Characterizing Genetic Mechanisms Involved in Measuring Day-length in *Drosophila melanogaster*

Harjit Khaira¹ and Kwangwon Lee^{1,2*}

¹ Department of Biology, Rutgers, The State University of New Jersey, Camden, New Jersey, United States of America

² Center for Computational & Integrative Biology, Rutgers, The State University of New Jersey, Camden, New Jersey, United States of America

Short Title: *PKCdelta* on Photoperiodism

* Corresponding author

E-mail: kwangwon.lee@rutgers.edu

Abstract

Many organisms are known to regulate seasonal behaviors and physiological processes in response to day length changes through photoperiodism. Extreme changes in photoperiods have detrimental effects on human health, which can impair development and serve as the origin of adult diseases. Since the seminal work by Bünning in 1936, there are studies supporting the view that organisms can measure the day length through an endogenous 24-hour cellular circadian clock. However, the mechanisms involved in measuring seasonal or day-length changes are not understood. In the current study, we performed a genome-wide association study (GWAS) on photoperiodism using the *Drosophila* Genetic Reference Panel. The GWAS identified 32 candidate genes responsible for photoperiodic regulations. The knockout mutants of the top four candidate genes (*Protein Kinase C delta* (*Pkcdelta*), *Glucuronyltransferase-P* (*GlcAT-P*), *Brain-specific homeobox* (*Bsh*), and *Diuretic hormone 31 Receptor* (*Dh31-RI*)) were analyzed for their photoperiod and circadian period phenotypes. *PKCdelta* and *GlcAT-P* mutants show a significantly different photoperiod response compared to that of the wild type strain, and also had an altered circadian period phenotype. Further molecular characterization revealed that the mutant two independent mutant alleles of *PKCdelta* with a defective catalytic domain had distinct photoperiod responses. Taken these data together, we concluded that there is overlap between the circadian clock and photoperiodic regulations in *Drosophila*, and *PKCdelta* is a component that is involved in both circadian and photoperiodic regulations. By identifying novel molecular components of photoperiod, the current study provides new insights into the genetic mechanisms of determining the seasonal changes.

Keywords Photoperiodism; Circadian Clock; GWAS; *Drosophila*

Introduction

Most organisms experience predictable environmental changes; daily, tidal, lunar, and seasonal. These changes mean different availability of resources and stressors at predictable time intervals. It is crucial for an organism to prepare for the incoming season for its survival in nature. Thus, it is not surprising that organisms have evolved time-measuring cellular devices. Among the known biological rhythms, seasonal rhythm plays a major role in many biological processes such as growth, development, reproduction, migration, and dormancy. A robust response to these rhythms can increase the survival success of an organism (Bradshaw and Holzapfel, 2010). However, little is known about the molecular and genetic mechanism of how an organism interprets and adapts to seasonal environmental changes.

Photoperiodism is the physiological response of an organism to changes in day length. Photoperiod is one of the most important factors that control seasonal breeding. Few species show extreme behaviors like hibernation while many species synchronize breeding and birth of their offspring to the most advantageous time of the year for better survival. Based on the gestation period, photoperiodic species can be categorized as long day (LD) or short day (SD) breeders. Animals like sheep and goats with long gestation periods begin to breed as day length decreases, while animals like mice and hamsters breed as day length increases (Ortavant et al., 1964). In humans there are some psychological and psychiatric consequences in response to the photoperiodic changes. There also is seasonality in birth rates among humans and onset of certain diseases like measles in the population (Martinez-Bakker et al., 2014). Patients with bipolar disorders have more depressive episodes during the late autumn and early winter, while there is an increase of manic episodes during spring and summer (Maruani et al., 2018). In flowering plants,

some are flowering at specific season of the year, and thus grouped as SD or LD flowering plants. The molecular and genetic pathways of photoperiodism have been studied in plants (Johansson and Staiger, 2015). It is also known that the circadian clock controls the activity of CONSTANS (CO), which is a key component of the photoperiod pathway. CO activates the transcription of *FLOWERING LOCUS T (FT)* which the ultimately leads to flowering in plants (Song et al., 2010).

Photoperiodic responses are well documented in insects (Bradshaw and Holzapfel, 2010). Although *Drosophila melanogaster* is the successful model species for many biological processes, the species was regarded as less optimal for studying photoperiodism due do a less pronounced photoperiod response (Tauber and Kyriacou, 2001). Most of the studies of photoperiodism in insects utilized the diapause as a phenotype. Diapause occurs when female insect ovaries shrink due to photoperiod changes as an indicator of photoperiodic phenotype. On the contrary to the concern in using *D. meanogaster* for studying photoperiodism, it also has been shown that *D. melanogaster* has a robust photoperiod response under semi-natural or SD conditions (Nagy et al., 2018; Saunders et al., 1989).

One compelling hypothesis regarding the seasonal rhythm is that an organism can ‘interpret’ the season by measuring the day length. A crucial assumption is that an organism knows the season by measuring photoperiods is due to the presence of an endogenous cellular daily clock, called the circadian clock. Circadian clocks are endogenous timekeeping networks that drive 24hr circadian rhythms. Bünning reported that the rate of leaf movement speed changes during different seasons and proposed the circadian clock controlled this via photoperiodism (Bünning, 1936). Since then, a number of studies supported Bunning’s hypothesis in certain organisms while rejecting his hypothesis in others. Studies on aphids and spider mites conclude that their

photoperiod mechanisms do not seem to depend on endogenous 24-h rhythms (Lees, 1960). Mutation of the classical clock gene *period* (*per*) still allowed *D. melanogaster* to distinguish between SD and LD . It was concluded that the clock gene *per* cannot be involved in night-length measurement mechanism (Saunders, 1990). However, another study in *Drosophila* suggests that the circadian clock is necessary for a photoperiodic response, as arrhythmic clock mutants *per⁰¹*, *tim⁰¹*, and *Clk^{JRK}* have altered photoperiodic responses from wild-type *Drosophila* (Pegoraro et al., 2014). For the last few decades, the genetic and molecular mechanisms of circadian clocks have been characterized in many organisms (Partch et al.2014, Nohales and Kay 2016). However, it is still not well understood on how the circadian clock is involved in measuring photoperiodism or day length changes.

There are at least two different hypotheses on how circadian clocks are involved in measuring seasonal changes: the external coincidence model and the internal coincidence model. In the external coincidence model, the photoperiodic response can be generated by the circadian clock that is sensitive to light at certain phases (Sawa et al., 2008). If the light signal is detected at this sensitive phase, then the photoperiodic response will be displayed such as flowering in plants. Light can entrain the clock and set the phase of photoperiodic response like flowering, but light can also directly induce flowering. In the internal coincidence model, two (or more) rhythms are produced by the clock. Under the internal coincidence model, responses, such as flowering, are only generated in conditions that bring both rhythms in phase with each other (Sawa et al., 2008). Various studies show one or the other hypotheses at work in a range of organisms. In *Arabidopsis*, external coincidence models can explain long day-specific dusk *FT* (a florigen gene) expression (Song et al., 2018). Studies using the parasitic wasp *Nasonia vitripennis* support the internal coincidence model. Using the Nanda-Hamner protocol it was found that *N.vitripennis* exhibit two

peaks in diapause, a dawn and dusk peak, which is suggestive of two circadian oscillators (Saunders, 2012). In some bacterial species the hourglass model is suggested to be responsible for timekeeping (Johnson et al., 2017). When bacteria exposed to more regular environmental conditions such as consistent light-dark cycles or constant photoperiod, hourglass model is at play. However, when the environmental conditions vary such as inconsistent light-dark cycles or photoperiods leads to a sustained oscillator (Johnson et al., 2017).

In our current study, we aim to better understand the genetic mechanisms of photoperiodism and the roles of circadian clocks in photoperiodism using the eukaryotic model organism *D. melanogaster*. *Drosophila* Genetic Reference Panel (DGRP) is a community resource of 205 inbred lines created from 40 wild type parents for studying population genomics and quantitative traits (Mackay et al., 2012). Because the DGRP is inbred to homozygosity, it is possible to identify SNP variants responsible for the trait of interest. DGRP is a fully sequenced population with annotated high-quality SNPs, indels, and chromosomal inversions (Huang et al., 2014; Mackay et al., 2012). DGRP has been used to study various complex traits such as starvation resistance, startle response, chill coma recovery (CCR) time, and circadian behavior (Mackay et al., 2012; Harbison et al., 2019).

Although diapause has been the phenotype of choice in characterizing the photoperiodic response, the diapause assay is challenging to perform in a high throughput manner. Therefore, we optimized the CCR assay as a phenotype for an organism's ability to respond to the photoperiodic changes to perform the initial DGRP Genome-Wide Association Study (GWAS). We propose that our approach would identify genes involved in both photoperiodism and the circadian clock or only in photoperiodism. Our current study identified 32 candidate photoperiod genes from which we assessed the photoperiod response of the top 4 genes using CCR. We further characterized the

role of allelic mutants of *PKCdelta* in photoperiodism using CCR, diapause, and molecular techniques.

Materials and Methods

Fly stocks

We purchased the DGRP strains from the Bloomington Drosophila Stock Center (Indiana University, Bloomington, IN). In the laboratory, the fly stock was maintained at room temperature (25°C) and in 12L:12D photoperiod conditions. The following are the *PKCdelta* mutants BDSC#18258 Pbac[RB]PKCdelta[e04408], BDSC#59433 Mi[MIC]PKCdelta[MI13377]. The BDSC#62582 is *GlcAT-P* knockout mutant Pbac[GlcAT-P]0027-G4, BDSC#27467 Bsh mutant EPbsh[G3075]. BDSC#22718 is the *Dh31-R1* knockout mutant ET1Dh31-R1[MB00175].

Photoperiod conditions and Chill coma recovery

Chill Coma Recovery (CCR) is a traditional method used to assess the photoperiod phenotype in *Drosophila* species. CCR is expressed as the time (in seconds) it takes an organism to recover from a paralysis like state after being kept under extremely low temperature. *D. melanogaster* enter the chill coma at 7.0 ± 0.9 °C (Hosler et al., 2000). It has been clearly demonstrated that flies raised under short-day (SD) photoperiod conditions compared to long-day (LD) photoperiod conditions have an advanced (or shorter) recovery time (Lanciani et al., 1992). DGRP fly lines were reared under LD and SD in a chamber, 25 ± 0.5 °C and 70% humidity. The LD consisted of 16 hours of light and 8 hours of dark (16h:8h), while the SD consisted of 8 hours of light and 16 hours of dark (8h:16h). The flies were reared under these conditions for two weeks. Chill coma was then performed by placing the flies in 4 °C for 60 minutes. Next, the flies were placed into 5 mm glass

tubes and into the *Drosophila* activity monitors (DAM2, Trikinetics) to measure the time it takes for the flies to recover. For an accurate measure of the recovery time, the flies were placed in the middle of the glass tube near the IR sensor to capture the exact time the fly recovered from the coma.

Diapause Assay

The diapause assay was performed by surgically removing the ovaries after rearing female flies under different photoperiods. Previous research suggested that female *Drosophila* placed under SD enter reproductive diapause which results in underdeveloped ovaries. The SD chamber condition mimics the winter-like days when most organisms cease to reproduce to conserve energy, therefore, diapause is induced in female flies (Saunders et al., 1989). Female *Drosophila* placed under LD condition do not enter diapause. The *PKCdelta* mutants, *PKCdelta*:Pbac and *PKCdelta*:MIC were reared under SD and LD photoperiodic conditions along with DGRP208 (WT) strain as a control. Newly hatched female flies were placed under the photoperiod conditions for 12 days, 25° C. Ovaries were dissected in 1x PBS, the image was taken with a light microscope (LEICA microsystems, LEICA EC3, Model:MEB126) under 4x magnification.

qPCR

The expression level of the candidate gene *PKCdelta* was characterized under SD and LD. The RNA was prepared using the Zymo Research Direct-Zol RNA extraction kit (Cat#R2050). The cDNA was prepared using the New England Biolabs (NEB) protoscript II first strand cDNA synthesis kit (Cat#E65605). For qPCR the Applied Biosystems PowerUp SYBER Green master mix was used (Cat#A25742) in the Applied Biosystems QuantStudio 6 Rlex Real-Time PCR

system according to the manufacture's manual. The following primer set was used; ACTIN served as a control, (Fw; GCAAGTACTCTGTCTGGATC) (Rv; CCAGCAGAATCAAGACCATCC). HK-5 primer set which targeted the RD transcript of *PKCdelta* (Fw; CGATCGATTGGGGTTTGCTGG) (Rv:GTGATGTGAGGATTTGTGTAGG), expected product size 283 bp. HK6 primer set targeted the RE transcript of *PKCdelta* (Fw; TCGAGCTGAAGACAACACCT) (Rv: CTTTCTTAGCTTCGACTTGGTT) expected to product size 160bp (Figure S4, Table S2).

Statistical Analysis

Cohen's D

CCR has been used successfully by other groups as the phenotype reflecting photoperiodic response (Pegoraro et al., 2014). Our primary interest is measuring the 'differential' response of each genotype to two different photoperiodic conditions. Cohen's D (CD) is used to measure the effect size of the photoperiodic response in each genotype (Lakens, 2013):

$$CD = \frac{(CCR_{SD} - CCR_{LD})}{Average(STDV_{CCR_{SD}}, STDEV_{CCR_{LD}})}$$

For each of the DGRP strains, the average CCR time was calculated from 20 replicates per strain in SD (CCR_{SD}) and LD (CCR_{LD}). The CD value of each genotype was calculated by subtracting CCR_{LD} from CCR_{SD} , and divided by the average of the two standard deviations of CCR_{LD} and CCR_{SD} . The CD values can be divided into three groups: large effect size, $CD > 0.8$; medium effect size, $0.8 > CD > 0.5$; and small effect size, $CD < 0.2$. The CD values for the mutants were obtained by subtracting equinox photoperiod CCR value from that of SD or that of LD respectively. This was done to ensure that the data was normalized to the EQ photoperiod.

$$CD_{SD} = \frac{(CCR_{SD} - CCR_{EQ})}{Average(STDV_{CCR_{SD\&EQ}})}, CD_{LD} = \frac{(CCR_{LD} - CCR_{EQ})}{Average(STDDEV_{CCR_{LD\&EQ}})}$$

CD value provides directionality by positive or negative values. Thus, CD value shows if a photoperiodic condition provided a positive effect on the strain or a negative effect in the CCR assay.

Genotype Effect Size

To measure the difference between the effect size of CD in different genotypes, we calculated the Genotype Effect Size (GES) by subtracting the CD value of the wild type from that of the mutant.

$$CD_{GenotypeX} - CD_{Wild\ type} = Genetic\ Effect\ Size\ (GES)$$

In calculating the GES values, DGRP208 (WT) was used as the baseline. GES value allowed us to determine which candidate gene had the most differential response to the photoperiod in comparison to that of the wild type.

GWAS Analysis

After calculating CD for all the strains under different photoperiodic conditions, we performed the GWAS using the web-based tool provided by DGRP2 (Mackay et al., 2012; Huang et al., 2014). The 1,881,979 select SNPs were analyzed for their association with the photoperiod response. All the variants from the GWAS were graphed in a manhattan plot with their chromosome on the X-axis and $-\log_{10}(p)$ of the p values on the Y-axis. There are two significant thresholds $-\log_{10}(5e-5)$ and $-\log_{10}(5e-8)$ (Figure 2). The SNPs above the $-\log_{10}(5e-8)$ were considered significant SNPs.

Heritability

The HiLMM R package was used to calculate narrow-sense heritability of photoperiod response in the DGRP population. The HiLMM package provides an estimation of heritability with confidence intervals in linear mixed models.

Results

Variation in CCR in the DGRP

Figure 1. Variation in chill coma recovery in the DGRP population. A. The overall distribution of the CD values within the DGRP population, Shapiro-Wilks normality test $p = 3.763e-09$, $p < 0.05$, and $W = 0.9194$. B. The Cohen's D (CD) of the DGRP population, $CD = (CCR_{SD} - CCR_{LD}) / \text{average}(\text{STDV_CCR}_{SD}, \text{STDV_CCR}_{LD})$. $CD > 0.8$ is large effect size, CD of 0.5 is medium effect size, and $CD < 0.2$ is small effect size. The majority of the DGRP strains fall under medium effect size total of 93 strains, while 63 strains fall under large effect size and 49 under small effect size. Thick horizontal lines represent mean and thin lines represent standard deviation. C. The CCR of the two representative strains from the DGRP population, DGRP28140 and DGRP28152 under Short Day (SD, circle) and Long Day (LD, square) photoperiods.

To test our hypothesis that genetic variations exist in photoperiodic response in *D. melanogaster*, we analyzed the CCR phenotype in DGRP, and found significant variations in CCR among DGRP

strains (Figure 1, S1). The distribution of Cohen's D values in DGRP, Shapiro-Wilks $p = 3.763 \times 10^{-9}$, and $W = 0.9194$ (Figure 1A). To determine the differential response of each DGRP genotype to the two different photoperiods, CD values were calculated for each genotype (Materials and Methods). The majority of the DGRP strains fall under medium effect size ($n = 93$), large effect size ($n = 63$), and small effect size ($n = 49$) (Figure 1B). There is more variation in CD values among the large effect size DGRP strains than the strains that fall under medium and small effect size. The overall average CCR in LD is 941 sec while CCR in SD is 1095 sec in the whole DGRP population. Two representative strains are shown along with their recovery time under SD and LD (Figure 1C). DGRP28140 has a delayed recovery time under SD in comparison to that of LD, while DGRP28152 has a faster recovery time in SD compared to LD. To determine how much of the variation is due to genetic factors in the DGRP population, we calculated the heritability. Heritability (h^2) for the long day condition was 0.51 and h^2 for the short-day condition was 0.45. It was previously reported that there is a sex difference in the CCR of *Drosophila* (Pegoraro et al., 2014). However, there was no strong correlation between sex and CCR in our experimental condition. The data from both sexes were grouped together for our GWAS analysis (Figure S2).

GWAS Analysis for Photoperiodism

256

257 We performed GWAS analysis on CCR using it as a readout for photoperiodic response (Materials
258 and methods) and identified genetic variants that were significant, $p = 1E-07$ (Figure 2). We also
259 identified the 32 most promising variants strongly associated with the photoperiod response (Table
260 S1). The enrichment analysis of these candidate genes shows a list of biological phenotypes
261 including jumping, flight phenotypes, and circadian rhythms (Figure S3).

We characterized the top four candidate genes that are strongly associated with photoperiod CCR responses, $p \geq 1E-07$; *Glucuronyltransferase-P* (*GlcAT-P*); $p = 6.63E-08$, *Brain-specific homeobox* (*Bsh*); $p = 8.46E-07$, *Protein Kinase C delta* (*Pkcdelta*); $p = 8.22E-07$, and *Diuretic hormone 31 Receptor* (*Dh31-RI*); $p = 3.49E-07$. *GlcAT-P* belongs to the glycosyltransferase family which are involved in catalyzing the conjugation of glucuronic acid from UDP-glucuronate to an acceptor. This particular gene is involved in the biosynthesis of L2/HNK-1 carbohydrate epitopes on both glycolipids and glycoproteins (Kim et al., 2003). *Bsh* encodes a putative transcription factor, and it is needed for the specification of neural identities in the adult brain. The rat ortholog of *Bsh* (*Bsx*) is suggested to be under circadian regulations and is expressed in the mature pineal gland of rat (Carstensen et al., 2019). *Dh31-RI* is *Diuretic hormone 31 Receptor* which encodes a protein that exhibits diuretic hormone receptor activity, and it is involved in G protein-coupled receptor signaling pathways. *Dh31-RI* is also known to be involved in time measurement and activity rhythms through the circadian clock system (Dubowy and Cavanaugh, 2014). *PKCdelta* belongs to a family PKC consisting of serine/threonine kinases which are activated by extracellular signaling. PKCs are divided into subgroups of Classical PKCs, Novel PKCs, and Atypical PKCs. *PKCdelta* belongs to the class of Novel PKCs which happen to be known for their two C1 domain and activation via 1,2-diacylglycerol (DAG) and phosphatidylserine (PS). While the cPKC family requires DAG and PS and aPKC only requires PS as a cofactor for activation (Shieh et al., 2002; Suh et al., 2003). We identified two independent SNP variants in *PKCdelta*, which strongly suggest that the alleles of *PKCdelta* is associated with the variation of CCR in the DGRP population.

Photoperiodic responses of the knockouts of the candidate genes

285

286 To further characterize the top 4 candidate genes, first, we analyzed CCR response of the knockout
 287 strains; *PKCdelta* mutant *PKCδ^{e04408}* (BDSC# 18258), *GlcAT-P* mutant *GLCAT – P^{0027–G4}*
 288 (BDSC#62582), *Bsh* mutant *Bsh^{G3075}* (BDSC#27467) and *Dh31-R1* mutant *Dh31 – R^{MB00175}*
 289 (BDSC#22718). As a control, one DGRP parent strain DGRP208 (WT) was used to compare

between the Cohen's D values. For this analysis equinox (EQ) condition, a 12:12 hr light:dark cycle was used along with SD and LD photoperiods. The CD was obtained by subtracting (or normalizing) the EQ photoperiod CCR from SD & LD CCR. The WT strain has a negative CD value of -0.69 (Figure 3A). The *Dh31-RI* mutant has a similar photoperiod response to the WT with negative CD values in both SD and LD photoperiod. *Bsh* mutant also has negative CD value though the effect is not as exaggerated. Both *GlcAT-P* and *PKCdelta;Pbac* mutants has positive CD values (Figure 3A). To determine how different these mutants' photoperiod responses were from the WT control, the CD value of the WT was subtracted from the CD values of the mutant, the final value is termed Genetic Effect Size (GES) (Figure 3B, Materials and Methods). *Bsh*, *GlcAT-P* and *PKCdelta:Pbac* have different photoperiod responses compared to the WT with *PKCdelta* being the most different. The GES of *PKCdelta:Pbac* is not significantly different between LD and SD. However, *GlcAT-P*, *Bsh*, and *Dh31-RI* have a difference in GES between LD and SD (Figure 3B).

Next, we tested the hypothesis that there are shared genetic regulators between the circadian clock and photoperiodism. We examined the period of the candidate photoperiod knock out genes using WT as a control (Figure 3C). Compared to the WT which had a period of 24 hr, *GlcAT-P* and *PKCdelta:MIC* mutants have a significantly longer period (t-test: $p < 0.001$, $p < 0.01$). We concluded that *GlcAT-P* and *PKCdelta:MIC* genes play a role in both circadian clock regulation and photoperiodism. We decided to investigate *PKCdelta* further for three reasons. The two *PKCdelta* mutant alleles showed a significantly different photoperiodic CCR response from that of the WT (Figure 3B), one of the allelic mutants of *PKCdelta* had an altered circadian period (Figure 3C), and GWAS analysis showed that there are two independent SNP variants associated with the photoperiodic CCR response for this gene (Figure 2 and File S3).

313

314 *PKCdelta* Candidate Gene

315 *PKCdelta* is a kinase known to be involved in cellular signaling pathways like cell growth,
 316 apoptosis, tumor promotion and carcinogenesis (Suh et al., 2003). We examined two parental
 317 DGRP strains one with a large positive CD value (DGRP379) and another with a large negative
 318 CD value (DGRP391). *PKCdelta* transcribes three transcripts, RC, RD, and RE (Figure S4). The
 319 transcripts RC and RD are similar to each other except with a slight difference in the UTR region;
 320 the RD isoform has a shorter UTR region than the RC isoform. The RC and RD transcript have
 321 regulatory domain which contain C2, C1 and pseudosubstrate domains toward the N-terminal.
 322 Toward the C-terminal there is a catalytic kinase domain which is shared with all three isoforms
 323 (Figure S5). There are two different transcription start sites, one that starts the beginning of
 324 isoforms RD and RC, and the other starts at the RE isoform. The RE isoform is a nested gene with
 325 its transcription start site residing in one of the introns of RC and RD transcripts (Kumar, 2009).

326 We found that there are two non-synonymous SNPs in the exon 6 which resides in the RE
 327 transcript in the DGRP379 strain but not in DGRP391. The first SNP position is at 12445745:
 328 Guanine to Adenosine, which changes Alanine to Threonine. The second SNP position is at
 329 12446453: Guanine to Cytosine, which changes Alanine to Proline. The first SNP position is in a
 330 ployQ region, which is conserved in the *Drosophila* species. The second SNP causes a repeat of
 331 the sequence surrounding the region (Figure S6A, S6B). We found that RE isoform is a nested
 332 gene which is only conserved in the *Drosophila* genus (Figure S7).

333

334 *Diapause and CCR in PKCdelta* mutant alleles

Figure 4. The photoperiod and period response of *PKCdelta*:MIC and *PKCdelta*:Pbac mutants compared to WT. A. The diapause response of candidate gene *PKCdelta* mutants compared to WT in LD (open) vs SD (filled). There is a significant difference in *PKCdelta*:Pbac mutant diapause response, compared to WT in both photoperiods (Z-test: $p = 0.0014$ (LD), 0.010 (SD)). *PKCdelta*:MIC diapause response compared to WT in each photoperiod is not significantly different (Z-test: $p = 0.27$ (LD), 0.30 (SD)). However, there is a significant difference in *PKCdelta*:Pbac LD and SD diapause response (Z-test $p = 0.002$) and in WT (Z-test: $p = 0.007$),

but not in *PKCdelta*:MIC (Z-test: $p = 0.29$). B. Cohen's D of *PKCdelta*:Pbac and *PKCdelta*:MIC compared to WT, the CD values were obtained by subtracting EQ CCR response from LD (open) and SD (filled) photoperiod response. WT has a negative CD, while *PKCdelta*:MIC and *PKCdelta*:Pbac are positive. C. GES of the mutants compared to the WT, both mutant alleles show significantly different photoperiod response from that of the WT, LD (open), SD (filled). D. Circadian period of the *PKCdelta* mutants compared to that of WT separated by homozygous females (open) and hemizygous males (filled). Error bars represent SD, $10 < n < 20$, homozygous vs hemizygous (t-test, $p > 0.05$).

335

336 To further explore the role of *PKCdelta* in photoperiodism, we performed the diapause assay using
337 the *PKCdelta*:Pbac and *PKCdelta*:MIC allelic mutants under different photoperiodic conditions.
338 The *PKCdelta*:Pbac mutant has an insertion of a *piggy bac* element in an intron close to the 3' end
339 of the gene. The *PKCdelta*:MIC mutant has an insertion in the 5' end of the gene and contains a
340 minos-based construct which was created by the Gene Disruption Project (GDP) (Venken et al.,
341 2011) (Table S4). We found that both mutant alleles have a mutation in the catalytic domain of the
342 enzyme *PKCdelta* (Figure S9, Table S3 and S4)

343 Using WT as a control, the diapause assay was carried out under SD and LD conditions
344 (Figure 4A). As expected, the WT has increased diapause incidences under SD conditions
345 compared to the LD condition (Z-test, $p < 0.05$). There was a significant difference in
346 *PKCdelta*:Pbac mutant's diapause response compared to that of WT in both photoperiods (Z-test,
347 $p = 0.0014$ (LD), $p = 0.01$ (SD)). The *PKCdelta*:MIC mutant's diapause response compared to WT
348 in each photoperiod is not significantly different (Z-test, $p = 0.27$ (LD), $p = 0.30$ (SD)). However,
349 there is a significant difference in *PKCdelta*:Pbac LD and SD diapause response (Z-test, $p =$

0.002), and in WT (Z-test, $p = 0.007$), but not in *PKCdelta*:MIC (Z-test, $p = 0.29$) (Figure 4A). The diapause response in *PKCdelta*:MIC mutant is not statistically significant from that of the WT. We interpreted this as the mutant allele *PKCdelta*:MIC cannot distinguish SD from LD because the diapause response between photoperiods is not significantly different (Z-test: $p = 0.29$).

To determine the necessity of *PKCdelta* for the photoperiodic CCR response, we preformed the CCR assay using the allelic mutants of *PKCdelta* (Figure 4B). The CD values of *PKCdelta*:Pbac and *PKCdelta*:MIC are positive compared to WT. However, *PKCdelta*:MIC have lower CD values from that of *PKCdelta*:Pbac. The GES of the mutants compared to the WT show that *PKCdelta*:Pbac have significantly different photoperiod response (Figure 4C).

Since *PKCdelta* is located on the X-chromosome, both of the *PKCdelta* mutants contain hemizygous males and homozygous females. The CCR was performed to determine whether there will be a difference in photoperiod response between hemizygous and homozygous flies under SD and LD photoperiod. Both *PKCdelta*:MIC and *PKCdelta*:Pbac mutants show a difference in photoperiod response between hemizygous and homozygous flies (Figure S8). To determine if there was a difference in period phenotype between hemizygous and homozygous flies in the *PKCdelta*:Pbac and MIC mutants, we performed the period assay. We found no significant difference between the period phenotype between hemizygous and homozygous flies (Figure 4D). WT flies have an expected period of 24 hr with no difference between hemizygous and homozygous flies (t-test, $p > 0.05$). *PKCdelta*:MIC hemizygous flies have a period of 30 hr and homozygous flies have a period of 29 hr but there was no statistically significant difference between hemizygous and homozygous flies (t-test, $p > 0.05$). There was no significant difference between hemizygous and homozygous flies of *PKCdelta*:Pbac in the circadian period (t-test, $p >$

0.05). Based on these data together, we concluded that MIC mutant allele affected the ability to distinguish between LD from SD, and also altered the circadian period.

Next, we predicted that *PKCdelta*:Pbac has knockdown of all three transcripts of *PKCdelta*, while *PKCdelta*:MIC has knockdown of RC & RD based on the reported locations of the insertion. The RE isoform is believed to be mainly responsible for photoperiod response because the *PKCdelta*:MIC mutant shows a decrease in the photoperiod response in the CCR and diapause assays. We also predicted that the transcripts RC and RD could play a role in interpreting photoperiod response, though they might not be major players in interpreting photoperiods. These predictions were then confirmed by performing a qPCR experiment using the *PKCdelta*:MIC and *PKCdelta*:Pbac mutants under SD and LD condition (Fig.5).

Expression of *PKCdelta*

is significantly reduced compared to WT (n=3-5, t-test, $p < 0.001$). However, there is a reduction of RE transcript under LD compared to SD in *PKCdelta*:MIC (n=3-5, t-test, $p < 0.001$). In the *PKCdelta*:Pbac background, the expression of RD and RE is overall reduced compared to WT (n=3-5, t-test, $p < 0.05$). There was no statistical difference in the expression levels under LD and SD photoperiods (n=3-5, t-test, $p > 0.05$).

385

386 We measured the expression levels of *PKCdelta* under different photoperiod conditions (Materials
387 and Methods). WT flies under LD and SD photoperiods did not show statistically significant
388 change in *PKCdelta* RD and RE transcript levels (t-test, $p > 0.05$). *PKCdelta*:MIC which was
389 predicted to have reduced expression of RD and normal expression of RE, show overall decrease
390 in expression of RD in both LD and SD (t-test, $p > 0.05$). However, there is an alteration in the
391 expression levels of the RE between the photoperiod conditions in *PKCdelta*:MIC. The
392 *PKCdelta*:MIC mutant under LD show a significant decrease in its expression level from that of
393 SD (t-test, $p < 0.001$). *PKCdelta*:Pac mutant show an overall decrease in RD and RE expression
394 from that of WT (t-test $p < 0.05$). Though there is no significant difference in the expression levels
395 of RD and RE between the photoperiods in *PKCdelta*:Pac mutant, (t-test $p > 0.05$). This suggests
396 that there is an alteration in the expression of the RE transcript in the *PKCdelta*:MIC mutant due
397 to long day photoperiod condition. Based on this data and the diapause data, we concluded that
398 RD transcript could be involved in distinguishing between SD and LD. Since *PKCdelta*:MIC
399 which has reduced expression of RD transcript show no difference in the diapause response
400 between photoperiods (Figure 4A & 5). Alternatively, the *PKCdelta*:Pbac mutant which has
401 reduced expression of both transcripts shows a greater proportion of diapause response compared

to WT in both LD and SD suggesting that expression *PKCdelta* and diapause response is negatively correlated in this mutant.

Next, we reasoned that the GWAS analysis of the DGRP CCR response itself as a trait under LD and SD conditions will lead us identifying the downstream genes involved in CCR. We performed the GWAS analysis on the CCR trait as a phenotype rather than the photoperiodic response (CD values) as a trait discussed earlier (Files S4 - S7). There were six genes that were highly associated with the CCR phenotypes both in SD and LD conditions: *Sif*, *Comm2*, *Bru-3*, *Fife*, *CG32111* and *Spab*. *Sif* is involved in actin cytoskeleton organization; visual perception and it localizes to the synapse. *Comm2* is essential for nerve cord development and axon guidance. *Bru-3* is known to mediate mRNA translation repression. While *Fife* is responsible for encoding synaptic scaffolding protein, which promotes neurotransmitter release. *CG32111* is a long non-coding RNA though much information is not known about this gene. *Spab* is a protein coding gene from the neuropeptide gene group, however not much is known about this gene (Flybase.org). As expected, most of the genes responsible for the CCR response play a role in the nervous system. However, *Sif* and *CG32111* are known to be a target of the circadian clock transcription complex CLK/CYC, suggesting that CCR photoperiod response could be regulated by the clock (Abruzzi et al., 2011). *Sif* is also known to interact with *Fas2* which than interacts with *PKCdelta* (Rees et al., 2011).

Discussion

In this study, we were able to establish that there is a variation in photoperiod response measured by the CCR phenotype within the DGRP population. By calculating the effect size (Cohen's D value) for each DGRP genotype we determined how each genotype's photoperiod response was

different from one another. The GWAS analysis on CD values allowed us to narrow down the 32 photoperiod candidate genes. We chose 4 top candidate photoperiod genes for genetic analysis. The allelic mutant of *PKCdelta* showed a significantly different photoperiod response from WT. To test the necessity of the gene for the photoperiodic response, we analyzed two different mutant alleles: *PKCdelta*:Pbac and *PKCdelta*:MIC. By performing a qPCR assay under LD and SD it was confirmed that *PKCdelta*:Pbac has decreased expression of RD and RE transcripts of the gene. The *PKCdelta*:MIC mutant which show reduced expression of the RD transcript, show a photoperiod effect under LD in the RE transcript levels. The CCR assay showed that *PKCdelta*:Pbac mutants have a greater effect size compared to *PKCdelta*:MIC. Finally, to test the hypothesis whether there is an overlap in the genetic mechanisms between the circadian clock and photoperiodism, the circadian periods of our candidate genes were analyzed. *GlcAT-P* knockout and *PKCdelta*:MIC mutants had a significantly longer period phenotype from the WT. Our data supports the view that *PKCdelta* is involved in both circadian clock and photoperiodism.

Diapause and CCR are two different output pathways that are under the photoperiod regulation in insects. Our data showed that the two different mutant alleles of *PKCdelta* manifested the altered photoperiod response in different ways. The flies with *PKCdelta*:MIC allele could not distinguish LD from SD, whereas the flies with *PKCdelta*:Pbac allele had an elevated instances of diapause under both photoperiods (Figure 4A). The CCR data suggests that the *PKCdelta*:Pbac mutation plays a role in increasing the sensitivity to photoperiod in the flies overall. Both the mutants have positive CD values indicating their recovery time was faster than that of WT. While *PKCdelta*:Pbac compared to *PKCdelta*:MIC has a greater effect size (Figure 4B).

Taken together, we have a working model for the photoperiodism in *D. melanogaster*. The photoperiod candidate gene *PKCdelta* is hypothesized to be regulated by the clock. There are

certain genetic components that are under clock control which interact with *PKCdelta*. The circadian transcription complex CLK/CYC is known to directly regulates the *Mef2* transcription which then has been identified to target *Fas2*. *Mef2* is known for its role in myogenesis and development (Sivachenko et al., 2013). However, *Mef2* is shown to be present in clock neurons and fluctuation in *Mef2* are rhythmic (Sivachenko et al., 2013). *Mef2* is a major component of the CLK/CYC, which is important for circadian behavior of neuronal morphology. A Chip-Chip analysis determined that *Fascilin2* (*Fas2*) is a direct target of *Mef2* (Sivachenko et al., 2013). *Fas2* is known to be involved in neural cell adhesion, and there is a rhythmic oscillation of *Fas2* in pigment-dispersing factor (PDF) cells types, suggesting that *Fas2* is under circadian clock control (Sivachenko et al., 2013). *PKCdelta* is proposed to interact with *Fas2* (Rees et al., 2011). PDF is known to effect *Drosophila* circadian clock where certain mutations in PDF can shorten the period of certain clock neurons while lengthening the period others neurons (Wülbeck et al., 2008). The diapause assay showed that *PKCdelta*:MIC mutants lose the ability to distinguish between day-lengths. *PKCdelta*:MIC mutant's loss of photoperiod sensitivity is similar to the well-studied *PDF*⁰¹ mutant (Ojima et al., 2018). Our data reject the model that circadian clock and photoperiodism are separate entities with no shared regulations. The current study sheds light onto potential genes that could be further studied to characterize the mechanisms of photoperiodism, and its relationship to the circadian clock.

Data Availability

All DGRP lines are available at the Bloomington Drosophila Stock Center, and the genotype information of all lines are available at <http://dgrp2.gnets.ncsu.edu>. The raw CCR and CD data for all DGRP lines are available under File S1. The Files S2, S4, & S6 contain the top GWAS

associations along with qq-plots, and input data used to obtain the GWAS results from CD, LD, and SD data sets. The Files S3, S5, & S7 contain all associations found in the GWAS study using the CD, LD and SD data sets. All Supplemental Materials are available at figShare portal, [gsajournals.figshare.com](https://www.figshare.com).

Acknowledgements

The authors are grateful for Drs. Nir Yakoby, Nathan Fried, and Charot Rodeget for critical reading and insightful suggestions. We are grateful for the Lee lab members, especially Sienna Casciato and Somya Patro, for helpful discussions. The authors would also like to thank Cody Stevens for helping us navigate through some of the challenging aspects of the project. The work was funded by the NIH MARC U-STAR, T34GM127154.

Supplemental Material

Figure S1. The distribution of core 38 set DGRP parent strains.

Figure S2. Sex difference in CCR assay among the select DGRP strains.

Figure S3. The Gene enrichment results of the 32 hits using flymodEnrichr.

Figure S4. Diagram of the two *PKCdelta* mutant alleles, Pbac and MIC.

Figure S5. Overview of the *PKCdelta* gene structure and the position of the oligos for PCR products.

Figure S6. Sequence variations in *PKCdelta* RE transcript.

Figure S7. Phylogenetic analysis of *PKCdelta* RE transcript.

Figure S8. Cohen's D and Genome Effect Size of *PKCdelta* mutants.

Figure S9. The regulatory and catalytic domains of *PKCdelta*.

- 494 **Table S1.** The top 32 hits of the GWAS analysis sorted by the significant p values.
- 495 **Table S2.** Oligos used in the current study.
- 496 **Table S3.** Predicted changes *PKCdelta*:Pbac mutant.
- 497 **Table S4.** Predicted changes in *PKCdelta*:MIC mutant.
- 498 **File S1.** Raw CCR and Cohen's D data for all the DGRP lines. It's divided into three tabs;
- 499 Cohen's values, CCR in LD and CCR in SD.
- 500 **File S2.** TOP GWAS hits along with images of QQ plots.
- 501 **File S3.** SNP associations found in the GWAS study.
- 502 **File S4.** TOP GWAS hits using the CCR data in LD.
- 503 **File S5.** All GWAS hits using the CCR data in LD.
- 504 **File S6.** TOP GWAS hits using the CCR data in SD.
- 505 **File S7.** All GWAS hits using the CCR data in SD.

506

507

508 **References**

- 509 Abruzzi, K.C., Rodriguez, J., Menet, J.S., Desrochers, J., Zadina, A., Luo, W., Tkachev, S., and
- 510 Rosbash, M. (2011). *Drosophila* CLOCK target gene characterization: implications for
- 511 circadian tissue-specific gene expression. *Genes & Development* 25, 2374–2386.
- 512 Bradshaw, W.E., and Holzapfel, C.M. (2010). What Season Is It Anyway? Circadian Tracking
- 513 vs. Photoperiodic Anticipation in Insects. *J. Biol. Rhythms* 25, 155–165.
- 514 Carstensen, M.B., Hertz, H., Bering, T., Møller, M., Rohde, K., Klein, D.C., Coon, S.L., and
- 515 Rath, M.F. (2019). Circadian regulation and molecular role of the *Bsx* homeobox gene in
- 516 the adult pineal gland. *J. Pineal Res.*

517 Chiu, J.C., Low, K.H., Pike, D.H., Yildirim, E., and Edery, I. (2010). Assaying Locomotor
518 Activity to Study Circadian Rhythms and Sleep Parameters in *Drosophila*. *JoVE J. Vis.*
519 *Exp.* e2157.

520 Dubowy, C.M., and Cavanaugh, D.J. (2014). Sleep: A Neuropeptidergic Wake-Up Call for Flies.
521 *Curr. Biol.* 24, R1092–R1094.

522 Harbison, S.T., Kumar, S., Huang, W., McCoy, L.J., Smith, K.R., and Mackay, T.F.C. (2019).
523 Genome-Wide Association Study of Circadian Behavior in *Drosophila melanogaster*.
524 *Behav. Genet.* 49, 60–82.

525 Hosler, J.S., Burns, J.E., and Esch, H.E. (2000). Flight muscle resting potential and species-
526 specific differences in chill-coma. *J. Insect Physiol.* 46, 621–627.

527 Huang, W., Massouras, A., Inoue, Y., Peiffer, J., Ramia, M., Tarone, A.M., Turlapati, L.,
528 Zichner, T., Zhu, D., Lyman, R.F., et al. (2014). Natural variation in genome architecture
529 among 205 *Drosophila melanogaster* Genetic Reference Panel lines. *Genome Res.* 24,
530 1193–1208.

531 Johnson, C.H., Zhao, C., Xu, Y., and Mori, T. (2017). Timing the day: what makes bacterial
532 clocks tick? *Nat Rev Microbiol* 15, 232–242.

533 Johansson, M., and Staiger, D. (2015). Time to flower: interplay between photoperiod and the
534 circadian clock. *Journal of Experimental Botany* 66, 719–730.

535 Kim, B.-T., Tsuchida, K., Lincecum, J., Kitagawa, H., Bernfield, M., and Sugahara, K. (2003).
536 Identification and Characterization of Three *Drosophila melanogaster*
537 Glucuronyltransferases Responsible for the Synthesis of the Conserved
538 Glycosaminoglycan-Protein Linkage Region of Proteoglycans. *J. Biol. Chem.* 278, 9116–
539 9124.

540 Kumar, A. (2009). An Overview of Nested Genes in Eukaryotic Genomes. *Eukaryot. Cell* 8,
541 1321–1329.

542 Lanciani, C.A., Lipp, K.E., and Giesel, J.T. (1992). The effect of photoperiod on cold tolerance
543 in *Drosophila melanogaster*. *Journal of Thermal Biology* 17, 147–148.

544 Larkin A, Marygold SJ, Antonazzo G, Attrill H, dos Santos G, Garapati PV, Goodman JL,
545 Gramates LS, Millburn G, Strelets VB, Tabone CJ, Thurmond J and the FlyBase
546 Consortium (2021) FlyBase: updates to the *Drosophila melanogaster* knowledge
547 base. *Nucleic Acids Res.* 49(D1) D899–D907

548 Lees, A.D. (1960). Some Aspects of Animal Photoperiodism. *Cold Spring Harb. Symp. Quant.*
549 *Biol.* 25, 261–268.

550 Mackay, T.F.C., Richards, S., Stone, E.A., Barbadilla, A., Ayroles, J.F., Zhu, D., Casillas, S.,
551 Han, Y., Magwire, M.M., Cridland, J.M., et al. (2012). The *Drosophila melanogaster*
552 Genetic Reference Panel. *Nature* 482, 173–178.

553 Martinez-Bakker, M., Bakker, K.M., King, A.A., and Rohani, P. (2014). Human birth
554 seasonality: latitudinal gradient and interplay with childhood disease dynamics. *Proc. R.*
555 *Soc. B Biol. Sci.* 281, 20132438.

556 Maruani, J., Anderson, G., Etain, B., Lejoyeux, M., Bellivier, F., and Geoffroy, P.A. (2018). The
557 neurobiology of adaptation to seasons: Relevance and correlations in bipolar disorders.
558 *Chronobiol. Int.* 35, 1335–1353.

559 Mukai, A., and Goto, S.G. (2016). The clock gene period is essential for the photoperiodic
560 response in the jewel wasp *Nasonia vitripennis* (Hymenoptera: Pteromalidae). *Applied*
561 *Entomology and Zoology* 51, 185–194.

562 Nagy, D., Andreatta, G., Bastianello, S., Martín Anduaga, A., Mazzotta, G., Kyriacou, C.P., and
563 Costa, R. (2018). A Semi-natural Approach for Studying Seasonal Diapause in
564 *Drosophila melanogaster* Reveals Robust Photoperiodicity. *J. Biol. Rhythms* 33, 117–
565 125.

566 Ojima, N., Hara, Y., Ito, H., and Yamamoto, D. (2018). Genetic dissection of stress-induced
567 reproductive arrest in *Drosophila melanogaster* females. *PLOS Genetics* 14, e1007434.

568 Ortavant, R., Mauleon, P., and Thibault, C. (1964). Photoperiodic Control of Gonadal and
569 Hypophyseal Activity in Domestic Mammals. *Ann. N. Y. Acad. Sci.* 117, 157–193.

570 Pegoraro, M., Gesto, J.S., Kyriacou, C.P., and Tauber, E. (2014). Role for circadian clock genes
571 in seasonal timing: testing the Bünning hypothesis. *PLoS Genet.* 10, e1004603.

572 Saunders, D.S. (1990). The Circadian Basis of Ovarian Diapause Regulation in *Drosophila*
573 *melanogaster*: Is the period Gene Causally Involved in Photoperiodic Time
574 Measurement? *J. Biol. Rhythms* 5, 315–331.

575 Saunders, D.S. (2012). Insect photoperiodism: seeing the light. *Physiol. Entomol.* 37, 207–218.

576 Saunders, D.S., Henrich, V.C., and Gilbert, L.I. (1989). Induction of diapause in *Drosophila*
577 *melanogaster*: photoperiodic regulation and the impact of arrhythmic clock mutations on
578 time measurement. *Proc. Natl. Acad. Sci.* 86, 3748–3752.

579 Sawa, M., Kay, S.A., and Imaizumi, T. (2008). Photoperiodic flowering occurs under internal
580 and external coincidence. *Plant Signal. Behav.* 3, 269–271.

581 Shieh, B.-H., Parker, L., and Popescu, D. (2002). Protein Kinase C (PKC) Isoforms in
582 *Drosophila*. *J. Biochem. (Tokyo)* 132, 523–527.

583 Song, Y.H., Ito, S., and Imaizumi, T. (2010). Similarities in the circadian clock and
584 photoperiodism in plants. *Curr Opin Plant Biol* 13, 594–603.

585 Song, Y.H., Kubota, A., Kwon, M.S., Covington, M.F., Lee, N., Taagen, E.R., Laboy Cintrón, D.,
586 Hwang, D.Y., Akiyama, R., Hodge, S.K., et al. (2018). Molecular basis of flowering under
587 natural long-day conditions in *Arabidopsis*. *Nat. Plants* 4, 824–835.

588 Sivachenko, A., Li, Y., Abruzzi, K.C., and Rosbash, M. (2013). The transcription factor Mef2, links the
589 *Drosophila* core clock to Fas2, neuronal morphology and circadian behavior. *Neuron* 79.

590 Suh, K.S., Tatunchak, T.T., Crutchley, J.M., Edwards, L.E., Marin, K.G., and Yuspa, S.H.
591 (2003). Genomic structure and promoter analysis of PKC- δ . *Genomics* 82, 57–67.

592 Tauber, E., and Kyriacou, B.P. (2001). Insect Photoperiodism and Circadian Clocks: Models and
593 Mechanisms. *J. Biol. Rhythms* 16, 381–390.

594 Vaziri, P., Ryan, D., Johnston, C.A., and Cripps, R.M. (2020). A Novel Mechanism for
595 Activation of Myosin Regulatory Light Chain by Protein Kinase C-Delta in *Drosophila*.
596 *Genetics* 216, 177–190.

597 Venken, K.J.T., Schulze, K.L., Haelterman, N.A., Pan, H., He, Y., Evans-Holm, M., Carlson,
598 J.W., Levis, R.W., Spradling, A.C., Hoskins, R.A., et al. (2011). MiMIC: a highly
599 versatile transposon insertion resource for engineering *Drosophila melanogaster* genes.
600 *Nat. Methods* 8, 737–743.

601 Wülbeck, C., Grieshaber, E., and Helfrich-Förster, C. (2008). Pigment-Dispersing Factor (PDF)
602 Has Different Effects on *Drosophila*'s Circadian Clocks in the Accessory Medulla and in
603 the Dorsal Brain. *J. Biol. Rhythms* 23, 409–424.

604

605

A

B

C