216037	Rho GTPase activating protein at 68F
DPOGS214375	Dihydropteridine reductase
DPOGS200286	Regulatory particle non-ATPase 10
DPOGS205508	CCAAT/enhancer binding protein zeta
DPOGS211776	rush hour
DPOGS214580	NOP10 ribonucleoprotein
DPOGS204316	Isocitrate dehydrogenase
DPOGS208154	uridine phosphorylase 1
DPOGS203757	Asparagine synthetase
DPOGS211134	Proteasome beta2 subunit
DPOGS200182	Dual oxidase
DPOGS212387	sans fille
DPOGS207188	Glutamate receptor IB
DPOGS215227	Cyclin-dependent kinase subunit 85A
DPOGS213746	Proteasome alpha6 subunit
DPOGS200285	Mesencephalic astrocyte-derived neurotrophic factor
DPOGS214164	Sec61 beta subunit
DPOGS209926	hattifattener
DPOGS211106	chickadee
DPOGS211051	FER tyrosine kinase
DPOGS202316	optic atrophy 3
DPOGS204777	Sec61 gamma subunit
DPOGS212707	potassium channel, subfamily K, member 18
DPOGS209613	tiptop
DPOGS213650	coiled-coil serine rich 2
DPOGS207200	His1:CG33861
DPOGS208768	Jrk-like
DPOGS212261	Kaz1-ORFB

Table 4 Continued

DPOGS204192	Glutamyl-prolyl-tRNA synthetase
DPOGS203404	highwire
DPOGS202523	Dynein heavy chain 64C
DPOGS209960	Vacuolar protein sorting 13
DPOGS209394	protein arginine N-methyltransferase 2
DPOGS213134	HECT and RLD domain containing protein 2
DPOGS212667	CG8740
DPOGS201500	PAS kinase
DPOGS207418	rugose
DPOGS205736	endosomal maturation defective
DPOGS210906	myopic
DPOGS203769	Ca[2+]-channel protein alpha[[1]] subunit D
DPOGS204111	sodium channel, voltage-gated, type X, alpha
DPOGS206600	radish
DPOGS204827	unc-13
DPOGS211160	Piezo
DPOGS205809	blue cheese
DPOGS205030	Rap GTPase activating protein 1
DPOGS204823	lymphocyte specific 1
DPOGS208356	DEAH (Asp-Glu-Ala-His) box polypeptide 35
DPOGS210629	APC-like
DPOGS212969	ring finger protein 144A
DPOGS207346	NA
DPOGS209535	meiosis specific with coiled-coil domain
DPOGS203506	CG15894
DPOGS216202	castor
DPOGS206932	RIM-binding protein
DPOGS212524	NA
DPOGS205581	Cullin 1
DPOGS206816	CG4658
DPOGS207452	Rab3 GDP-GTP exchange factor
DPOGS210020	inositol polyphosphate-5-phosphatase F
DPOGS215603	rapamycin-insensitive companion of Tor
DPOGS211143	WD repeat domain 59
DPOGS206164	Brahma associated protein 170kD
DPOGS212332	Adenylyl cyclase 76E
DPOGS206043	ftz transcription factor 1

Table 4 Continued

DPOGS209239	zinc finger and BTB domain containing 40
DPOGS204972	ring finger protein 165
DPOGS201030	Marcal1
DPOGS207938	still life
DPOGS203622	meiosis regulator and mRNA stability 1
DPOGS209769	septin interacting protein 1
DPOGS206403	nicotinic Acetylcholine Receptor alpha6
DPOGS205084	RNA-binding Fox protein 1
DPOGS203803	c-Maf inducing protein
DPOGS200827	Formin-like
DPOGS202986	Gigyf
DPOGS203978	predicted gene 11639
DPOGS209780	family with sequence similarity 160, member B1
DPOGS204256	tramtrack
DPOGS204257	tramtrack
DPOGS214346	wunen
DPOGS215331	Ecdysone-induced protein 74EF
DPOGS201750	grappa
DPOGS202907	centrosomal protein 162
DPOGS203253	Hepatocyte growth factor regulated tyrosine kinase substrate
DPOGS201547	X11Lbeta
DPOGS213525	Shaker cognate b
DPOGS202684	Wnk kinase
DPOGS212980	armadillo repeat containing 8
DPOGS211471	collagen, type XXIII, alpha 1
DPOGS201244	CAP
DPOGS202471	pseudouridylate synthase 7
DPOGS211197	SET domain containing 1A
DPOGS203274	Rho GTPase activating protein at 100F
DPOGS214481	Wnk kinase
DPOGS204391	female sterile (1) homeotic
DPOGS207677	R3H domain and coiled-coil containing 1 like
DPOGS211802	skywalker
DPOGS206516	proteoglycan 4 (megakaryocyte stimulating factor, articular superficial zone protein)
DPOGS201245	CAP

Table 4 Continued

DPOGS208933	Nna1 carboxypeptidase
DPOGS211301	Synaptosomal-associated protein 25kDa
DPOGS205950	lethal (3) malignant brain tumor
DPOGS212207	pinstripe
DPOGS202457	PDGF- and VEGF-receptor related

Table 5: List of genes with similar temporal expression patterns in the brains of monarchs raised in LP and in summer-like monarchs and/or in the brains of monarchs raised in SP and in wild-caught migrants. R-R: Genes rhythmic in all four conditions; R-AR: Genes rhythmic in LP and summer-like monarchs, and arrhythmic in SP monarchs and migrants; AR-R: Genes arrhythmic in LP and summer-like monarchs, and rhythmic in SP monarchs and migrants. The annotation is based on *Drosophila* (no shading) or mouse (colored shading) orthologues. NA, monarch genes without orthologues in *Drosophila* or the mouse.

Gene ID	R-R
DPOGS200190	similar
DPOGS201012	Leucine-rich pentatricopeptide repeat containing 2
DPOGS201013	high mobility group box transcription factor 1
DPOGS202237	DnaJ heat shock protein family (Hsp40) member B4
DPOGS203797	Heat shock factor
DPOGS203908	period
DPOGS204253	calsyntenin 2
DPOGS204552	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS204644	phospholipid phosphatase 1
DPOGS205105	ATP binding cassette subfamily G member 4
DPOGS205549	G protein-coupled receptor 158
DPOGS207058	choline dehydrogenase
DPOGS207730	CG10082
DPOGS208606	vrille
DPOGS209175	phosphatidylinositol-specific phospholipase C, X domain containing 3
DPOGS209797	Vacuolar H[+] ATPase 100kD subunit 2
DPOGS209874	solute carrier family 2, (facilitated glucose transporter), member 8
DPOGS209925	clockwork orange
DPOGS210128	PAPS synthetase

Table 5 Continued

DPOGS211121	ubiquitin-conjugating enzyme E2Q family-like 1
DPOGS212590	neither inactivation nor afterpotential B
DPOGS213114	Choline transporter-like 2
DPOGS213900	Heat shock protein 68
DPOGS213901	Heat shock protein 68
DPOGS213925	Heat shock protein 68
DPOGS214070	stumps
DPOGS214162	Na[+]/H[+] hydrogen antiporter 1
DPOGS214179	timeless
DPOGS214215	Na[+]/H[+] hydrogen antiporter 1
DPOGS215159	Trehalose transporter 1-1
DPOGS215160	Trehalose transporter 1-2
DPOGS215419	dumpy
DPOGS215738	eiger
Gene ID	**R-AR**
DPOGS201241	CAP
DPOGS201544	Organic anion transporting polypeptide 74D
DPOGS201894	Maltase A4
DPOGS202126	very low density lipoprotein receptor
DPOGS202609	Acetylcholine esterase
DPOGS202827	scavenger receptor acting in neural tissue and majority of rhodopsin is absent
DPOGS203228	Cln7
DPOGS203808	arc
DPOGS203890	scavenger receptor acting in neural tissue and majority of rhodopsin is absent
DPOGS205077	neurotrimin
DPOGS205727	Glycoprotein hormone alpha 2
DPOGS207157	Formin homology 2 domain containing
DPOGS207651	UDP glucuronosyltransferase 2 family, polypeptide B1
DPOGS207671	branchless
DPOGS208868	Membrin
DPOGS208999	retinol dehydrogenase 13 (all-trans and 9-cis)
DPOGS210627	carbonic anhydrase 3
DPOGS212492	CG31324
DPOGS213594	Glycogen binding subunit 76A
DPOGS214045	transmembrane and tetratricopeptide repeat containing 2

Table 5 Continued

DPOGS214921	Multicopper oxidase-1
DPOGS215237	Mid1
Gene ID	**AR-R**
DPOGS200182	Dual oxidase
DPOGS202243	ribosomal RNA processing 15 homolog (S. cerevisiae)
DPOGS203217	Chaperonin containing TCP1 subunit 8
DPOGS204316	Isocitrate dehydrogenase
DPOGS205508	CCAAT/enhancer binding protein zeta
DPOGS209768	Signal recognition particle receptor beta
DPOGS210429	Rab23
DPOGS210994	Cellular Repressor of E1A-stimulated Genes
DPOGS211622	purine-nucleoside phosphorylase 2
DPOGS212331	lethal (1) G0469
DPOGS213101	Collagen type IV alpha 1
DPOGS213604	family with sequence similarity 205, member A2

Table 6: RNA-seq mapping summary for LP, SP, summer-like (summer), and migrant monarchs. R: replicate.

Samples	Number of reads	Mapped reads	Mapping percentage
LP ZT01 R1	13,522,929	12,569,145	92.9
LP ZT01 R2	13,678,471	12,641,190	92.4
LP ZT04 R1	15,821,730	14,679,157	92.8
LP ZT04 R2	16,370,285	15,082,230	92.1
LP ZT07 R1	18,287,614	16,960,023	92.7
LP ZT07 R2	18,547,268	17,043,095	91.9
LP ZT10 R1	13,373,316	12,350,949	92.4
LP ZT10 R2	16,462,573	15,232,242	92.5
LP ZT13 R1	15,378,742	14,253,836	92.7
LP ZT13 R2	18,124,188	16,640,088	91.8
LP ZT16 R1	16,141,542	14,943,063	92.6

Table 6 Continued

LP ZT16 R2	17,448,980	15,981,062	91.6
LP ZT19 R1	14,581,672	13,501,921	92.6
LP ZT19 R2	15,306,522	14,013,930	91.6
LP ZT22 R1	15,207,486	14,104,449	92.7
LP ZT22 R2	16,168,931	14,936,831	92.4
SP ZT01 R1	10,939,912	10,068,478	92.0
SP ZT01 R2	11,026,339	10,288,670	93.3
SP ZT04 R1	11,957,374	11,060,333	92.5
SP ZT04 R2	11,984,004	11,214,567	93.6
SP ZT07 R1	14,661,160	13,598,461	92.8
SP ZT07 R2	11,503,579	10,754,899	93.5
SP ZT10 R1	11,591,287	10,850,911	93.6
SP ZT10 R2	11,747,485	10,928,317	93.0
SP ZT13 R1	13,794,223	12,895,039	93.5
SP ZT13 R2	11,163,015	10,402,241	93.2
SP ZT16 R1	12,703,894	11,832,060	93.1
SP ZT16 R2	12,174,241	11,372,549	93.4
SP ZT19 R1	11,931,961	11,165,608	93.6
SP ZT19 R2	11,276,364	10,498,869	93.1
SP ZT22 R1	9,356,277	8,741,841	93.4
SP ZT22 R2	9,315,476	8,698,696	93.4
Summer ZT01 R1	8,329,736	7,474,898	89.7
Summer ZT01 R2	11,307,582	9,768,700	86.4
Summer ZT04 R1	8,308,713	7,353,942	88.5
Summer ZT04 R2	9,752,315	8,477,008	86.9
Summer ZT07 R1	7,489,482	6,551,179	87.5
Summer ZT07 R2	9,802,883	8,436,946	86.1
Summer ZT10 R1	8,292,323	7,183,739	86.6
Summer ZT10 R2	10,067,093	8,634,538	85.8

Table 6 Continued

Summer ZT13 R1	7,414,155	6,488,395	87.5
Summer ZT13 R2	9,848,313	8,840,377	89.8
Summer ZT16 R1	8,373,408	7,151,758	85.4
Summer ZT16 R2	10,074,990	9,041,720	89.7
Summer ZT19 R1	8,023,590	6,893,871	85.9
Summer ZT19 R2	8,636,414	7,766,293	89.9
Summer ZT22 R1	8,279,696	7,136,863	86.2
Summer ZT22 R2	10,286,399	9,071,636	88.2
Migrant ZT01 R1	16,093,489	15,043,013	93.5
Migrant ZT01 R2	15,823,203	14,866,974	94.0
Migrant ZT04 R1	14,592,336	13,293,343	91.1
Migrant ZT04 R2	17,808,260	16,668,099	93.6
Migrant ZT07 R1	17,429,487	16,324,273	93.7
Migrant ZT07 R2	17,125,519	16,034,231	93.6
Migrant ZT10 R1	15,979,866	14,920,559	93.4
Migrant ZT10 R2	14,289,759	12,589,285	88.1
Migrant ZT13 R1	17,412,821	16,261,315	93.4
Migrant ZT13 R2	18,173,157	17,048,724	93.8
Migrant ZT16 R1	19,273,821	18,000,493	93.4
Migrant ZT16 R2	15,910,647	14,849,245	93.3
Migrant ZT19 R1	18,186,048	16,980,976	93.4
Migrant ZT19 R2	17,892,892	16,129,214	90.1
Migrant ZT22 R1	15,167,938	14,404,139	95.0
Migrant ZT22 R2	15,017,821	13,750,935	91.6
Average	13,312,703	12,230,335	91.5

Table 7: Expressed and cycling genes in the brains of LP, SP, summer-like and migrant monarchs.

Samples	Genes expressed	Cycling genes	% of cycling genes
LP	10,333	272	2.70
SP	10,113	268	2.66
Summer-like	10,079	282	2.80
Migrant	10,353	306	3.04