

Pegoraro, M, Rezek, ES, Fishman, B and Tauber, E

Nucleotide Variation in *Drosophila* cryptochrome Is Linked to Circadian Clock Function: An Association Analysis

<http://researchonline.ljmu.ac.uk/id/eprint/16336/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Pegoraro, M, Rezek, ES, Fishman, B and Tauber, E (2022) Nucleotide Variation in *Drosophila* cryptochrome Is Linked to Circadian Clock Function: An Association Analysis. *Frontiers in Physiology*, 13. ISSN 1664-042X

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk



Nucleotide Variation in *Drosophila* *cryptochrome* Is Linked to Circadian Clock Function: An Association Analysis

Mirko Pegoraro^{1,2}, Emily Sayegh Rezek³, Bettina Fishman³ and Eran Tauber^{2,3*}

¹School of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, United Kingdom,

²Department of Genetics and Biology, University of Leicester, Leicester, United Kingdom, ³Department of Evolutionary and Environmental Biology, Institute of Evolution, University of Haifa, Haifa, Israel

OPEN ACCESS

Edited by:

Charalambos P. Kyriacou,
University of Leicester,
United Kingdom

Reviewed by:

Rodolfo Costa,
University of Padua,
Italy
John Ewer,
Universidad de Valparaíso, Chile

*Correspondence:

Eran Tauber
etauber@univ.haifa.ac.il

Specialty section:

This article was submitted to
Chronobiology,
a section of the journal
Frontiers in Physiology

Received: 22 September 2021

Accepted: 12 January 2022

Published: 17 February 2022

Citation:

Pegoraro M, Sayegh Rezek E,
Fishman B and Tauber E (2022)
Nucleotide Variation in *Drosophila*
cryptochrome Is Linked to Circadian
Clock Function: An Association
Analysis.
Front. Physiol. 13:781380.
doi: 10.3389/fphys.2022.781380

Cryptochrome (CRY) is a conserved protein associated with the circadian clock in a broad range of organisms, including plants, insects, and mammals. In *Drosophila*, *cry* is a pleiotropic gene that encodes a blue light-dedicated circadian photoreceptor, as well as an electromagnetic field sensor and a geotaxis behavior regulator. We have generated a panel of nearly-isogenic strains that originated from various wild populations and which carry different natural alleles of *cry*. Sequencing of these alleles revealed substantial polymorphism, the functional role of which was elusive. To link this natural molecular diversity to gene function, we relied on association mapping. Such analysis revealed two major haplogroups consisting of six linked nucleotides associated with circadian phase (haplotypes All1/All2). We also generated a maximum-likelihood gene-tree that uncovered an additional pair of haplogroups (B1/B2). Behavioral analysis of the different haplotypes indicated significant effect on circadian phase and period, as well on the amount of activity and sleep. The data also suggested substantial epistasis between the All and B haplogroups. Intriguingly, circadian photosensitivity, assessed by light-pulse experiments, did not differ between the genotypes. Using CRISPR-mediated transgenic flies, we verified the effect of B1/B2 polymorphism on circadian phase. The transgenic flies also exhibited substantially different levels of *cry* transcription. We, moreover, analyzed the geographical distribution of the B1/B2 haplotypes, focusing on a 12 bp insertion/deletion polymorphism that differentiates the two haplotypes. Analysis of *cry* sequences in wild populations across Europe revealed a geographical cline of B1/B2 indel frequency, which correlated with seasonal bioclimatic variables. This spatial distribution of *cry* polymorphism reinforces the functional importance of these haplotypes in the circadian system and local adaptation.

Keywords: association mapping, circadian clock, cryptochrome, *Drosophila*, genetic cline, genetic variation, molecular polymorphism

INTRODUCTION

The circadian clock system is one of the best examples of a molecular circuit that can generate an organized output without any external cues, yet which is amenable to environmental modulators, such as temperature and light. Entrainment of the pacemaker to daily temperature or light-dark cycles adjusts the period to 24h and sets the system to a phase, which is ecologically adaptive (Johnson et al., 2003). Given the broad range of temperature and light conditions to which wild populations are exposed, it is not surprising that molecular adaptations in circadian clock genes are pervasive and accordingly, have been reported in various phyla, such as plants (Endo, 2016), insects (Numata et al., 2015), and vertebrates (Johnsen et al., 2007). Furthermore, temperature and light vary systematically with latitude and consequently may lead to latitudinal clines in circadian clock phenotypes and in molecular polymorphisms in clock genes (Kyriacou et al., 2008; Hut et al., 2013). The threonine-glycine (TG) repeat polymorphism within the *Drosophila period* (*per*) gene was the first such cline to be described (Costa et al., 1992). In northern Europe, the frequency of the two major alleles, TG17 and TG20, follows a cline. Specifically, TG20 is more abundant at high latitudes, whereas TG17 is more common in the south. The possibility of a similar cline in Australia has been debated (Weeks et al., 2006; Kyriacou et al., 2007). Laboratory experiments showed that TG polymorphism is associated with thermal adaptation (Sawyer et al., 1997). Flies carrying the TG17 allele have a free-running period (FRP) very close to 24h at high temperatures and are, therefore, well adapted to the warmer regions of Europe. Flies with the TG20 variant show superior temperature-compensation, which fits well with the higher thermal variability seen in the northern latitudes.

A polymorphism in the *timeless* (*tim*) gene also follows a latitudinal cline (Tauber et al., 2007). This polymorphism consists of a single base insertion/deletion in the 5' UTR region located between two translation start codons. The allele with the deletion yields only the short isoform (from the downstream ATG) and is called *s-tim*, whereas the allele with the insertion is translated to both the long and short isoforms (*ls-tim*). Rather than affecting thermal adaptation, the *tim* polymorphism has been shown to impact circadian photosensitivity and photoperiodism (Sandrelli et al., 2007). The frequency of *s-tim* in European populations is higher in northern regions, whereas *ls-tim* prevails in southern regions. However, analysis of the variation pattern suggested that cline can explain the fact that *ls-tim* is the derived allele that has recently emerged in southern Italy and is currently spreading across Europe by directional selection (Tauber et al., 2007; Zonato et al., 2018). Interestingly, the study of *tim* polymorphism was vital for the discovery of another protein named JETLAG, an F-box protein, which binds to TIM and induces TIM ubiquitination (Koh et al., 2006). The phenotype of JETLAG mutants, which are rhythmic in continuous light, is expressed only in strains carrying the *ls-tim* allele and not in those carrying the *s-tim* allele (Peschel et al., 2006). These discoveries

demonstrate how understanding natural genetic variation in clock genes can help explain the evolution and molecular mechanisms that drive the clock.

Cryptochrome (CRY) is an evolutionary conserved blue-light photoreceptor associated with light entrainment of the circadian clock in *Drosophila* (Emery et al., 1998; Thompson and Sancar, 2002). Research in recent years suggested that *cry* is a pleiotropic gene involved in multiple cellular processes. Light-activated CRY targets clock proteins for proteasomal degradation (Naidoo et al., 1999; Lin et al., 2001; Busza et al., 2004). It also drives membrane depolarization in clock neurons (l-LNVs) through modulation of potassium ion channels (Fogle et al., 2011). Additionally, CRY is involved in the response of *Drosophila* to low-intensity electromagnetic fields (Gegear et al., 2008; Yoshii et al., 2009; Fedele et al., 2014) and in the negative geotaxis exhibited by adult flies (Toma et al., 2002; Fedele et al., 2014).

We previously carried out a survey of natural allelic variation in *cry* (Pegoraro et al., 2014). The screen, based on a small number of coding DNA sequences, focused on the widespread missense single-nucleotide polymorphism (SNP) L232H. The intermediate frequencies of the two alleles across many wild populations suggested that the variation observed was driven by balancing selection. The pattern of variation associated with this SNP and the high level of linkage disequilibrium in the *cry* locus supported this interpretation. We further identified circadian phenotypes associated with the L232H SNP, namely, those in which the phase of locomotor activity and adult eclosion were affected.

In the present study, we generated a panel of 33 near-isogenic lines (NIL) carrying *cry* alleles that were introgressed from various wild population strains. These lines were generated by backcrossing wild strains to deficiency line for eight generations and therefore theoretically share 99.79% of their genomes, excluding the *cry* locus (although the possibility remains that residual genetic variability in other loci also contributes to the phenotypes of interest). We used these lines for fine association mapping of polymorphic sites in the gene with circadian clock properties.

MATERIALS AND METHODS

Fly Strains

To generate *Drosophila melanogaster* isogenic lines with natural *cry* alleles, (P) single males were isolated from 33 isofemales lines (Supplementary Table S1) and crossed with deficiency [Df(3R)Dl-BX12, *ss*¹ *e*⁴ *ro*¹/TM6B, *Tb*¹; stock number 3012, cytological position 91F1-2; 92D3-6] virgin females. The deficiency overlaps with the *cry* gene (cytological position 91F11). A single non-*Tb* virgin female (F1) was then isolated and crossed with Df males. This back-cross was repeated for eight generations (F1–F8). To remove the Df from the isogenic line, *Tb* males and females (F9) were crossed, and non-tubby pupae were isolated in the following generation (F10). Flies were maintained for the entire experiment in a 12h light/12h dark cycle at 25°C (LD 12:12). Lines were maintained in plastic vials (10×2 cm) in standard cornmeal food.

Cry Allele Sequencing

DNA was isolated from 5 to 10 flies per line using 300 µl of extraction buffer (0.25 mM EDTA, pH 8, 25 mM NaCl, 10 mM Tris-HCl, pH 8) supplemented with proteinase K to a final concentration of 0.2 mg/ml. The flies were homogenized and incubated at 37°C for 45 min. Proteinase K was deactivated by incubation at 95°C for 5 min. The isolated DNA was either used immediately or stored at -20°C. The *cry* genomic region was PCR-amplified (35 cycles: 92°C for 30 s, 58°C–62°C for 30 s, 72°C for 2 min and 30 s) using a combination of different primers (**Supplementary Table S2**). The reaction mix contained the primers (0.5 µM), 4.5 µl of PCR buffer (dNTPs (1 mM each), 11 mM ammonium sulfate, 4.5 mM MgCl₂, 6.7 mM 2-mercaptoethanol, 4.4 mM EDTA, pH 8, 113 µg/ml BSA, 45 mM Tris-HCl, pH 8), 1.5 µl Taq-pfu mix (M7841 and M7741, Promega) and 2 µl of DNA in a final volume of 50 µl. The amplification was separated in a 1% agarose gel and the *cry* amplicon was gel purified using a Zymoclean Gel recovery kit (D4002) following the manufacturer's instructions. The purified bands were Sanger-sequenced (Beckman Coulter Genomics) using the primers listed in **Supplementary Table S2**. All sequences were deposited at GenBank under accession numbers MW758991–MW759024.

Polymorphism Analysis

DnaSP (Ver5.10.01) was used to analyze *cry* genetic variation (Librado and Rozas, 2009). Three lines (HOJ12_133, REN12.1, and HOJ34_111) were excluded from the polymorphism analysis because their sequence was not complete. For the McDonald Kreitman test (McDonald and Kreitman, 1991) and maximum likelihood phylogenetic analysis, *cry* genomic sequences of *Drosophila simulans* were included (Hu et al., 2013; Lack et al., 2016).

Locomotor Activity and Circadian Behavior

Locomotor activity was recorded as previously described (Rosato and Kyriacou, 2006) at 25°C. Using the TriKinetics system, we recorded the locomotor activity of 3–4 day-old males maintained initially for 3–5 days in LD 12:12, followed by 5 days in constant dark (DD) at 25°C. The flies were then re-entrained for 4 days in LD 12:12. A 20 min light pulse (saturating white light >500 lux) was given at Zeitgeber Time 15 (ZT15; 3 h after lights-off) during the last night, before releasing the flies to DD for a further 5 days.

The phase of activity in LD 12:12 was analyzed on the third day of LD 12:12. Given the bimodal locomotor activity profiles of *Drosophila*, phase was determined for both the morning and evening activity peaks. Circular statistic software (Oriana, Kovach Computing Services, United Kingdom) and the Watson-Williams *F*-test were used to analyze the phase data. To test whether population mean ranks differed, we conducted the Wilcoxon Signed-rank test and a one-way permutation test based on 9,999 Monte Carlo simulations (Wilcoxon, 1945).

The amount of sleep, defined as 5 min of inactivity or longer (Hendricks et al., 2000; Tononi, 2000), and total amount of

activity per day were also recorded during the LD 12:12. Sleep amount in LD 12:12 was analyzed separately during the 12 h of light (min/12 h) and the 12 h of dark (min/12 h) using nested ANOVA. We re-sampled across all groups to calculate 95% confidential limits for *F* scores.

The activity data collected for the first 5 days in DD were used to calculate the free-FRP and acrophase (ϕ) using cosinor-rhythmometry (Nelson et al., 1979). Changes in the FRP were analyzed using nested ANOVA and re-sampling to calculate the *F* score 95% confidential limit.

The difference between the phase on the third day after the light pulse (ϕ_{lp}) and the reference phase (ϕ_{ref} : phase of DD before the light pulse) was calculated with a custom-made Excel macro. Circular statistic software (Oriana) and a Watson-Williams *F*-test were used to analyze the phase data.

Genotype–Phenotype Association

To estimate associations between phenotypic traits and genotypes, we used TASSEL software (Trait Analysis by aSSociation, Evolution and Linkage; Bradbury et al., 2007). From the full alignment of 33 wild-type *cry* sequences, a file containing only the most common 53 SNPs and indels (present in 10 or more lines, cut-off 30%) was generated. This file was used to test associations between the different circadian phenotypes and genetic variation in *cry*. The phenotypes were represented by average values per line. The association was tested using a General Linear Model, a function that uses a fixed effects regression model. A main effect-only model was automatically built using all variables in the input data. Each trait by marker combination was tested and trait by marker *F*-tests were produced. The algorithm runs a permutation test which calculates the predicted and residual values of the reduced model, permutes the residuals (1,000 permutations) and adds them to the predicted values (Anderson et al., 2003).

CRISPR-Mediated Homology-Directed Repair

We used CRISPR-mediated homology-directed repair (HDR) to insert All2 or B1 alleles into the same genetic background (*act-cas9: y1 P(act5c-cas9, w+) M(3xP3-RFP.attP)ZH-2A w**; Bloomington 24,480), as previously described (Gratz et al., 2013). To drive the HDR, two sgRNAs for the All haplotypes and two for the B haplotypes were cloned into plasmid pCFD4 using the primers listed in **Supplementary Table S2** and a Gibson Assembly kit (E5510 NEB). The sgRNAs complemented the PAM site at positions 1,744–1,746 (5' All) and 2,108–2,110 (3' All) and 2,850–2,852 (5' B) and 3,238–2,852 (3' B). To promote HDR, the *act-cas9* line used was All1B2 and single-stranded oligonucleotides (ssODNs) encoding All2 and B1 (both synthesized by Genescript, Hong Kong) were employed. Each ssODN included 60 nucleotide homology arms left of the 5' PAM site and right to the 3' PAM site. In the ssODNs design, the sequence of the All2 and B1 haplotypes was the most common nucleotide per significant site per haplotype. For the not informative site, the *act-cas9* sequence was maintained. Injection of the sgRNA

pCDF4-expressing vectors and ssODH was done using the microinjection service of the Department of Genetics, University of Cambridge. This procedure yielded two recombinant lines (All2B2 and All1B1).

RNA Extraction and qPCR Analysis

Transgenic (All1B1, All2B2) and *act-cas 9* (All1B2) flies were collected in liquid nitrogen at ZT01 and ZT13 (1 h after light on and light off respectively). About 50 heads from each sample were used for total RNA extraction using the PureLink RNA Mini kit (Invitrogen). First-strand cDNA was synthesized using 500 ng RNA of each sample, with the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). The cDNA samples were diluted (10 \times), and 1 μ l was used in quantitative PCR (qPCR) (FastSYBRTM Green Master Mix, Thermo Fisher). The primer sequences are listed in **Supplementary Table S2**. The *Rpl32* gene served as a reference for gene expression. We used the QuantStudioTM 3 Real-Time PCR System (Applied Biosystem). A standard curve was plotted for each of the genes, and four biological replicates were analyzed for each condition. The cycle crossing point (Cp) of each reaction was calculated by QuantStudio software version 1.5.

RESULTS

Genetic Variation

Analysis of the *cry* allele sequences in 33 NIL strains revealed substantial genetic variation, including 32 synonymous and 16 non-synonymous changes (**Supplementary Table S3**). Nucleotide diversity was estimated using θ s to be 0.011 ± 0.0034 (SD) and π to be 0.0104 ± 0.00065 .

We used Tajima's D statistics (Tajima, 1989) to test for the signature of selection at this locus. A sliding window analysis revealed both negative and positive significant peaks (**Figure 1**). The first part of the sequence, in particular the first intron, showed significant negative values of Tajima's D, while significant positive peaks were present at the end of the sequence, suggesting that different selection scenarios act across the gene.

To test for adaptive evolution, we used the McDonald-Kreitman test (McDonald and Kreitman, 1991). As this test compares the ratio of non-synonymous to synonymous variation within and between different species, we compared *cry* sequences from four *D. simulans* strains. We found 127 fixed and 215 polymorphic silent changes (Ds, Ps) and one fixed and 20 polymorphic (Dn, Pn) replacement changes.

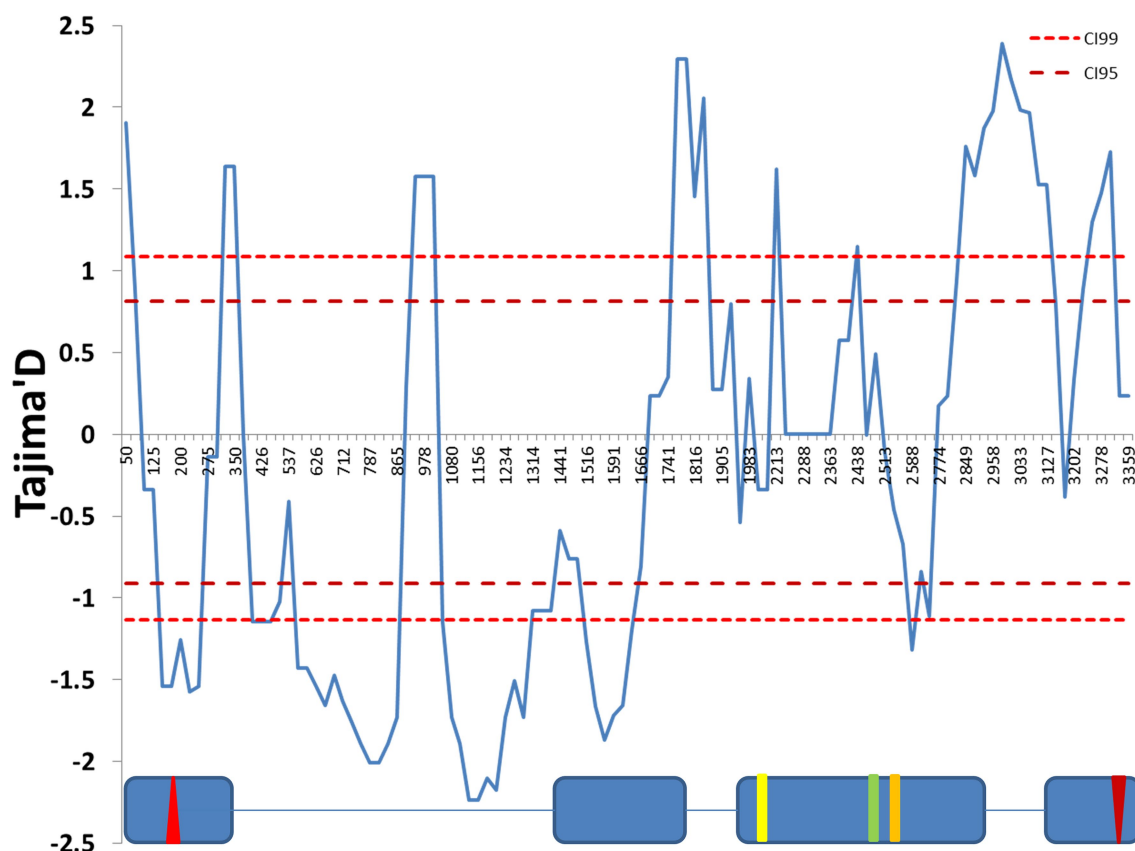


FIGURE 1 | Sliding window analysis of Tajima's D neutrality test statistics. The figure shows a sliding window of Tajima's D calculated every 100 bp, with a window size of 25 bp. The 99% and 95% CIs were produced by coalescent simulation using $\theta = 31.04$, $n = 30$, number of replications = 1,000 and estimated recombination per gene $R = 41.199$. The exon (blue boxes) and intron (lines) structure of the *cry* genomic region is depicted below. The red up- and down-oriented triangles indicate the start and stop codons, respectively. Yellow, green, and orange vertical lines indicate three non-synonymous substitutions (L232H, E335D, and D348N, respectively). Nucleotide positions are numbered beginning from 10 nucleotides before the first exon, relative to the *cry* sequence (FBgn0025680).

The neutrality index was significantly higher than unity ($Ni=11.81$, Fisher's exact test $p=0.002$), indicating departure from neutrality due to an excess of non-silent polymorphism.

We also have found an excess of linkage-disequilibrium between the 69 informative sites (70 significant pairwise comparisons, after Bonferroni correction; **Supplementary Table S4; Figure 2**). The significant linkage-disequilibrium loci were organized in four main blocks, a group of 10 sites within intron II, a small cluster of four sites around two missense mutations in exon III (E335D and D348N), a large cluster of 18 sites within intron III and a small group of three sites within exon IV, within 150 bp of the stop codon (**Supplementary Table S4; Figure 2**).

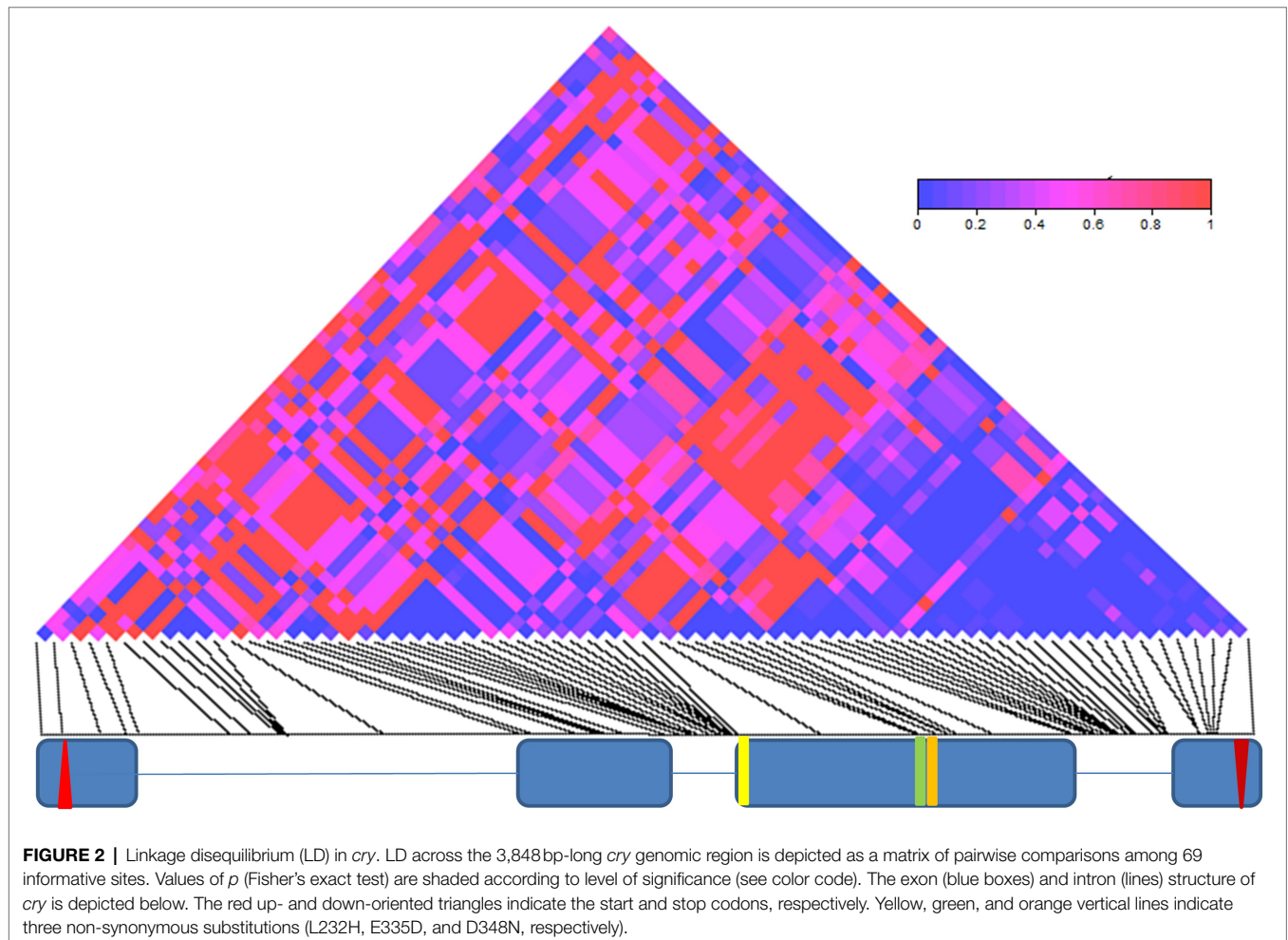
A maximum-likelihood gene tree for *cry* revealed two major clades (**Figure 3**), which define two broad haplotypes that we named B1 and B2. Nine SNPs and a single indel (12 bp-long) underlie the differences between the clades. These SNPs form an LD block (**Supplementary Table S3; Figure 2**) in the intron III-beginning of exon IV (from position 2,964 to position 3,209; **Figure 3**), including the 12 nucleotide-long insertion/deletion (indel) at position 3,091 (**Figure 3**).

Phenotypic Analysis

SNPs across the entire *cry* region were tested for association with various circadian traits. We identified a significant association between a group of six SNPs within intron II and the DD acrophase (**Figure 4**). These SNPs form a linkage cluster that defines two haplotypes, which we named All1 and All2. These haplotypes are distinct from the B1/B2 haplotypes identified in the maximum-likelihood gene tree. It is noteworthy that the Tajima D analysis indicated significant positive values in both the A and B regions, suggesting that both polymorphisms are driven by balancing selection (**Figure 1**).

We re-analyzed the phenotypic data from the NIL, grouping the lines according to both haplotypes All1/All2 and B1/B2. The FRP of All2 flies was shorter than that of All1 flies ($F_{1,1,048}=14.33$, $p=1.42E-04$), while the FRP of B1 flies was shorter than that of B2 flies (both at about 0.2h, $F_{1,1,048}=7.75$, $p=5.46E-03$; **Figure 5**). This difference increased to 0.3h when comparing the genotypes of both haplotypes ($F_{3,1,048}=6.99$, $p=1.17E-04$; **Figure 5**).

The acrophase of All1 flies was significantly delayed, as compared to that of All2 flies, while there was no significant effect seen with the B haplotypes ($F_{1,1,079}=38.91$, $p=6.35E-10$; **Figure 5**). Comparing the full haplotypes revealed intricate effects (**Figure 5**). The most advanced acrophase was presented



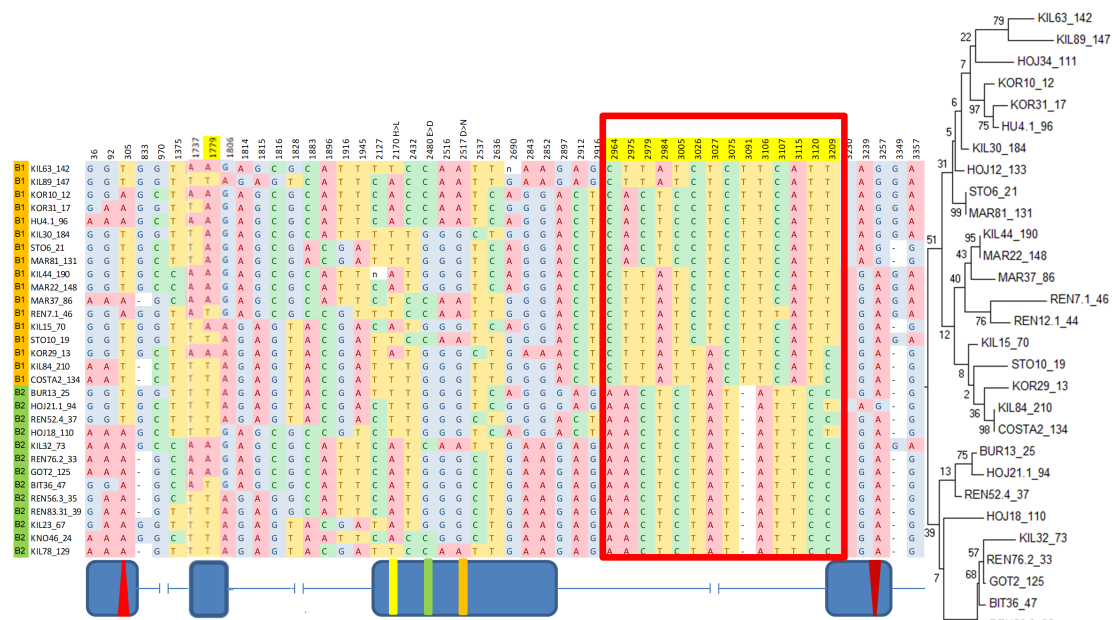


FIGURE 3 | B1/B2 haplogroups revealed by a maximum-likelihood gene tree of *cry*. Red boxes indicate the polymorphism underlying the two clusters in the phylogenetic tree. Fly strains are listed according to their gene tree position. Top branches: haplogroup B1 (orange); Bottom branches: haplogroup B2, (green). Position 3,091 represents a 12 bp indel (truncated for clarity). Shown are polymorphic sites with minor allele frequency >30%. The maximum-likelihood gene tree is shown on the right. Schematic diagram (not to scale) indicating the exon-intron is shown below.

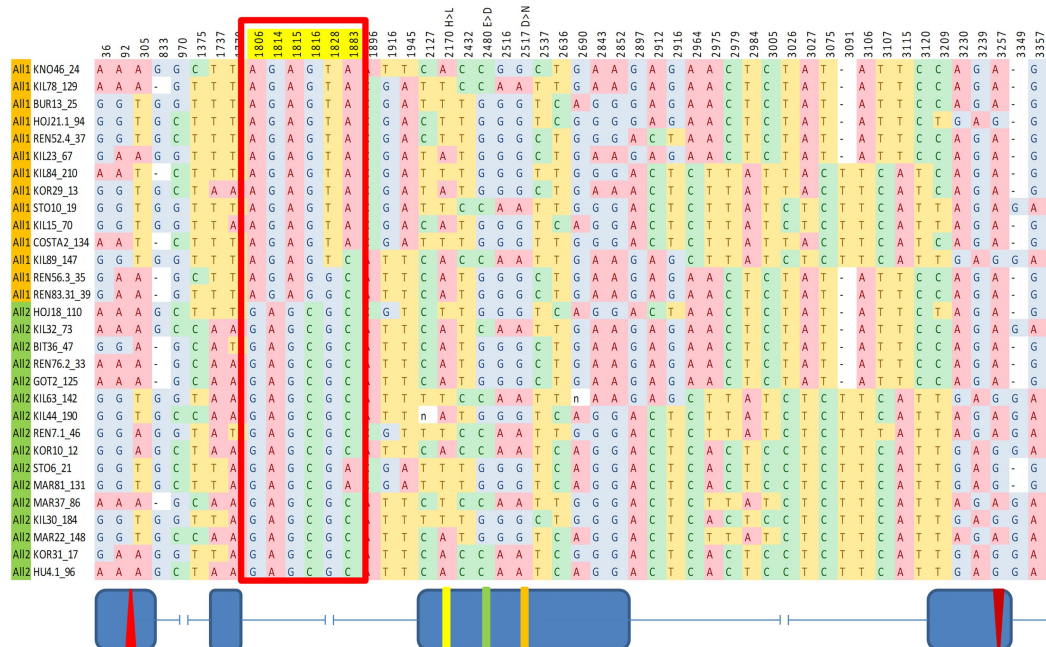


FIGURE 4 | Fine mapping of SNPs in *cry* associated with circadian acrophase. The red box indicates six SNPs showing a significant association. Fly strains are organized according to two haplotype groups (All1, orange; All2, green). These SNPs form a linkage group within intron II. Schematic diagram (not to scale) indicating the exon-intron is shown below.

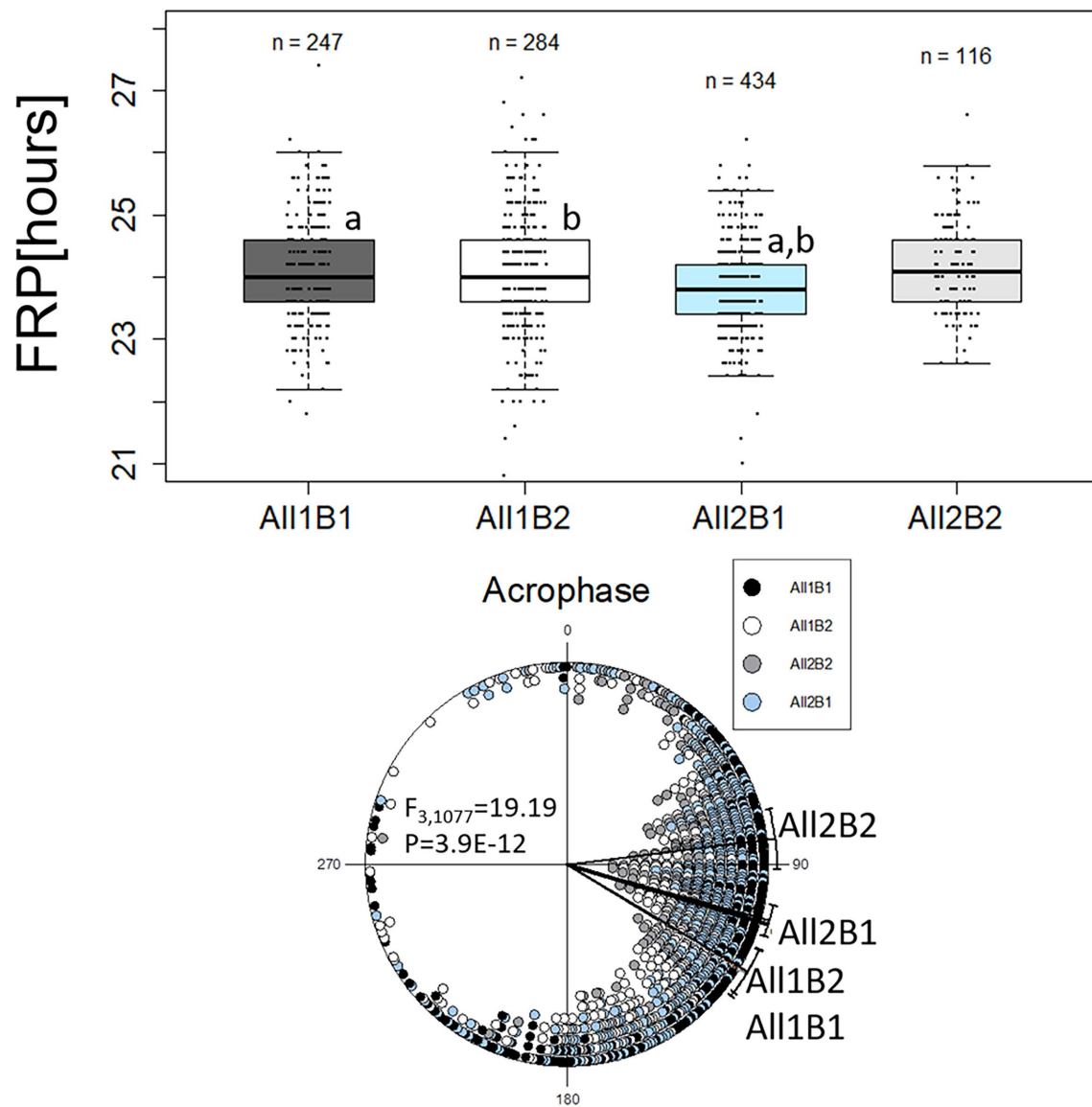


FIGURE 5 | Circadian free-running period (FRP) and acrophase of *cry* haplotypes. The FRP (**upper panel**) is depicted using a boxplot, where the middle line represents the median, and the box indicated the interquartile range (IQR). Whiskers signify $1.5 \times \text{IQR}$. White circles are potential outliers. Similar letters on boxplots indicate significant differences revealed by *post-hoc* tests (a, b, c, $p < 0.05$). FRP: All1B1 $24 \text{ h} \pm 0.07 \text{ h}$ (median \pm SEM), All1B2 $24 \text{ h} \pm 0.1 \text{ h}$, All2B1 $23.8 \text{ h} \pm 0.03 \text{ h}$ and All2B2 $24.1 \text{ h} \pm 0.07 \text{ h}$. The circadian acrophase (phase in DD; **lower panel**) was calculated over 5 days and is depicted by circular plots. Each symbol represents a single observation. 0° and 180° correspond to subjective light on (CT0) and light off (CT12), respectively ($15^\circ = 1 \text{ h}$). Acrophases: All1B1 $121.16^\circ \pm 57.56^\circ$ (mean vector \pm circular standard deviation, $n=247$); All1B2 $120.61^\circ \pm 56.17^\circ$ ($n=284$); All2B1 $105.46^\circ \pm 46.50^\circ$ ($n=434$); and All2B2 $82.91^\circ \pm 46.21^\circ$ ($n=116$).

by All2B2 flies that was significantly different from all other haplotypes. Acrophase in All2B2 flies was more advanced than that of All2B1 ($F_{1,548} = 21.17$, $p = 5.22\text{E-}06$), All1B2 ($F_{1,398} = 41.7$, $p = 3.10\text{E-}10$) and All1B1 flies ($F_{1,361} = 39.95$, $p = 7.70\text{E-}10$). Acrophase in All2B1 flies was significantly advanced, relative to that of All1B1 flies ($F_{1,679} = 14.02$, $p = 1.97\text{E-}04$) and All1B2 ($F_{1,716} = 14.64$, $p = 1.42\text{E-}04$). These results suggest that epistasis between the All and B haplotypes contributed to acrophase variation.

Analysis of circadian phases under a light–dark cycle revealed no differences in the morning activity peak (**Figure 6**). In the

evening peak, however, there was a small but significant difference between the haplotypes. In the NIL flies, All1B1 exhibited the most delayed evening peak, as compared to All2B2 ($7.25^\circ \sim 30 \text{ min}$; $F_{1,212} = 17.47$, $p = 4.26\text{E-}05$), while the smallest difference was between All1B2 and All2B1 ($3.4^\circ \sim 14 \text{ min}$; $F_{1,397} = 6.05$, $p = 0.014$). All1B1 was delayed, as compared to All2B1 ($5.19^\circ \sim 21 \text{ min}$; $F_{1,374} = 12.34$, $p = 4.97\text{E-}04$), whereas All2B2 was advanced, as compared to All1B2 ($5.46^\circ \sim 22 \text{ min}$; $F_{1,235} = 10.855$, $p = 0.001$; **Figure 6**).

To delineate the contributions of the All and B haplogroups to circadian behavior, we produced CRISPR-mediated transgenic

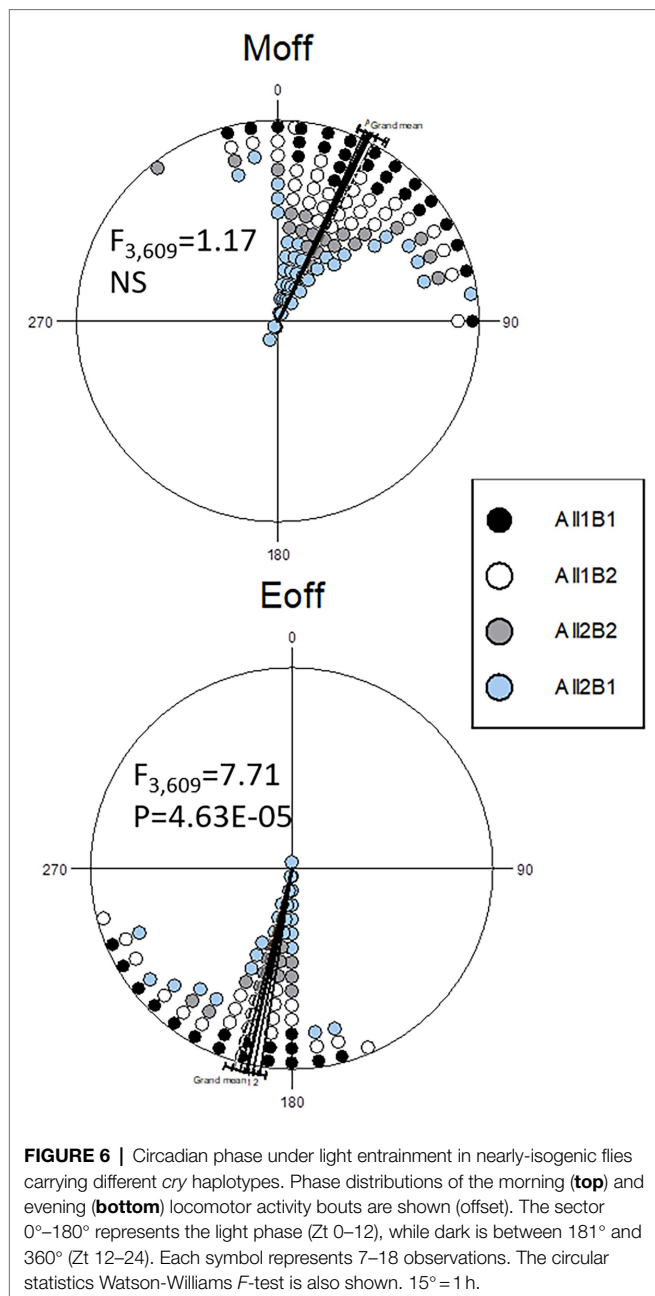


FIGURE 6 | Circadian phase under light entrainment in nearly-isogenic flies carrying different *cry* haplotypes. Phase distributions of the morning (**top**) and evening (**bottom**) locomotor activity bouts are shown (offset). The sector 0°–180° represents the light phase (Zt 0–12), while dark is between 181° and 360° (Zt 12–24). Each symbol represents 7–18 observations. The circular statistics Watson-Williams *F*-test is also shown. 15° = 1 h.

strains carrying the All1B1 or All2B2 haplotypes. The *act-cas9* strain used for the transformation carried the All1B2 haplotype. We found significant differences between haplotypes in the evening phase ($F_{2,76}=11.29$, $p=5.08E-05$; **Figure 7**). The All1B1 strain was significantly more advanced than were All2B2 ($18.55^\circ \sim 74$ min, $F_{1,52}=19.28$, $p=5.56E-05$) and the *act-cas9* All1B2 line ($17.71^\circ \sim 71$ min, $F_{1,51}=18.04$, $p=9.18E-05$). However, there were no differences between All2B2 and the All1B2 haplotypes ($F_{1,49}=0.03$, NS), suggesting that phase differences were due to the B haplotypes (**Figure 7**).

The transgenic lines DD activity phase differed significantly between haplotypes, with All1B2 being the most advanced, All1B1 most delayed, and All2B2 showing intermediate phase

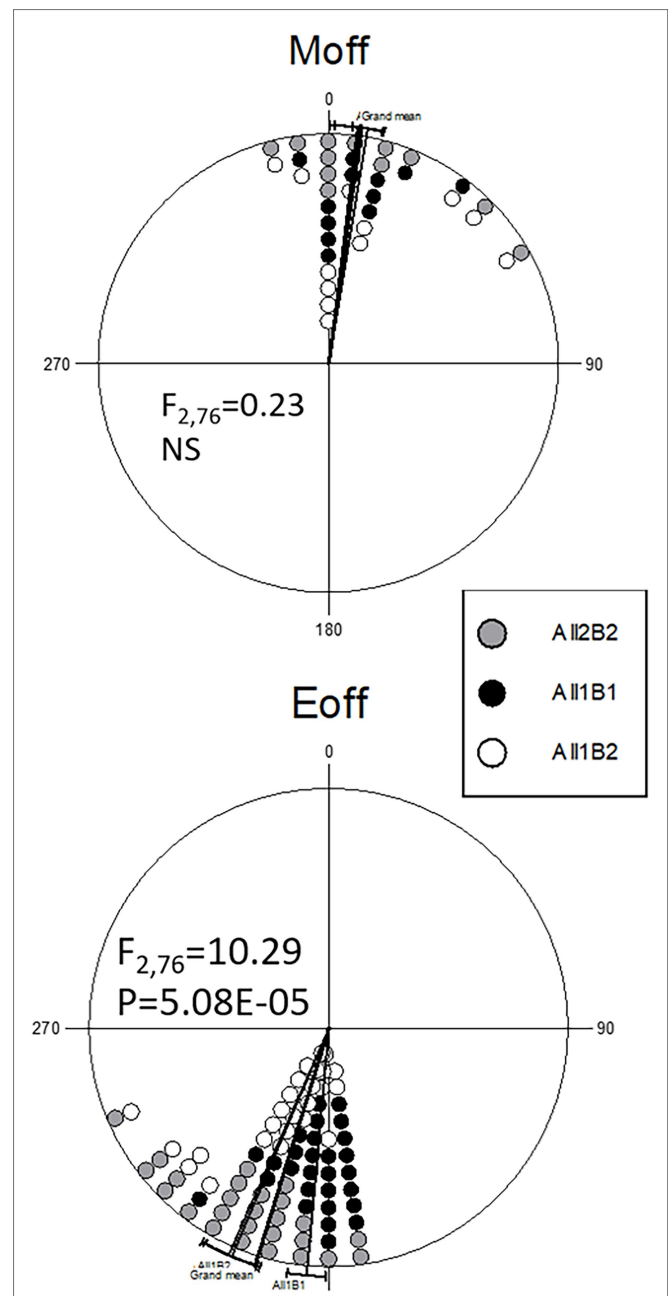


FIGURE 7 | Circadian phase in transgenic flies carrying different *cry* haplotypes. Phase distributions of the morning (**top**) and evening (**bottom**) locomotor activity bouts are shown (offset). The sector 0°–180° represents the light phase (Zt 0–12), while dark is between 181° and 360° (Zt 12–24). Each symbol represents 1–3 observations. The circular statistics Watson-Williams *F*-test is also shown. 15° = 1 h.

(**Supplementary Figure S3**). Interestingly, the phase differences were not due to a difference in the FRP, which did not differ between the haplotypes ($F_{2,93}=0.22$, $p=0.81$; **Supplementary Figure S3**).

Given the contribution of CRY to circadian photosensitivity (Sandrelli et al., 2007), we tested the NIL flies in light pulse experiments. We measured the phase delay after an early-night

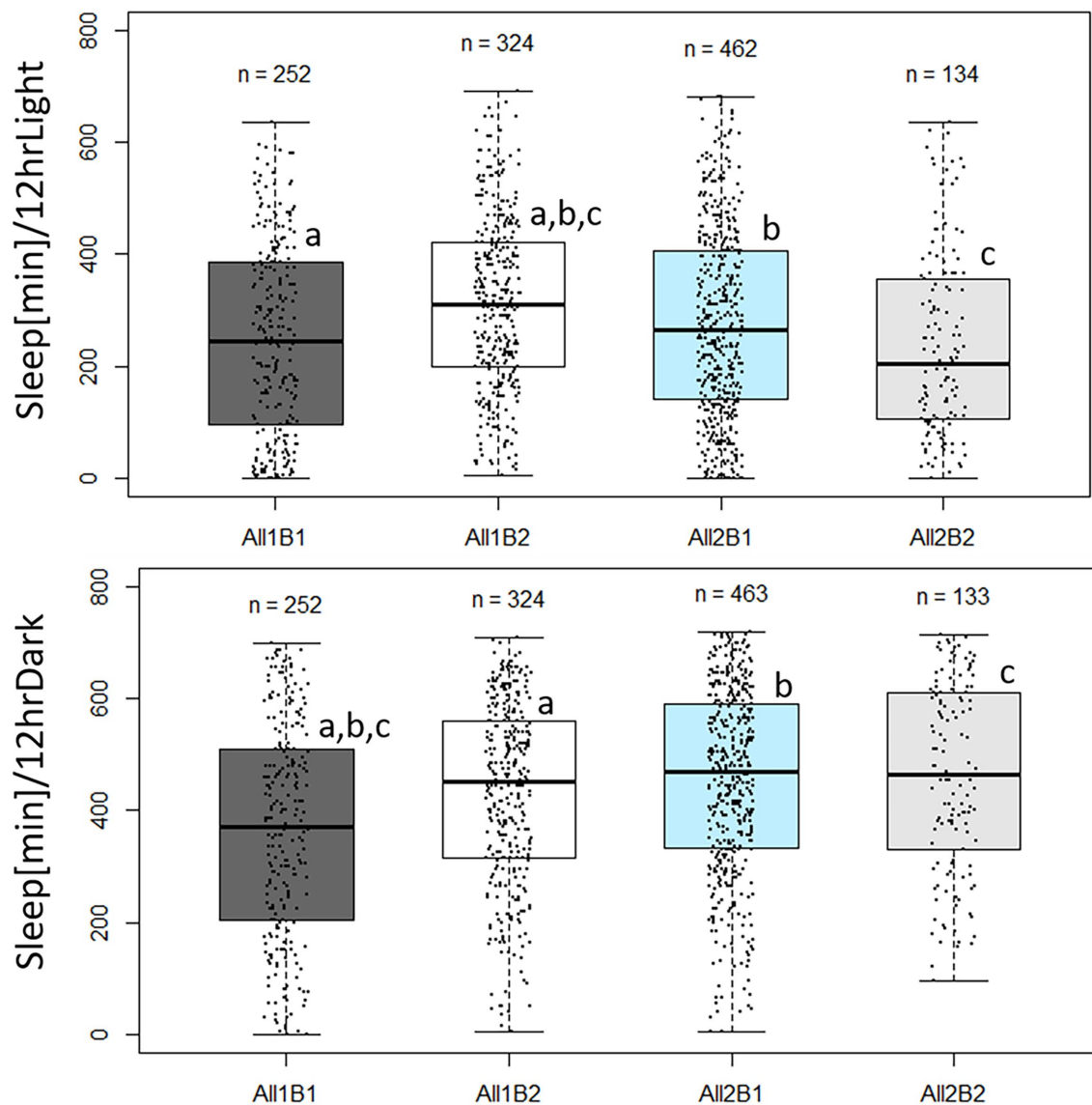
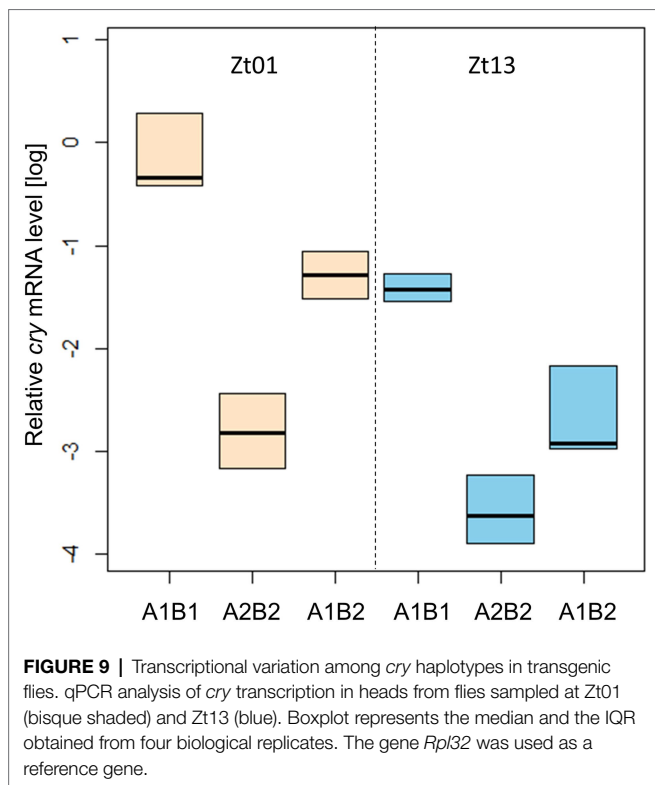


FIGURE 8 | Sleep duration in *cry* haplotypes. Boxplot showing minutes of sleep during photophase (upper panels) and scotophase (lower panels). The middle line represents the median, and the box indicates the interquartile range (IQR). The whiskers signify $1.5 \times \text{IQR}$. The number of flies in each haplotype is indicated. Dots represent individual values. Sleep was quantified during the third day in LD. Letters indicate significant *post-hoc* comparisons (a, b, c, $p < 0.05$). Photophase sleep: All1B1 245 min \pm 10.6 min (median \pm SEM); All1B2 310 min \pm 8.6 min; All2B1 265 min \pm 7.8 min; and All2B2 205 min \pm 14.0 min. Scotophase sleep: All1B1 370 min \pm 12 min; All1B2 452.5 min \pm 9 min; All2B1 470 min \pm 7.7 min; and All2B2 465 min \pm 14.5 min.

light pulse and found no differences among the haplotypes in terms of response to the light pulse (All: $F_{1,1,073}=0.12$, NS; B: $F_{1,1,073}=0.13$, NS; AllB: $F_{3,1,071}=1.29$, NS; **Supplementary Figure S1**).

As *cry* affects arousal locomotor activity (Fogle et al., 2015), we also analyzed locomotor activity and sleep and associated these traits with the different haplotypes. Haplogroup All1 flies were more active than were All2 flies ($F_{1,1,141}=29.5$, $p=6.84\text{E-}08$), while B1 flies were more active than were B2 flies ($F_{1,1,141}=9.42$, $p=0.0022$; **Supplementary Figure S2**). All1B1 flies were more active than all other haplotypes ($F_{3,1,141}=24.07$, $p=4.16\text{E-}15$).

Sleep pattern also varied between the *cry* haplotypes. During daytime, All2B2 flies slept the least and All1B2 flies slept the most ($F_{3,1,139}=14.69$, $p=2.19\text{E-}09$; **Figure 8**). All1 flies slept ~35 more minutes than did All2 flies ($F_{1,1,139}=4.52$, $p=0.034$), while B1 flies slept ~27 min less than did B2 flies ($F_{1,1,139}=8.822$, $p=0.003$). Nighttime sleep also differed among haplotypes ($F_{3,1,139}=25.03$, $p=1.1\text{E-}15$; **Figure 8**) mostly due to All1B1 flies sleeping less than did the other genotypes. This is perhaps an expected result, given that All1B1 flies showed the highest amount of activity during the 24h period (**Supplementary Figure S2**). All2 flies slept 45 min more than did All1 flies during nighttime, compensating for the



sleep difference seen during daytime ($F_{1,139} = 40.92$, $p = 2.31E-10$). The B2 flies slept 15 min more during nighttime than did B1 flies, replicating the daytime difference ($F_{1,139} = 5.48$, $p = 0.0194$).

We determined the transcription of *cry* under light–dark cycles at Zt01 and Zt13 by qPCR analysis (Figure 9). The abundance of *cry* mRNA differed significantly between the transgenic strains carrying the *cry*-haplotypes, at both time-points. Expression in All1B1 flies was the highest compared to the other haplotypes and was the lowest in A2B2; in All1B2 flies, expression level was intermediate ($F_{2,15} = 35.1$, $p < 0.001$).

Geographical Distribution

Since geographical clines in allele frequencies are strong indicators of local adaptation, we sought to determine the spatial distribution of the B1/B2 haplotypes. We took advantage of recently published genomic sequence data collected across Europe (Kapun et al., 2020) and determined the frequency of the 12bp indel that typifies the B1/B2 haplotypes. A generalized linear model revealed both latitude, altitude and longitude had significant, albeit small, effects on haplotype frequency. The frequency of the B2 allele (with the deletion) tended to decrease at higher latitudes ($z = -3.182$, $p = 0.015$, $df = 165$) and higher altitudes ($z = -4.39$, $p = 1.14E-05$) and increased toward Eastern longitudes ($z = 5.3$, $1.16E-07$). To further test the signature of local adaptation caused by climatic differences, we tested association of B1/B2 polymorphism with 19 climatic variables recorded at the WorldClim database (Hijmans et al., 2005). Two compound variables, PC1 and PC2, were previously generated to sum

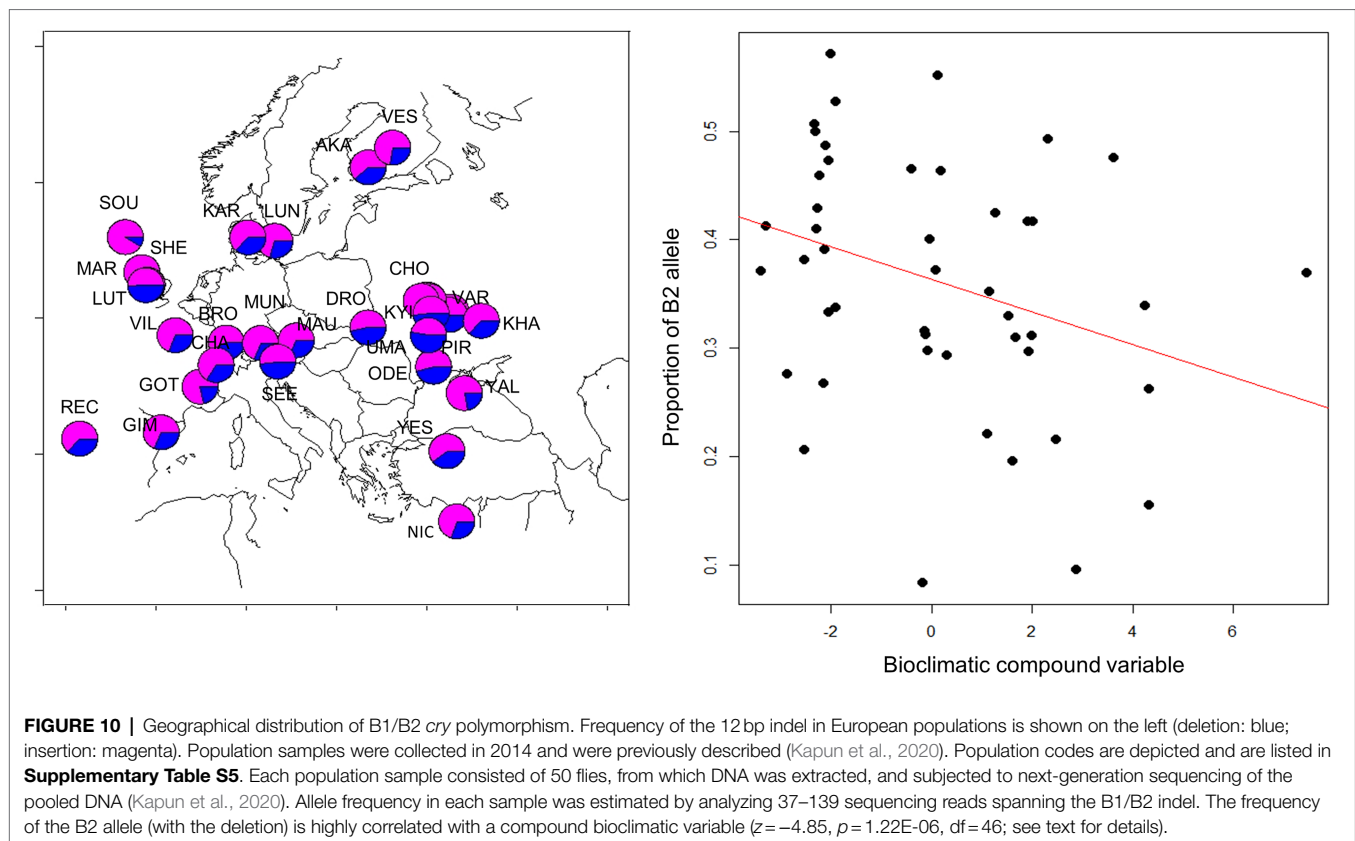
the climatic variable set (Kapun et al., 2020). PC1 represents mostly climatic variables that distinguish warm and cold climates, while PC2 corresponds to climate seasonality. We found significant correlation of B1/B2 frequency with climatic variable PC2 (but not with PC1). The B2 allele was more frequent in climates of high “temperature seasonality standard deviation,” “Mean temperature of wettest quarter,” and “temperature annual range” (Figure 10).

DISCUSSION

While previous studies of molecular polymorphisms in circadian genes focused on a single variation (Costa et al., 1992; Tauber et al., 2007; Pegoraro et al., 2014), the present study aimed at simultaneous testing multiple variations across the *cry* gene. Accordingly, we generated 33 *Drosophila* nearly isogenic lines by introgressing natural *cry* alleles into the same genetic background. Using two different approaches, we identified two major haplogroups that influence circadian function. The first haplogroup (All1/2), identified by the TASSEL algorithm, forms an LD group within intron II. The second haplogroup (B1/B2) emerged from a maximum-likelihood gene tree that we constructed. This haplogroup represents a large LD cluster covering most of intron III and part of exon IV (Figures 3, 4).

Several circadian and sleep phenotypes were affected by the ALL1/2 B1/2 haplotype, including the FRP, phase (in DD and LD) and total sleep. The multiple traits with which these haplotypes are associated are consistent with *cry* being a pleiotropic gene (Damulewicz and Mazzotta, 2020). It is conceivable that these variations and other SNP across the *cry* genomic region are associated with variations in other non-circadian traits, such as geotaxis and magnetoreception. The pleiotropy of *cry* could be mediated by various molecular mechanisms, such as multiple products of the *cry* gene, different spatiotemporal patterns of expression, and different interacting proteins. A combination of these mechanisms has been shown to operate in expression of the *foraging* gene (*for*), a classic example of a behavioral pleiotropic gene (Anreiter and Sokolowski, 2019). The various *for* phenotypes, such as larval nociception, larval foraging path length, and feeding behavior of adults, are regulated by four different promoters of the gene, each dedicated to a different specific function.

The pleiotropism of *cry* may also underlie the variable pattern of polymorphism that we observed across the *cry* region. This, in turn, may reflect the compound signature of natural selection that targets different loci within the gene, each associated with a different trait. For example, Tajima’s D statistics showed both extreme positive and negative values, suggesting different selection scenarios, such as balancing and directional selection. A similar pattern was observed for *Dopa decarboxylase* (*Ddc*), another pleiotropic gene which also exhibited a variable Tajima D across its genomic region (De Luca et al., 2003). Interestingly, the excessive polymorphism in *Ddc* was associated with antagonistic pleiotropism, where traits controlled by the same gene may



increase or reduce fitness under different conditions (for example, at different life stages). Antagonistic pleiotropism may increase genetic variation *via* the maintenance of intra-locus polymorphism, and lead to a balanced polymorphism (Mérot et al., 2020).

It is intriguing that circadian photosensitivity, often conceived as the canonical function of *cry* in *Drosophila*, was not associated with any genetic variation. We made a similar observation in a previous study (Pegoraro et al., 2014), where a pervasive missense SNP was associated with the FRP, circadian phase and eclosion, yet not with circadian photosensitivity. Since other light inputs contribute to light resetting of the clock (Yoshii et al., 2016), it is conceivable that evolution of this trait was mediated by genetic variations in genes other than *cry*. Indeed, QTL mapping of *Drosophila* identified a major QTL on chromosome 2 associated with circadian photosensitivity (Adewoye et al., 2017). The *tim s/l*s polymorphism was shown to be a major contributor to this QTL.

Our study suggests that from an evolutionary perspective, *cry* was the target of natural selection, largely due to its role as setting the circadian activity phase. Indeed, changes in the preferred time of activity during the day serves as a powerful adaptation of populations to local or season-specific climates and defines the ecological niche of a species (Hut et al., 2012). Previous studies have already shown that CRY co-expression with pigment-dispersing factor receptor (PDFR) in specific clock neurons is instrumental in setting the amplitude and phase of the circadian clock (Im et al.,

2011). For example, testing the behavior of *cry^bPd^f* double mutants in constant darkness revealed defective rhythms that were much severe than found in each single mutant. Given the importance of the spatial expression of *cry* in the brain, it is conceivable that polymorphism in *cry* regulatory regions may lead to the generation of different subsets of neurons in which *cry* alleles are expressed, resulting in phenotypic circadian variations.

Variations in CRY expression in clock neurons may also underlie the macro-evolution of the neuronal clock network across *Drosophila* species (Helfrich-Förster et al., 2020). While *Siphodora* fly species, such as *Drosophila hydei* and *Drosophila mercatorum* that reside in tropical/temperate regions show CRY (and PDF) expression like *Drosophila melanogaster*, sub-arctic species (e.g., *Drosophila ezoana* and *Drosophila littoralis*) show a distinct neural pattern; l-LNv neurons in these species do not express CRY. It is therefore plausible that adaptive molecular variations in *cry* within species populations are the foundation for species differentiation.

The geographical distribution of the B1/B2 haplotypes, specifically the 12bp indel associated with this haplogroup, provides hints to the adaptive function of this polymorphism. Using two compound climatic variables established in a previous study (Kapun et al., 2020), we found strong correlation with PC2, a variable that reflects seasonality differences between localities. This variable strongly correlates with longitude and likely tracks the transition from an oceanic to a continental climate. The frequency of the B1/B2 polymorphism was not correlated with PC1, a climatic

variable that tracks the typical temperature gradient and is highly correlated with latitude. Thus, the B1/B2 polymorphism in *cry* is likely to be driven by balancing selection that retains haplotypes that are important for seasonal adaptation and adjustment of circadian function, particularly phase of activity and sleep, to annual fluctuations of the environment. How the B1/B2 12 bp indel variation induces circadian phenotype variation has yet to be determined. An analysis of RNA binding protein motifs may hint the underlying mechanism. Using the ScanForMotifs server (Biswas and Brown, 2014), we found a reduced number of binding sites for the Elav and qkr58E-1 RNA-binding proteins in the deletion allele-containing flies. This allele also harbors a unique site for the Tis11 zinc finger protein that is not present in the insertion allele. These three proteins play roles in controlling mRNA post-transcriptional processing and catabolism (Herold et al., 2009; Choi et al., 2014; Lee et al., 2021), suggesting that B1/B2 polymorphism is associated with post-transcriptional maturation of *cry* mRNA.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.ncbi.nlm.nih.gov/genbank/>, MW758991–MW759024.

REFERENCES

- Adewoye, A. B., Nuzhdin, S. V., and Tauber, E. (2017). Mapping quantitative trait loci underlying circadian light sensitivity in *Drosophila*. *J. Biol. Rhythm.* 32, 394–405. doi: 10.1177/0748730417731863
- Anderson, M., Ter Braak, C., Anderson, M. J., Ter Braak, F., and C. J., (2003). Permutation tests for multi-factorial analysis of variance. *J. Stat. Comput. Simul.* 73, 85–113. doi: 10.1080/00949650215733
- Anreiter, I., and Sokolowski, M. B. (2019). The *foraging* gene and its behavioral effects: pleiotropy and plasticity. *Annu. Rev. Genet.* 53, 373–392. doi: 10.1146/annurev-genet-112618-043536
- Biswas, A., and Brown, C. M. (2014). Scan for motifs: a webserver for the analysis of post-transcriptional regulatory elements in the 3' untranslated regions (3' UTRs) of mRNAs. *BMC Bioinform.* 15, 1–6. doi: 10.1186/1471-2105-15-174
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Busza, A., Emery-Le, M., Rosbash, M., and Emery, P. (2004). Roles of the two *Drosophila cryptochrome* structural domains in circadian photoreception. *Science* 304, 1503–1506. doi: 10.1126/science.1096973
- Choi, Y.-J., Lai, W. S., Fedic, R., Stumpo, D. J., Huang, W., Li, L., et al. (2014). The *Drosophila* Tis11 protein and its effects on mRNA expression in flies. *J. Biol. Chem.* 289, 35042–35060. doi: 10.1074/jbc.M114.593491
- Costa, R., Peixoto, A. A., Barbuji, G., and Kyriacou, C. P. (1992). A latitudinal cline in a *Drosophila* clock gene. *Proc. R. Soc. B Biol. Sci.* 250, 43–49. doi: 10.1098/rspb.1992.0128
- Damulewicz, M., and Mazzotta, G. M. (2020). One actor, multiple roles: the performances of Cryptochrome in *Drosophila*. *Front. Physiol.* 11:99. doi: 10.3389/FPHYS.2020.00099
- De Luca, M., Roshina, N. V., Geiger-Thornsberry, G. L., Lyman, R. F., Pasyukova, E. G., and Mackay, T. F. C. (2003). *Dopa decarboxylase (Ddc)* affects variation in *Drosophila* longevity. *Nat. Genet.* 34, 429–433. doi: 10.1038/ng1218

AUTHOR CONTRIBUTIONS

ET contributed to the conception and design of the study and wrote sections of the manuscript. MP, ES, and BF carried out the experiments. MP performed the statistical analysis and wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

Financial support has been provided by the Biotechnology and Biological Sciences Research Council (BBSRC, United Kingdom) grant BB/G02085X/1, which funded the postdoctoral fellowship of the lead author and supplies to conduct this research project. The Israel Science Foundation grant (1737/17) to ET also funded research supplies. The funders had no involvement in the study design, the collection, analysis and interpretation of data, the writing of the report, or in the decision to submit the article for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2022.781380/full#supplementary-material>

- Emery, P., So, W. V., Kaneko, M., Hall, J. C., and Rosbash, M. (1998). CRY, a *Drosophila* clock and light-regulated cryptochrome, is a major contributor to circadian rhythm resetting and photosensitivity. *Cell* 95, 669–679. doi: 10.1016/S0092-8674(00)81637-2
- Endo, M. (2016). Tissue-specific circadian clocks in plants. *Curr. Opin. Plant Biol.* 29, 44–49. doi: 10.1016/j.pbi.2015.11.003
- Fedele, G., Edwards, M. D., Bhutani, S., Hares, J. M., Murbach, M., Green, E. W., et al. (2014). Genetic analysis of circadian responses to low frequency electromagnetic fields in *Drosophila melanogaster*. *PLoS Genet.* 10:e1004804. doi: 10.1371/journal.pgen.1004804
- Fogle, K. J., Baik, L. S., Houli, J. H., Tran, T. T., Roberts, L., Dahm, N. A., et al. (2015). CRYPTOCHROME-mediated phototransduction by modulation of the potassium ion channel β -subunit redox sensor. *Proc. Natl. Acad. Sci.* 112, 2245–2250. doi: 10.1073/pnas.1416586112
- Fogle, K. J., Parson, K. G., Dahm, N. A., and Holmes, T. C. (2011). CRYPTOCHROME is a blue-light sensor that regulates neuronal firing rate. *Science* 331, 1409–1413. doi: 10.1126/science.1199702
- Gegear, R. J., Casselman, A., Waddell, S., and Reppert, S. M. (2008). Cryptochrome mediates light-dependent magnetosensitivity in *Drosophila*. *Nature* 454, 1014–1018. doi: 10.1038/nature07183
- Gratz, S. J., Cummings, A. M., Nguyen, J. N., Hamm, D. C., Donohue, L. K., Harrison, M. M., et al. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194, 1029–1035. doi: 10.1534/genetics.113.152710
- Helfrich-Förster, C., Bertolini, E., and Menegazzi, P. (2020). Flies as models for circadian clock adaptation to environmental challenges. *Eur. J. Neurosci.* 51, 166–181. doi: 10.1111/ejn.14180
- Hendricks, J. C., Finn, S. M., Panckeri, K. A., Chavkin, J., Williams, J. A., Sehgal, A., et al. (2000). Rest in *Drosophila* is a sleep-like state. *Neuron* 25, 129–138. doi: 10.1016/S0896-6273(00)80877-6
- Herold, N., Will, C. L., Wolf, E., Kastner, B., Urlaub, H., and Lührmann, R. (2009). Conservation of the protein composition and electron microscopy structure of *Drosophila melanogaster* and human Spliceosomal complexes. *Mol. Cell. Biol.* 29, 281–301. doi: 10.1128/MCB.01415-08

- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. doi: 10.1002/joc.1276
- Hu, T. T., Eisen, M. B., Thornton, K. R., and Andolfatto, P. (2013). A second-generation assembly of the *Drosophila simulans* genome provides new insights into patterns of lineage-specific divergence. *Genome Res.* 23, 89–98. doi: 10.1101/gr.141689.112
- Hut, R. A., Kronfeld-Schor, N., van der Vinne, V., and De la Iglesia, H. (2012). In search of a temporal niche: environmental factors. *Prog. Brain Res.* 199, 281–304. doi: 10.1016/B978-0-444-59427-3.00017-4
- Hut, R. A., Paolucci, S., Dor, R., Kyriacou, C. P., and Daan, S. (2013). Latitudinal clines: an evolutionary view on biological rhythms. *Proc. R. Soc. B Biol. Sci.* 280:20130433. doi: 10.1098/rspb.2013.0433
- Im, S. H., Li, W., and Taghert, P. H. (2011). PDFR and CRY signaling converge in a subset of clock neurons to modulate the amplitude and phase of circadian behavior in *Drosophila*. *PLoS One* 6:e18974. doi: 10.1371/journal.pone.0018974
- Johnsen, A., Fidler, A. E., Kuhn, S., Carter, K. L., Hoffmann, A., Barr, I. R., et al. (2007). Avian clock gene polymorphism: evidence for a latitudinal cline in allele frequencies. *Mol. Ecol.* 16, 4867–4880. doi: 10.1111/j.1365-294X.2007.03552.x
- Johnson, C. H., Elliott, J. A., and Foster, R. (2003). Entrainment of circadian programs. *Chronobiol. Int.* 20, 741–774. doi: 10.1081/CBI-120024211
- Kapun, M., Barrón, M. G., Staubach, F., Obbard, D. J., Wiberg, R. A. W., Vieira, J., et al. (2020). Genomic analysis of European *Drosophila melanogaster* populations reveals longitudinal structure, continent-wide selection, and previously unknown DNA viruses. *Mol. Biol. Evol.* 37, 2661–2678. doi: 10.1093/molbev/msaa120
- Koh, K., Zheng, X., and Sehgal, A. (2006). JETLAG resets the *Drosophila* circadian clock by promoting light-induced degradation of TIMELESS. *Science* 312, 1809–1812. doi: 10.1126/science.1124951
- Kyriacou, C. P., Peixoto, A. A., and Costa, R. (2007). A cline in the *Drosophila melanogaster* period gene in Australia: neither down nor under. *J. Evol. Biol.* 20, 1649–1651. doi: 10.1111/j.1420-9101.2007.01352.x
- Kyriacou, C. P., Peixoto, A. A., Sandrelli, F., Costa, R., and Tauber, E. (2008). Clines in clock genes: fine-tuning circadian rhythms to the environment. *Trends Genet.* 24, 124–132. doi: 10.1016/j.tig.2007.12.003
- Lack, J. B., Lange, J. D., Tang, A. D., Corbett-Detig, R. B., and Pool, J. E. (2016). A thousand fly genomes: an expanded *Drosophila* genome nexus. *Mol. Biol. Evol.* 33, 3308–3313. doi: 10.1093/molbev/msw195
- Lee, S., Wei, L., Zhang, B., Goering, R., Majumdar, S., Wen, J., et al. (2021). ELAV/Hu RNA binding proteins determine multiple programs of neural alternative splicing. *PLoS Genet.* 17:e1009439. doi: 10.1371/journal.pgen.1009439
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Lin, F.-J., Song, W., Meyer-Bernstein, E., Naidoo, N., and Sehgal, A. (2001). Photoc signaling by Cryptochrome in the *Drosophila* circadian system. *Mol. Cell. Biol.* 21, 7287–7294. doi: 10.1128/MCB.21.21.7287-7294.2001
- McDonald, J. H., and Kreitman, M. (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351, 652–654. doi: 10.1038/351652a0
- Mérot, C., Laurens, V., Normandeau, E., Bernatchez, L., and Wellenreuther, M. (2020). Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat. Commun.* 11, 670–611. doi: 10.1038/s41467-020-14479-7
- Naidoo, N., Song, W., Hunter-Ensor, M., and Sehgal, A. (1999). A role for the proteasome in the light response of the timeless clock protein. *Science* 285, 1737–1741. doi: 10.1126/science.285.5434.1737
- Nelson, W., Tong, Y. L., Lee, J. K., and Halberg, F. (1979). Methods for cosinor-rhythmometry. *Chronobiologia* 6, 305–323
- Numata, H., Miyazaki, Y., and Ikeno, T. (2015). Common features in diverse insect clocks. *Zoological Lett.* 1, 10–17. doi: 10.1186/s40851-014-0003-y
- Pegoraro, M., Noreen, S., Bhutani, S., Tsolou, A., Schmid, R., Kyriacou, C. P., et al. (2014). Molecular evolution of a pervasive natural amino-acid substitution in *Drosophila cryptochrome*. *PLoS One* 9:e86483. doi: 10.1371/journal.pone.0086483
- Peschel, N., Veleri, S., and Stanewsky, R. (2006). *Veela* defines a molecular link between Cryptochrome and Timeless in the light-input pathway to *Drosophila's* circadian clock. *Proc. Natl. Acad. Sci. U. S. A.* 103, 17313–17318. doi: 10.1073/pnas.0606675103
- Rosato, E., and Kyriacou, C. P. (2006). Analysis of locomotor activity rhythms in *Drosophila*. *Nat. Protoc.* 1, 559–568. doi: 10.1038/nprot.2006.79
- Sandrelli, F., Tauber, E., Pegoraro, M., Mazzotta, G., Cisotto, P., Landskron, J., et al. (2007). A molecular basis for natural selection at the *timeless* locus in *Drosophila melanogaster*. *Science* 316, 1898–1900. doi: 10.1126/science.1138426
- Sawyer, L. A., Hennessy, J. M., Peixoto, A. A., Rosato, E., Parkinson, H., Costa, R., et al. (1997). Natural variation in a *Drosophila* clock gene and temperature compensation. *Science* 278, 2117–2120. doi: 10.1126/science.278.5346.2117
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi: 10.1093/genetics/123.3.585
- Tauber, E., Zordan, M., Sandrelli, F., Pegoraro, M., Osterwalder, N., Breda, C., et al. (2007). Natural selection favors a newly derived *timeless* allele in *Drosophila melanogaster*. *Science* 316, 1895–1898. doi: 10.1126/science.1138412
- Thompson, C. L., and Sancar, A. (2002). Photolyase/cryptochrome blue-light photoreceptors use photon energy to repair DNA and reset the circadian clock. *Oncogene* 21, 9043–9056. doi: 10.1038/sj.onc.1205958
- Toma, D. P., White, K. P., Hirsch, J., and Greenspan, R. J. (2002). Identification of genes involved in *Drosophila melanogaster* geotaxis, a complex behavioral trait. *Nat. Genet.* 31, 349–353. doi: 10.1038/ng893
- Tononi, G. (2000). Correlates of sleep and waking in *Drosophila melanogaster*. *Science* 287, 1834–1837. doi: 10.1126/science.287.5459.1834
- Weeks, A. R., McKechnie, S. W., and Hoffmann, A. A. (2006). In search of clinal variation in the *period* and *clock* timing genes in Australian *Drosophila melanogaster* populations. *J. Evol. Biol.* 19, 551–557. doi: 10.1111/j.1420-9101.2005.01013.x
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biom. Bull.* 1, 80–83. doi: 10.2307/3001968
- Yoshii, T., Ahmad, M., and Helfrich-Förster, C. (2009). Cryptochrome mediates light-dependent magnetosensitivity of *Drosophila's* circadian clock. *PLoS Biol.* 7, e1000086–e1000089. doi: 10.1371/journal.pbio.1000086
- Yoshii, T., Hermann-Luibl, C., and Helfrich-Förster, C. (2016). Circadian light-input pathways in *Drosophila*. *Commun. Integr. Biol.* 9:e1102805. doi: 10.1080/19420889.2015.1102805
- Zonato, V., Vanin, S., Costa, R., Tauber, E., and Kyriacou, C. P. (2018). Inverse European latitudinal cline at the *timeless* locus of *Drosophila melanogaster* reveals selection on a clock gene: population genetics of *ls-tim*. *J. Biol. Rhythm.* 33, 15–23. doi: 10.1177/0748730417742309

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Pegoraro, Sayegh Rezek, Fishman and Tauber. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.