A Tale of Two Chronotypes:

Behavioral and Genetic Characterization of Chronotype Divergence in *Drosophila melanogaster*

Thesis submitted
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

by

Arijit Ghosh

JNCASR

Chronobiology and Behavioral Neurogenetics Laboratory, Neuroscience Unit
Jawaharlal Nehru Centre for Advanced Scientific Research
Jakkur, Bangalore – 560064, India
June 2022

Dedicated to Maa, Didi, and the memories of my father

Table of contents

Declaration vii

Certificate viii

Acknowledgements ix

Abbreviations and Glossary xiii

Synopsis xvi

Chapter 1: Introduction …………………………………………………………...... 1

1.1 On the concept of time and timekeeping 2

1.2 On biological clocks and their nature 4

1.3 On the importance of timing, chronotypes,
and entrainment of circadian clocks 8

1.4 Study of chronotypes in laboratory populations 14

1.5 GATE populations at JNCASR 17

1.5.1 Generation and maintenance protocol of early and late populations 17

1.5.2 Direct response to selection, evolution of the emergence waveform, and associated changes in clock properties 20

1.6 Summary of the present study 26

Chapter 2: Characterization of masking responses to light and temperature in eclosion rhythm 29

2.1 Introduction 30

2.2 Materials and methods 36

2.2.1 Selection protocol and fly maintenance 36

2.2.2 Behavioral experiments 37

2.2.3 Analysis of data and statistics 38

2.3 Results 39

2.3.1 Lights-ON elicits an immediate response in the early chronotypes 39

2.3.2 early flies attenuate and delay their emergence under DD 42

2.3.3 early flies consistently emerge close to lights-ON

under both short and long T-cycles *43*
2.3.4 early flies show phase lability under skeleton T-cycles *48*
2.3.5 early flies show negative masking to warm temperature *52*
2.4 Discussion 54
2.4.1 Evolution of masking in early populations along with advanced phase of entrainment *55*
2.4.2 Higher range of "apparent" entrainment achieved by masking *58*
2.4.3 The duration and intensity of light dramatically alters magnitude of response in eclosion *59*
2.4.4 Plausible mechanisms driving higher masking in early populations *60*
2.4.5 Evolution of negative masking response to high temperature in early flies *61*
Chapter 3: Characterization of masking responses in locomotor activity rhythm and sleep 63
3.1 Introduction 64
3.2 Materials and methods 68
3.2.1 Selection protocol and fly husbandry *68*
3.2.2 Behavioral experiments *68*
3.2.3 Data analysis and statistics *69*
3.3 Results 70
3.3.1 No difference in masking under low intensity photoperiods between chronotypes *71*
3.3.2 Difference in masking between chronotypes under high light intensity photoperiods *72*
3.3.3 High light intensity during the day disrupts nighttime sleep amount and quality in extreme chronotypes *77*
3.3.4 High light intensity photoperiods negatively affect nighttime sleep amount and quality of early and late flies *79*
3.4 Discussion 82
3.4.1 Enhanced masking in early flies under high intensity photoperiods *82*
3.4.2 Extreme chronotypes have lesser and poor-quality sleep under high intensity photoperiods *86*
Chapter 4: Genetic characterization of divergent chronotype populations by pooled sequencing 90

4.1 Introduction 91

4.2 Materials and methods 96

4.2.1 Sample preparation and sequencing 96

4.2.2 Variant identification and filtering 96

4.2.3 Calculation of Tajima's Pi (π), D, and Waterson's Theta (θ_w) 97

4.2.4 Allele frequency-based differentiation tests 98

4.2.5 Identification of genome-wide selective sweeps 99

4.2.6 Analysis of identified variants 99

4.3 Results 100

4.3.1 Identification of SNPs and estimating population genomic parameters 101

4.3.2 Identification and characterization of significant SNPs by quantifying differential allele frequency changes between populations 105

4.3.3 Signatures of selective sweeps in populations 115

4.4 Discussion 118

4.4.1 Photosensitivity, sleep, and clock related changes in early populations 119

4.4.2 Clock, mRNA splicing, and ecdysis related changes in late populations 122

Chapter 5: Screening of the DrosDel collection for differential phasing of eclosion rhythm 127

5.1 Introduction 128

5.2 Materials and methods 132

5.2.1 Deficiency line procurement and fly husbandry 132

5.2.2 Eclosion assay, analyses, and statistics 133

5.3 Results 134

5.3.1 Detection of rhythmicity and entrainment in deficiency lines 136

5.3.2 Identification of "hits" – lines showing significantly advanced/delayed phase of emergence 136

5.3.3 Description of the early chronotype (advanced Ψ_{CoM}) lines 138

5.3.4 Description of the late chronotype (delayed Ψ_{CoM}) lines 140

5.3.5 Identification of genes affected by the deficiencies and their characterization 141

5.4 Discussion 142

Chapter 6: Summary, conclusions, and future directions 147

Appendix 1: VANESSA – Shiny apps for accelerated time-series analysis and visualization of Drosophila circadian rhythm and sleep data 156

Appendix 2: Supplementary information for Chapter 2 170

Appendix 3: Supplementary information for Chapter 3 178

Appendix 4: Supplementary information for Chapter 4 198

Appendix 5: Supplementary information for Chapter 5 211

References 218

List of publications 262

Declaration

The work presented in this thesis, entitled "**A Tale of Two Chronotypes: Behavioral and Genetic Characterization of Chronotype Divergence in *Drosophila melanogaster***" is the result of investigations carried out by me in the Neuroscience Unit of Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR), Bangalore, India, under the supervision of Professor Sheeba Vasu. I hereby declare that the work reported here has not been submitted elsewhere for any other degree in this or any other university.

Any part of the presented content, if an outcome of collaborative research or adopted from other studies, has been duly acknowledged within the text and in the references. Any omission which may have occurred is likely due to oversight or an error in judgment and is highly regretted.

Arijit Ghosh

Arijit Ghosh
Bangalore
February 2022

Certificate

It is certified that the work contained in this thesis, titled “**A Tale of Two Chronotypes: Behavioral and Genetic Characterization of Chronotype Divergence in *Drosophila melanogaster***,” is the result of investigations undertaken by **Mr. Arijit Ghosh** under my supervision in the Neuroscience Unit of Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560064 (India) and that the results presented in the thesis have not previously formed the basis for the award of any diploma, degree or fellowship.

Sheeba.Vs

Dr. Sheeba Vasu
Neuroscience Unit
JNCASR
February 2022

Acknowledgements

First and foremost, I would like to thank my PhD advisor, Professor Sheeba Vasu. She believed in me when it was hard for me to believe in myself, she supported me, and when I thought it was the end, she pulled me through in situations I felt I couldn't be rescued from. She has been an incredible source of inspiration, and support, for issues both personal and professional. I wish every graduate student to get an advisor as kind, compassionate, sympathetic, and brilliant as her. There are many other small but effective things I learned from her, which I plan to apply in my journey deeper into academia.

I am incredibly thankful to the late Professor Vijay Kumar Sharma (VKS) for taking me in as a graduate student to the Chronobiology group at JNCASR. He introduced me to the realm of classical chronobiology and helped me appreciate the power of harnessing natural genetic variations to study chronotype divergence, one avenue I actively want to pursue in the future too. His untimely death was a monumental loss to the global chronobiology community and to the broader scientific community. Surely his legacy will live on through the students he trained, already at different corners of the globe and making their mark in clock research.

I would like to thank various scientists I met and interacted with over the years, to discuss my work, getting inputs, and sometimes, just getting advice – Professors Mary Harrington, John Ewer, Amita Sehgal, Joanna Chiu, Tanya Leise, Erik Herzog, Patrick Emery, Robert Koffler, Eran Tauber, Mark Phillips, Neda Barghi, Simon Boitard, Chandan Goswami, Ravi Manjithaya, Kushagra Bansal, and Ms. Marta Pelizolla. These interactions helped me structure my interpretations of various results and also pursue newer and better analysis methods. I also want to thank Dr. Shahnaz R Lone and Dr. Aishwarya Ramakrishnan for sharing their data freely to include as demo files in the open-source R-Shiny applications I built, and I am indebted to Dr. Quentin Giessmann for his help in creating these apps. I am grateful to my Graduate Student Advisory Committee (GSAC) members – Professors Hassan Annegowda Ranganath and Anuranjan Anand, for their constant constructive criticisms, which helped me improve as a researcher. I would also like to thank Professor Amitabh Joshi for the courses which I took in the first few years. His classes made me appreciate population genetics, the nuances of evolution, and the importance of organismal biology. I am also grateful to the Council of Scientific & Industrial

Research, Government of India, New Delhi, for providing me with a Junior Research Fellowship and a Senior Research Fellowship during my tenure as a graduate student, the Department of Science & Technology, and the Department of Biotechnology, Government of India, for their continuous support to our lab through consumable grants, which made the work done in this thesis possible. Parts of the work presented in this thesis are either published or under review, and I thank the anonymous reviewers for their time and useful comments. I also extend my sincere gratitude to my thesis examiners for their time and effort in reviewing my work.

The first person you work with as a new graduate student has a significant impact on you. I could not be luckier in this regard; I got to work with two of the most outstanding seniors ever – Nikhil KL and Abhilash Lakshman. Both of them were working with divergent chronotype populations at that time, albeit on different aspects of their clocks. They taught me most of what I know about classical chronobiology and the most important of all – scientific and experimental rigor; they also have been fantastic collaborators for a few projects. Nikhil, Abhilash, and Vishwanath were seniors with whom I became extremely comfortable outside the lab also, dabbled in philosophy, politics, music, wandered the city with, and discovered unique places and interests we did not know would be so important to us. They became some of the closest friends I have, and hopefully, in the future we will again work together. The lab, which we endearingly call the clock club, has been a home away from home for the past five years. All past and present members with whom I overlapped have been nothing short of excellent. Over the years, VKS and Sheeba have gathered the best minds in the country here at their labs in JNCASR. Aishwarya Ramakrishnan and Pragya Sharma have been two rocks in the lab for me for the past 2.5 years; we have become terrific friends, I have shared anything and everything with them, and they have been highly supportive. Chitrang Dani, my batchmate in the lab, who joined with me, and is soon going to finish, has been a very good friend, and we have been through a lot together in the lab, and throughout it all, always supported each other. Other lab members, Aishwariya Iyengar, Viveka, Rutvij, Manishi, Ankit, Sheetal, Mansi, Roshan, Sushma, have been great to interact with and always gave me very constructive criticisms during work presentations in lab meetings.

I have to thank the many people who, over the years, have created and maintained the Drosophila populations I worked with. These populations (GATEs) have been under selection for ~20 years and ~350 generations now. The people involved till now were – Shailesh, Koustubh, Nikhil, Abhilash, Ratna, Swathi, Shephali. Pragya and Mansi will be maintaining these populations after me, and I am grateful to them for that.

I have mentored and taught a few short-term students at the lab – Sanath, Sankeert, Vaishnavi, Sowmya, Shephali, and these interactions have been an incredibly enriching experience. I have also taught and mentored several POBE students and Integrated PhD students, which was a similarly gratifying experience.

Superhumans do exist – Rajanna, Muniraju, and Samuel are three lab assistants who have helped the lab over the years by washing, cleaning, and maintaining huge numbers of glass vials, cages, petri dishes, clothes, etc. They have become integral parts of the lab and are absolutely necessary for the lab to function the way it does.

My life at JNCASR would not have been as memorable as it is without all the friends outside the lab, at the campus, or from around town. When I first arrived in the city, a few of seniors from my college (NISER) were already in Bangalore for a few years, carrying out their doctoral work at other institutes, and they kind of adopted me, took care of me during some of the darkest moments of my adult life, and brought me back to action. I am indebted to Arun, Abhash, Asim, Somarupa, Nikhil Bansal, Niraj, Prasad Ji, Joshua, and Gagan for their help and companionship. My graduate school batchmates – Priya, Varghese (whom I have known for 11 years now), Rajarshi, Resmi, Preeti, Vijay, who were in different departments, have been instrumental in keeping my sanity in check for the past five years. They have been extremely good friends, in thick and thin, and I realized I could rely on them for anything and everything, and no matter how bad things get in life, your friends can pick you up from the worst and make you feel better. There have been many more people in and around JNCASR who have gone out of their way to help me; I cannot possibly name them all here, and neither can I list what they have done for me. I am extremely thankful to Priya, who has been by my side for the past four years and supported me through all things good and bad and has been a constant source of inspiration to make myself better. These people made the whole graduate school experience worth it, and I urge any graduate student reading this section to make close and wonderful friends outside the lab, even outside the campus; they make us see our lives from a different perspective, and honestly, can give a pretty lousy day a fantastic ending.

Finally, I am grateful for a supportive family – my mother and my sister. They have supported me in each of my decision, no matter how perplexing it seemed to them or how actual stupid the decisions were. They stood by me and cheered for me, even when I thought I did not deserve it. I left home when I was 8 years old, and today, I am 28 years old and have spent 20 years away from home, but my mom and didi never let me feel that I was not at home. I can not even begin to imagine what my life would have been like if they had stopped caring, but they did not, they accepted me for all my quirks and stupidity, and they loved me unconditionally. For me, it made a huge difference and made me the person I am today. I would do all this again in a heartbeat!

Abbreviations and Glossary

After-effects: Change in free-running period (FRP) as a consequence of the entraining regime.

Amplitude expansion: Increase in amplitude of rhythm under entrainment relative to amplitude under constant darkness.

Circadian (rhythm and clock): An endogenous biological rhythm with a natural (or free-running) period (τ) close to, but not necessarily the same as that of the earth's rotation (i.e., 24-h; Latin, *circa*–"about", *dian*–"day"). Any set of mechanisms within the organism that drives such rhythms are called circadian clocks.

Chronotype: Heritable variations in the timing of a rhythmic behaviors are known as chronotypes. Chronotypes can be the result of an interaction of varying strength of masking and circadian entrainment.

CMH test: Cochran-Mantel-Haenszel (CMH) test for repeated tests of independence for every SNP. The CMH test is used for detecting significant and consistent changes in allele frequencies, when independent measurements of allele frequencies are available, as in the case of this thesis where the replicate populations can be treated as genetically independent samples.

DD: Constant darkness.

DRC: Dose Response Curve; a plot of phase-shifts incurred by a circadian system in response to different doses of stimuli (either intensity or duration) at different phases of the circadian system

Entrainment: The coupling of a circadian rhythm to a zeitgeber such that both have the same period ($\tau = T$) resulting in a stable and reproducible Ψ_{ent}. Entrainment can only occur within a range of T values and this range is referred to as the "range of entrainment".

F_{ST}: F_{ST} is a measure of population divergence. It measures the reduction in heterozygosity in a subdivided population compared to that in large, random mating population with the same allele frequencies.

FRP: Free-Running Period; also represented as τ.

Gate-width: Traditionally defined as the 'allowed' zone for adult emergence to occur within a day. It is calculated as the difference between phases of onset and offset of the emergence rhythm.

GWAS: A genome-wide association study (GWAS) is an approach used in genetics research to associate specific genetic variations with diseases. The method involves scanning the genomes from many different people and looking for genetic markers that can be used to predict the presence of a disease.

LD: Light/Dark.

Masking: A non-clock phenomenon, immediate effects of zeitgebers on initiation/cessation rhythmic behavior, physiological processes bypassing the circadian clock.

Phase response curve (PRC): A plot of the shift in an instantaneous state (phase) of the circadian rhythm/oscillation caused due to a perturbation (using a zeitgeber) at different times of the rhythm under constant conditions.

Phase-control: The phenomenon wherein rhythms free-run under constant conditions post entrainment from the phase determined by the last entraining cycle.

Phase-relationship/Phase of entrainment (Ψ_{ent}): Difference in time (either in hours or degrees or any other unit of time) between any instantaneous state (phase) of the circadian rhythm/oscillation and that of a reference phase of the environmental oscillation.

Photophase: Duration of the day when light is present.

Scotophase: Duration of the day when light is absent.

Skeleton photoperiod: An entraining regime wherein two pulses, one mimicking dawn and the other mimicking dusk are provided to act as a "skeleton" to a full photophase.

SNP: Single nucleotide polymorphisms, frequently called SNPs (pronounced "snips"), are the most common type of genetic variation among people. Each SNP represents a difference in a single DNA building block, called a nucleotide.

VRC: Velocity response curve. Theoretically, a plot of the change in angular velocity of the circadian rhythm/oscillation caused due to a perturbation (using a zeitgeber) at different times of the rhythm under constant conditions.

Zeitgeber (German, *zeit*–"time", *geber*–"giver"): Any forcing oscillation (with period '*T*') in the environment that can entrain a biological oscillation, for instance, light/dark or temperature cycles. Zeitgeber cycles with T different from 24-h are referred to as *T*-cycles.

ZT: Zeitgeber Time; ZT00 refers to the time at which lights turn ON. For other zeitgebers, it is time at which the zeitgeber value starts to increase from its lowest value.

Synopsis

Circadian clocks adaptively schedule behavior and physiology to occur at a specific time of the day. Such scheduling is believed to be critical for maintaining our general health and well-being. Heritable variations in the timing (or phasing) of rhythmic events with respect to daily time cues result in what is referred to as chronotype variation. A clear example is that of variation in mid-sleep timing among humans on free days –while most humans will fall in the category of 'normal' chronotype, some individuals are of "early" or "late" chronotypes. Chronotypes are primarily controlled by circadian clocks. Studies, including those on humans, have suggested that variation in entrained phases arise due to differences in underlying clock properties such as length of the intrinsic period of circadian clocks, phase/ velocity response curves (PRCs/ VRCs), amplitude of the circadian clock, inter-oscillator coupling, and amplitude of the zeitgeber. Several studies exploring the genetic basis of chronotypes have reported varying degrees of heritability across human populations; interestingly, some of the identified genes have also been implicated with similar functions in mice and Drosophila, thus suggesting that the genetic architecture underlying chronotype variation may at least partly be conserved across organisms.

A laboratory artificial selection approach was initiated ~20 years ago to understand the evolution of clock properties in divergent chronotypes and the natural genetic correlations of chronotype evolution. We now have populations of *Drosophila melanogaster* exhibiting early and late chronotypes with respect to their adult emergence rhythm. Over the course of ~350 generations, chronotypes were shown to be associated with differences in circadian period, zeitgeber sensitivity, amplitude, coupling, phase and period responses, and also molecular clocks.

I chose to address questions regarding contributions of non-clock mechanisms to phase divergence among chronotypes, sleep differences among chronotypes, and the genetic basis of chronotype. I have briefly summarized below the proposed chapters of my PhD thesis:

1. Characterization of masking responses to light and temperature in eclosion rhythm:

I hypothesized that our selection protocol has inadvertently resulted in selection for masking, a non-clock phenomenon, in the *early* chronotype due to the placement of our selection window (which includes the lights-ON transition). Based on theoretical predictions and previous studies on our populations, I designed experiments to discriminate between enhanced masking to light versus circadian clock mediated changes in determining enhanced emergence in the morning window in our *early* chronotypes. Using a series of phase-shift protocols, LD-DD transition, and T-cycle experiments, I found that our *early* chronotypes have evolved positive masking, and their apparent entrained phases are largely contributed by masking. Through skeleton T-cycle experiments, I found that in addition to the evolution of greater masking, our *early* chronotypes have also evolved advanced phase of entrainment. Furthermore, our study systematically outlined experimental approaches to examine relative contributions of clock versus non-clock control of an entrained behavior. Although it has previously been suggested that masking may confer an adaptive advantage to organisms, here I provided experimental evidence for the evolution of masking as a means of phasing that can complement clock control of an entrained behavior. Additionally, using temperature cycles, I found that *early* flies emerged exclusively during the cold part of the day (subjective night), whereas *control* and *late* flies had significant emergence under high temperature. I also observed that *early* flies show high emergence immediately after the high temperature is alleviated (end of subjective day, early hours of the subjective night), which clearly showed that the high temperature actively suppresses emergence in *early* flies as a negative masking effect. I concluded that *early* flies have evolved significantly more negative masking to high temperature compared to the *control* and *late* flies.

2. Characterization of masking responses in locomotor activity rhythm and sleep:

I also investigated the possibility of heightened masking in locomotor activity rhythm. Although there is no difference in the masking response among the stocks under LD12:12 (and other photoperiods) and ~70 lux light, I found that higher light intensity photoperiods were able to separate the clock-controlled and masking peak efficiently in our stocks. *early* and *control* flies showed significantly higher masking compared to *late* flies under long photoperiod(s), and under short photoperiods, the morning peak was negligible but well separated from the clock-controlled peak. I also observed that *early* flies showed a trend of higher masking with longer day length(s) compared to other stocks.

Sleep is perhaps the most versatile physiological state that is regulated by the circadian clock. Chronotypes have long been studied from the context of sleep timings (mid-sleep time), and the differences in different sleep parameters among different chronotypes are well-established. However, studies from humans and other organisms fall short in the aspect of evolution of sleep differences among chronotypes. With our experimentally evolved divergent chronotype populations, I have a model system that is uniquely suited to study evolution of sleep differences. With the goal to understand the evolution of sleep differences among divergent chronotypes, I characterized sleep in our populations. There was no difference in sleep amount or other sleep parameters (bouts, latency) among stocks under low light intensity (~70 lux). Under high light intensity (~500 lux, comparable to indoor office lighting), *early* flies showed significantly lower nighttime sleep and lower nighttime sleep quality (more fragmented sleep) compared to that of low light intensity. These differences became prominent under shorter day lengths, which are supposed to increase nighttime sleep, but extreme chronotypes (*early* and *late*) failed to sleep more compared to *control* flies. *early* flies also exhibited significantly more fragmented sleep compared to *control* and *late* flies under shorter day lengths. I speculated that extreme chronotypes experience significantly more disruptions in nighttime sleep amount and quality, which can possibly be overcome by using artificial light in later hours of the day.

3. Genetic characterization of divergent chronotype populations by pooled sequencing:

Experimental evolution in laboratory populations followed by whole-genome sequencing, commonly called "Evolve and Resequence" (E&R), is an attractive alternative for investigating the genetic basis of a selected trait. Genome-wide sequencing of genetic variation present in experimentally evolving sexual populations after many generations sheds light on the relative importance of selective sweeps, particularly alleles being driven to fixation.

The four genetically independent replicate populations of each *early*, *control* and *late* flies in my experimental design make it uniquely suited to understand the genomic basis of chronotype evolution in *Drosophila melanogaster*. The genomes of the *early*, *control*, and *late Drosophila melanogaster* populations were subjected to high-depth sequencing after pooling genomic DNA from ~500 individuals from each population to identify putative loci that are likely to be associated with entrainment-phase/chronotype differences. I performed various population genomic analyses and identified genomic regions undergoing positive selection in *early* and *late* populations compared to the *control* populations. I also identified various SNPs significantly enriched in either *early* or *late* populations, compared to the *control* populations. I further predicted different pathways plausibly under differential selection in *early* and *late* populations. I concluded that *early* and *late* chronotypes are not just two sides of the same coin, i.e., there is a minimal overlap among genomic regions where *early* and *late* flies have diverged. I observed that *early* flies have more changes in genes responsible for both negative and positive loop of the TTFL, learning, and memory, calcium channels, photoreception, ecdysone pathway, immunity. In contrast, *late* flies have more changes in genes responsible for only the negative loop of the TTFL, temperature sensitivity, metabolism of insect hormones, sleep, synapse formation, splicing mechanism, ETH signaling, light-sensitive calcium channels. Overall, I established a database of putative variations associated with divergent chronotypes, which are plausibly natural genetic correlations, and hope that this work will pave the way to investigate novel targets regulating phase divergence in Drosophila.

4. Screening of the DrosDel collection for differential phasing of eclosion rhythm:

I took another approach to discover loci affecting chronotype divergence in *Drosophila melanogaster* by screening several lines from the DrosDel collection for differential phasing of the eclosion rhythm. I chose 114 non-overlapping deletions covering ~55% of the euchromatin. Among the lines assayed for eclosion rhythms, only 40 could be entrained to a LD12:12 regime. By using a stringent cutoff of ±3 SD of the phase angle of mean phase of emergence and normalized amplitude (mean vector length in polar coordinates) of the control background, I find only 10 lines spanning over four major chromosomal arms (2L, 2R, 3L, and 3R) to be non-control like (early-like and late-like). These 10 lines cover a total of 595 genes and are enriched in genes involved in nucleosome and chromatin assembly, chromatin and gene silencing, ecdysone induced genes, etc. I proposed that these genes may act as fine regulators of phases of the eclosion rhythm in *Drosophila melanogaster*, but not to have a large role to play in the generation of the eclosion rhythm per se. Further narrower deletions will need to be screened to get a better idea of genes directly involved in regulation of phases.

5. Appendix: Development of open-source tools for timeseries analysis and visualization:

Chronobiologists and sleep researchers often need to estimate various rhythm and sleep parameters from locomotor activity data from different organisms. The available open-source or expensive paid tools do not offer consolidated analysis and visualization options in one bundle, are often cumbersome for users unfamiliar with coding, offer very low customization options, introduce sources of human errors by requiring users to manually pick period and power values from periodogram plots, and do not generate reproducible reports. I present VANESSA, a family of cross-platform apps written in R, which, in our opinion, have several advantages compared to available. I will continue to develop VANESSA with more valuable features, and version control will be done via archiving versions with significant changes on GitHub (https://github.com/orijitghosh/VANESSA-DAM) and Zenodo.

In conclusion, in my thesis, I have

1. established masking as a feasible mechanism governing phase of emergence rhythms
2. shown effect of high light intensity photoperiods on masking in locomotor activity rhythms and nighttime sleep amount and quality.
3. established a database of putative variations associated with divergent chronotypes which are plausibly natural genetic correlations.

4. found 10 large deletions spanning over four major chromosomal arms having regulators (~595 genes) of phase divergence in *Drosophila melanogaster*.

Future work will mostly focus on the effect of different light intensity and quality on locomotor activity rhythms and sleep in these divergent chronotype populations, creating point mutations to mimic variations associated with divergent chronotypes and establishing their effects on chronotype divergence, and further characterization of narrower deletion lines to find causal regulators of phase divergence in *Drosophila melanogaster*.

Chapter 1

Introduction

1.1 On the concept of time and timekeeping:

> "*In January 1906, several thousand cotton-mill workers rioted on the outskirts of Bombay. Refusing to work at their looms, they pelted factories with rocks, their revolt soon spreading to the heart of the city, where more than 15,000 citizens signed petitions and marched angrily in the streets. They were protesting the proposed abolition of local time in favor of Indian Standard Time, to be set five-and-a-half hours ahead of Greenwich. To early 20th-century Indians, this looked like yet another attempt to crush local tradition and cement Britannia's rule. It wasn't until 1950, three years after Indian independence, that* a single time zone was adopted *nationwide. Journalists called this dispute the "Battle of the Clocks." It lasted nearly half a century.*"
>
> – A Brief History of (Modern) Time (Ian P. Beacock, The Atlantic)

For those workers, preservation of local time may have been secondary to their desire to oppose their British oppressors, but this story reminds us that the global time as we know it today did not just emerge; it had to be invented, and moreover,, it had to be imposed. Naturally, the question arises why it was so important to have global standard times and, more broadly, a united understanding of time at all. Time, as a concept, has been a matter of debate among the scholarly circles of ancient philosophers and modern scientists alike. While the "reductionism/relationism" views of time have been supported by the likes of Aristotle and Leibniz, which argue that time is not independent of the events that occur in time; the other "absolute" views of time have been defended by Newton, Plato, and others, which view time as an empty container in which you can place events, but this container itself is independent of what is placed in it (Emery et al., 2020). These two views of time as a concept also affect defining the topology (structure/shape) of time itself. It is intuitive to represent time by a line, but is the line straight? Does the line have a beginning and end? Is it a straight line or a circular one? – Aristotle had vividly argued against there being a "first moment of time". Likewise, it is also worth asking if time can be represented by one single line, the same question that has fueled a tremendous amount of beloved science fiction through the years. However, no matter how enticing or intuitive it feels to have different notions about time, in the last century, Albert Einstein's theories of relativity popped the time bubble. These theories showed us that time is relative and is not absolute, as Plato or Newton had vehemently proposed and supported. Although some theoretical physicists, like Carlo Rovelli, speculate that time does not even exist, what is certain is our perception of time does exist. The human perception of time has been shaped by different cultures at different periods throughout history. In fact, an Amazonian tribe, the *Amondawa*, has no abstract concept of time (Sinha et al., 2011); however, there have been debates over if this is purely a linguistic argument or not. The usage of time as a concept is so inconspicuous that "time" is the most used noun in the English language, whereas different measures of time – "year", is the third most used noun; "day", fifth most, and "week" stands as the 17th most used noun. The oldest manmade timekeeping mechanism can be traced back to 8000 BC in eastern Scotland, where hunter-gatherers dug a row of 12 pits to track lunar cycles. In 2013, a sundial was unearthed in Egypt's valley of kings, which could be traced back to 1250 BC. The most accurate form of a timekeeping device in the ancient world was the water clock of *clepsydra*, in Egypt (1500 BC), and then had widespread usage in Alexandria, and later worldwide. In 1090 AD, a Chinese civil servant named Su Song built the first-known mechanical clock, knows as the "Heavenly Clockwork" (Richard et al., 1997), which was an elaborate water-powered machine; traditional mechanical clocks appeared two centuries later in Europe. The influence of religion was also evident in history regarding the evolution of timekeeping – The Julian calendar, even though it incorporated leap years, still fell short of a day every 134 years, which posed confusion for the Roman Catholic Church because the important holidays like Easter, were drifting from spring into summer. This forced them to consult mathematicians and astronomers, and finally, Pope Gregory XIII implemented the Gregorian Calendar, which subtracts one day every 3300 years. Around 1650, Galileo first suggested pendulums as a physical mechanism for regulating clocks, though these were implemented much later by Dutch craftsmen, inspired by the designs of astronomer Christiaan Huygens. Hourglasses were also a major device to measure the flow of time, especially in navigation, e.g., Ferdinand Magellan, when circumnavigating the globe (1522), had 18 hourglasses on each of his ships. While western cultures embraced a linear concept of time, many eastern cultures viewed time as recurring cycles.

> *"Hinduism and Buddhism, for example, adopted a cyclic view of time that suggested the eventual return of the world to its former state; nothing is permanent, and even death is merely a passage to rebirth and renewal."*
>
> – Dan Falk (Arrows of time, Quanta Magazine)

Fast forward to 1729, an 18th-century astronomer, Jean-Jacques d'Ortous de Mairan, showed and recorded that the leaf movement of the *Mimosa* plant had a periodicity of ~24 hours in their leaf opening and closure, even under constant darkness in a "dark place". Scientists slowly realized there is a clock inside us, inside most living organisms – biological clocks.

1.2 On biological clocks and their nature:

Our ancestors must have been aware of the periodic changes in their environment – starting from changes of seasons to the daily blooming of a flower, their own sleep-wake rhythms; perhaps these rhythms were so obvious, that they attracted virtually no experimentation or justification, just became facts. Aristotle wrote about sleep-wake cycles in animals, but it took de Mairan 2000 years from then to experimentally show that the daily leaf movements in the *Mimosa* plant persist even in the absence of a solar day-night cycle and hinted toward an endogenous "clock" in organisms (De Mairan, 1729). de Mairan and his colleague, Jean Marchant at the Royal Academy of Science, even drew parallelism between the phenomena of plant leaf opening and closing under constant darkness to sick bed-ridden people feeling the day or night from their beds without ever actually seeing daylight.

> "*It may have a probable connection with the ability of sick people to differentiate day and night from their bed.*"
>
> – Marchant and de Mairan (De Mairan, 1729)

With this observation, they hinted toward the possibility of the existence of an endogenous clock in humans as well. In conclusion of the essay which Jean Marchant wrote in lieu of de Mairan's interest in writing a paper on a mere plant, they stated that true science is experimental and a very slow process indeed it took scientists from different fields ~200 years to establish the endogenous nature of the biological clock conclusively. While in de Mairan's experiments, the only cycling environmental factor considered was the solar day-night cycle, there could as well be many more cycling biotic or abiotic factors, e.g., temperature or humidity. Later, it was reported that leaf movements of plants could persist in a dark cave environment, with constant temperature, along with constant darkness (du Monceau M., 1758). An American endocrinologist, Frank Brown, did not believe that biological clocks are endogenous in nature. The major argument he had was that there might be a factor present, generated by the earth's axial rotation, which experimental biologists were not accounting for while performing experiments under constant conditions to prove the endogenous nature of biological clocks (Brown, 1970). The cynicism of scientists like Brown led to extensive experimentations in the second half of the twentieth century to settle this debate and ruled out the presence and contribution of the majority of the geophysical cycles towards the generation of biological rhythms. These experiments even led scientists to run experiments in the "Spacelab", far away from the earth's influence, and they still found the conidiation rhythm of Neurospora to be present in the absence of absolutely all geophysical cycles (Sulzman et al., 1984). During the race to prove the endogenous nature of the biological clocks, in the 1920s, Anthonia Kleinhoonte and Erwin Bünning proved conclusively using light-dark cycles and showing plants under constant light conditions had a sustained rhythm of leaf movement which significantly deviated from 24 hours respectively, that the leaf movement rhythms could not be attributed to any geophysical cycles, but had to be endogenous (Bunning, 1930; Daan, 2010; Kleinhoonte, 1929). This was also shown by Augustus Pyramus de Candolle much before Kleinhoonte and Bünning, stating Mimosa leaf movement rhythm had a periodicity of ~22 hours under constant light (de Candolle, 1832). Around the same time, data supporting non-24 hour period endogenous rhythms started pouring in from animals – first in the crustacean *Hippolyte varians* (pigment change rhythm), and then from experiments of psychologist Curt Richter, who studied locomotor activity rhythms of rats under constant conditions (Gamble and Keeble, 1900; Richter, 1922). These rhythms, which appeared and were sustained under constant conditions with non-24-hour periodicity, are now known as free-running rhythms and are considered definitive proof of endogenous generation of these rhythms. The word "circadian" (from latin *circa* – approximately, *dian* – a day) was coined by Franz Halberg to emphasize the nature of deviation of the free-running period from 24 hours (Halberg, 1959).

Jürgen Aschoff, a German medical doctor by training, entered the research on biological rhythms in the 1950s. He raised seven generations of rats under constant light and found that in each generation, the robustness of the daily rhythms was preserved (Aschoff and Meyer-Lohmann, 1954). Fearing that embryos would have been exposed to the mother's rhythms *in-utero* (Davis and Mannion, 1988), he also raised chickens developing inside eggs under constant light and observed normal daily rhythms, showing this information is not transferred from mothers but are generated from within the organism (Aschoff, 1955). Although these experiments did not prove that prior experiences do not play any role in rhythm generation, they definitely demonstrated prior experience of a rhythmic environment was not needed for rhythm generation, thus lending evidence to support the endogenous nature of the circadian rhythms.

Ingeborg Beling, along with her supervisor Karl von Frisch did extensive experiments with honeybees and discovered a "time memory" in them. von Frisch later reported that the honey bees must use a clock system to know the passage of time (Beling, 1929; Frisch, 1950). Their observations were supported by another German scientist, Gustav Kramer, who arrived at similar conclusions about starlings using a time-compensated sun-compass for seasonal migrations (Daan, 2010; Kramer, 1950). Maynard Johnson, during his studies under constant light with white-footed mice noted a very important observation that the free-running period of the animals under constant light changed with the intensity of the light (Johnson, 1939) – higher intensity increased and constant darkness reduced the periodicity. These observations were supported by Hans Kalmus, who showed the periodicity of the adult emergence rhythm in Drosophila was strongly dependent on the ambient temperature (Kalmus, 1940). These observations posed a question – if the circadian clock is susceptible to changes in environmental light or temperature conditions so much, how it can maintain stable phases/timing under natural conditions, where most environmental cues vary tremendously over the day. Colin Stephenson Pittendrigh, a British scientist, rose to this challenge and focused on the Drosophila adult emergence rhythm. He soon showed that Kalmus mistook a transient change in the first cycle after temperature change for a true change in the periodicity of the rhythm. By observing multiple subsequent cycles of emergence after temperature change, Pittendrigh proved that the free-running period of the emergence rhythm was not affected by the temperature change (Pittendrigh, 1954). This phenomenon of circadian clocks resisting change in their period as a response to temperature change/perturbation soon came to be known as "temperature compensation" and was confirmed in other systems (Hastings and Sweeney, 1957).

The term "entrainment" is used widely today in research of circadian rhythms and clocks. In the early days of circadian clock research, it was immediately evident that the circadian rhythms were synchronized by at least one of the major geophysical cycles – the alternation of light and darkness in 24 hours. Aschoff coined the term "Zeitgeber" (German for "time giver") for these signals or time cues that can synchronize the internal circadian clock of organisms. Zeitgeber includes periodic signals that can entrain the circadian rhythm to its own periodicity and can establish a stable phase relationship with it. In the 1950s and 1960s, the study of entrainment became the focus of much of the work on circadian rhythms and clocks. Most of the theories of entrainment were developed by Aschoff and Pittendrigh.

1.3 On the importance of timing, chronotypes, and entrainment of circadian clocks:

> "*Under ordinary conditions, the cycling of this innate biological clock is synchronized by the overwhelmingly greater geophysical clock on which it is modeled. But the inner clock*

> *still beats and plays a vital role. Our task is to listen for the inner beating of the biological clock in those rare situations where it can be heard independently*."
>
> – Arthur Winfree (Winfree, 1987)

Recording biological rhythms under free-running (constant) conditions has established some of the fundamental properties associated with circadian clocks today, namely, endogeneity and temperature compensated periodicity (Brady, 1987; Pittendrigh, 1954). However, Winfree also says how the biological clock is modeled on the geophysical clock. He wrote, "*A clock is not much good if you can't pull out its stem and reset it*" – the critical property of biological clocks is the ability to reset when presented with time cues, the ability to entrain itself to periodic environmental time cues (Winfree, 1987). It is obvious that circadian clocks must have evolved under cycling conditions. Constant conditions or nearly constant conditions are rare in nature. Various environmental factors can act as time cues for the circadian clock. Measuring the passage of time internally can provide organisms an upper hand compared to merely reacting to these changes in the environment. Externally, it can help time physiological and behavioral processes optimally by anticipating daily changes; internally, it can maintain a temporal order so that processes can be timed with respect to each other and/or provide a reference timepiece under conditions that lack time information (Vaze and Sharma, 2013). The process by which a biological oscillation adopts the periodicity of the environmental oscillation is termed entrainment. A key difference between this entrainment of the circadian clock with mere synchronization is that the waveforms of the two oscillations are necessarily not the same. Two mechanisms of entrainment have been proposed till now – propagated by the two stalwarts of the field, Pittendrigh, and Aschoff. The phasic model suggested that a time cue has an instantaneous effect in setting the phase of the circadian clock, and that transitions were crucial in entraining a biological rhythm

(Daan, 2000; Daan and Pittendrigh, 1976). On the other hand, the tonic model suggested that time cues would have a cumulative effect on the speed of the oscillation (Daan, 2000; Swade, 1969). How entrainment occurs remains a widely studied question, and how entrainment is achieved might differ among rhythms and/or organisms. All things considered, it is not a far-fetched assumption that under natural conditions, natural selection would act on the timing (phase) of a biological rhythm but not on the free-running period. As a result of this selection, other clock properties might also undergo changes. The timing of a biological process with respect to the environmental cycle is referred to as the phase of entrainment (Ψ_{ent}). Ψ_{ent} had been defined in the field quite early and is a measurable outcome of the effect of a time cue (zeitgeber) on the biological oscillator (Daan and Aschoff, 2001). It is measured as the time relation between a phase of the biological oscillation, like the time when activity starts and a phase of the environmental oscillation, e.g., lights-ON. So, the question arises – what can affect this Ψ_{ent}?

1) Free-running period of the circadian clock (τ) – the difference between the free-running period of the rhythm and the period of the environmental cycle (T) determines the phase relationship (Bordyugov et al., 2015; Pittendrigh and Daan, 1976).
2) Strength of the zeitgeber – Physical characteristics of the zeitgeber, especially its range of intensities (amplitude), are referred to as its strength. The dependence of phase angle on the amplitude of a zeitgeber has been discussed and demonstrated using empirical data and theoretical frameworks (Aschoff, 1960; Eelderink-Chen et al., 2015; Roenneberg et al., 2019).
3) Sensitivity of the clock to the zeitgeber – Measured as the phase response to a zeitgeber of fixed strength and duration, the amount of phase shift incurred when presented with the zeitgeber affects the phase angle of a rhythm (Aschoff, 1960; Aschoff and Pohl, 1978).

Both 2 and 3 constitute how strongly the clock and zeitgeber will be coupled to each other.

4) Coupling within the system (e.g., coupling between the central and peripheral/driving oscillators or coupling among the components of the central pacemaker) – Stronger coupling within the system translates to greater amplitude of the whole which under theoretical and experimental framework has been shown to have more rigidity in the face of resetting signals. Thus, differences in coupling will also alter the phase angle of entrainment (Schmal et al., 2018).
5) Mechanisms that can bypass the clock and induce direct responses to zeitgebers (often referred to as masking) will also influence the overt/observed phase angle of a rhythm (Mrosovsky, 1999).

The Ψ_{ent} may also differ depending on the biological rhythm being measured even within the same individual because the driving (peripheral) oscillator and/or predominant mechanism of entrainment may be different. In humans, for example, Ψ_{ent} may be different for the sleep-wake cycle, DLMO (dim light melatonin onset), and body temperature rhythms. Within the individuals of a species for a given circadian rhythm, there exist substantial variations in the Ψ_{ent}, giving rise to chronotypes. The variations may arise due to differences in one or more of the above-stated factors. The term "chronotype" was first proposed in 1974 as "an organism's temporal organization" or "a temporal phenotype" (Ehret, 1974; Samis, 1978).

> *"We like the term "construct" because chronotype actually pertains to the organization of an entire system and not to one of its subparts, like the suprachiasmatic nucleus (SCN) or liver (the "temporal program", as Colin Pittendrigh called it,* (Pittendrigh, 1993)*). It is thus virtually impossible to directly assess an individual's phase of entrainment, i.e., her*

> *or his internal time, since there is no single circadian phase of entrainment of an organism."*
>
> – Till Roenneberg and colleagues (Roenneberg et al., 2019)

Though estimating the Ψ_{ent} of the complete circadian system of an organism is difficult, we can use the timing of different biological processes under the control of the circadian clock as a proxy/biomarker for the underlying clock (Roenneberg et al., 2019). Such biomarkers used in humans can be acrophase of activity, DLMO, etc. Entrainment to the solar day-night cycle has been shown as more prevalent than social timing in population-scale country-wide "chronotype questionnaire" studies in Germany (Roenneberg et al., 2007). A student of Aschoff, Till Roenneberg, has devoted a significant portion of his career pioneering these large-scale internet-based human chronotype research. His renowned Munich ChronoType Questionnaire (MCTQ) uses mid-sleep on free days (MSF), a variable derived from the users' self-reported sleep timing on the questionnaire. Individuals with earlier phase of entrainment have traditionally been referred to as early chronotype or "larks" while those with later phase of entrainment as late chronotype or "owls". It should be noted that early and late is relative and should be defined after considering the chronotype distribution and central tendencies of the phase angle of the sample/population. Another questionnaire – the Horne-Östberg Morningness-Eveningness Questionnaire (MEQ) produces similar results to MCTQ; MCTQ additionally collects information such as sleep-wake behavior under natural conditions (Horne and Ostberg, 1976; Zavada et al., 2005). The distribution of chronotypes in a population is primarily contributed by the genetic makeup of the individuals, the age structure of the population, and environmental conditions (Carskadon et al., 2004; Hsu et al., 2015; Roenneberg et al., 2004, 2013). Chronotypes become later in winter, probably due to later sunrise and sunset timings and lower light intensities throughout the day (Kantermann et al., 2007). This variation in chronotypes demonstrates the remarkable plasticity of the entrainment process of the circadian clock (Roenneberg and Merrow, 2016). Three separate clocks affect our daily lives: a) the Social clock, which represents the local time, b) the Sun clock, related to the axial rotation of earth, creating day-night cycles and providing one of the strongest zeitgebers to organisms; and c) our Biological clock (circadian clock) (Roenneberg et al., 2019). Before global time zones were implemented in the late 19th century (International Meridian Conference, 1884), the social clocks were aligned to sun clocks, but due to centralized time zones, the social clocks became different than the sun clocks. This dissociation between social clock and sun clocks has become even more prominent with the introduction of daylight-saving times and depends on local government and civil work culture. Circadian alignment/misalignment can be explained as – "*if circadian rhythms in the different organ clocks are not only synchronized to 24 h but also adopt normal phase relationships to one another, they are aligned; if they are synchronized to 24 h but adopt unusual phase relationships to one another, they are misaligned. Traditionally, this applies to the relationship between the circadian program of an individual on one hand and the timing of the physical (e.g., light and darkness) and the social (e.g., school and work times) environment on the other.*" (Roenneberg and Merrow, 2016). Circadian misalignments have become rampant, and if characterized and diagnosed correctly, may actually take the shape of an epidemic (Roenneberg and Merrow, 2016). The clinical guideline to identify Advanced or Delayed Sleep Phase is two or more hours earlier or later, respectively, relative to desired or socially customary sleep times (Auger et al., 2015). According to data from Till Roenneberg (unpublished, derived from the MCTQ database as of November 2015), approximately 25% of the population satisfies this criterion according to the distribution of chronotypes (Roenneberg et al., 2012). The circadian alignment/misalignment problem is basically related to variation in chronotypes in populations, and thus it makes it an imperative to study what gives rise to variations in chronotype, how chronotypes evolve and the genetic basis of chronotype diversity, to make clinical interventions possible.

1.4 Study of chronotypes in laboratory populations:

The idea of chronotypes and that selection would act on the timing of biological rhythms has been around since the conception of the field of chronobiology. Correlation between Ψ_{ent} and the factors that affect it (listed in the section above) became apparent with empirical data as well as theoretical frameworks. However, a robust experimental system would be needed to make strong conclusions regarding the same. Colin Pittendrigh, one of the founders of the field, made efforts to study the adaptive significance of the timing of biological rhythm as a trait and how selection for this trait could affect clock properties (like free-running period, intrinsic amplitude, coupling in the system, and sensitivity of the system). Laboratory selection can help one make causal arguments because the selection is imposed on a trait of choice under controlled conditions. Thus, one can eliminate several potentially uncontrolled variables. Changes in the trait in response to selection is suggestive of its adaptive significance, and subsequently, correlated changes can be studied. With replicate populations and systematic selection, this framework of study has the potential to unveil which clock properties would change upon selection on timing/chronotypes (of eclosion) and how. The advantages of laboratory selection approaches are many – a) it gives the experimenter control over the experimental setup and levels of replication, b) independent population-level replicates along with their ancestral unselected controls allow one to dissect the contribution of selection pressure and genetic drift to the evolved trait, c) it allows one to make causal arguments in favor of the evolved trait being an adaptation to the imposed novel ecology (novel ecology being selection for divergent chronotypes in this case), d) it allows quantitative estimation of trajectories of the evolution of traits, and the prospect of detailed genetic analyses (Garland and Rose, 2009; Hartl et al., 1997). Like any other experimental method, laboratory selection also has certain disadvantages – a) due to rigor (population-level replicates, strict maintenance regime, application of selection pressure at consistent intervals, etc.) involved in these experiments, they are best suited for laboratory conditions than wild, b) lack of ecological realism, in these experiments, generally one ecological factor is varied, whereas in the wild, multiple factors acting in synergy are likely to affect traits and their evolution, and thus adaptive values of those traits may actually differ from what we infer manipulating one factor at a time in the lab. Nevertheless, if we weigh the advantages and limitations of all other methods of studying adaptive significance and traits underlying divergent circadian programs, it becomes clear that laboratory selection is among the few ideal strategies currently available (Abhilash and Sharma, 2016). To carry out selection experiments, it becomes essential that the model organism has a short generation time and can be maintained with large standing genetic variation (large population size). This limits the model organisms that can be used for studies. Fruit flies emerged as an apt model system that can be used for such studies with selection on timing of rhythm because not only does it satisfy the mentioned criteria but also because its eclosion rhythm was among the first systematically studied and characterized circadian rhythms (Chandrashekaran, 1967; Chandrashekaran and Loher, 1969; Pittendrigh, 1954, 1967). Eclosion is the process of the emergence of an adult from the pupa in the life cycle of a holometabolous insect. Thus, selection for "early" and "late" chronotypes was carried out in *Drosophila melanogaster* as well as *Drosophila pseudoobscura* using the eclosion rhythm (Pittendrigh, 1967). The selection was carried out on wild-type population of the species under LD12:12 conditions. After 50 generations of selection, the phase difference between the two strains was ~4 hours. For *D. pseudoobscura* the period of eclosion and locomotor rhythms had changed in the same direction, indicating that the central pacemaker had undergone a change due to selection on the phase of eclosion. Such changes in phase could be attributed to changes in the free-running period of the pacemaker, the difference in sensitivity of the central pacemaker to the zeitgeber (measured by differences in the phase response curve), or modification of the coupling of the pacemaker to the downstream mechanism that controls eclosion. It was found that “early” had a longer free-running period (FRP) than “late”; this difference was small and did not explain the phase difference. The phase response was not found to differ in the two selected strains compared to their parent unselected population. Selection for “early” and “late” eclosion in *D. auraria* resulted in “early” having free-running period greater than “late” by almost 2.5 hours. However, when *D. auraria* were selected for early and late eclosion under an extremely short photoperiod of 15 minutes, it resulted in “early” having a shorter free-running period than “late” (Pittendrigh and Takamura, 1987). Information regarding the exact time windows and levels of replication is missing in these reports. The data did not add up to the phasic or non-parametric model of entrainment, and even in the cases when it did, the differences in the free-running period did not explain the divergence in phase. In *D. melanogaster* Oregon R and W2 strains were used to select for “early” and “late” eclosion; the time window for early selection was ZT 18 to ZT 22, and that for late was ZT 2 to ZT 6 (Clayton and Paietta, 1972). This study was focused on the comparison between the strains than on evolved clock properties, yet they reported that only after 16 generations of selection, the percentage of flies emerging in the morning and evening selection windows was significantly higher than control flies. However, in all the above studies, it became obvious that flies respond well to selection for the timing of eclosion. Taking the aforementioned limitations, inconsistent results of previous studies, and lack of detailed analyses of evolved clock properties into account, our research group employed a laboratory selection approach to study the evolution of different clock properties and divergence of genetic trajectory along with the evolution of divergent chronotypes.

1.5 GATE populations at JNCASR:

To study the evolution of clock properties and genetic architecture along with chronotype divergence, our lab had initiated a long-term experimental evolution study and continues to maintain populations of *Drosophila melanogaster* that show divergent timing of the eclosion rhythm. These populations have been used for over ~350 generations (~20 years) of selection to examine different clock properties underlying divergent chronotypes. We call these populations the "GATE" populations. This term does not have any scientific merit and is more of an idiomatic reference to our restricting the gate of their emergence to certain times of the day. I will describe the results from these past studies on the GATE populations and their maintenance regime here.

1.5.1 Generation and maintenance protocol of *early* and *late* populations:

Four replicates of *early* ($early_{1\text{-}4}$), *control* ($control_{1\text{-}4}$), and *late* ($late_{1\text{-}4}$) populations were derived from four common ancestral, large, and outbred populations approximately 20 years ago (Kumar et al., 2007). The *early* and *late* populations have been subjected to selection for the timing of adult emergence phases since then and are maintained as independent populations for more than 350 generations now. The $early_{1\text{-}4}$, $control_{1\text{-}4}$ and $late_{1\text{-}4}$ populations that share the same subscript (referred to as "blocks" or "replicates" in this thesis) share a common ancestry, and the populations with different subscripts indicate independent genetic substructure. All the 12 populations, four each of *early, control,* and *late,* are maintained on banana-jaggery (B-J) medium under conditions of LD12:12 (with ~70lux light intensity during the photophase) at 25±0.5°C and ~65-70% RH on a 21-day discrete, non-overlapping generation cycle. Only the flies emerging between ZT21 to

ZT01 (Zeitgeber Time 00, or ZT00 is the time of lights-ON in any LD cycle) on days 9th to 13th post egg collection are collected to form the breeding population for the next generation of the *early* populations (*Fig. 1.1A*). Similarly, only the flies emerging during ZT09 to ZT13 on the same days as that of *early* populations are collected to form the breeding population for the next generation of *late* populations (*Fig. 1.1A*). On the other hand, flies emerging throughout the day are collected to form the next generation of *control* populations (*Fig. 1.1A*). On the 18th day after egg collection, flies are provided with a petri-plate full of B-J medium covered with live yeast paste as a protein supplement for three days. On day 21, cut plates of B-J medium are provided to all the fly populations for ~6-h to lay eggs. These eggs are collected and dispensed into vials in a batch of ~300 eggs/vial to initiate the next generation; we collect 24, 16, and 48 such vials each for the *early*, *control* and *late* populations, respectively, owing to inherent differences in the emergence in their respective windows and as a way of ensuring that there are sufficient flies in the cages for initiating the next generation.

Before performing any of the assays reported in this thesis, all populations were subjected to one generation of common rearing (standardization) to minimize maternal and non-genetic inheritance effects on the trait being measured (Bonduriansky and Day, 2009). The offspring of the standardized populations are, henceforth, referred to as standardized flies/populations (Abhilash, 2020).

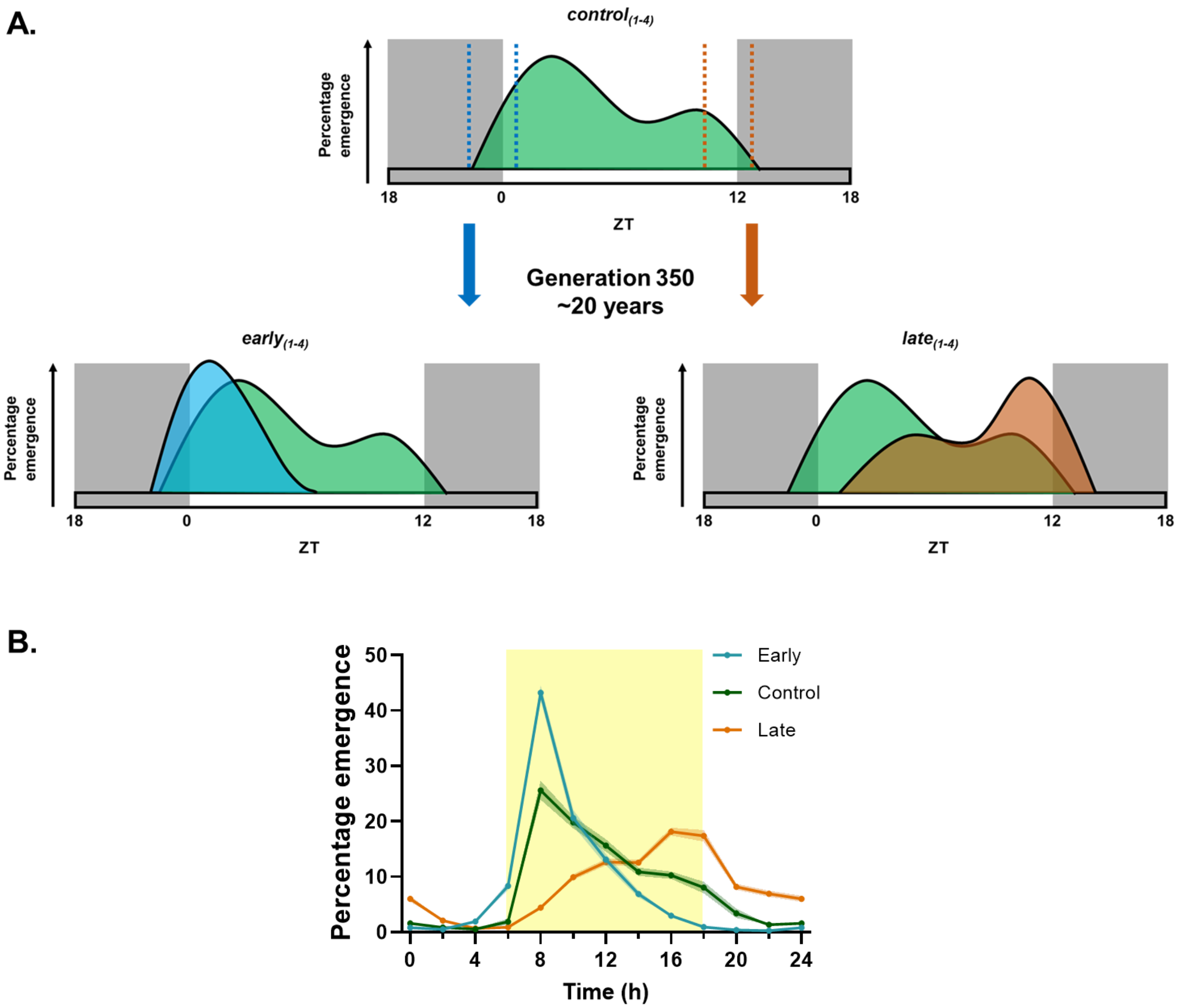


Figure 1.1: Schematics of selection protocol and emergence profiles of *early*, *control*, and *late* populations. A. Green filled schematic waveforms depict emergence waveform of the control population, blue and orange filled schematic waveforms depict emergence waveforms of *early* and *late* populations after ~350 generations of selection. The blue dotted lines show the selection window for early population (ZT21-01) and orange dotted lines show the selection window for late population (ZT09-13). **B.** Empirically obtained emergence profiles of *early*, *control*, and *late* flies at generation 320 under LD12:12 and constant 25°C. Error bands are ±SEM. Yellow shaded region is light part of the day.

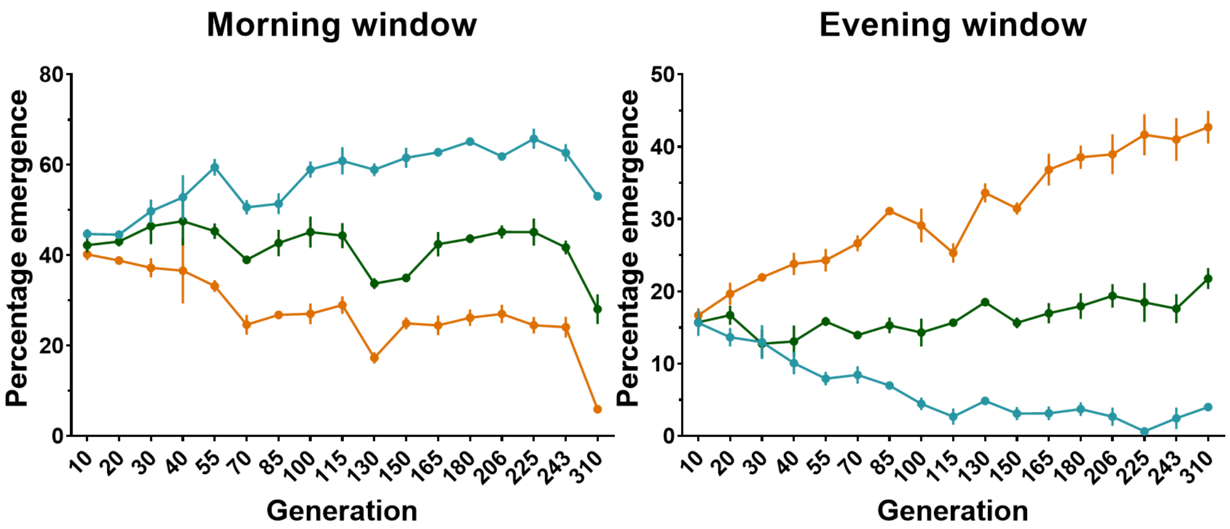


Figure 1.2: Percentage in the morning window and evening window of *early*, *control*, and *late* flies. Percentage of flies emerging in the morning window was estimated as the sum of all flies emerging after ZT20 and up to ZT02, and in the evening window as the sum of all flies emerging after ZT08 and up to ZT14. Error bars are ±SEM. Color codes: blue – *early*, green – *control*, and orange – *late*.

1.5.2 Direct response to selection, evolution of the emergence waveform, and associated changes in clock properties:

I analyzed direct responses to selection using data collected from eclosion rhythm assays performed under LD12:12 cycles (25°C) at different times through the course of selection. I estimated the percentage of flies emerging in the morning window as the sum of all flies emerging after ZT20 and up to ZT02, and in the evening window as the sum of all flies emerging after ZT08 and up to ZT14. These windows (different from the selection windows as described above) are chosen because eclosion rhythms have been assayed in the laboratory in 2-hour intervals at even time-points. Percentage eclosion in each window is estimated for each vial, and these are averaged to provide estimates of block means. The *early* populations continued to exhibit an increase in percentage eclosion till the 100th generation, beyond which the percentage does not change much. The *late* populations showed a continued reduction in percentage eclosion in the morning window till about generation 70, beyond which percentage remained stable. In the most recent experiment, I found that the *late* populations showed a further reduction in percentage eclosion in the morning window compared to all previous generations. As all three populations show a reduction in percentage eclosion in the morning window, I think that this could be an experiment-specific feature at that generation. However, the reduction in both the *late* populations and the *early* populations is indicative of this reduction being a possible outcome of population-specific changes to selection, at least in part. Assays in subsequent generations are required to gain clarity on this issue. In the evening window also, the *control* populations continued to show fairly low emergence in the window, the *early* chronotypes had reduced their percentage eclosion to almost 0, around the 100th generation onwards (*Fig. 1.2*). The *late* chronotypes showed steadily increasing percentage eclosion in the evening window, with the most recent generation (310) showing a nearly 100% increase (*Fig. 1.2*).

Waveforms are simply the shape of the oscillation over the entire length of one cycle. As a consequence of the evolution of percentage eclosion in the morning and evening windows, the *early* and *late* chronotypes have also evolved to exhibit very divergent eclosion waveforms (*Fig. 1.1 & 1.2*) under standard maintenance conditions described above. As a consequence of the direct effects of selection, there appears to be concomitant evolution of phases of the behavior – mean phase of emergence (phase of the center of mass of the emergence waveform – Ψ_{CoM}) (Ghosh et al., 2021). To analyze this parameter of the waveform after 320 generations of selection, I used block means of Ψ_{CoM}. The *early* populations evolved an advanced Ψ_{CoM} relative to both *control* and *late* populations (detailed in chapter 2). Previous studies have shown that the *early* populations evolved an advanced phase of onset of emergence relative to both *control* and *late* populations while the latter two did not differ from each other (Abhilash, 2020; Abhilash et al., 2019). In the case of the phase of the peak of emergence, the *early* and *control* populations did not differ from each other, but the *late* populations evolved a delayed phase relative to both (Abhilash, 2020; Abhilash et al., 2019; Ghosh et al., 2021). Further, the *early* populations evolved a significantly advanced phase of offset of emergence relative to *control* populations, whereas the *late* populations evolved a delayed phase (Abhilash, 2020; Abhilash et al., 2019). Previous studies also found that the *early* populations evolved a significantly higher peak of emergence compared to the *control* populations and the *late* populations evolved a diminished peak (Abhilash, 2020; Abhilash et al., 2019). In accordance with the emergence profiles (*Fig. 1.1B*), previously we found that the *early* populations showed a highly narrow gate width relative to the *late* populations, which showed a much wider allowed zone for emergence (Abhilash, 2020).

The first published study describing these populations and associated changes in clock properties underlying such divergent evolution of emergence waveforms found that the FRP (τ) of the emergence rhythm under constant conditions also evolved in divergent directions in the *early* and *late* populations (Kumar et al., 2007). While the *early* populations showed a shorter period (22.51 h) than *control* populations (22.94 h), the *late* populations showed a longer period (23.86 h) than the *control* populations, a result expected from the non-parametric model of entrainment (see previous sections). Kumar and colleagues also found that the photic PRCs of these stocks had also evolved such that the *early* populations showed larger phase advances and the *late* populations, larger phase delays (Kumar et al., 2007) – this lends support to the relationship between Ψ_{ent}, and FRP observed before, as this relationship will hold true only if the photic PRCs of the populations are similar in shape. This implied that temporal light utilization by the two divergent chronotype populations to achieve their characteristic evolved phases (Ψ_{ent}), must be different. To test this, another study was carried out, wherein emergence waveforms were assessed under a skeleton photoperiod regime with a 15-minute light pulse given starting at the time of lights-ON to indicate dawn and another 15-minute light pulse given starting 15-min before lights-OFF to indicate dusk (the regime is referred to as skeleton photoperiod because it provides a skeleton to a full photoperiod, with one pulse indicating dawn and the other indicating dusk) (Vaze et al., 2012a). The authors found from this experiment that the evolved emergence waveforms under LD12:12 were not replicated under the skeleton photoperiod for any population. Waveforms under the two regimes are expected, from theory, to be similar under the assumptions of the non-parametric model of entrainment. This result suggested that perhaps parametric effects of light contribute to entrainment and, therefore characteristic waveform shape under LD12:12. To test this, they provided the populations with two asymmetric skeleton photoperiods (Vaze et al., 2012b). Interestingly, they found that the *early* populations showed waveforms identical to those under LD12:12 when lights were ON during the second half of the day, whereas the *late* populations showed similar waveforms to those under LD12:12 when lights were ON during the first half of the day. As the *early* populations need delays to entrain and had smaller delay shifts in their PRC, the authors argued that they must require light for a longer duration in the second half of the day (part of the day corresponding to phase delays). Reciprocally, they also argued that the *late* populations must require a longer duration of light during the first half of the day when they will incur phase advances, as they have smaller advance zones. These results provided evidence suggestive of the dominant role of parametric effects of light on entrained behavior of the emergence rhythm in *early* and *late* populations. Further, Kumar and colleagues showed that in addition to changes in the FRP of the adult emergence rhythm, FRP of the locomotor activity rhythm also evolved such that FRP of *early* was shorter than FRP of *late* (Kumar et al., 2007). This result was also observed in another study performed much later along the course of selection (Nikhil et al., 2016b) and persists even after 320 generations of selection. It is interesting here to note that while such differences in the clock period exist for the locomotor activity rhythm, these rhythms do not show any difference in the entrained phase of their rhythm under LD12:12. This was suggestive of divergent photic PRCs of the locomotor activity rhythm. Therefore, yet another study from our laboratory assayed phase-responses of the locomotor activity rhythm and found that there was no difference in the PRCs of *early* and *late* populations (Nikhil et al., 2016c). This indicated that perhaps, the locomotor activity rhythm, like that of the emergence rhythm, also predominantly utilized parametric effects of light to entrain to LD cycles. This idea also gained anecdotal evidence from other experiments that examined light sensitivity of the locomotor activity rhythm clock using a wide repertoire of experiments (Nikhil et al., 2016c). Along a different line of investigation to examine the sensitivity of *early* and *late* populations to different zeitgebers, the emergence waveform of these populations was examined under an outdoor enclosure, wherein light and temperature cycled as in nature (referred to as semi-natural conditions). Authors found that the divergence between the emergence rhythm of the two populations drastically increased under seminatural conditions relative to divergence under laboratory LD12:12 (Vaze et al., 2012b). To further understand the role of light and temperature, other experiments in our laboratory were performed, wherein emergence waveforms were observed under temperature cycles alone and light and temperature cycles in-phase and out-of-phase. It was found that while light reduced chronotype divergence, temperature enhanced the same. Further, light appeared to have an overall phase-delaying effect and temperature, a phase-advancing effect (Nikhil et al., 2014). The authors, based on these results, argue that under natural conditions, optimal phase-relationships are driven by a synergistic effect of light and temperature, thereby giving rise to observed chronotype divergence. A study was then undertaken to understand the effect of low and high amplitude temperature cycles on the phase divergence of *early* and *late* populations. This study (Abhilash et al., 2019) showed that the *early* populations do not vary their Ψ_{ent} under different temperature regimes, whereas the *late* population showed high phase lability (up to ~5 hours), which implies the existence of a genetic correlation between Ψ_{ent} and temperature sensitivity of the circadian clock. Further, to study if the molecular clockwork has also diverged between *early* and *late* populations, mRNA expression profiles of core clock genes such as *per*, *tim* and *clk* (Hardin, 2011), and two more circadian clock components representing input (circadian photoreceptor, *cry*) and output (*vri*) pathways (Hardin, 2011) were assessed. Pigment Dispersing Factor (PDF) levels (neuropeptide orchestrating circadian rhythms) were also assessed to test for the hypothesis of a weakly coupled oscillator network in *late* populations. It was found that, in accordance with their emergence chronotypes, the phase of *per, tim*, *clk*, *vri* and PDF oscillation in *early* and *late* populations have diverged with the mRNA and neuropeptide levels peaking earlier in *early* populations relative to *late* populations (Nikhil et al., 2016b). Furthermore, amplitude and levels of mRNA and PDF oscillations have also diverged between these two populations (Nikhil et al., 2016b). These results were taken to suggest that since *vri* apart from being an output molecule, also regulates *per* and *tim* mRNA expression (Hardin, 2011), selection on timing of eclosion probably drove the divergence of *vri* oscillation which in turn may have caused the divergence of the core molecular clockwork between *early* and *late* populations (Nikhil et al., 2016b); thus, highlighting the possible ways in which selection on Ψ_{ent} of a circadian behavior might drive the evolution of underlying circadian clocks. The above discussed studies highlight that features, such as FRP, amplitude, PRC and network level properties of circadian clocks co-evolve in response to selection on timing of behavior; therefore implying that various aspects of circadian organization and their relative responses to different zeitgebers may, in principle, affect the ways in which organisms entrain to light and temperature, thereby influencing timing of behavior (also discussed in the previous sections) (Abhilash, 2020).

1.6 Summary of the present study:

To advance our understanding of how the divergent chronotype populations in the lab have achieved their advanced and delayed Ψ_{ent} under their maintenance regime, divergence in other coevolved phenotypes (sleep), and the genetic underpinnings of this divergence, I used the *early*, *control,* and *late* populations, and other inbred *Drosophila melanogaster* flies (explained in detail in chapter 5) in a series of experiments, which will be discussed in the following paragraphs.

In the second chapter of my thesis, I present my findings on the role of a non-clock phenomenon, masking, in the Ψ_{ent} of *early* flies. I hypothesized that our selection protocol has inadvertently resulted in selection for masking in the *early* chronotype due to the temporal placement of our selection window (which includes the lights-ON transition). Based on theoretical predictions and previous studies on our populations, I designed experiments to discriminate between enhanced masking to light, and high temperature versus circadian clock mediated changes in determining enhanced emergence in the morning window in our early chronotypes. I concluded that *early* flies have evolved significantly more positive masking to lights-ON and negative masking to high temperature compared to the *control* and *late* flies.

Results of experiments carried out to understand the role of masking to light in locomotor activity rhythm and the effect of light intensity and duration on the sleep of different chronotypes are reported in chapter three. I first established that higher light intensity photoperiods enhance masking in locomotor activity rhythm of *early* populations significantly more than *late* populations. I also found evidence that higher light intensity during the daytime significantly affects night sleep duration adversely in divergent chronotype flies than the *control* populations. I also show that shorter day lengths affect nighttime sleep amount and quality in divergent chronotypes significantly more adversely than in the *control* populations.

Subsequently, to understand the genetic architecture of the *early* and *late* populations, compared to the *control* populations, I performed a pool-sequencing of all 12 populations (*early*, *control*, and *late*, and their four replicates each). I performed various population genomic analyses and identified genomic regions undergoing positive selection in *early* and *late* populations compared to the *control* populations. I also identified various SNPs significantly enriched in either *early* or *late* populations compared to the *control* populations. I further predicted different pathways plausibly under differential selection in *early* and *late* populations. I concluded that *early* and *late* chronotypes are not just two sides of the same coin, i.e., there is a minimal overlap among genomic regions where early and late flies have diverged. These results are reported in chapter four.

In the fifth chapter of my thesis, I conducted a genome-wide screening of chromosomal deficiency lines of *Drosophila melanogaster* (the core DrosDel kit) for differential phasing of the eclosion rhythm. I found only 10 lines spanning over four major chromosomal arms to be non-control like (early-like and late-like). These 10 lines cover a total of 595 genes and are enriched in genes involved in nucleosome and chromatin assembly, chromatin and gene silencing, ecdysone induced genes, etc. I proposed that these genes may act as fine regulators of phases of the eclosion rhythm in *Drosophila melanogaster*, but not have a large role to play in the generation of the eclosion rhythm *per se*. Further narrower deletions will need to be screened to get a better idea of genes directly involved in the regulation of phases.

In my sixth chapter, I discuss the implications of results from all the previous chapters and make some general remarks and propose a hypothesis regarding how masking along with circadian clock-driven entrainment may adjust Ψ_{ent} in different chronotypes. I discuss the repercussions of the genetic associations I propose in chapters 4 and 5. I also discuss future experiments that could be done to further understand the inter-relationship between timing of behavior, sleep, and underlying genetic components.

At the end of my thesis, I have multiple appendices. The first appendix is the description of two software I created for high throughput and easy analysis and visualization of circadian rhythm and sleep in organisms. The subsequent 4 appendices are supplementary information, tables, and figures for chapters 2, 3, 4, and 5.

Chapter 2

Characterization of masking responses to light and temperature in eclosion rhythm

This section has been published in Journal of Biological Rhythms as "Evidence for Co-Evolution of Masking With Circadian Phase in *Drosophila melanogaster*, Arijit Ghosh, Pragya Sharma, Shephali Dansana, Vasu Sheeba, 2021, JBR".

2.1 Introduction:

Circadian clocks adaptively schedule behavior and physiology to occur at a specific time of the day. Such scheduling is believed to be critical for maintaining our general health and well-being (Horn et al., 2019; Roenneberg, 2012; Vaze et al., 2014). Heritable variations in the timing (or phasing) of rhythmic events with respect to daily time cues result in what is referred to as chronotype variation (Infante-Rivard et al., 1989; Roenneberg et al., 2007). A clear example is that of variation in mid-sleep timing among humans on free days. While most humans will fall in the category of 'normal' or 'neither' chronotype, some individuals tend to fall asleep relatively early in the evening and wake up early in the mornings, hence referred to as '*early'* chronotypes or '*larks*' while there are those among us who prefer very late sleep timings and associated late wake timings, also referred to as '*late*' chronotypes or '*owls*' (Randler et al., 2017). Chronotypes are primarily controlled by circadian clocks. Studies, including those on humans, have suggested that variation in entrained phases arise due to differences in underlying clock properties such as length of the intrinsic period of circadian clocks, phase/ velocity response curves (PRCs/ VRCs), amplitude of the circadian clock, inter-oscillator coupling, and amplitude of the zeitgeber (Aschoff and Pohl, 1978; Bordyugov et al., 2015; Granada et al., 2013; Johnson et al., 2003; Pittendrigh and Daan, 1976; Roenneberg, 2012; Swade, 1969).

While light can bring about a change in phase of an entrained rhythm by influencing the phase of circadian clock (Saunders et al., 1994; Schlichting and Helfrich-Förster, 2015), one cannot rule out aspects of direct effects of light. Such exogenous environmental influences on endogenously generated circadian rhythms which obscure aspects of circadian clock expression is referred to as masking (Aschoff, 1960; Fry, 1947; Mrosovsky, 1999). Traditionally, non-involvement of the circadian clock has been considered an essential criterion defining masking, however, Terry Page (Page, 1989) and Nicholas Mrosovsky (Mrosovsky, 1999) make a strong case for the importance of masking as a complement to circadian clock regulation of daily rhythms. While there are studies which suggest that masking responses are present in animals without a functional clock (clock mutants or surgical ablation of the central clock) (Redlin and Mrosovsky, 1999a; Wheeler et al., 1993), Aschoff had argued that there might be time-of-day dependence in certain masking responses (eliciting activity in blind male hamsters in the presence of mates) (Aschoff and Honma, 1999; Aschoff and von Goetz, 1988). Positive masking refers to the masking response which elicits the beginning of a behavior, while negative masking refers to the inhibition or ceasing of the behavior. Traditionally studied in mammals with respect to locomotor activity rhythms, masking has also received some attention in insects, in both locomotor activity and eclosion rhythms (Hamblen-Coyle et al., 1992; Kempinger et al., 2009; Lu et al., 2008; Rieger et al., 2003; Sheppard et al., 2015; Thakurdas et al., 2009; Wheeler et al., 1993).

The act of emergence of a pharate adult fly from its pupal case or eclosion is developmentally gated, and a population level rhythm. It was also among the earliest circadian rhythms to be studied systematically (Chandrashekaran, 1967; Chandrashekaran and Loher, 1969; Engelmann, 1969; Harker, 1965; Pavlidis, 1967; Pittendrigh, 1954, 1967; Pittendrigh et al., 1958; Skopik and Pittendrigh, 1967; Zimmerman et al., 1968). The anatomical and physiological processes underlying eclosion have also been extensively investigated (Johnson and Milner, 1987; Krüger et al., 2015; Peabody and White, 2013; Selcho et al., 2017; Thummel, 2001). It is deemed to be amongst one of the most critical events in the lifetime of a holometabolous insect (Mcmahon and Hayward, 2016; Zitnan and Adams, 2005). The *Drosophila* eclosion rhythm has been shown to be regulated by the circadian pacemakers previously implicated in activity rhythms in addition to prothoracic gland clocks (Morioka et al., 2012; Myers et al., 2003; Pittendrigh and Bruce, 1959;

Selcho et al., 2017; Zimmerman et al., 1968). Additionally, in contrast to several behavioral rhythms, the act of eclosion is free from any motivational state, other behaviors, or interactions among individuals, but susceptible to disruption when mutations are introduced in core clock genes like *period* (*per*) or *timeless* (*tim*) (Qiu and Hardin, 1996; Ruf et al., 2019; Sehgal et al., 1994) both under constant and cycling conditions. Thus, relative to other rhythms, it appears to be more reliable indicator of perturbations in the core clock. Eclosion rhythms, their entrainment to light/ dark cycles, temperature cycles, synergistic light and temperature cycles as well as molecular mechanisms of their entrainment have been studied in great details in Drosophilid species (Emery et al., 1997; Kumar et al., 2006, 2007; Morioka et al., 2012; Myers et al., 2003; Nikhil et al., 2015, 2014, 2016b; Pittendrigh and Minis, 1972; Pittendrigh, 1966; Pittendrigh and Bruce, 1959; Prabhakaran et al., 2013; Qiu and Hardin, 1996; Vaze and Sharma, 2013). The rhythm in eclosion is modulated by the lights-ON signal (Chandrashekaran and Loher, 1969; Engelmann, 1969; Pittendrigh, 1967; Thakurdas et al., 2009), even though wing-expansion, the last behavioral event of the *Drosophila* adult eclosion sequence (Fraekkel, 1935), is not affected by the same (McNabb and Truman, 2008), suggesting that the lights-ON signal may have a role specific to the act of eclosion itself. It was suggested previously that the lights-ON signal may have two distinct effects on the timing of eclosion of flies - a) stimulation of eclosion hormone release, and b) reduction in the latency of eclosion relative to eclosion hormone release (Baker et al., 1999; McNabb and Truman, 2008).

With reference to *Drosophila melanogaster*, Hamblen-Coyle and colleagues first reported a lights-ON peak in locomotor activity rhythm which they designated as a ‘*startle*’ effect. It was seen that even though flies carrying core circadian clock mutations adopt distinct phases in evening peak timings, their morning peak (lights-ON peak) phases were not very different (Hamblen-Coyle et al., 1992). Later it was shown that under LD12:12, even without a functional clock, animals exhibited this lights-ON peak, but it was absent when shifted to constant darkness (DD) (Wheeler et al., 1993). Rieger and colleagues showed that under laboratory LD12:12 conditions, the lights-ON peak and the morning peak are indistinguishable as they overlap with each other (Rieger et al., 2003). These two peaks can be separated from each other under different photoperiods, where the lights-ON peak is still phase-locked to the dark-to-light transition, but the morning peak is advanced or delayed under short or long photoperiod respectively (Rieger et al., 2003). Artificial moonlight can make fruit flies nocturnally active, and this nocturnal light is known to induce strong locomotor activity in flies via masking (Kempinger et al., 2009). Recent detailed genetic dissections of masking in *Drosophila* have revealed complex pathways mediating light-induced masking of locomotor activity (Rieger et al., 2003). Furthermore, Lu and colleagues demonstrated a circadian rhythm in light-induced locomotor activity against a background of DD and showed that the circadian clock genes *timeless* and *clock* are involved in regulation of this masking response (Lu et al., 2008; Sheppard et al., 2015). Taken together, the above studies on effect of light on locomotor activity rhythms and eclosion suggest that masking can affect the timing of a circadian rhythm in *Drosophila*.

I reasoned that one set of our long-term laboratory-selected populations of *D. melanogaster* could potentially have evolved a masking response due to the nature of the selection regime that had been imposed on them for ~320 generations ($early_{1\text{-}4}$ – described below). In this chapter, I investigated the immediate effects of light and temperature on the timing of their eclosion rhythm. These populations are part of an on-going long-term experimental evolution study at the Chronobiology laboratory, JNCASR. The primary goals of creating these populations were – a) to demonstrate the adaptive significance of phasing of circadian rhythm (here, eclosion rhythm) b) to then study the associated clock properties mediating phase divergence. Indeed, our populations selected for morning emergence (*early*$_{1-4}$) exhibit significantly shorter (22.51 h, ±95% CI = 0.106) mean free-running period (FRP) compared to those selected for evening emergence (*late*$_{1-4}$ mean FRP 23.86 h, ±95% CI = 0.106) and the *control*$_{1-4}$ populations (mean FRP 22.94 h, ±95% CI = 0.106), which did not undergo any selection for timing of emergence (Kumar et al., 2007; Nikhil et al., 2016b). Over the years several studies from our lab have shown that entrainment of these divergent chronotypes to light-dark (LD) cycles cannot be fully explained by exclusively invoking either the parametric or non-parametric models of entrainment (Abhilash and Sharma, 2020; Kumar et al., 2007; Vaze et al., 2012a). It was also hypothesized that differences in inter-oscillator (A / master / central oscillator and B / slave / peripheral oscillator) coupling might explain the chronotype divergence in eclosion rhythm (Abhilash et al., 2019; Nikhil et al., 2016b).

I reasoned that our selection lines may provide material to examine some aspects of masking and circadian control of a rhythmic phenomenon because the design of our selection regime is such that flies of the *early* populations have over generations been forced to emerge in a window around lights-ON (3 hours prior to lights-ON till 1-hour post lights-ON). It is possible that this protocol has inadvertently resulted in selection for the phenomenon of masking. Therefore, among other clock related factors that have been uncovered to have changed in these populations previously (Abhilash and Sharma, 2020; Nikhil et al., 2015, 2016c; Vaze et al., 2012b), I hypothesized that the phase divergence among *early*, *control* and *late* chronotypes are partly due to differences in masking – specifically, that *early* populations exhibit a high degree of masking. I designed several experiments to examine whether *early* flies exhibit enhanced positive masking in eclosion rhythm compared to *control*, and *late* flies as shown in *Fig. 2.1*. In addition to this, I also observed from a previous study in the lab that early flies showed predominant emergence in the cool part of the day (subjective night part) when assayed under a temperature cycle of 12:12 hours (TC12:12) and avoid emerging in the warm part (subjective day part) (Nikhil et al., 2014). In that study the flies were kept under LD12:12 along with TC12:12 for the first five days of development so the effect of temperature cycles alone could not be inferred. In nature, Drosophila generally emerges during early morning, and one possible explanation can be presence of higher humidity and cool air which favors the wing expansion in newly emerged adults (Palaksha and Shakunthala, 2014; Pittendrigh, 1954; Tanaka and Watari, 2009). Different insects have their peak of emergence rhythm during the early morning, primarily because the process of wing expansion is severely affected by higher temperatures later in the day (Tanaka and Watari, 2009; Watari, 2002). I hypothesized that our *early* flies, as they have been selected for emergence in the early morning, may have also evolved temperature preference for emergence and will emerge only during the cool temperature part of the day. I designed experiments to verify this hypothesis as depicted in *Fig. 2.7B.*

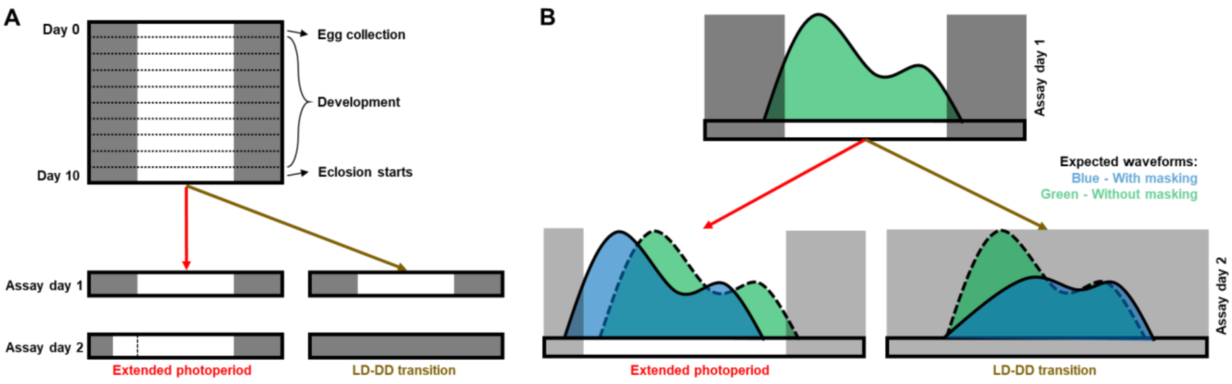

A
Day 0
Day 10
Egg collection
Development
Eclosion starts
Assay day 1
Assay day 2
Extended photoperiod
LD-DD transition
B
Assay day 1
Expected waveforms:
Blue - With masking
Green - Without masking
Extended photoperiod
LD-DD transition
Assay day 2

Figure 2.1: Schematics of experimental design and hypotheses. A. After egg collection, vials were kept under LD12:12 till 9th day. Emergence assay was started on the 10th day (*Assay day 1*), and flies were counted every two/half an hour interval depending on experiments and time points. On the 11th day (*Assay day 2*), vials were placed under a 3-hour phase-advanced light schedule (Extended photoperiod) or complete darkness (LD-DD transition). **B.** The schematic shows expected waveforms in case of complete circadian control (in green) or masking (blue). Dark rectangular shades depict duration of darkness.

Here I demonstrate that – a) *early* chronotype flies have indeed evolved significantly more positive masking compared to *control* and *late* flies, b) under full photoperiod, apparent entrained phases of *early* flies are largely contributed by masking, c) under skeleton photoperiod, *early* flies do show phase lability, and retain advanced phase of entrainment compared to *control* and *late* flies to different *T*-cycles, suggesting that our selection indeed has selected for greater masking alongside selection for advanced phase of entrainment, and d) *early* flies have evolved significantly more negative masking to warm temperature compared to *control* and *late* flies.

2.2 Materials and methods:

2.2.1 Selection protocol and fly maintenance:

Four sets of genetically independent *Drosophila melanogaster* populations were used to artificially select for morning eclosion (*early* populations) and evening eclosion (*late* populations) timing. Henceforth, denoted as $early_{(1\text{-}4)}$, $control_{(1\text{-}4)}$ (no selection imposed) and $late_{(1\text{-}4)}$. At each generation, ~300 eggs are collected and placed in glass vials which are maintained in a light-proof, temperature-controlled cubicle in a light-dark cycle of 12/12 hours (LD12:12), 25±0.5 °C, and 65±5% RH. Flies emerging from ZT21-ZT01 (ZT0 is Zeitgeber Time 0, when light comes ON) are collected to form the breeding pool for the next generation of *early* flies, while flies emerging from ZT9-ZT13 form the breeding pool for the next generation of *late* flies. Flies emerging throughout the day are collected to make up the breeding pool for the next generation of the *control* flies. This collection goes on for 3-4 days and total ~1200 adult flies with ~1:1 sex ratio of each of the 12 populations are maintained in Plexiglas™ cages (25 cm×20 cm×15 cm) with petri plates with banana-jaggery culture media. The flies are maintained on a 21-days discrete generation cycle, and all experiments were done with progeny of flies that experienced one generation of common rearing (standardized) to avoid confounding factors due to maternal effects (Bonduriansky and Day, 2009). For more details of the selection regime, see Kumar *et al.* (Kumar et al., 2007), and Abhilash *et al.* (Abhilash et al., 2019). All experiments were done between generations 320 – 330.

2.2.2 Behavioral experiments:

Before each experiment, ~300 eggs were collected and placed in 10 vials each for all 12 standardized populations. After egg collection, the vials were maintained in different regimes specific to each experiment. Emerged flies were counted every two hours or half an hour after the assay started, depending on the experimental regime. Briefly, all rhythm assays in *Fig. 2.2 and 2.3* were carried out with half hour resolution 12 hours around lights-ON and in assays depicted in *Fig. 2.4 and 2.5* (full and skeleton *T*-cycle experiments), 2 hours resolution was used due to logistic constraints. To account for differences in development time, if any, fly counts from the first emergence cycle were excluded from analysis. All 12 populations were assayed in parallel. In all experiments, temperature (25±.5 °C), and light intensity (~70 lux, from a white LED source) were kept constant. Rhythm assays performed and presented in *Fig. 2.7* were done under a TC12:12 cycle of 19°-28°C and constant darkness (DD). All assays in *Fig. 2.7A* were performed for at least 4 days with 2 hours resolution and assays in *Fig. 2.7B* were performed for three days with half hour resolution 12 hours around the warm temperature ON signal. Other details of the light regime are mentioned in each experiment separately, and light ON-OFF times (step wise) were programmed with a TM619 timer (Frontier Timer, Pune, India) in the incubators (DR-36VL, Percival Scientific, Perry, USA) in which the experiments were performed. For temperature cycle experiments, cool and warm temperature timings were programmed in the incubators (DR-36VL, Percival Scientific, Perry, USA).

2.2.3 Analysis of data and statistics:

Average profiles were constructed first by averaging over multiple cycles for a vial, and then by averaging over vials for each population. All statistical comparisons were made using either a 2-way or 3-way randomized block design ANOVA with selection regime and *T*-cycle (as applicable) as fixed factors and blocks (replicates) as the random factor. Results were deemed significant for main effect or interaction as applicable at $\alpha < 0.05$. Post-hoc comparisons were carried out by a Tukey's Honest Significant Difference (HSD) test. In the main text, I report significant main effects or interaction effects while supplementary tables provide the full details of the ANOVA. All mean values and the 95% CI used for post-hoc comparisons are also reported in supplementary tables. All statistics were performed in Statistica 7 (StatSoft, Tulsa, USA). Standard errors of means have been plotted as error bars in average profiles for ease of visualization. 95% confidence intervals from the Tukey's HSD are plotted in all other graphs and used for quantitative comparisons for visual hypothesis testing. Basic data processing and calculations were done with Microsoft Excel 365, and all graphs were plotted with Graphpad Prism 8. Criteria for "apparent entrainment" was $T_{observed}$ must equal to $T_{environment}$ ($T_{observed}$ is the observed period of the eclosion rhythm and $T_{environment}$ is the duration of the light/ dark cycle). $T_{observed}$ was calculated with JTK-cycle (Hughes et al., 2010) employed in MetaCycle2d (Wu et al., 2016) with percentage eclosion data for each vial. Vials showing a "JTK_pvalue" less than 0.05 were considered to be rhythmic, and among the rhythmic vials, the ones with "JTK_period" within a range of ±1 hour of $T_{environment}$ were considered to be "apparently entrained". All calculated phases are essentially phase relationships with the lights-ON signal. Peak phase is denoted by Ψ_{Peak} and measured in hours, and phase of Centre of Mass is denoted by Ψ_{CoM} and measured in degrees. Ψ_{Peak} was calculated as the time where maximum number of flies emerged in each vial and Ψ_{CoM} was calculated as a measurement of mean phase of emergence in polar coordinates for each vial (corrected for different lengths of *T*-cycles). Consolidation of emergence/ normalized amplitude (R) was also calculated in a polar coordinate system for eclosion data averaged over cycles, details of computation and usefulness of which can be found in a recent publication from our lab (Abhilash et al., 2019). For temperature cycle experiments depicted in *Fig. 2.7C & D*, quantifications were performed as mentioned in the figure legend.

2.3 Results:

2.3.1 Lights-ON elicits an immediate response in the *early* chronotypes:

Previously, all eclosion rhythm assays on our populations were carried out with a maximum resolution of 2 hours and thus far no difference was detected in the peak phase (Ψ_{Peak}) of *early* and *control* flies, both of which were found to emerge maximally at (ZT2) (Kumar et al., 2007; Nikhil et al., 2016b). To examine whether there are subtle changes in the onset or peak of emergence among stocks I increased the resolution to 0.5 hours for the first half of the day (*Assay day 1* and *2, Fig. 2.2A & 2.3A*). In fact, I now find that peak (Ψ_{Peak}) emergence for both *early* (~30%) and *control* (~15%) flies occur at ZT0.5, i.e., immediately after lights-ON. Onset of emergence is similar as that obtained by 2-hours resolution assays done previously (Kumar et al., 2007; Nikhil et al., 2016b) across stocks.

When lights-ON was advanced by 3 hours on the following day (*Assay day 2, Fig. 2.2A*), *early* flies exhibited high emergence in the first 0.5 hour, whereas very few flies emerged in the early part of the light phase in both the *control* and *late* populations – they mostly emerged at similar time as the previous cycle (*Assay day 1, Fig. 2.2A*). Peak emergence of *early* flies occurred at ZT2 on *Assay day 2*. On both assay days, majority of *early* flies emerged immediately after lights-ON (*Fig. 2.2A*). The phase of peak emergence for *control* flies remained similar to *Assay day 1*, even after the light phase was advanced on *Assay day 2* (*Fig. 2.2A*). Emergence waveform of *late* flies did not change from *Assay day 1* to *Assay day 2* (*Fig. 2.2A*).

Based on the emergence profile of the *early* populations I quantified the difference in emergence during two specific time windows on the two experimental days. The first window (*solid rectangle, Fig. 2.2B*) depicts the time window of maximum emergence for *early* flies on *Assay day 1* (LD) and the second window (*dashed rectangle, Fig. 2.2B*) depicts the time window for maximum emergence (immediately after lights-ON) for *early* flies on *Assay day 2* (Extended photoperiod). If *early* flies show more masking than circadian control of their emergence, one expects a large difference in levels of eclosion between the two days because they are strongly modifying their waveforms in response to lights-ON. I observe that the difference in percentage emergence during the first window (between *Assay days 1* and *2*) is significantly higher for *early* flies than for *control* and *late* flies, the latter two showing almost no change across days (*solid rectangle, Fig. 2.2B; Appendix table A2.1*). Similarly, in the second window, *early* flies show significantly higher change across *Assay days 1* and *2* compared to *control* and *late* populations (*dashed rectangle, Fig. 2.2B; Appendix table A2.2*). This suggests that on *Assay day 2*, the circadian response of *early* flies is overridden by the immediate lights-ON response, thus exhibiting higher positive masking response compared to *control* and *late* flies. When considering two larger windows each of 2.5 hours (around lights-ON of pre-shift day) similar differences prevail (*Fig. 2.6C*), showing that high emergence immediately after lights-ON is specific to *early* flies.

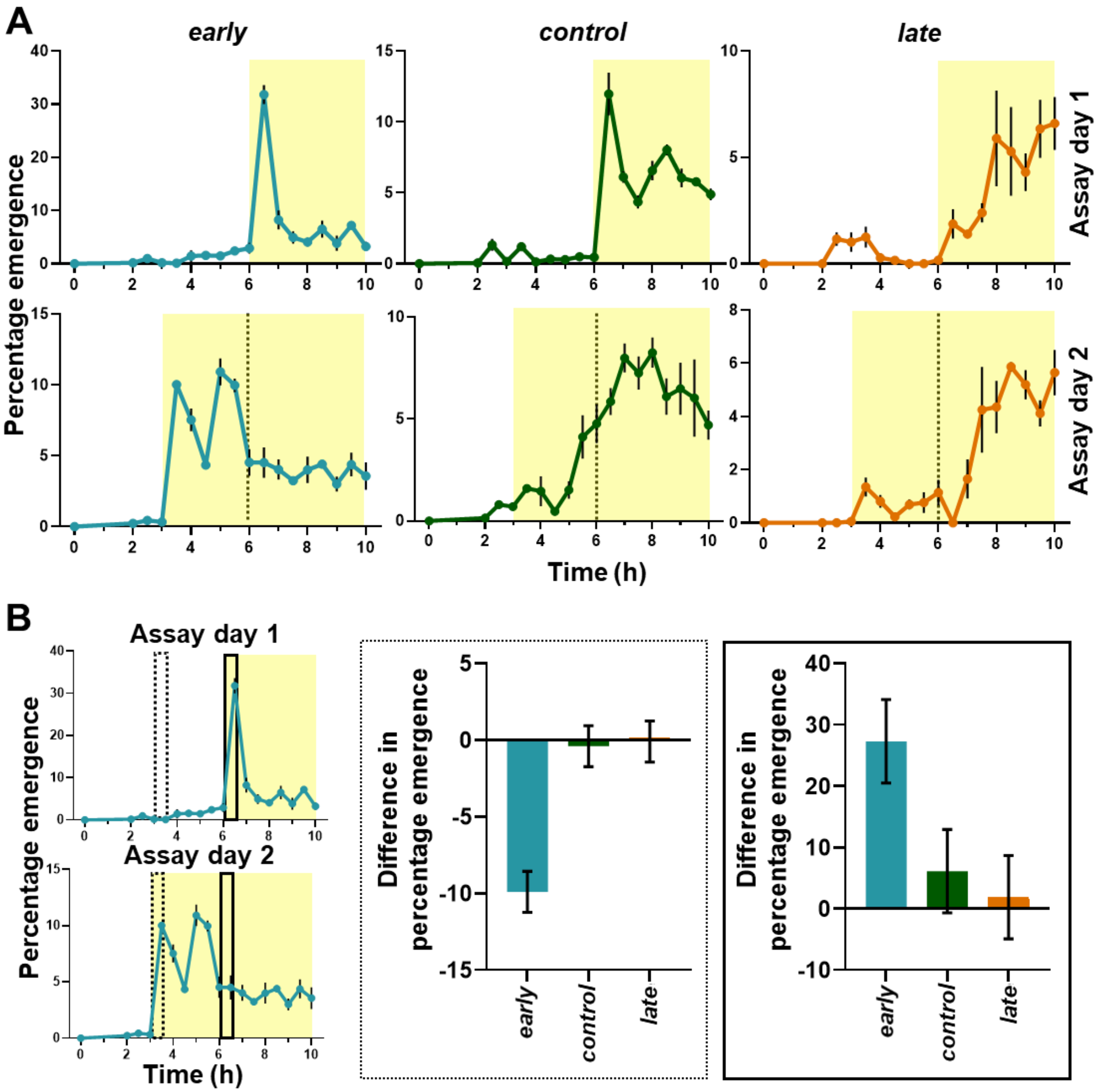


Figure 2.2: Emergence profile during first 10 hours on day 10 (post egg collection) in pre-shift (*Assay day 1*) and post-shift (*Assay day 2*) days and difference in percentage emergence between pre-shift (*Assay day 1*) and post-shift (*Assay day 2*) days in two 0.5-hour time windows. A. *early* flies (blue line) advance their emergence waveform to emerge immediately after lights-ON on *Assay day 2*. The *control* (green line) and *late* (orange line) flies conserve their waveform on both days by not advancing their emergence in response to sudden advancement of the lights-ON stimulus. The yellow shading indicates the photophase in the LD cycle each day. Error bars are ± SEM. Grey dotted lines in the bottom panel indicates lights-ON time on previous day. **B.** The left panels show time windows used for analysis (dashed and solid rectangles respectively). In same time windows *early* flies show significantly higher (dashed rectangle region) or lower (solid rectangle region) emergence than that of *control* and *late* flies on *Assay day 2* than *Assay day 1*, showing this high emergence immediately after lights-on is specific to *early* flies.

2.3.2 *early* flies attenuate and delay their emergence under DD:

Since the above experiments suggested that masking plays a prominent role in the timing of eclosion of *early* flies, I then attempted to parse the relative contribution of circadian clock control on the emergence profile of *early* flies. Therefore, one set of cultures were shifted to constant darkness (DD – *Assay day 2*, *Fig. 2.1A right panels; LD-DD transition*). On *Assay day 2*, very few flies of the *early* populations emerged in the 2-hour duration corresponding to immediately after lights-ON of *Assay day 1*, resulting in a large difference in emergence between two consecutive cycles (*Fig. 2.3A*). This difference in emergence between cycles was lesser in *control* and almost non-existent in *late* flies (*Fig. 2.3B right panel*). If the circadian clock has a strong control over the eclosion rhythm, the *early* flies should show high emergence in the early part of the subjective photophase in DD, which was not observed. The sharp morning peak, an identifiable marker for our *early* flies was absent in DD (*bottom left panel, Fig. 2.3A*).

I quantified the difference in percentage emergence from ZT0-0.5 of first day (LD, *Assay day 1*) and the same phase in second day (DD, *Assay day 2*), and find that *early* flies show significantly larger difference than *control* and *late* flies, suggesting attenuated emergence immediately after starting of subjective photophase in DD (*Fig. 2.3B; Appendix table A2.3*). While a reduction in amplitude is expected under DD, it is significantly greater for *early* flies than *control* flies. This suggests the high emergence of *early* flies immediately after lights-ON is largely a masking response and less of a clock-controlled phenomenon.

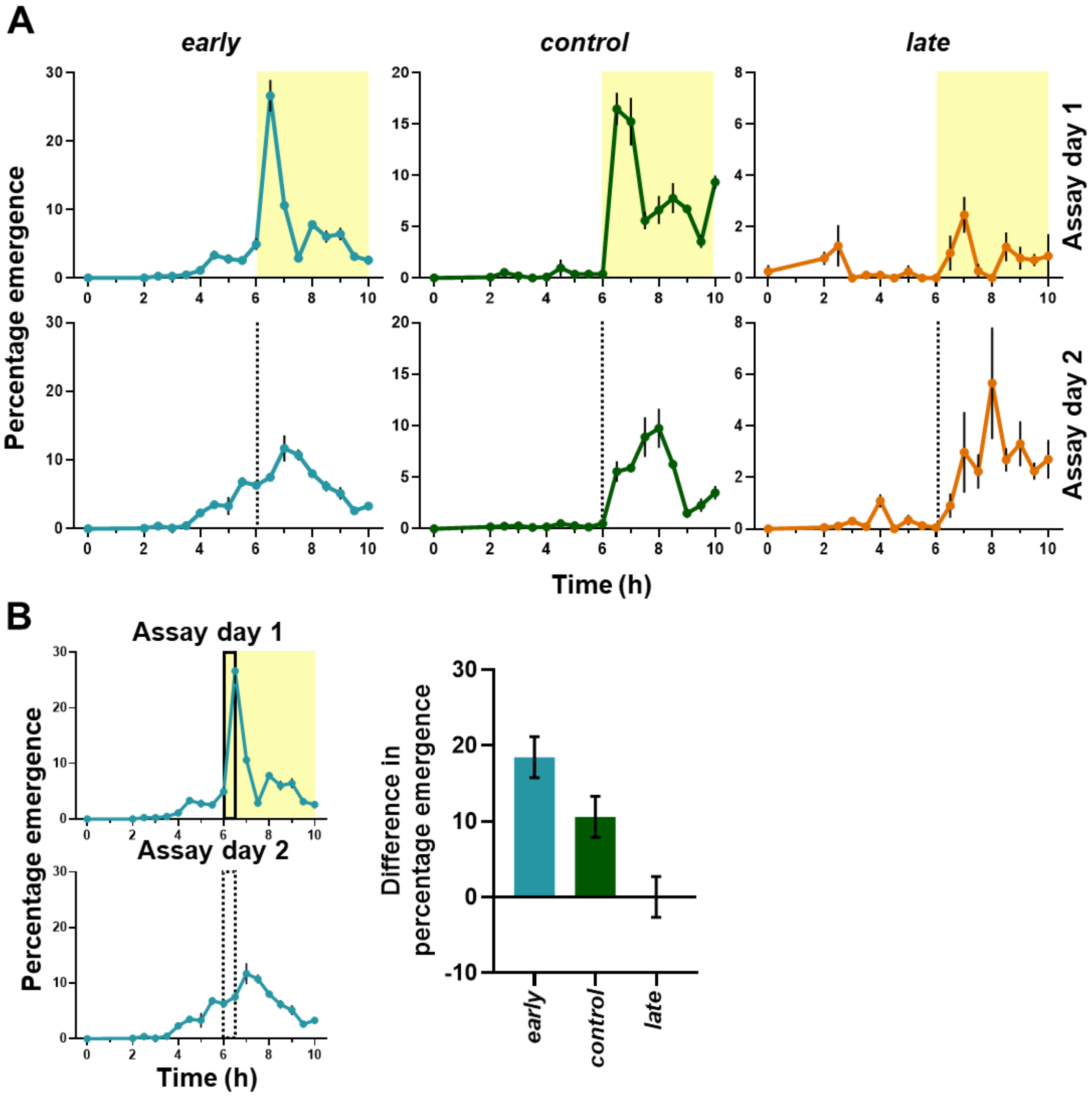


Figure 2.3: Emergence profile during first 10 hours on day 10 (post egg collection) (after 10 days) in LD (*Assay day 1*) and DD (*Assay day 2*) days and quantification of difference of percentage emergence between pre-shift and post-shift days from ZT0-0.5 (lights-ON phase followed from LD cycle). A. *early* flies attenuate their emergence to a great extent in absence of light stimulus in the earliest part of the subjective photophase, unlike their waveform in LD12:12. However, *control* and *late flies*, emerge with waveforms like under LD12:12, albeit with lower amplitude. Error bars are ± SEM. Grey dotted lines in the bottom panel indicates lights-ON time on previous day. **B.** *early* flies show significantly higher difference than *control* and *late* flies. This suggests the high emergence of *early* flies immediately after lights-on was a masking response rather than a clock mediated response.

2.3.3 *early* flies consistently emerge close to lights-ON under both short and long *T*-cycles:

T-cycles shorter than the FRP of the organism are expected to delay phases of circadian rhythms (phase relationships with Zeitgeber) while *T*-cycles longer than the FRP of the organism advance phases of the circadian rhythms (Aschoff, 1965; Wheeler et al., 1993; Yadav et al., 2015). Therefore, I subjected our populations to a series of short (T20; LD10:10 and T22; LD11:11) and long (T28; LD14:14 and T26; LD13:13) *T*-cycles (*Fig. 2.4A-C*). Under the two extreme *T*-cycles T20 and T28, although *early* flies showed higher percentages of “apparent” entrainment, very small fraction of *control* and *late* flies entrained, thus rendering comparison among stocks inappropriate. Under T20, significantly few vials of *control* and late flies show entrained rhythms compared to *early* flies, and under T28, significantly low number of vials of *late* flies entrained compared to *control* and *early* flies (*left panel Fig. 2.4C; Appendix tables A2.17 and A2.18*), hence the results of these extreme *T*-cycles T20 and T28 were not considered for further analyses and quantification of phases. Under T20 and T28 cycles, *early* flies showed high emergence immediately after lights-ON compared to *control* and *late* flies (*middle and right panel Fig. 2.4C*) suggesting that, *early* flies mask successfully and show an apparent entrained behavior. I expected that eclosion rhythm of *control* and *late* flies, owing to strong control by circadian clocks will shift their phases in the predicted directions. I hypothesized that eclosion rhythm of *early* flies, if strongly governed by masking and less by the circadian clock, or controlled by parametric changes in the circadian clock, will not modify phases of the eclosion rhythm in the predicted directions – delay under T>24 hours and advance under T<24 hours.

Under T22, *early* flies maintain similar waveform as under T24 (*left panel top row, Fig. 2.4A*). The *control* flies showed clear phase delay (*middle panel top row, Fig. 2.4A*), whereas *late* flies showed even greater phase delay in response to a T22 cycle (*right panel top row, Fig. 2.4A*) compared to their waveforms in T24 cycles.

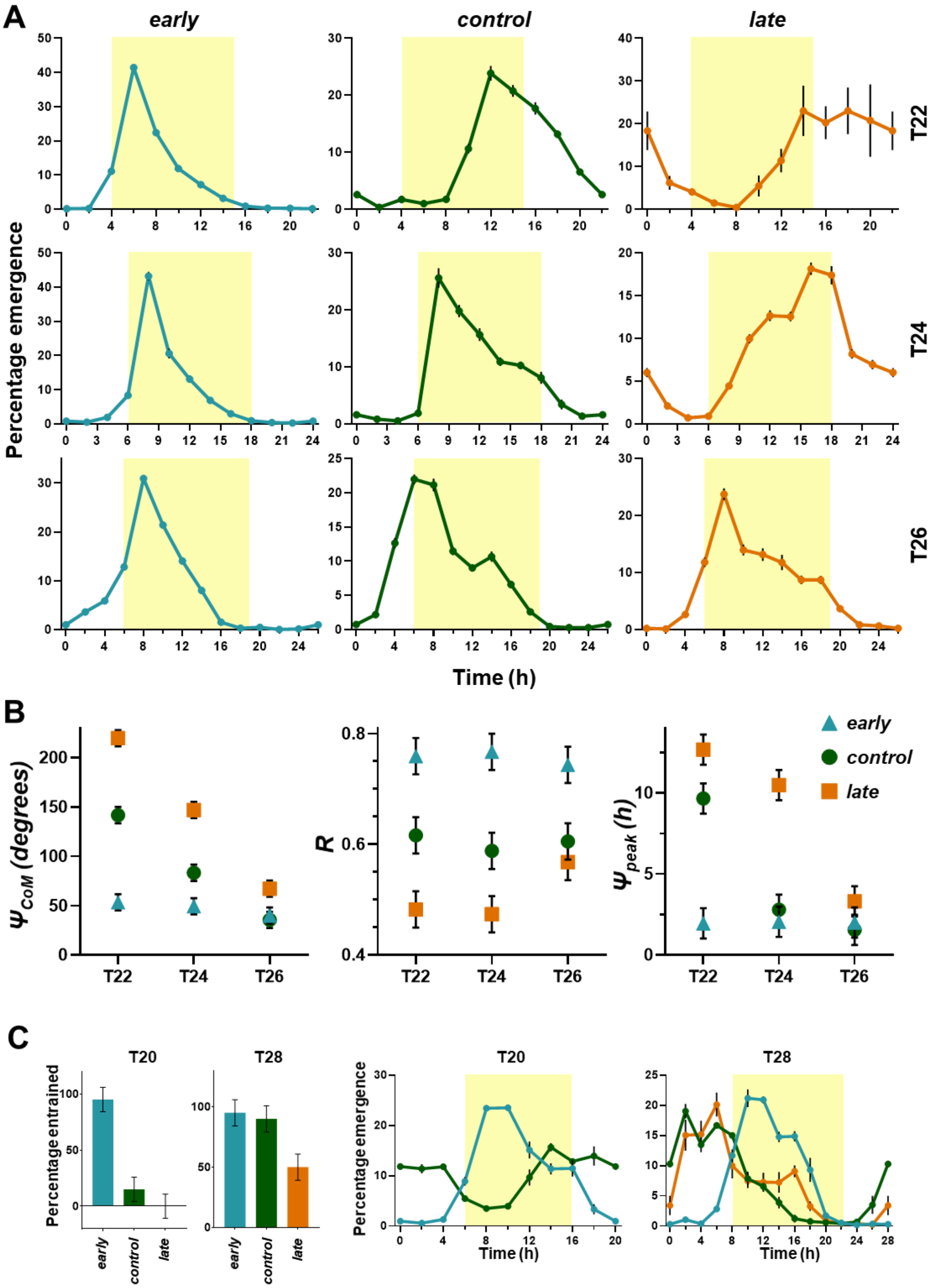


Figure 2.4: Emergence profile of *early*, *control*, and *late* flies under T20, T22, T24, T26 and T28 cycles and eclosion rhythm parameters under *T*-cycles. A. *early* flies do not change their emergence profile depending on *T*-cycles, whereas *control* and *late* flies shift their emergence profile in expected directions (phase advance in longer

than 24-hour cycle and phase delay in shorter than 24-hour cycle). Yellow shading depicts the photophases of respective *T*-cycles. Error bars are ± SEM. **B.** Left panel: Centre of Mass (Ψ_{CoM} – in degrees), middle panel: Consolidation of emergence (R), right panel: Peak in ZT (Ψ_{Peak}, in hours). Error bars are ±95% CI. *early* flies do not change their Ψ_{CoM}, Ψ_{Peak}, and R under any *T*-cycle, showing they are phase-locked to the lights-ON stimulus. Color codes: *early* – blue, *control* – green, *late* – orange. **C.** Percentage entrainment and emergence profiles of *early*, *control* and *late* flies under T20 and T28. Error bars are ±95% CI for percentage entrainment and ± SEM for profiles.

Under T26, *early* flies showed similar emergence waveform as observed under T24 (*left panel third row, Fig. 2.4A*). The *control* flies showed phase advance (*mid panel third row, Fig. 2.4A*) and *late* flies showed even larger phase advance in response to T26 cycle (*right panel third row, Fig. 2.4A*) compared to their waveforms under T24. There is some degree of anticipation in *early* flies to lights-ON under all three *T*-cycles as there is a gradual rise in emergence prior to lights-ON (*Fig. 2.4A*). This issue is addressed in *Fig. 2.6D*. Briefly, the underlying clock of *early* flies delay or advance under T22 and T26 respectively, still the high emergence immediately after lights-ON is conserved under all three *T*-cycles due to masking.

I quantified three parameters of the emergence waveform under all three *T*-cycles: a) Centre of Mass (Ψ_{CoM}), which is an estimate of mean phase angle of emergence in a circular scale, b) Peak in ZT (Ψ_{Peak}), and c) R, which is a measure of normalized amplitude of the eclosion rhythm and comprehensively describes consolidation of emergence. All these parameters have been previously used to describe the eclosion rhythm waveform of *Drosophila melanogaster* (Abhilash et al., 2019).

As seen in *Fig. 2.4B*, Ψ_{CoM} for *early* flies did not change across the *T*-cycles, whereas Ψ_{CoM} for *control* and *late* flies shifted in expected directions for T22 (delayed) and T26 (advanced) compared to T24 cycle (*left panel, Fig. 2.4B; Appendix table A2.4*). Ψ_{Peak} for *early* flies remained similar across all three *T*-cycles (*right panel, Fig. 2.4B; Appendix table A2.5*). *control* and *late* flies showed expected trends in change of direction of peak phase shift in T22 and T26 cycles (*right panel, Fig. 2.4B; Appendix table A2.5*). In case of *control* flies, although there was a trend of advancing peak phase in T26 cycle, compared to T24, this difference was not significant, whereas the delay in peak phase in T22 was much larger when compared to T24 cycle (*right panel, Fig. 2.4B; Appendix table A2.5*). In *late* flies, the delay in peak phase in T26 was significantly larger when compared to T24 cycle, but the phase advance in T26 was not significant compared to T24 (*right panel, Fig. 2.4B; Appendix table A2.5*). Next, I quantified the consolidation of emergence/ normalized amplitude of the eclosion rhythm (R) of all populations under all *T*-cycles. A large R value is characteristic of the *early* population as evident by their narrow gate-width of emergence (Nikhil et al., 2016b). If the shorter and longer *T*-cycles had indeed shifted the circadian clock of *early* flies and the masking component is only responsible for the high emergence in the earliest part of the photophase, then value of R is expected to change among different *T*-cycles. Also, decrease in R may indicate one clock-controlled peak and one masking peak, as previously reported in case of locomotor activity rhythm in Drosophilid species (Prabhakaran and Sheeba, 2012; Rieger et al., 2003). I observed that R is significantly higher in *early* populations in all three *T*-cycles which suggests that they maintain constant high consolidation of emergence as observed previously under T24 (*middle panel, Fig. 2.4B; Appendix table A2.6*). *control* flies do not change R among *T*-cycles, but *late* flies show significantly higher R under T26 compared to T24 and T22, primarily because of high amplitude of their emergence under T26.

Taken together, these results indicate that peak phase of *early* flies do not change when they entrain to short or long *T*-cycles. The fact that emergence occurs maximally at similar ZT across *T*-cycles supports the idea of larger masking component to this synchronization to *T*-cycles than a circadian clock mediated entrainment.

2.3.4 *early* flies show phase lability under skeleton *T*-cycles:

Previous experiments (*Fig. 2.2, 2.3, 2.4*) suggest some degree of anticipation (a hallmark of circadian clock-controlled rhythms) under full *T*-cycles. Therefore, I asked, to what extent circadian clock controls the phase of emergence in our stocks. The non-parametric model of circadian entrainment posits that lights during the dawn and dusk transitions entrain the clock and even short light pulses at these phases are sufficient to reproduce the waveform of the rhythm seen under full photoperiod. Masking responses to light depend both on the duration of the illumination, and the intensity of light (Aschoff and von Goetz, 1988; Mrosovsky, 1999). I carried out skeleton photoperiod experiments to examine the extent of circadian clock control over the phase of emergence in our stocks, with short duration (0.25 hours) light pulses of ~70 lux and asked if these pulses elicit a masking response as well.

Although all populations showed entrainment under all skeleton *T*-cycles provided, their eclosion waveforms did not closely mimic those under full photoperiods of respective *T*-cycles (*Fig. 2.4A & 2.5A*). Under skeleton *T*-cycles, *early* flies became phase labile (*left and right panel, Fig. 2.5B*). With long *T*-cycles they advanced their phases, just as *control* and *late* flies (*Fig. 2.5A & 2.5B*). The change in Ψ_{CoM} across different *T*-cycles was small for *early* flies, compared to *control* and *late* flies (*left panel, Fig. 2.5B*).

To distinguish the effects of non-parametric or parametric entrainment/ masking, I estimated the Sum of Square Differences (SSD) between full and skeleton *T*-cycles. I also compared the difference between waveforms under full and skeleton *T*-cycles across *T*-cycles of different lengths (*Fig. 2.5C*). If the SSD values are not significantly different from zero, it can be assumed that the light pulses non-parametrically entrained the populations with waveforms similar to full *T*-cycles whereas SSD values significantly different from zero will hint towards parametric entrainment or of masking under skeleton *T*-cycles. Eclosion waveform of *early* flies showed large changes (in SSD) under skeleton *T*-cycles, whereas the changes were very small in case of *control* and *late* flies (*Fig. 2.5C; Appendix tables A2.7, A2.9* – T22 and T26 respectively). I also compared the Ψ_{Peak} and Ψ_{CoM} of the eclosion rhythm of all populations under different lengths of skeleton *T*-cycles. Ψ_{Peak} of the eclosion rhythm of all populations showed trends similar to Ψ_{CoM} but was not significantly different among stocks across *T*-cycles (*right panel, Fig. 2.5B; Appendix table A2.12*), mostly due to the fact that the eclosion waveforms were not strictly unimodal in all individual vials. Although, *early* flies showed phase lability under skeleton *T*-cycles, they also showed a significantly advanced phase of entrainment (Ψ_{CoM}) than *control* and *late* flies, across all three regimes (*left panel, Fig. 2.5B; Appendix table A2.11*), suggesting that in addition to the masking response, advanced phase of circadian entrainment has also been selected for in our *early* flies. The extent of change in eclosion waveform between skeleton and full *T*-cycles can be seen in the change in R values (*Fig. 2.5B*). The *early* flies show significantly lower R value under T26 skeleton *T*-cycle, compared to the same under T22 and T24 skeleton *T*-cycles (*middle panel, Fig. 2.5B; Appendix table A2.10*).

These results suggest that, in *early* flies, although positive masking responses to lights-ON strongly influences Ψ_{CoM} and Ψ_{Peak} of the eclosion rhythm under full photoperiods, under skeleton photoperiods, these phases are clock controlled and significantly advanced (only Ψ_{CoM}) compared to *control* and *late* flies.

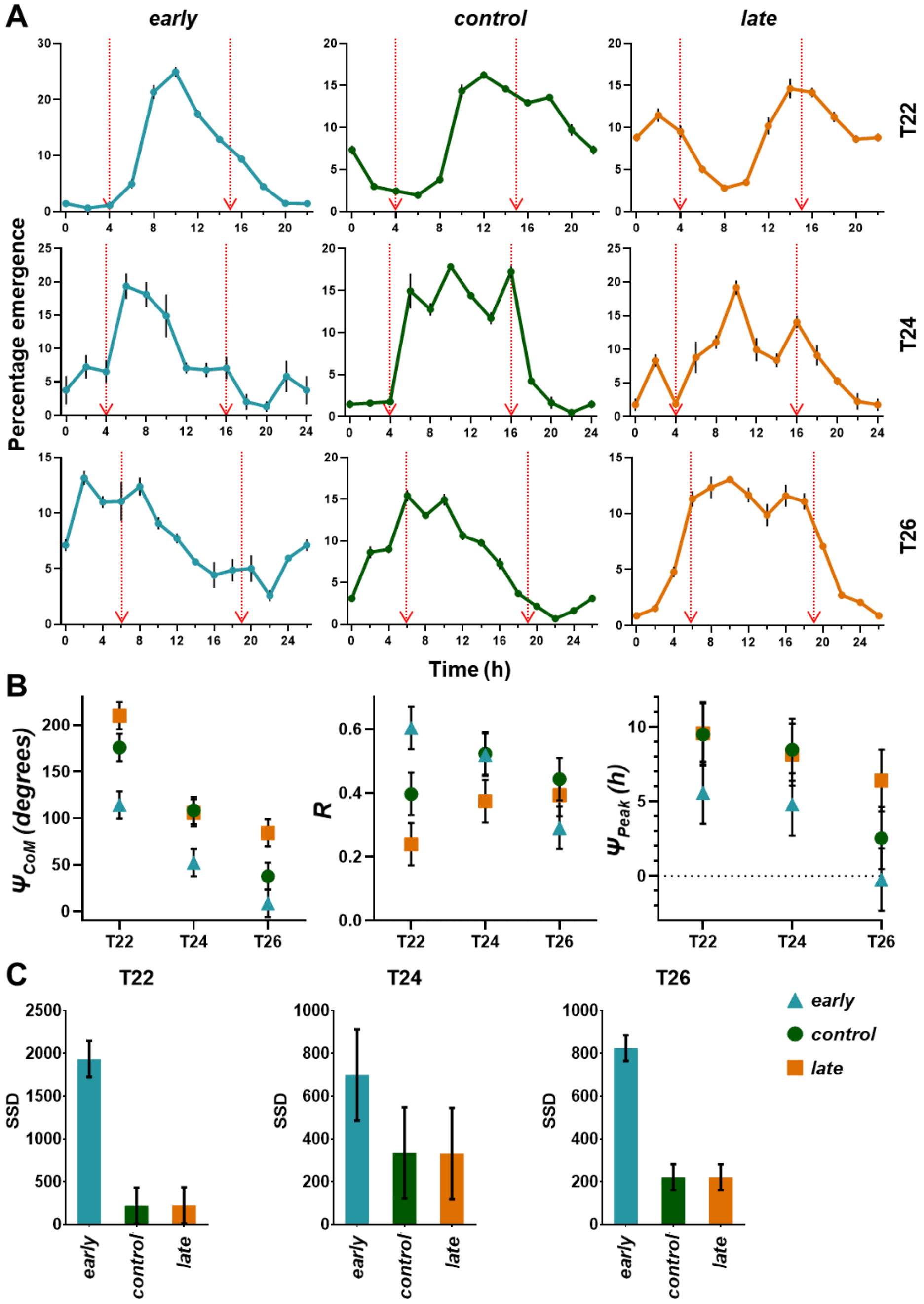


Figure 2.5: Emergence profile of *early*, *control*, and *late* flies in T22, T24, and T26 skeleton photoperiods and quantification of different eclosion rhythm parameters under different skeleton *T*-cycles. A. *early* flies change their emergence profile depending on skeleton *T*-cycles as do *control* and *late* flies. Red arrows depict the 15 minutes of the day when flies get light in respective skeleton photoperiods. Error bars are ± SEM. **B.** Left panel: Centre of Mass (Ψ_{CoM} – in degrees), middle panel: Consolidation of emergence (R), right panel: Peak in ZT (Ψ_{Peak}, in hours). Error bars are ±95% CI. **C.** Sum of Square Difference (SSD) of waveforms under full and skeleton photoperiods. Significantly high SSD in *early* flies depict large changes in waveform and phases between full and skeleton photoperiods. Error bars are ±95% CI.

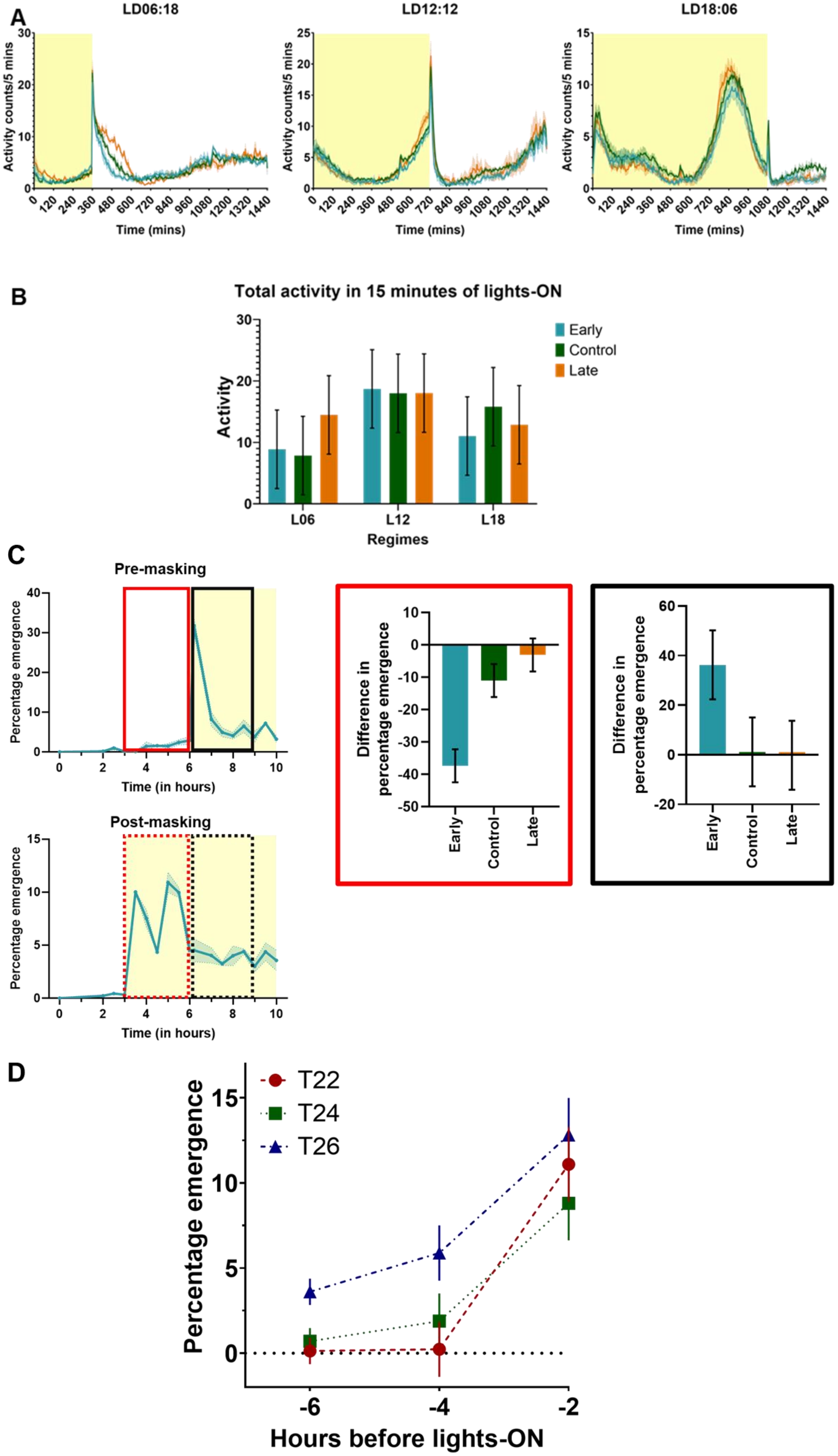

A
LD06:18
LD12:12
LD18:06
Activity counts/5 mins
Time (mins)
0
120
240
360
480
600
720
840
960
1080
1200
1320
1440
B
Total activity in 15 minutes of lights-ON
Activity
Early
Control
Late
L06
L12
L18
Regimes
C
Pre-masking
Post-masking
Percentage emergence
Time (in hours)
Difference in percentage emergence
Early
Control
Late
D
T22
T24
T26
Percentage emergence
Hours before lights-ON
-6
-4
-2

Figure 2.6: Average activity profile of *early*, *control* and *late* flies under LD06:18, LD12:12 and LD18:06 and total activity within 15 minutes of lights-ON of different stocks under different photoperiods from *Nikhil et al., 2016*; Difference in percentage emergence between pre-masking and post-masking days in two 2.5-hour time windows; Percentage emergence of early flies under full *T*-cycles (T22, T24 and T26) 2, 4, and 6 hours before lights-ON. **A.** Average activity profiles of *early*, *control* and *late* flies under ~70 lux LD06:18, LD12:12 and LD18:06 regimes over 4 days. Error bands are ± SEM. Red arrows indicate the time where masking is expected to occur. Purple arrow in right panel indicates a possibly delayed clock-controlled peak. **B.** Total activity within 15 minutes of lights-ON of stocks under different photoperiods. Error bars are ±95% CI (Selection × Regime effect on total activity – *Appendix table A2.19*). Details of experiment are in *Nikhil, Abhilash et al., 2016*. **C.** The left panels show time windows used for analysis (red and black rectangles respectively). In same time windows *early* flies show significantly higher (red region) or lower (black region) emergence than that of *control* and *late* flies on post-masking day than pre-masking day, showing this high emergence immediately after lights-ON is specific to early flies. Error bars are ±95% CI. **D.** As postulated, the putative clock-controlled peak for early flies occurs at ZT1 under LD12:12 (*Fig. 2.2A and 2.3A*). The anticipation to this clock-controlled peak can be seen under all *T*-cycles in form of emergence before lights-ON (*Fig. 2.4A, left column*). Though the masking-induced high emergence immediately after lights-ON at ZT2 is conserved under all three *T*-cycles (*Fig. 2.4A, left column*), the emergence starts much earlier under T26 (compared to under T22 and T24) as evident by significantly high emergence 6 hours before lights-ON, as though the underlying clock has advanced. On the other hand, under T22, the emergence 2 hours prior to lights-ON is similar to T24 and T26, while it is almost zero prior to this, as if the underlying clock is delayed compared to the other *T*-cycles. Error bars are ±95% CI (main effect of *T*-cycle, three separate ANOVAs for three different time points – *Appendix tables A2.14, A2.15 and A2.16*).

2.3.5 *early* flies show negative masking to warm temperature:

Previous experiments from the lab suggested that the early flies preferably emerge under cool temperature if provided a TC12:12 cycle during development (Nikhil et al., 2014). But these experiments were done with a presence of LD12:12 cycle along with TC12:12 during the first five days of development. I raised the flies under DD and a TC12:12 immediately after eggs were collected and thus the emergence rhythm phase should be only attributed to the TC12:12 in this case, instead of an interaction of light and temperature cues.

I observed that under TC12:12 and DD, *early* flies emerge almost exclusively in the cool part of the day (subjective night), whereas *control* and *late* flies show considerable emergence in the warm part of the day (subjective day) (*Fig. 2.7A*). When I designed experiments (*Fig. 2.7B*) to understand if this aversion of *early* flies to emerge during the warm part of the day, I observed that on *Assay day 1*, under TC12:12, *early* flies show majority of their emergence in the cool part of the day, whereas on *Assay day 2*, when the onset of warm temperature was advanced by 6 hours to cover the peak of emergence of *early* flies on *Assay day 1*, *early* flies suppress their emergence

(*Fig. 2.7C*) significantly more than *control* and *late* flies. (*Fig. 2.7D; left pane; Appendix table A2.20*). Even more interesting was the fact that *control* and *late* flies also suppressed their emergence due to the advancement of onset of warm temperature, but they still emerge later in the day under warm temperature, whereas *early* flies do not emerge in the later warm part of the day (*Fig. 2.7C*). *early* flies emerge in high numbers immediately after warm temperature was withdrawn, in the beginning of *Assay day 3* (beginning of subjective night). This high temperature induced active suppression of eclosion was exclusive to *early* flies and significant when compared to *control* and *late* flies (*Fig. 2.7D; right panel; Appendix table A2.21*). These results suggest that *early* flies show significantly more negative masking to high (warm) temperature compared to *control* and *late* flies, and actively suppress their emergence till cool temperature comes on (*Fig. 2.7*).

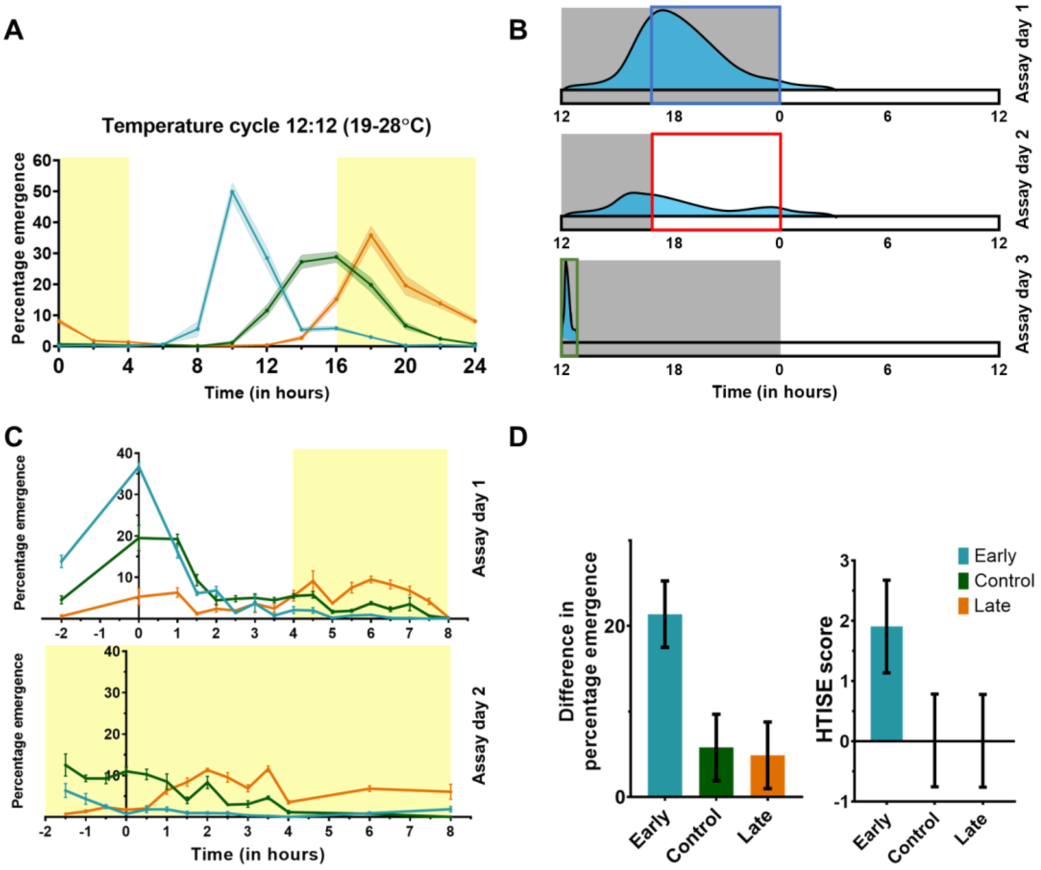


Figure 2.7: Schematics and average activity profile of *early*, *control* and *late* flies under temperature cycle (19°-28°C) and phase advanced temperature cycle (warm temperature on advanced by 6 hours), and quantification of the difference in percentage emergence in specific windows. **A.** Average emergence profiles of *early*, *control* and *late* flies under temperature cycle of 19°-28°C of 12:12 hours (TC:12:12) and constant darkness (DD) over 4 days. Error bands are ± SEM. Yellow shaded area – warm temperature (28°C) duration. **B.** Schematic of experimental details. Assay Day 1: emergence of early flies in TC12:12 regime. Assay Day 2: emergence of early flies when the warm temperature started 6 hours earlier than that of Assay Day 1 to cover the peak in emergence; note the reduced emergence after warm temperature started. Assay Day 3: high emergence of early flies as soon as warm temperature on Assay Day 2 was withdrawn and cool temperature started on the next day, as if flies were ready to emerge throughout the day and did not emerge because of the presence of warm temperature but emerge as soon as warm temperature is withdrawn. **C.** Average emergence profiles of early, control, and late flies during select 10 hours of the day during the cool to warm temperature transitions in Assay Day 1 and 2. The high emergence in early flies on Assay Day 1 was not visible on Assay Day 2 when warm temperature started 6 hours earlier. Grey shaded area depicts cool temperature (19°C). Error bars are ± SEM. **D.** Left panel: Quantification of suppression of emergence due to the advancement of warm temperature on Assay Day 2 compared to Assay Day 1 – quantified by the difference in percentage emergence in 6 hours depicted by blue rectangle in Assay Day 1 (**C**) and red rectangle in Assay Day 2 (**C**). Right panel: Quantification of High Temperature Induced Suppression of Eclosion (HTISE) score. HTISE score is a measurement of active suppression of emergence by high (warm) temperature. This measurement actually quantifies what proportion of flies did not emerge because of the advancement of warm temperature on Assay Day 2 but emerged immediately after warm temperature was withdrawn and cool temperature came on as if they were ready to emerge and their emergence was blocked/not allowed by the warm temperature. HTISE score was quantified as the proportion of emergence in hours marked by green rectangle on Assay Day 3 (**C**) and red rectangle on Assay Day 2 (**C**). Error bars are 95% CI.

2.4 Discussion:

While there have been several cases for the existence of masking as a phenomenon mediating biological rhythmicity (Aschoff, 1960; Aschoff and von Goetz, 1988; Binkley et al., 1983; Redlin and Mrosovsky, 1999b), it has been long neglected by circadian biologists; Nicholas Mrosovsky (1999) wrote "*Nevertheless, as a phenomenon worth study in itself, masking generally has been neglected by circadian biologists. Their attention has been focused more on making sure that interpretations involving masking can be excluded in research on rhythms and on devising ways of eliminating masking effects from measurements of clock phase; these include the use of constant darkness and skeleton photoperiods*" (Mrosovsky, 1999).

For the most part, masking has received attention in study of locomotor activity rhythm (Fry, 1947; Hamblen-Coyle et al., 1992; Kempinger et al., 2009; Prabhakaran and Sheeba, 2012; Rieger et al., 2003), although recently it has been invoked to explain aspects of eclosion rhythm in *Drosophila* (McNabb and Truman, 2008; Thakurdas et al., 2009) and the silk moth *Bombyx mori*

(Ikeda et al., 2019). It is known that the lights-ON signal mediated masking of *Drosophila* eclosion rhythm is brought about by the release of eclosion hormone as removal of eclosion hormone neurons attenuates the masking response (McNabb and Truman, 2008). A sudden temperature change can also induce masking response in eclosion via the eclosion hormone neurons (Jackson et al., 2005). There is tangential evidence that masking provides some adaptive value (Bloch et al., 2013; Lu et al., 2010), however to the best of my knowledge, there is no experimental evidence for masking evolving in response to selection.

2.4.1 Evolution of masking in *early* populations along with advanced phase of entrainment:

I hypothesized that due to the temporal placement of our selection window for the *early* flies (ZT22-ZT1), we may have inadvertently selected for individuals evolving a strong masking response to lights-ON (ZT0-ZT1), as I see in *Fig. 2.2A & 2.3A*, with very few (<5%) flies emerging in the window of ZT22-ZT0. Also, I observed 25-35% flies of *early* populations emerge immediately within half an hour after lights-ON (Ψ_{Peak} = ZT0.5) (*Fig. 2.2A & 2.3A*), which led me to believe there might be a significant masking component regulating the phase of the eclosion rhythm of *early* flies.

Unlike *control* and *late* populations, upon advancement of lights-ON, many of the *early* flies eclose immediately after lights-ON rather than what I would have expected if their eclosion rhythm was strongly circadian clock driven (*Fig. 2.2A & 2.2B*). Alternatively, this result suggests that the light-sensitive A-oscillator in *early* flies is stronger or dominant and facilitates re-entrainment in the very next cycle. However, as described in a later part of the discussion on duration and intensity of light, the lights-ON response in eclosion is not as immediate as for other behaviors such as locomotor activity rhythm (as fast as under 5 minutes of lights-ON, Prabhakaran and Sheeba, 2012), so the high emergence immediately after lights-ON on *Assay day 2* (*Fig. 2.2A*) being a result of faster re-entrainment seems highly unlikely. My hypothesis of strong masking for the *early* flies was further validated by lower emergence in *early* population at subjective dawn in DD (*Fig. 2.3A & 2.3B*) in contrast to the flies from the other two sets of populations. This result may also be interpreted as – *early* flies showing higher amplitude expansion under entrainment for eclosion rhythms followed by amplitude reduction in response to transition from LD to DD. However, the high emergence at ZT0.5 (presumably the masking induced peak) under LD is absent in DD, whereas the next highest peak at ZT1 (presumably the clock-controlled peak) under LD is maintained at similar phase under DD (~10% emergence), showing a phase control and is seen in previous experiments as well (*Fig. 2.2A*). While *control* and *late* flies phased eclosion in the expected direction based on circadian clock control under short and long *T*-cycles *(Fig. 2.4A & 2.5A)*, the *early* flies consistently respond by positive masking with high emergence immediately after lights-ON. However, the response of *early* flies under skeleton pulse-induced *T*-cycles was remarkable because similar to the *control* and *late* flies, they too delayed and advanced emergence phases (Ψ_{CoM}) under T22 and T26 respectively (*Fig. 2.5*).

In the phase advance experiments (*Fig. 2.2*), although *early* flies exhibited high emergence immediately after lights-ON at ZT0.5, there was considerable emergence even around ZT1.5 – 2 – this peak was comparable to the peak at ZT1 on *Assay day 2* of LD-DD transition experiments (*bottom panel, Fig. 2.2A & bottom panel*, *Fig. 2.3A*). I reasoned that the phase advance of lights-ON on *Assay day 2* exposed the late night (advance zone) of the PRC of *early* flies to light (Kumar et al., 2007), thus advancing the clock and hence producing an advanced peak (at ZT 2 in phase advance experiments and ZT1 in LD-DD transition experiments – both at *Assay day 2*). The exposure to three skeleton *T*-cycles revealed that indeed *early* flies maintained significantly advanced Ψ_{CoM} compared to *control* and *late* flies (*Fig. 2.5B*), suggesting that our selection protocol has indeed resulted in advanced phase of emergence in *early* flies along with greater masking which is revealed under full photoperiods. The fact that the *control* flies show some extent of masking (*Fig. 2.3B*), and upon selection for advanced phase of entrainment, can give rise to significantly higher masking (*early* population), suggests that masking-inducing variations are present in populations, albeit in low frequency. Recently, Pegoraro and colleagues have shown that negative masking could evolve when flies were selected for nocturnality (Pegoraro et al., 2019). However, it is important to note that not all short-period flies are expected to show this masking response to lights-ON, e.g., per^s flies with a period of ~19.5 hours, have such an advanced phase of emergence compared to their controls, that their Ψ_{Peak} occurs several hours before the lights-ON transition (Qiu and Hardin, 1996). In our populations, this evolved masking response is associated with the early chronotype and their advanced phases after selection under LD12:12 and is unlikely to be driven by the shorter period of the *early* flies (barely 0.5 hours < *control*).

Interestingly, the locomotor activity rhythm of *early* flies does not show heightened masking response to lights-ON compared to the *control* and *late* flies (up until ~ generation 260, reanalyzed data from *Nikhil, Abhilash et al., 2016*; *Fig. 2.6*). This may be partly due to the fact that, in general, flies show high startle response to lights-ON stimulus (Rieger et al., 2003). Previous studies on these populations have shown that there are no differences in phasing of locomotor activity rhythms under LD12:12, even though there is a significant difference in free running period (Kumar et al., 2007; Nikhil et al., 2016b). For eclosion rhythm, both phase and period have changed significantly among *early, control* and *late* flies (Kumar et al., 2007; Nikhil et al., 2016b). Since period changes are similar for both rhythms while phase divergence among chronotypes is seen only in eclosion rhythm, I speculate that this may be a result of differences in output pathways governing phases of these two rhythms.

2.4.2 Higher range of "apparent" entrainment achieved by masking:

My results suggest that apparent phase of an entrained rhythm may be contributed by both masking and the circadian clock, and in some regimes, more by the masking component than the clock (under full photoperiods, *Fig. 2.2, 2.3 & 2.4*). I introduce the term "apparent entrainment" because under skeleton *T*-cycles I see that the *early* flies have evolved an earlier phase of emergence which is clock-driven, along with positive masking. I reason that the *early* flies have evolved "apparently entrained" phases via changes in both the clock and the masking response. I do not equate this "apparent entrained" state with "entrainment" which is a core property of the circadian clock and its oscillators (Bittman, 2020), but merely mean a "synchronized state" of the overt rhythm with an environmental cyclic cue. I caution that "apparent entrained" states may present as clock driven entrained states, and without targeted experimental designs and adequate temporal sampling resolution, be hard to identify. Our own previous studies with 2-hour resolution suggested that the *early* strains are entrained with an advanced phase of emergence, however, by increasing sampling frequency, I uncovered the effect of masking on this "apparently entrained" emergence profile.

My studies also show that a spectacularly high range of "apparent entrainment" can be achieved just by masking, at least for the eclosion rhythm. Under extreme *T*-cycles of T20 and T28 *early* flies still have high emergence immediately after lights-ON and show an apparent entrained behavior, whereas percentage entrainment of *control* and *late* flies under these regimes are significantly lower than *early* flies. Altogether my experiments make a very strong case that in the process of creating the *early* chronotypes, in addition to them evolving a faster clock with advanced phase of entrainment, they have also evolved robust positive masking responses to lights-ON. This facilitates the idea that in nature, organisms may use masking – a clock-independent phenomenon, as a mechanism to phase themselves to appropriate times of the day, e.g., eclosion being gated to the early part of the day to prevent desiccation and enhance survival rate. It can be argued that masking may be an evolutionary disadvantage in the sense that to compensate for a highly "noisy" clock, masking may evolve to maintain specific phases locked to a zeitgeber(s) and does not have the flexibility of phase lability in complex environments. Indeed, previous work from our lab has shown that under complex zeitgeber conditions (light:dark 12:12 + warm:cool 12:12 ; in-phase and out-of-phase), *early* flies show remarkable resilience to shift phases and stay phase locked to the lights-ON signal, whereas *control* and *late* flies change their phases more readily (Abhilash et al., 2019). Moreover, empirical evidence suggests that a fully functional clock and masking can co-exist in an animal and in all probability, these are complementary mechanisms for organisms to maintain specific phases (Aschoff and von Goetz, 1988; Mrosovsky, 1994; Redlin and Mrosovsky, 1999b; Rensing, 1989).

2.4.3 The duration and intensity of light dramatically alters magnitude of response in eclosion:

The light used in all my experiments was of ~70 lux (to match the long-term selection maintenance regime) and can be considered low intensity compared to majority of the eclosion rhythm studies in *Drosophila* (which typically used ~750 lux). It is especially so in those studies where relationship between the lights-ON signal and downstream pathways have been analyzed and linked to timing of eclosion (Baker et al., 1999; McNabb and Truman, 2008). They show that lights-ON signal can rapidly induce eclosion of up to ~20% of the waveform within about 10 minutes. This suggests that a skeleton photoperiod should be able to induce high masking responses, which I did not observe in any of the skeleton *T*-cycles. This can be explained by the different light intensity used in my experiments (at least 10 times lower). Nevertheless, the regime

I used, induces ~35% emergence in a mere 30 minutes when the lights-ON was advanced (*Fig. 2.2B*). I propose that similar number of photons integrated over time could induce high emergence in *early* populations, compared to *control* and *late* populations.

Another important difference from previous studies which were mostly aimed to understand the mechanisms of eclosion hormone release and downstream pathways leading to the act of eclosion is, that they used cultures containing similar developmental stages. To observe light-induced emergence, typically experiments were set up such that flies were expected to emerge in the next ~60 minutes after light pulses were given (McNabb and Truman, 2008). My studies together with this information from previous reports lead me to hypothesize that photon integration is a part of this masking response, and that as soon as a threshold is crossed, masking response to lights-ON is observed.

2.4.4 Plausible mechanisms driving higher masking in *early* populations:

Some hint of possibility of the evolution of masking mechanisms are seen in an *in-silico* study, where gene regulatory networks were allowed to evolve under selection for correct prediction of phases under light-dark cycles (Troein et al., 2009). It was seen that under a simple LD12:12 condition, only delayed light responses were selected for, while no oscillatory mechanism was found to evolve. This suggests that in a minimal environment, where the only cycling cue is a light-dark cycle, oscillators may not be necessary to achieve particular phases, and a delayed light response may as well do the trick. In our selection regime, environmental conditions are similarly minimal, with only one cycling cue, a ~70 lux LD12:12 square light cycle (abrupt transitions) regime with constant ambient temperature of 25±0.5 °C, and results from Troein and colleagues supports the idea of evolution of simpler phasing mechanisms, such as strong masking. This evolved masking response may also be governed by co-evolved non-circadian photosensitivity. Thus, there are the possibilities of (a) coupling of strong photosensitivity and eclosion hormone expressing neurons (when ablated, lights-ON response absent; (McNabb et al., 1997)) or (b) heightened photosensitivity of the prothoracic gland itself having evolved in the *early* flies.

2.4.5 Evolution of negative masking response to high temperature in *early* flies:

I hypothesized that *early* flies prefer the cool temperature part of the day to emerge under temperature cycles in absence of any light cue. Indeed, I observed that *early* flies emerge exclusively under the cool part of the day (19°C) even though this part of the day generally depicts subjective night (*Fig. 2.7A-D*). This behavior was not observed for *control* and *late* flies (*Fig. 2.7C-D*). Previously, I observed that *early* flies show significantly more positive masking responses to the lights-ON signal, which in nature depicts the start of the subjective day (*Fig. 2.2, 2.3, 2.4*). I speculated that both the positive masking behavior to lights-ON and negative masking to high temperature can act together to ensure phasing of the emergence timing to the early part of the day, particularly around dawn in the *early* flies. This is shown previously where LD12:12 cycle was present along with a TC12:12 with a 4 hour phase difference (warm temperature started 4 hours after lights-ON) and *early* flies still remain phase locked onto the lights-ON signal and emerge under cool temperature (Abhilash et al., 2019; Nikhil et al., 2014). I concluded that during selection for advanced phase of emergence, *early* flies evolved a strong positive masking response to the lights-ON signal, and coevolved negative masking response to warm temperature, both of which may help the *early* flies phase their emergence strictly around dawn.

To the best of my knowledge this work is the first experimental demonstration that masking can evolve as a response to selection for phase of entrainment. I propose that masking can be a valid mechanism by which organisms show *early* chronotype in an environment where strong light transitions are present. I also suggest that light masking and temperature masking can co-evolve in an organism to phase themselves accordingly in presence of multiple strong zeitgebers (light and temperature). Further investigations into the molecular mechanisms and neuronal control of this masking response are needed, for which some plausible targets have been mentioned above. Overall, this work highlights the complex mechanisms of “apparent” light entrainment and provides an experimental framework to dissect out relative contributions of masking and the circadian clock regulating timing of a behavior.

Chapter 3

Characterization of masking responses in locomotor activity rhythm and sleep

3.1 Introduction:

"*The masking factors were defined as those factors which prevent a second identity from operating on the organism to the extent that it would if the masking factor were not present.*"

– F.E.J. Fry, in "Effects of environment on animal activity" – first mention of masking in the context of locomotor activity rhythm (Fry, 1947).

Jürgen Aschoff later described "masking" as "*Certain (sometimes overlooked) experimental condition* (sic) *can obscure the real zeitgeber-mechanism. We may call them masking conditions.*" (Aschoff, 1960). In context of an organism's locomotor activity rhythm, masking can be defined as a sudden increase or cessation of the locomotor activity in response to the sudden appearance or withdrawal of an environmental cue. These sudden changes in behavior/behavioral rhythms may conceal the circadian clock that governs these rhythms (Mrosovsky, 1999). Masking may be important for organisms living in simple environments where sharp transitions in environmental cues are present and help supplement the circadian clocks in phasing different behaviors to the transitions, and thus provide some adaptive value (Bloch et al., 2013; Ghosh et al., 2021; Lu et al., 2010). It has also been postulated that masking may also play some role, and even be a predictor of chronotypes (Refinetti et al., 2019). Chronotypes have long been studied in humans and other organisms, however, the source of these inter-individual differences in phases of entrainment is largely unknown. These differences are known to be affected by individual genotypes and environment, which suggests that differences in masking responses (non-circadian photic/non-photic sensitivity) can also alter chronotypes. *Octodon degus* (common degus) are dual-phasing rodents, some individuals show a nocturnal chronotype, while others show a diurnal chronotype under laboratory-housing conditions, and following a change of light regimes, an intermediate chronotype is also observed. It has been shown that in these animals, the nocturnal animals are characterized by a strong masking response, the diurnal animals do not show any masking response, and the intermediate chronotypes show variable levels of masking, which indicates a high probability of masking affecting chronotypes, at least for the locomotor activity rhythm (Vivanco et al., 2009).

Light is known to elicit a strong positive masking response in the *Drosophila melanogaster* locomotor activity rhythm (Hamblen-Coyle et al., 1992; Kempinger et al., 2009; Rieger et al., 2003; Wheeler et al., 1993). However, till now, there has been no systemic investigation on the effect of masking in different chronotypes of *Drosophila melanogaster*. Perhaps, this lack of literature on the topic stems from unavailability of suitable models of Drosophila chronotypes. As described in previous chapters, we have divergent chronotype populations of *Drosophila melanogaster* based on their timing of adult emergence, with population-level replicates, which enable us to study the evolution of different clock/non-clock properties in different chronotypes. I attempted to address the following questions related to the evolution of masking in the locomotor activity rhythms in these divergent chronotype fly populations:

1. Can masking of locomotor activity evolve differently in divergent chronotypes?
2. Is the activity masking response important for adjustments under different day lengths in flies?

Sleep is probably the most versatile physiological state which interacts with the circadian clock, and daily sleep timing is controlled by the individual’s circadian clock to a large extent (Shafer and Keene, 2021). Chronotypes in humans have been described majorly on the basis of their mid-sleep timing, and it is no surprise that different chronotypes have different sleep timings (Randler et al., 2017; Roenneberg et al., 2007, 2012). Depending on the organism’s circadian clock and phase, light or any other zeitgeber can either delay (when falling at dusk) or advance (when falling at dawn), or not have an effect at all (when falling at midday). These phase shifts establish the organism's circadian clock at a particular phase angle with a zeitgeber cycle, which is called the phase of entrainment (Roenneberg et al., 2003). Variations in the phases of entrainment in a population lead to divergent chronotypes (Roenneberg et al., 2007). Roenneberg and colleagues found that the distribution of chronotypes within a population is almost gaussian in nature, and chronotypes were defined by their mid-sleep timing on weekends (non-working free days). They found a large percentage of the population were either early type (35.02%) or late type (50.38%), and the phase difference between extreme early and extreme late types can as well span over ~18 hours (Roenneberg et al., 2007). Though they found no significant correlation between sleep timing and sleep duration, there was a significant correlation between sleep duration in weekdays (work days) and weekends (free days) and chronotype – late chronotypes had shorter work day sleep and longer free day sleep (Roenneberg et al., 2007). It has also been proposed that both early and late chronotypes have a mismatch in their sleep duration between work days and free days, while "neither" or "normal" chronotypes do not show this mismatch, thus suggesting that the misalignment between the biological and social clocks can affect both chronotypes, though late chronotypes were more affected on weekdays (Roenneberg et al., 2007). A recent study found that sleep duration is actually predicted by their average bedtime, not average wake time, and women sleep more than men (Walch et al., 2016). They also predicted later sunrise times may decrease sleep duration, and later sunset times and higher brightness may increase the duration of sleep (Walch et al., 2016). Genetic influences can explain a significant percentage of the chronotype variation within a population; some heritability estimates suggest up to 50% (Kalmbach et al., 2017). Polymorphisms in core clock genes such as *PER3* and *CLOCK* have been associated with extremely delayed sleep timing and duration (Ebisawa et al., 2001; Katzenberg et al., 1998). Similarly, polymorphisms in genes like *PER3* and *ARNTL2* have been found to be associated with earlier sleep timing (Carpen et al., 2005, 2006). Furthermore, the near-normal distributions of chronotypes indicate a highly polygenic basis for their wide divergence, and detection of these small genetic effects would require very large population level sample sizes and genome-wide association studies (Kalmbach et al., 2017; Roenneberg et al., 2007).

How sleep amount and quality are affected in different chronotypes is generally studied in different cohorts of human societies where multiple zeitgebers are present and "social jetlag" plays a huge role (Allebrandt et al., 2013; Gao et al., 2019; Martin et al., 2012), and isolating humans to study effects of singular modalities prove to be difficult. Model organisms can be very powerful tools for understanding how sleep is affected by different strengths and length of particular zeitgebers in different chronotypes. Due to the availability of *Drosophila melanogaster* populations evolved to have divergent chronotypes, it presented a unique opportunity for me to study how sleep is affected in these populations under different day lengths and light intensities, and I asked the following specific questions:

1. Can differences in light intensity during daytime impact sleep properties in different chronotypes?
2. Is sleep different among extreme chronotypes (*early* and *late* flies) compared to their *control*?
3. Do sleep parameters differ among chronotypes under different day lengths?

In this chapter, I present the results of studies examining the effect of masking on different chronotypes and show that the *early* populations have evolved significantly higher masking compared to the *late* populations. I also show that extreme chronotypes (*early* and *late*) have significantly lower nighttime sleep amount and quality under high light intensity at daytime and under shorter daylengths compared to the *control* populations.

3.2 Materials and methods:

3.2.1 Selection protocol and fly husbandry:

Selection protocol and fly husbandry have been described in detail in previous chapters. All experiments pertaining to this chapter were done between generations 330 to 350 with standardized populations of flies.

3.2.2 Behavioral experiments:

Before each experiment, ~300 eggs were collected and placed in 10 vials each for all 12 standardized populations. Virgin males were collected for 3 days and aged for 2-4 days after final collection. These flies were then loaded into 5 mm locomotor activity tubes with standard cornmeal food, and their activity was recorded at 1-minute intervals with the Trikinetics DAM systems (Trikinetics, Waltham, MA) under 25±0.5° C. For different experiments, light intensity inside the Percival incubators (Percival Scientific Inc, Perry, IA) was set manually using a handheld luxmeter (LI-COR Biosciences, Lincoln, NE) to either ~70 lux or ~500 lux to ensure uniform light distribution at each monitor. In different photoperiod experiments, flies were first entrained to a minimum of 3 days LD12:12, and at the end of the third day, upon a visual inspection confirming their entrainment to LD12:12, the light regimes were changed to respective photoperiods (LD04:20, LD08:16, LD16:08, and LD20:04). All light regimes were set by programming the incubators.

3.2.3 Data analysis and statistics:

All locomotor activity data were collected, scanned through the Trikinetics DAMScan program (Trikinetics, Waltham, MA) to get rid of any error in recording, and data was binned to 1-minute intervals. For activity rest analysis, data were analyzed with VANESSA-DAM-CRA (Ghosh and Sheeba, 2022). Only rhythmic individuals with at least 4 days of uninterrupted data were considered for further analyses. Rhythmicity was determined using Lomb-Scargle periodograms employed in VANESSA-DAM-CRA, and rhythmicity was considered significant at $\alpha = 0.05$. Both raw activity profiles and normalized activity profiles (normalized by the day's total activity) were constructed after averaging over cycles for each individual and then averaging over individuals of a genotype. While I have only plotted raw activity profiles, all comparisons were done for both raw activity counts and normalized activity, and results were considered significant only when both comparisons yielded similar statistical results. Raw activity counts and profiles are not always reliable as Drosophila locomotor activity counts may vary between experiments and genotypes, and without normalization, comparison among different experiments and genotypes may lead to inaccurate conclusions. Masking responses were calculated as total raw activity counts or normalized activity within 5, 10, and 15 minutes of the lights-ON signal in Microsoft Excel 365. Raw activity profiles of 1-minute bin were smoothened using kernel smoothing with a bandwidth of 30 to visualize the clock-controlled and masking peak clearly by getting rid of the noise associated with VANESSA-DAM-CRA. Sleep was calculated with the raw locomotor activity data of 1-minute bin using VANESSA-DAM-SA, adhering to the traditional definition of ≥ 5 minutes of continuous immobility (Hendricks et al., 2000; Shaw et al., 2000). Each day's sleep parameters (total sleep amount, number of sleep bouts, mean bout length, and latency to sleep) were calculated for day and night phases separately for each photoperiod and then were averaged over days for each individual and then across individuals of a genotype respectively. All statistics were done on four block means (replicate populations, each replicate population had ≥ 25 flies). For masking quantifications, I performed a two-way randomized block design mixed-model ANOVA, where blocks were treated as a random factor and selection (genotype) was treated as fixed factor. The Tukey's HSD test was done for multiple comparisons following the ANOVA. For sleep parameter quantifications, I performed a three-way randomized block design mixed-model ANOVA, where blocks were treated as a random factor, and selection (genotype) and regimes (either light intensity or photoperiods) were treated as fixed factors. Tukey's HSD test was done for multiple comparisons following the ANOVA. The same dataset was used for activity rest and sleep calculations. The ~70 lux photoperiod data was taken from a previous publication from the lab (Nikhil et al., 2016b), and Figure 3.1 was published as it is in one of my earlier publications (Ghosh et al., 2021). Though I have presented data here from one set of experiments, all high light intensity experiments were repeated at least twice at different generations and in different incubators and showed similar results. Tables with these results are provided in *Appendix table A3.54-A3.58*.

3.3 Results:

I reanalyzed locomotor activity data from previously conducted experiments under light-dark cycles of 12 hours each (LD12:12), short photoperiod (LD06:18), long photoperiod (LD18:06), and ~70 lux of light intensity (intensity similar to their maintenance regime), and quantified masking responses (extremely immediate and immediate). Next, I quantified masking responses under high light intensity (~500 lux) photoperiods systematically – LD04:20, LD08:16, LD12:12, LD16:08, and LD20:04. I found that high light intensity photoperiods were able to separate the clock-controlled peak and the masking peak efficiently. I also quantified sleep and sleep architecture under these photoperiods.

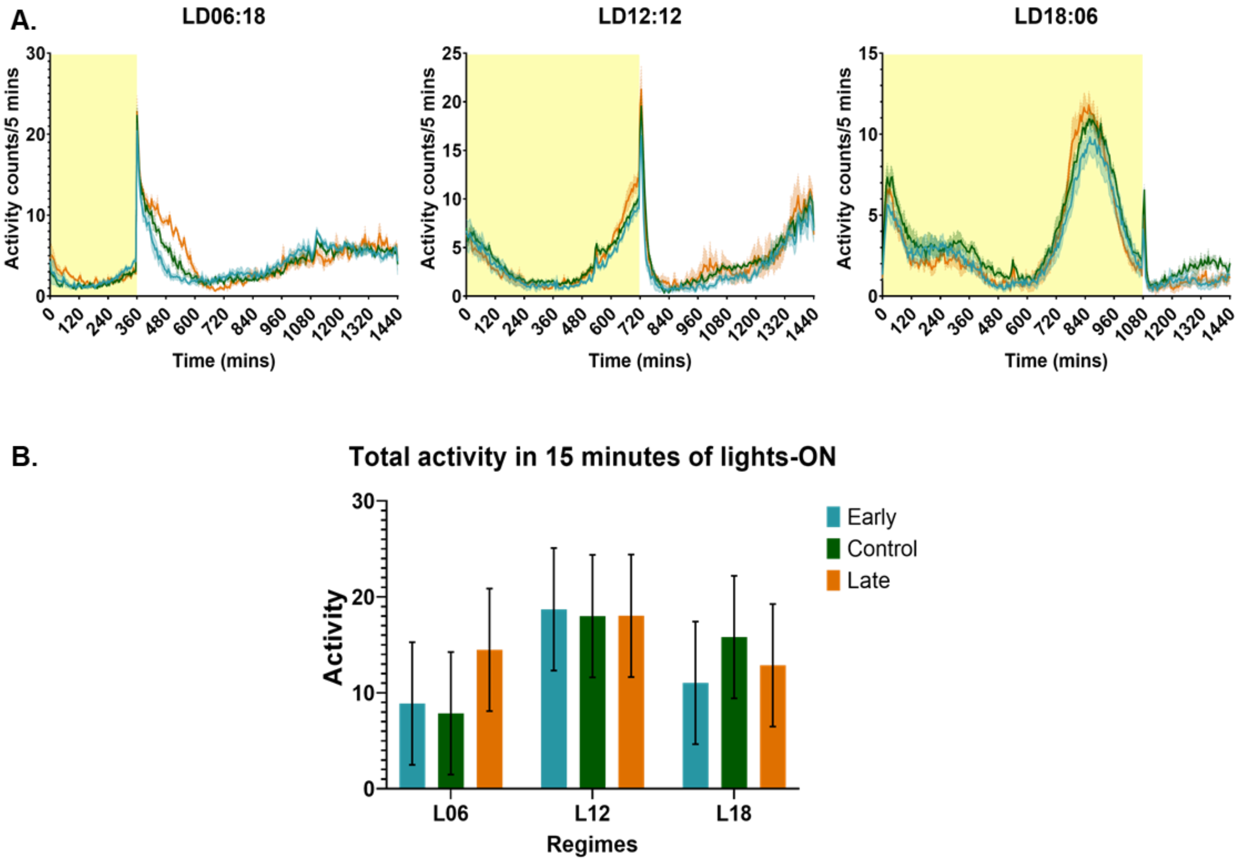


Figure 3.1: Activity profiles and quantification of lights-ON mediated masking responses under low light intensity. A. Activity profiles of *early*, *control* and *late* flies under ~70 lux photoperiods (*left* – LD06:18, *middle* – LD12:12, *right* – LD18:06). Clock-controlled peak advanced or delayed under short and long photoperiods, respectively. Yellow shaded regions are light part of the day. **B.** Quantification of masking by summing total activity within 15 minutes of lights-ON under different photoperiods (L06 – LD06:18, L12 – LD12:12, L18 – LD18:06). There was no difference in masking among populations and photoperiods. Error bands: ±SEM, error bars: 95% CI from a two-way ANOVA followed by Tukey's HSD.

3.3.1 No difference in masking under low intensity photoperiods between chronotypes:

Positive masking response in *Drosophila melanogaster* locomotor activity rhythm is generally quantified by total activity immediately after the lights-ON transition in different time windows (Prabhakaran and Sheeba, 2012). When I quantified positive masking responses as the total amount of activity within 5, 10, and 15 minutes of lights-ON, I did not find any main effect of selection or selection × regime interaction (*Fig. 3.1B; Appendix table A3.1*). However, for all the chronotypes, the clock-controlled peak appeared to advance or delay under short and long photoperiods, respectively, whereas the masking peak was phase-locked onto the lights-ON signal (*Fig. 3.1A*). Under short photoperiod (LD06:18), the masking peak was almost negligible, and under long photoperiod (LD18:06), the masking peak was identifiable and was comparable to the peak at lights-OFF (*Fig. 3.1A*).

3.3.2 Difference in masking between chronotypes under high light intensity photoperiods:

Under high intensity LD12:12, the total activity of the flies was higher than under low intensity LD12:12 (*Fig. 3.2A, B, E, F* – note *y*-axis differences). *early* flies showed a trend of higher activity than *control* and *late* flies immediately after lights-ON, but this difference was not significant (*Fig. 3.2C, D, E, F; Appendix table A3.2-A3.13*). To account for differences and avoid bias due to differences in total activity level among populations and among experiments, I also calculated percentage activity in addition to raw activity counts (*Fig. 3.2 C, D*). However, when I assayed flies under high intensity photoperiods, the morning clock-controlled peak separate from the masking peak as expected – it advanced and delayed under shorter and longer photoperiods, respectively (*Fig. 3.3*).

Under different photoperiods, clock-controlled phases of *early* flies were more advanced than those of *control* and *late* flies (most noticeable in shorter photoperiods – *Fig. 3.3A, B*) – reaffirming that their locomotor rhythm chronotype differences mostly stem from their circadian clock driven phases. Under long photoperiods, the clock-controlled peaks separate well, but the phase differences were not apparent (*Fig. 3.3D, E*). After confirming that high intensity photoperiods could clearly separate the clock-controlled and masking peaks around dawn, I quantified extremely immediate and immediate masking responses by quantifying raw activity counts and normalized activity counts within 5, 10, and 15 minutes of lights-ON. Under short photoperiods (*Fig. 3.4A, B*), I did not find any differences in the masking response to the lights-ON signal (*Fig. 3.4 C, D, E, F; Appendix table A3.14-A3.25*).

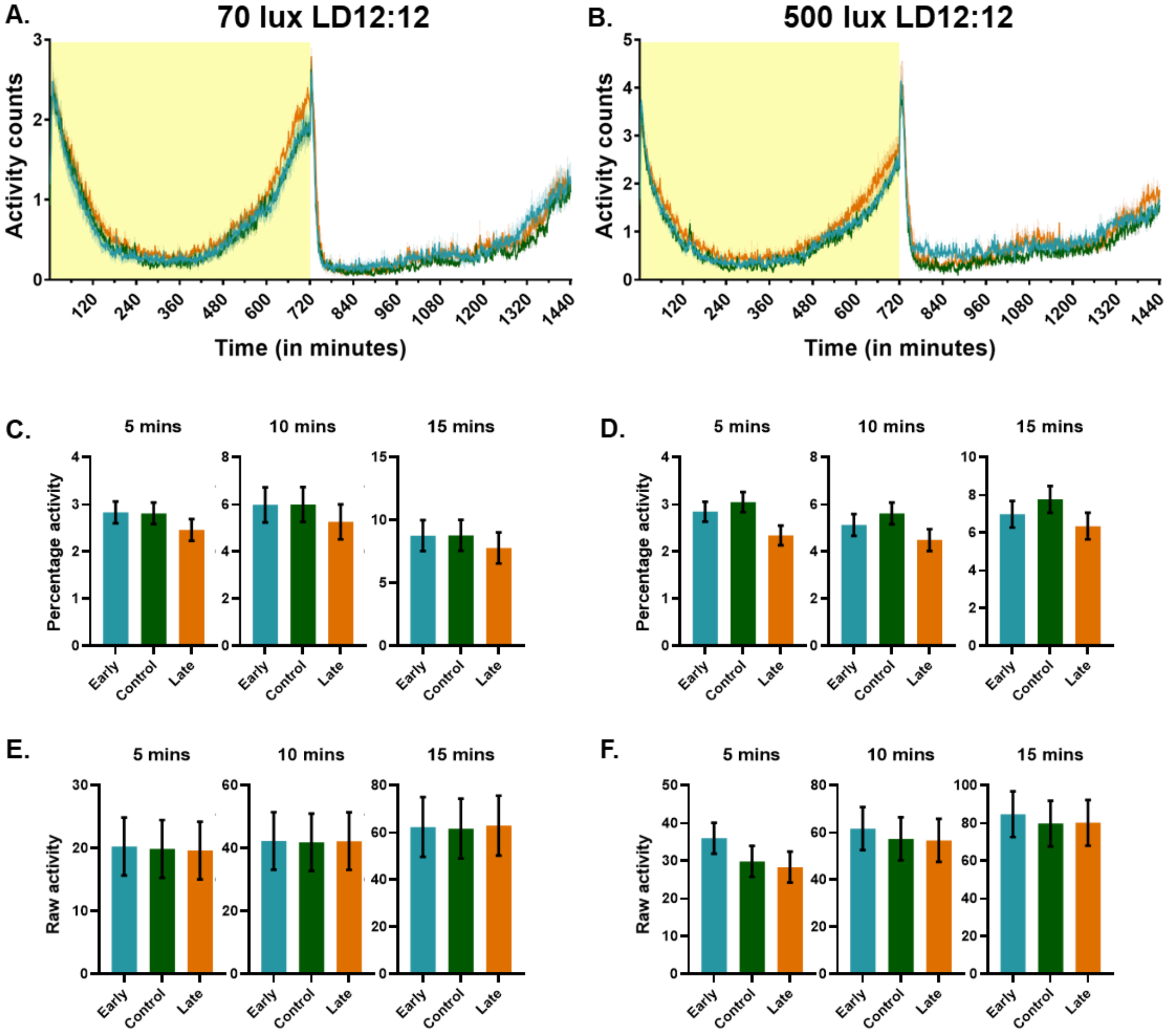


Figure 3.2: Activity profiles and quantification of lights-ON mediated masking responses under low and high light intensity LD12:12. A, B. Activity profiles of *early*, *control* and *late* flies under low (~70 lux) and high light (~500 lux) intensity LD12:12. **C, E.** Quantification of masking by summing total activity within 5, 10, and 15 minutes of lights-ON under low light intensity LD12:12 (***C*** – percentage activity, ***E*** – raw activity counts). **D, F.** Quantification of masking by summing total activity within 5, 10, and 15 minutes of lights-ON under high light intensity LD12:12 (***D*** – percentage activity, ***F*** – raw activity counts). There are no differences in masking responses among populations under different light intensities. Error bands: ±SEM, error bars: 95% CI from a one-way ANOVA followed by Tukey's HSD. Color codes: blue – *early*, green – *control*, orange – *late*.

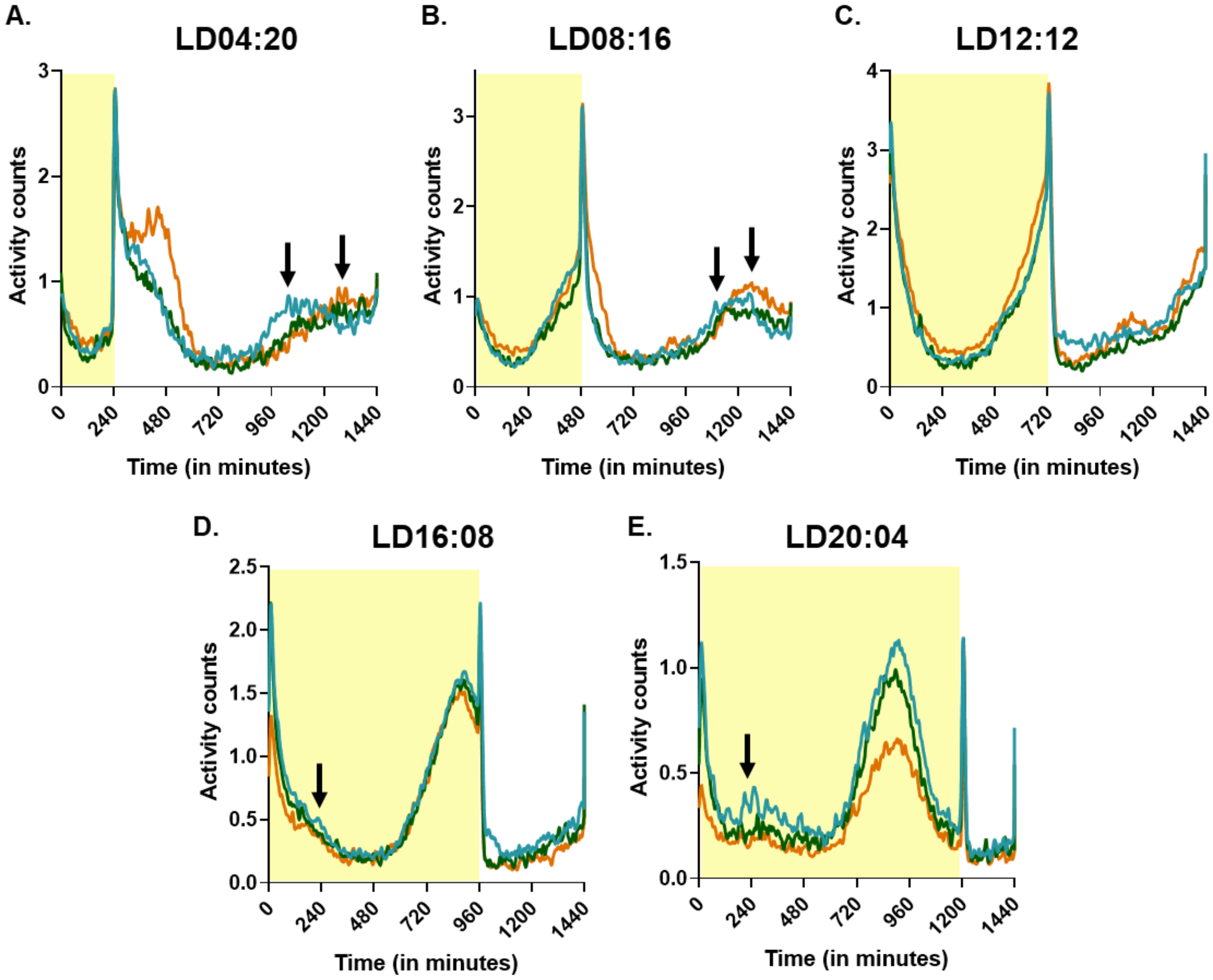


Figure 3.3: Smoothened activity profiles under high light intensity photoperiods. Activity profiles were smoothened using a kernel smoothing filter with bandwidth 30 and then plotted to reduce noise in high resolution (1-minute bin) activity data. Clock-controlled peaks were advanced under shorter photoperiods (***A*** – LD04:20 and ***B*** – LD08:16) and delayed under longer photoperiods (***D*** – LD16:08 and ***E*** – LD20:04), whereas clock-controlled peaks were not separated from the masking peak under LD12:12 (***C***). Error bands: ±SEM.

However, this can be explained by the absence of any observable masking peak at lights-ON under shorter photoperiods. I found stark differences in the morning masking response under long photoperiods (*Fig. 3.5A, B* – insets). *early* and *control* flies showed significantly higher masking responses – both extremely immediate and immediate responses, compared to the *late* flies (*Fig. 3.5C, D, E, F; Appendix table A3.26-A3.37*). Under the longest photoperiod assayed (LD20:04), *early* flies showed even more significant masking differences, compared to both *control* and *late* flies (*Fig. 3.5D*), especially for normalized activity levels immediately after lights-ON. The significantly higher masking response seems to stem from their total exposure to light, as a lower amount of light exposure under short photoperiods essentially abolished the masking response to the lights-ON signal, whereas the sharp negative masking response to the lights-OFF signal during the evening is conserved under all photoperiods (*Fig. 3.2A & B, 3.3, 3.4A & B, 3.5A & B*).

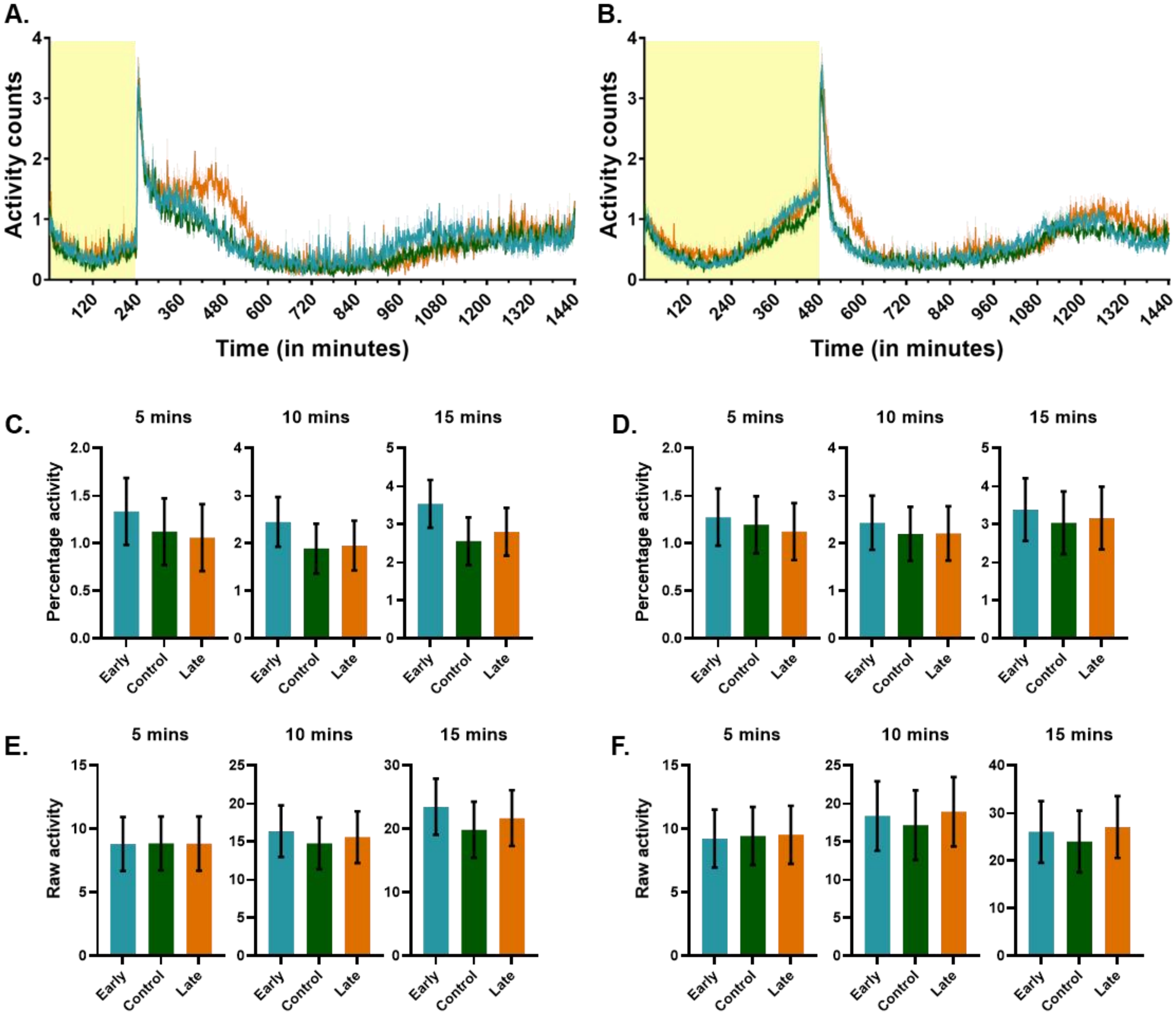


Figure 3.4: Activity profiles and quantification of lights-ON mediated masking responses under high light intensity shorter photoperiods. A, B. Activity profiles of *early*, *control* and *late* flies under short photoperiods (***A*** – LD04:20, ***B*** – LD08:16). **C, E.** Quantification of masking by summing total activity within 5, 10, and 15 minutes of lights-ON under LD04:20 (***C*** – percentage activity, ***E*** – raw activity counts). **D, F.** Quantification of masking by summing total activity within 5, 10, and 15 minutes of lights-ON under LD08:16 (***D*** – percentage activity, ***F*** – raw

activity counts). There are no differences in masking responses among populations under different photoperiods. Error bands: ±SEM, error bars: 95% CI from a one-way ANOVA followed by Tukey's HSD.

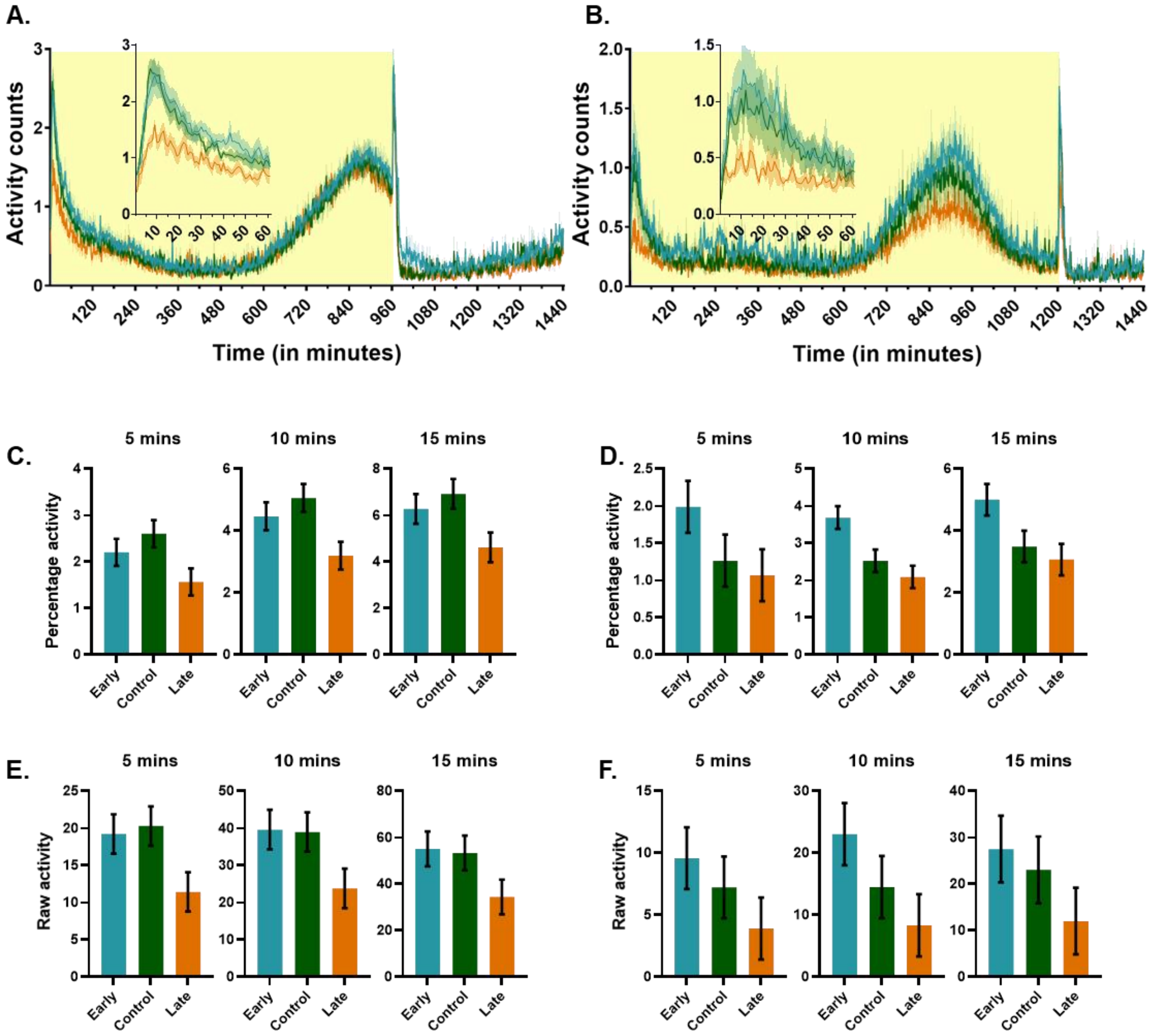


Figure 3.5: Activity profiles and quantification of lights-ON mediated masking responses under high light intensity longer photoperiods. A, B. Activity profiles of *early*, *control* and *late* flies under long photoperiods (***A*** – LD16:08, ***B*** – LD20:04). **C, E.** Quantification of masking by summing total activity within 5, 10, and 15 minutes of lights-ON under LD16:08 (***C*** – percentage activity, ***E*** – raw activity counts). **D, F.** Quantification of masking by summing total activity within 5, 10, and 15 minutes of lights-ON under LD20:04 (***D*** – percentage activity, ***F*** – raw activity counts). *early* and *control* flies showed significantly higher masking responses than *late* flies under both longer photoperiods, additionally, *early* flies showed significantly more masking than both *control* and *late* flies under the longest photoperiod LD20:04. Error bands: ±SEM, error bars: 95% CI from a one-way ANOVA followed by Tukey's HSD.

3.3.3 High light intensity during the day disrupts nighttime sleep amount and quality in extreme chronotypes:

I assayed the sleep of the populations under low light intensity (~70 lux) and high light intensity (~500 lux) LD12:12 regimes (*Fig. 3.6A, B*). Under ~70 lux, there was no difference among populations in total sleep amount, number of sleep bouts, mean bout length, and latency to sleep in the daytime (*Fig. 3.6C, D, G, H; Appendix table A3.38-A3.41*). However, *early* flies slept a significantly lower amount of time than the *control* flies at night (~41.33 mins – *Fig. 3.6F; Appendix table A3.44*). Under ~500 lux, these differences became much more prominent. Both *early* and *late* flies slept significantly lower amount of time than *control* flies at nighttime (~106.4 mins for *early* and ~38.6 mins for *late* – *Fig. 3.6F; Appendix table A3.44*). *early* flies also showed significantly more fragmented nighttime sleep under ~500 lux compared to *control* and *late* flies (~5.54 bouts more than *control* and ~4.45 bouts more than *late* – *Fig. 3.6E; Appendix table A3.42*).

When I compared different sleep parameters between low and high intensity light regimes, I found there were no differences during daytime (*Fig. 3.6C, D, G, H; Appendix table A3.38-A3.41*), but at nighttime, *early* flies showed significantly more fragmented sleep under ~500 lux than ~70 lux (~5.5 bouts higher under ~500 lux – *Fig. 3.6E; Appendix table A3.42*), whereas *control* and *late* flies did not show any such trend. Overall nighttime sleep was significantly lower under ~500 lux than under ~70 lux for all populations (*Fig. 3.6F; Appendix table A3.44*), whereas daytime sleep was not affected (*Fig. 3.6D; Appendix table A3.40*). Mean nighttime sleep bout lengths were lower under ~500 lux than ~70 lux for all populations, and latency to night sleep was higher under ~500 lux than ~70 lux for all the populations, though these differences were not statistically significant (*Fig. 3.6I, J; Appendix table A3.43, A3.45*). I observed no such trend in daytime sleep. This indicates that higher light intensity specifically affects nighttime sleep amount and quality as I saw lesser sleep amount, more fragmented sleep, smaller bout lengths, and higher latency under ~500 lux when compared to ~70 lux. I also observed that higher light intensity affects different chronotypes (*early* and *late* flies) significantly more than the *control* flies, especially for nighttime sleep quality and quantity.

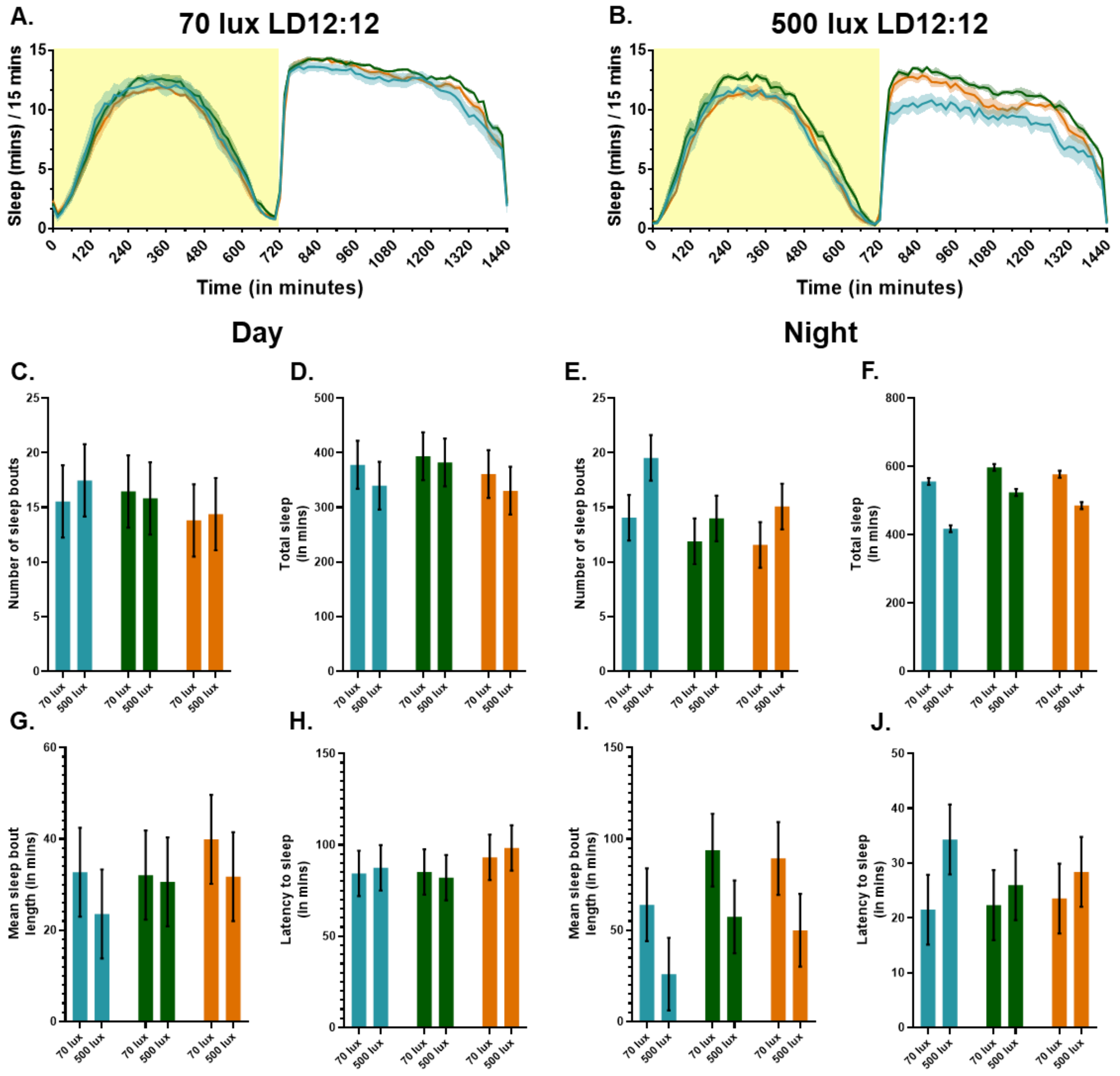


Figure 3.6: Sleep profiles and quantification of different sleep parameters under low and high light intensity LD12:12. A, B. Sleep profiles of *early*, *control* and *late* flies under LD12:12 (***A*** – ~70 lux, ***B*** – ~500 lux). **C, D, G, H.** Quantification of number of sleep bouts (***C***), total sleep (***D***), mean sleep bout length (***G***), and latency to sleep (***H***) in daytime under low and high light intensity LD12:12. **E, F, I, J.** Quantification of number of sleep bouts (***E***), total

sleep (***F***), mean sleep bout length (***I***), and latency to sleep (***J***) in nighttime under low and high light intensity LD12:12. Error bands: ±SEM, error bars: 95% CI from a two-way ANOVA followed by Tukey's HSD.

3.3.4 High light intensity photoperiods negatively affect nighttime sleep amount and quality of *early* and *late* flies:

I assayed sleep under high light intensity photoperiods as different daylengths have been known to change nighttime sleep amount and quality and these changes are known to affect chronotypes adversely (Allebrandt et al., 2013; Gao et al., 2019; Martin et al., 2012). Under shorter photoperiods of LD04:20 and LD08:16, there was visible lower nighttime sleep in the *early* and *late* flies compared to *control* (*Fig. 3.7A, B, C, D*). All daytime sleep parameters – total sleep amount, mean bout length, number of sleep bouts and latency to sleep did not differ among populations in any of the photoperiods (*Fig. 3.8A, C, E, G; Appendix table A3.46-A3.49*).

I noticed an incremental trend in total daytime sleep and daytime number of sleep bouts among different photoperiods for all populations, as expected, since daytime sleep is supposed to increase with increasing day length (*Fig. 3.8A, C; Appendix table A3.46, A3.48*). Similar patterns, but with decremental trends were also observed for nighttime total sleep amount and bout numbers (*Fig. 3.8B, D; Appendix table A3.50, A3.52*) along with decreasing night length from shorter to longer photoperiods. Different photoperiods did not have any effect on mean bout length, either among populations or different regimes (*Fig. 3.8E, F; Appendix table A3.47, A3.51*). However, sleep latency showed an interesting trend – daytime sleep latency was highest under LD12:12 and decreased with either shorter or longer photoperiods, and this decrease in day sleep latency was statistically significant for all populations (*Fig. 3.8G; Appendix table A3.49*). Deviation from a standard LD12:12 regime induced a ~50% decrease in daytime sleep latency. However, nighttime sleep latency was significantly higher for all populations under the shortest photoperiod assayed

(LD04:20) compared to all the other regimes, and nighttime sleep latency decreased with longer photoperiods (*Fig. 3.8H; Appendix table A3.53*).

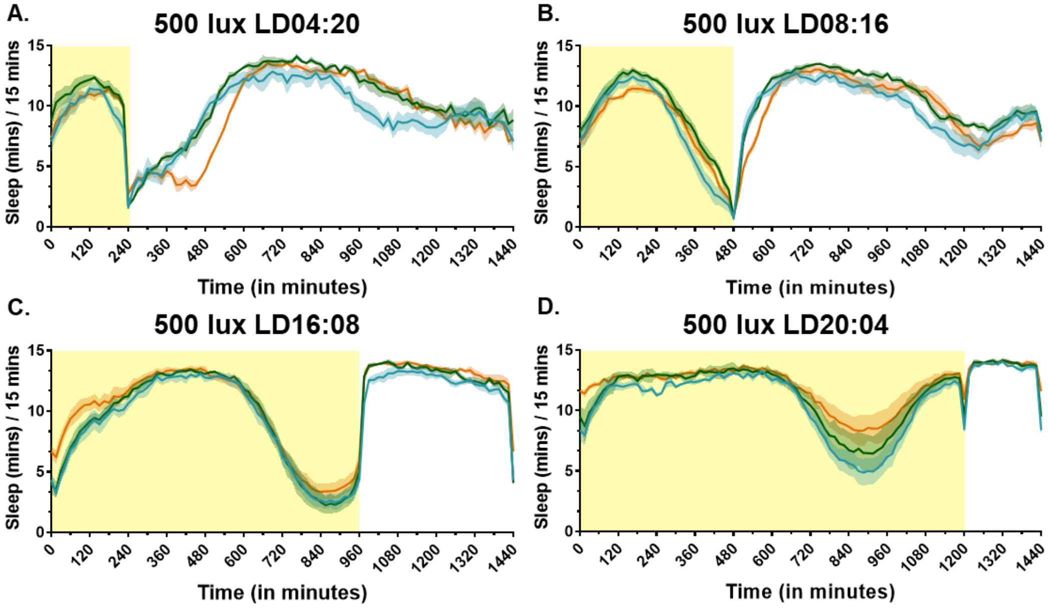


Figure 3.7: Sleep profiles under high light intensity photoperiods. A, B, C, D. Sleep profiles of *early*, *control* and *late* flies under different ~500 lux photoperiods (A – LD04:20, B – LD08:16, C – LD16:08, D – LD20:04). Error bands: ±SEM.

More interestingly, both *early* and *late* flies had significantly lower nighttime sleep under shorter photoperiods compared to *control* flies, *early* flies also showed significantly lower nighttime sleep compared to *control* flies under LD12:12 (*Fig. 3.8B; Appendix table A3.52*). These differences in nighttime sleep among populations were ameliorated under longer photoperiods where flies were exposed to more light at the later parts of the day, and under long photoperiods the daytime sleep was also not different among populations (*Fig. 3.8A, B; Appendix table A3.48, A3.52*). *early* flies also showed significantly more nighttime sleep compared to *control* and *late* flies under LD12:12 and shorter photoperiods (~5-7 more bouts compared to *control – Fig. 3.8D; Appendix table A3.50*). Overall, extreme chronotypes showed significantly lower nighttime sleep and more fragmented sleep compared to *control* flies and these differences were not present under longer photoperiods.

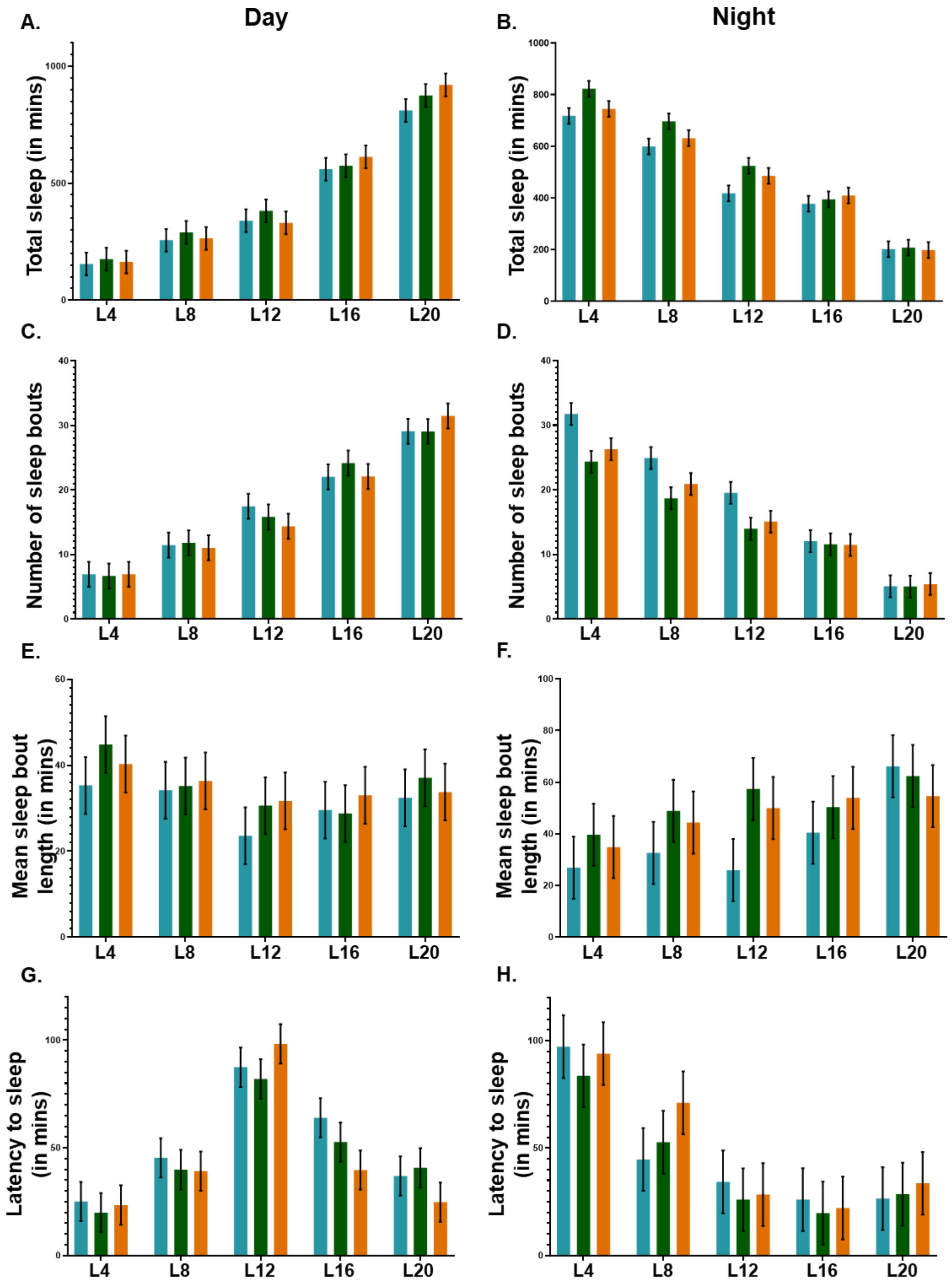


Figure 3.8: Quantification of different sleep parameters under high light intensity photoperiods. A, B, C, D, E, F, G, H. Quantification of number of sleep bouts (***C*** – daytime, ***D*** – nighttime), total sleep (***A*** – daytime, ***B*** – nighttime),

mean sleep bout length (***E*** – daytime, ***F*** – nighttime), and latency to sleep (***G*** – daytime, ***H*** – nighttime). *x*-axis codes: L4 – LD04:20, L8 – LD08:16, L12 – LD12:12, L16 – LD16:08, and L20 – LD20:04. Error bars: 95% CI from a two-way ANOVA followed by Tukey's HSD.

3.4 Discussion:

The role of masking in Drosophila locomotor activity rhythms has been given some attention occasionally since the past three decades (Hamblen-Coyle et al., 1992; Kempinger et al., 2009; Prabhakaran and Sheeba, 2012; Rieger et al., 2003). *Drosophila melanogaster* shows two canonical masking driven peaks in their locomotor activity rhythm under laboratory LD12:12 conditions – one is at dawn (presumably the lights-ON mediated positive masking response), and one at dusk, in the form of suppression of activity immediately after the lights-OFF signal (presumably a negative masking response to the lights-OFF signal). There is no systematic study till date investigating masking responses of the locomotor activity rhythm in divergent chronotypes in *Drosophila melanogaster*, especially shedding light on the evolution of masking in divergent chronotypes. There is also a lack of literature on sleep differences of different chronotypes in *Drosophila melanogaster*, barring one article which addresses sleep differences of artificially selected nocturnal or diurnal strains (Pegoraro et al., 2019) – which can be considered as extremely divergent chronotype flies. Equipped with one of the most appropriate models to address evolution of masking and sleep in divergent chronotypes of *Drosophila melanogaster*, I carried out carefully designed experiments to address the questions stated in the introduction section of this chapter.

3.4.1 Enhanced masking in *early* flies under high intensity photoperiods:

In our populations, low light intensity photoperiods did separate the clock-controlled and the masking peaks, but the differences were not very pronounced (*Fig. 3.1A*), and there was no difference in the masking responses among populations (*Fig. 3.1C*, *E*). Under high intensity LD12:12, neither there was a difference in masking responses among populations, nor did the clock-controlled and masking peaks separate (*Fig. 3.2B, D, F*). When I assayed the populations under high light intensity photoperiods, the clock-controlled and masking peaks separated readily – clock-controlled peaks advanced under shorter photoperiods (*Fig. 3.3A, B*) and delayed under longer photoperiods (*Fig. 3.3D, E*). I observed significantly enhanced masking in *early* and *control* flies compared to the *late* flies only under longer photoperiods of LD16:08 and LD20:04 (*Fig. 3.4, 3.5*), whereas *early* flies also showed significantly higher masking compared to both *control* and *late* flies under the longest photoperiod of LD20:04 (*Fig. 3.5B, D*). This difference in the masking response to the lights-ON signal was dependent on the total daylength experienced by flies. I observed a trend of a higher masking component in the *early* and *control* flies under shorter photoperiods also, though this was not significantly different under shorter photoperiods and LD12:12 but became significant under longer photoperiods – ultimately resulting in highest masking responses in *early* flies under the longest photoperiod.

Drosophila melanogaster is a cosmopolitan species originating from Africa and is now present from Chile and New Zealand in the southern hemisphere to Siberia in the northern hemisphere, where they experience a multitude of temperature differences and photoperiods (Eggleston et al., 1988; Godoy-Herrera et al., 1997; Seebens et al., 2017). Though *Drosophila melanogaster* would not experience a daylength of 4 or 20 hours in nature, these regimes (LD04:20 and LD20:04) can shed light on the circadian clock properties of different chronotypes, and all our populations entrained to these extreme short or long photoperiods. Importantly, the *early* and *late* flies did not differ in their phases under either low or high intensity LD12:12, but under different photoperiods the separation of the clock-controlled and masking peaks demonstrated chronotype differences clearly, particularly under shorter photoperiods where the evening activity of *late* flies was delayed compared to both *early* and *control* flies (*Fig. 3.4A, B*).

The clock-controlled morning peak under different photoperiods also clearly demonstrated the underlying clock differences among the chronotypes – *early* flies show an advanced phase compared to *late* flies and vice versa. These chronotype differences are masked by the high morning peak under LD12:12, where the masking responses take over the clock-controlled peaks. The higher masking responses under long photoperiods suggested that there may be a requirement to start activity in the day part after a significantly shorter period of quiescence/sleep, and the masking response gives that indication to the organism to initiate activity immediately when they day starts. Whereas, under short photoperiods where the night length is greater and flies have a longer period of quiescence/sleep before the lights-ON transition, the circadian clock driven activity starts much before the lights-ON signal, and the organism does not need an indication to initiate activity immediately after lights-ON. Being a predominantly diurnal organism, *Drosophila melanogaster* performs the majority of its locomotor activity under daylight. Under extreme short photoperiods, they start their activity much before dawn (*Fig. 3.4A, B*), and under long photoperiods, they limit all their activity to the day part (*Fig. 3.5A, B*). I observed that their inability to restrict activity to very short durations – under only 4 or 8 hours of daylight was evident by the initiation of their activity much before dawn, and I speculated as activity for the day has already started, they did not show the masking response to the lights-ON signal as it is no more required. However, under long photoperiods, this restriction of performing all their activity within the small amount of daylight was not present as the daylight was present for longer periods of time (16 or 20 hours). Thus, the masking response under short photoperiods was absent, and the major masking related changes were observed under long photoperiods – which might have been needed to indicate the initiation of activity at dawn.

The canonical masking responses are known to be mediated by the compound eyes, and not by cryptochrome, ocelli, nor the H-B eyelets as *cry*^{b}, *cli*^{eya} and *so*^{1} mutants show intact masking peaks (Rieger et al., 2003). Previously it has been reported that in the absence of several photoreceptors, *Drosophila melanogaster* can also show paradoxical masking (decrease in activity after lights-ON and increase in activity after lights-OFF) (Rieger et al., 2003). In my investigations, I did not encounter any paradoxical masking and thus did not attempt to quantify this. In chapter 2, I have shown that *early* flies show significantly stronger positive masking responses to the lights-ON signal compared to *control* and *late* flies in eclosion rhythm, a population level rhythm, and proposed that *early* flies evolved stronger masking for the eclosion rhythm, while also evolving advanced phase of entrainment. Next, I observed that *early* flies also seemed to evolve significantly stronger masking responses in locomotor activity rhythm, albeit detectable only under high intensity longer photoperiods. In a separate study, where I analyzed genomic regions under selection in *early* flies, I found several genes with significantly differentiated SNPs compared to *control* flies – genes involved in light mediated (induced), photoreceptor regulation, phototransduction genes like *nord*, *Arrestin 1*, *reduced ocelli*, *ninaD*, *notch*, *Basigin*, and *Leucyl-tRNA synthetase*. Previous studies from our lab on these populations have shown *early* populations may have evolved heightened light sensitivity; in addition to that, there may be photosensitivity differences in the circadian clock between *early* and *late* populations (Ghosh et al., 2021; Kumar et al., 2007; Vaze et al., 2012a). These differences may explain the heightened light sensitivity of the *early* flies and, in turn, high masking to the lights-ON signal in locomotor activity rhythm.

3.4.2 Extreme chronotypes have lesser and poor-quality sleep under high intensity photoperiods:

While it is known that sleep in humans can be affected by both chronotypes and photoperiods, most of the sleep differences are observed in the late (evening) chronotypes (Allebrandt et al., 2013; Gao et al., 2019; Martin et al., 2012). This may stem from "social jetlag" – the misalignment of social timing and biological timing, that late chronotypes face more than the early (morning) chronotypes in a human society (Wittmann et al., 2006). However, all chronotype differences majorly stem from genetic variations and environmental influences. In our case, *early*, *control*, and *late* flies were kept under similar environmental conditions (under a ~70 lux LD12:12 cycle, 25±0.5° C), and apart from the selection on their timing of emergence, they did not undergo any other selection pressure in their adult lives. As environmental influences were kept constant and the same for all populations, I speculated the differences in their chronotypes stemmed majorly from genetic differences. Also, the *early* and *late* flies did not undergo any form of social jetlag under their selection regime, so the differences in adverse effects, as observed in human chronotypes (which could arise because of asymmetric social jetlag on early and late chronotypes), may not be observed in our populations.

I used high light intensity (~500 lux, higher than their normal maintenance regime of ~70 lux) for most of my experiments in this chapter. This relatively "higher" light intensity was actually comparable to general indoor office lighting (Blume et al., 2019; Spitschan et al., 2016) and much lower than what they would actually experience in nature. When provided with high light intensity in daytime, *early* and *late* flies both showed a significant decrease in nighttime sleep compared to *control* flies, and this decreased nighttime sleep was not observed for *late* flies under low light intensity (~70 lux; *Fig. 3.6B, F*). As different photoperiods are known to affect sleep duration and quality, and particularly shorter daylengths are known to increase nighttime sleep (Blume et al., 2019), I quantified and compared sleep under different photoperiods – ranging from extreme short daylength to extreme long daylength (LD04:20, LD08:16, LD16:08, LD20:04). I observed that *early* and *late* flies sleep significantly less at night under shorter photoperiods of LD04:20 and LD08:16, suggesting extreme chronotypes failed to increase their nighttime sleep under short photoperiods as compared to *control* flies (*Fig. 3.7A, B, 8B*). Additionally, I observed that *early* flies had significantly higher number of night sleep bouts compared to *control* flies, indicating their low sleep quality under shorter photoperiods and LD12:12 (*Fig. 3.8D*). *late* flies also had higher number of night sleep bouts compared to *control* flies under shorter photoperiods, as well as under LD12:12, but this difference was not statistically significant (*Fig. 3.8D*). These differences in nighttime sleep amount and quality in the extreme chronotypes were ameliorated when flies were provided with more light at the later part of the day as evident under longer photoperiods (*Fig. 3.8B, D*). Most of these differences in nighttime sleep under short photoperiods seemed to stem from and can be explained by their advanced or delayed circadian rhythms of locomotor activity. Under short photoperiods, *late* flies have much higher evening activity compared to *control* and *early* flies, which reflects in their sleep profile too – *late* flies sleep less in early night (*Fig. 3.3A, B, 3.7A, B*). *early* flies advance their clock-controlled peak under short photoperiods and start their activity much before dawn, which leads to them sleeping less in the late night (*Fig. 3.3A, B, 3.7A, B*). However, under longer photoperiods, *late* flies slept more, and *early* flies slept less compared to *control* flies in the daytime, but these differences were not statistically significant (*Fig. 3.7C, D, 3.8A*). I conclude that extreme chronotypes fail to increase their nighttime sleep under shorter daylengths. In humans, these sleep differences generally get reported in the late chronotypes; I speculated this is because early chronotypes do not face social jetlag and thus do not show significant differences from their controls.

Overall, I concluded that the clock-controlled and the masking peak around dawn can be well-separated under high light intensity photoperiods. As I observed both *early* and *control* flies showing significantly more positive masking response to the lights-ON signal compared to the *late* flies, it may be possible that the masking response has evolved to be attenuated in the *late* flies. However, I did observe *early* flies show significantly more masking under the longest photoperiod, which also suggested that *early* flies evolved significantly higher masking compared to *control* and *late* flies. I had also shown previously (chapter 2) that *early* flies have evolved significantly higher positive masking response to the lights-ON signal in eclosion rhythm (Ghosh et al., 2021). I also concluded that extreme chronotypes evolved to have significantly lower nighttime sleep under shorter photoperiods, whereas only *early* flies evolved to have more fragmented nighttime sleep, though a trend of more fragmented sleep was observed in *late* flies also when compared to *control* flies. While I observed overall sleep levels decreasing under high intensity LD12:12, compared to low intensity LD12:12, for the extreme chronotypes, the decrease in nighttime sleep was significant and increased with shorter photoperiods (*Fig. 3.8B*). This suggested that extreme chronotypes have evolved to have lower nighttime sleep amount and quality, and these differences can be overcome by exposure to more light in later hours of the day. Without the effects of social jetlag or any other environmental variations, probably all human (or other organisms) chronotypes would show similar kinds of differences. I also concluded that these evolved differences in sleep amount and quality in *early* and *late* flies stem from the difference in their circadian locomotor activity waveforms under different photoperiods.

Future work will need to focus on the following – (i) effect of light intensities and qualities on the sleep of female flies (as a sexual dichotomy is observed in Drosophila sleep phenotypes), (ii) effect of sleep induction and sleep deprivation under regimes where sleep is different in extreme chronotypes, (iii) delineating the effects of circadian clock circuits and sleep circuits governing the defect in sleep in the extreme chronotypes, (iv) devising strategies to mitigate the loss of sleep in the extreme chronotypes, (v) assaying sleep under semi-natural conditions in the presence of multiple zeitgebers.

Chapter 4

Genetic characterization of divergent chronotype populations by pooled sequencing

The work presented in this section has been done in collaboration with Dr. Nikhil KL, who is an alumnus of our group, and is currently a post-doctoral researcher at the Herzog lab, WUSTL, USA. Some data from these experiments had been analyzed and presented by Dr. Abhilash Lakshman in his thesis. However, all the data presented in this thesis have been analyzed, visualized, and interpreted solely by me.

4.1 Introduction:

All living organisms – from cyanobacteria to modern-day humans, exhibit robust 24-hour circadian (*circa* – approximately, *dian* – a day) rhythms in diverse behaviours and physiological processes, and these rhythmic behaviours are controlled by a set of self-sustaining, endogenous circadian clocks. Chronotypes are a result of heritable inter-individual variation in the timing of daily rhythmic behaviours or processes with respect to external time cues. They are genetically encoded but are known to be affected by both genetic and environmental conditions and gene × environment interaction (Roenneberg et al., 2007). Chronotypes essentially reflect differences in the phases of entrainment of the underlying circadian rhythm/clock; therefore, exploring mechanisms driving differential phases of entrainment can help us better understand chronotype regulation. Although the vast majority of studies on chronotypes are in the context of the timing of sleep/wake behaviour in humans, there have been a few studies on other organisms as well (Aschoff and Wever, 1962; Dominoni et al., 2013; Frías-Lasserre et al., 2019; García-Allegue et al., 1999; Helm and Visser, 2010; Ocampo-Garcés et al., 2006; Refinetti et al., 2016; Schwartz and Smale, 2005; Sheeba et al., 2001; Stuber et al., 2015; Vivanco et al., 2009). Among humans, while some individuals have an inherent propensity to wake up and sleep earlier in the day; referred to as 'early' chronotypes, 'late' chronotypes constitute individuals who inherently wake up and sleep later (Randler et al., 2017). Various studies have suggested that chronotypes are driven by differences in circadian clocks, which not only drive rhythms in sleep/wake but also in other aspects of behaviour and physiology.

Chronotypes have attracted considerable attention in the recent past, with various studies reporting the association of chronotype differences with a myriad of psychological, metabolic, and physiological dysregulations (Bullock, 2019; Kivelä et al., 2018; Koren et al., 2016; Manfredini et al., 2018; Roenneberg and Merrow, 2016; Roenneberg et al., 2019) thus highlighting the importance of understanding the functional underpinnings of chronotype regulation. Various studies have suggested chronotypes may affect psychological wellbeing – late chronotypes have reported more frequent psychological and psychosomatic problems than early chronotypes (Giannotti et al., 2002; Mecacci and Rocchetti, 1998; Wittmann et al., 2006). Daytime tiredness and low night-time sleep quality had also been associated with late chronotypes more than early chronotypes (Giannotti et al., 2002; Taillard et al., 2003; Volk et al., 1994; Wittmann et al., 2006). This disparity between chronotypes majorly stems from the differential interaction of the biological clock and the "social clock" in different chronotypes and can lead to severe "social jetlag" depending on the chronotype of the individual. This "social jetlag" can be explained as a chronic form of jetlag where the biological clock and social clock are not in sync, e.g., in late chronotypes, the late sleep onset time may be regulated by their biological clock, but they also have to wake up earlier because of the social clock, and thus on work days, late chronotypes may accumulate more sleep debt than early chronotypes, who can somewhat better synchronize their biological clock with the social clock (Wittmann et al., 2006). Repercussions of this were observed in a recent study where the authors found early chronotype students outperformed late chronotypes in all subjects in school, especially mathematics, when school timing was in the morning shift, while late chronotype students performed better in evening classes (Goldin et al., 2020). In another study, where the Seattle school district delayed their secondary school start time by ~1 hour, there was an increase in median sleep duration by 34 minutes, and median grades of the students increased by 4.5% (Dunster et al., 2018). Several studies exploring the genetic basis of chronotypes have reported varying degrees of heritability (~50% in the USA, UK, Scandinavia, Brazil, 14% in Hutterites, and 23% in Amazonians, 12-19.4% in UK BioBank data) across human populations (Aguiar et al., 1991; Hur et al., 1998; Jones et al., 2016, 2019; Klei et al., 2005; Koskenvuo et al., 2007; Lane et al., 2016; Von Schantz et al., 2015) while others have reported association of chronotypes with various polymorphisms in clocks genes such as *CLOCK* (Katzenberg et al., 1998; Mishima et al., 2005), *PER1-3* (Archer et al., 2003; Carpen et al., 2005, 2006; Ebisawa et al., 2001; Pereira et al., 2005), and *ARNTL2* (Parsons et al., 2014). However, some of these reports remain inconclusive as other studies either failed to replicate these findings or observed conflicting or no association of these polymorphisms with chronotypes (An et al., 2014; Barclay et al., 2011; Drake et al., 2015; Iwase et al., 2002; Kunorozva et al., 2012; Osland et al., 2011; Pedrazzoli et al., 2007; Perea et al., 2014; Pereira et al., 2005; Robilliard et al., 2002; Viola et al., 2007), possibly due to differences in statistical powers of detection because of variation in size and genetic backgrounds of sampled populations. In recent years, genome wide association studies on human cohorts have identified multiple circadian clock genes associated with sleep chronotypes, and also several other genes known to be involved in neuronal signalling, sleep homeostasis, and light input pathways to the clock (Gottlieb et al., 2007; Hu et al., 2016; Jones et al., 2016, 2019; Lane et al., 2016). Interestingly, some of the identified genes have also been implicated with similar functions in mice and Drosophila, thus suggesting that the genetic architecture underling chronotype variation may partly be conserved across organisms. Although genome wide association study (GWAS) is a powerful strategy to identify genes that are unlikely to be identified by other approaches, the list of candidate genes may vary across studies for multiple reasons (Kalmbach et al., 2017); thus necessitating further validation of the candidates, which is often not possible in humans and requires a well-established model system.

In this respect, experimental evolution in laboratory populations followed by whole-genome sequencing, commonly called “Evolve and Resequence (E&R)” (Turner et al., 2011), is an attractive, cost-effective alternative to individual sequencing for investigating the genetic basis of a selected trait (Kawecki et al., 2012). E&R has been applied in Drosophila with varying degrees of success to investigate the genetic basis of bristle development (Cassidy et al., 2013), longevity and aging (Burke et al., 2010; Remolina et al., 2012), adaptation to novel environments (Orozco-Terwengel et al., 2012), hypoxia tolerance (Zhou et al., 2011), body size (Turner et al., 2011), temperature (Tobler et al., 2014), courtship song (Turner and Miller, 2012), and diet (Reed et al., 2014). Genome wide sequencing of genetic variation present in experimentally evolving sexually reproducing populations after many generations shed light on the relative importance of selective sweeps, particularly alleles being driven to fixation (Phillips et al., 2016). Although successive selective sweeps, along with continuous hitchhiking, could purge genetic variation, no widespread purging of genetic variation was detected in laboratory evolved fly populations even after a few decades (Burke et al., 2010). Pooled-sequencing has been used to identify genotype-phenotype correlation and mechanisms of evolution successfully in *Drosophila* in recent years (Barghi et al., 2019; Burke et al., 2016; Fabian et al., 2012; Graves et al., 2017; Mueller et al., 2018; Phillips et al., 2016, 2018; Rose et al., 2014; Tobler et al., 2014). With recent advances in sophisticated algorithms, the processing power of computers and mass scale high-depth sequencing techniques, pooled-sequencing large number of individuals from a population in a high coverage (50X-100X) depth gives reliable variant identification (Burke et al., 2010, 2014; Cutler and Jensen, 2010; Futschik and Schlötterer, 2010; Phillips et al., 2016).

To this end, we adopted a laboratory selection approach (Garland and Rose, 2009). Using 4 independent populations of *Drosophila melanogaster*, each selected for exhibiting *early* or *late* emergence chronotypes, we examined the end product of the evolution of clock properties in divergent chronotypes and the natural genetic correlations of chronotype evolution (Kumar et al., 2007). The advantages of using such an approach to study the evolution of traits, especially those related to circadian clocks, have been laid out in a recent review (Abhilash and Sharma, 2016). Over the course of ~20 years (~350 generations), we have reported that *early* and *late* emergence chronotypes in these populations are associated with differences in circadian period (Kumar et al., 2007), zeitgeber sensitivity (Ghosh et al., 2021; Nikhil et al., 2014; Vaze et al., 2012c, 2012a), amplitude, coupling, phase and period responses (Nikhil et al., 2016c), temperature sensitivity (Abhilash et al., 2019, 2020) and also molecular clocks (Nikhil et al., 2016b). These are in accordance with studies on other organisms, including humans reporting that chronotypes are a complex trait that may stem from differences in various clock properties and their interaction with zeitgebers (Aschoff and Pohl, 1978; Duffy and Wright, 2005; Duffy et al., 2001; Kerkhof and Van Dongen, 1996; Lehmann et al., 2012; Vivanco et al., 2010).

Having established a well-characterized model of chronotypes that can serve as a system for further molecular-genetic studies, I sequenced the genomes of the *early* and *late Drosophila melanogaster* populations by pooling genomic DNA from ~500 individuals from each population to identify putative loci that are likely to be associated with entrainment-phase/chronotype differences. Here I present results from pooled sequencing of our *early*, *control*, and *late* populations. I performed various population genomic analyses and identified genomic regions undergoing positive selection in *early* and *late* populations compared to the *control* populations. I also identified various SNPs significantly enriched in either *early* or *late* populations, compared to the *control* populations. I further predicted different pathways plausibly under differential selection in *early* and *late* populations. The four genetically independent replicate populations of each *early*, *control*, and *late* populations in my experimental design make it uniquely suited to understanding the genomic basis of chronotype evolution in *Drosophila melanogaster*.

4.2 Materials and methods:

4.2.1 Sample preparation and sequencing:

Approximately 250 male and 250 female flies (4-5 days old) were randomly sampled from each population (total 12 samples with *early*$_{1-4}$, *control*$_{1-4}$, and *late*$_{1-4}$) at generation 260. Total genomic DNA was extracted from them with the DNeasy Blood and Tissue Kit (#69504, QIAGEN, MD, USA). These DNA samples were subjected to quality check (Nanodrop QC and Qubit QC), and vast paired-end libraries of each sample were prepared for sequencing with tags (Illumina HiSeq manufacturer's protocol). The Illumina HiSeq generated 150 paired-end reads (~80-100 million per sample), which were quality checked using FastQC (Andrews, 2010). These raw reads were processed by Cutadapt (Martin, 2011) to remove adapters and for low-quality base trimming. A ~100X depth of coverage was aimed for as this was a pooled-sequencing experiment for reliable variant identification (Futschik and Schlötterer, 2010). The reads from all populations were then aligned to the *Drosophila melanogaster* reference genome (Reference genome: BDGP6) using bowtie2 (Langmead et al., 2013), which is a variant of the well-known Burrows-Wheeler Aligner (BWA) algorithm (Li and Durbin, 2010) using default parameters.

4.2.2 Variant identification and filtering:

SNPs and INDELs were identified using SAMtools1.2 and BCFtools1.2 (Li et al., 2009). Potential variants were identified using a read depth threshold $\geq$ 20 and a mapping quality threshold $\geq$ 30. mpileup files were generated using SAMtools1.2 from the bam files. The mpileup files were further processed using VarScan2 (Koboldt et al., 2009) for the identification of line-specific markers between the samples in a group. Next, VCF files were created for each of the 12 samples with BCFtools1.2. For downstream analyses, both VCF files and mpileup files were used.

I only analyzed SNPs because they were more numerous and partly because INDEL analyses are not standardized for pooled sequencing studies.

4.2.3 Calculation of Tajima's Pi (π), *D*, and Waterson's Theta (θ_w):

Population genomic parameters (Tajima's Pi (π), *D*, and Waterson's Theta (θ_w)) were calculated using popoolation_1.2.2 (Kofler et al., 2011a) and the mpileup files for each sample. π is a measure of DNA sequence variation based on the average pairwise distance between all samples, and θ_w is a measure of DNA sequence variation based on the observed number of segregating sites (polymorphic sites) and number of chromosomes per sample. *D* is a statistic that measures the difference between estimators, θ_w (expected heterozygosity) and π (observed heterozygosity) (Biswas and Akey, 2006; Tajima, 1989). Positive values of *D* arise from an excess of intermediate frequency alleles (higher heterozygosity than expected) and can result from population bottlenecks, structure and/or balancing selection. Negative values of *D* indicate an excess of low frequency alleles (lower heterozygosity than expected) and can result from population expansions or positive selection (Biswas and Akey, 2006). Discovery of regions with significantly low *D* was determined using the Two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, with Q = 5% using GraphPad Prism 8 (13213 independent t-tests for each comparison). Since π, *D*, and θ_w are sensitive to variation in coverage, reads were subsampled to a uniform coverage of 20 to remove any bias in calculations before I calculated π, *D*, and θ_w for 50Kb windows with a 10Kb step size using the *Variance-sliding.pl* script (corrected for pooled-sequencing) (Kofler et al., 2011a).

4.2.4 Allele frequency-based differentiation tests:

I used the sorted bam files to create population-wise mpileup files – one mpileup file contained information from *control*$_{1\text{-}4}$ and *early*$_{1\text{-}4}$, and another mpileup file contained information from *control*$_{1\text{-}4}$ and *late*$_{1\text{-}4}$. These two mpileup files were further converted into Popoolation2 compatible SYNC files using a JavaScript (*mplieup2sync.jar*), keeping a minimum quality of 30. These two SYNC files were then used to calculate pairwise F_{ST} values between each SNP of matched populations (*control*$_1$ vs. *early*$_1$, *control*$_1$ vs. *late*$_1$, *control*$_2$ vs. *early*$_2$, *control*$_2$ vs. *late*$_2$ and so on). F_{ST} is a statistic that quantifies levels of differentiation between subpopulations and can be used to identify SNPs whose allele frequency varies significantly between two subpopulations (Cormack et al., 1990). I applied a Z-transformation on the F_{ST} values from four different replicates of *control-early* and *control-late* comparisons and used a 5% threshold using the *quantile* function in R (version 4.0.2) to identify significant SNPs (R Core Team, 2020). All significantly differentiated SNPs common among all four replicates for each pairwise comparison were considered for further analyses (686 SNPs in *early* and 79 SNPs in *late* populations). Only SNPs from all replicates were considered for further analyses because SNPs changing in allele frequency in all four replicates are much more likely to be due to the similar selection pressure they experience than just by genetic drift, which might have been acting differentially in different replicate populations. Next, I performed a Cochran-Mantel-Haenszel (CMH) test for repeated tests of independence for each SNP. The CMH test is used for detecting significant and consistent changes in allele frequencies when independent measurements of allele frequencies are available, as in my case, where I can treat the replicate populations as genetically independent samples. CMH test was implemented by the *cmh-test.pl* code with *min-count* 2 per population, *min-coverage* of 30, and *max-coverage* of 400 between *control-early* and *control-late* (Kofler et al., 2011b). The traditional CMH tests are known to be less conservative and may give rise to numerous false positives. I took a permutation-based approach to determine significance thresholds for respective CMH tests (Graves et al., 2017). I used permutations to generate null distributions and then set a genome-wide significance threshold based on the smallest p values in that distribution (1000 permutations, 5% threshold). Using this strict threshold, I identified 333 SNPs in *early* populations ($\mathrm{CMH}p_{corrected} = 1.396817^{-58}$) and only 3 SNPs in *late* populations ($\mathrm{CMH}p_{corrected} = 4.556171^{-55}$), where allele frequencies changed significantly from *control* populations. Therefore, I lowered the threshold for *late* populations ($\mathrm{CMH}p_{corrected} = 3.046445^{-27}$) to include SNP numbers similar to that of the *early* populations by choosing the top 333 SNPs with the lowest CMHp value. Manhattan plots of results from pairwise F_{ST} comparisons and CMH tests were plotted using the CMplot (LiLin-Yin, 2020) package in R (version 4.0.2) (R Core Team, 2020).

4.2.5 Identification of genome-wide selective sweeps:

I took a more direct approach to identify genomic regions that have likely undergone selective sweeps in the *early*, *control*, and *late* populations using *Pool-hmm* (Boitard et al., 2013). I ran *Pool-hmm* for each chromosomal arm separately with respective populations size for pooling individuals with k value of 1^{-10} and θ value as calculated before for each sample (θ_w). Further data processing was done using the data.table package (Dowle and Srinivasan, 2020) and plotted using the karyoploteR package (Gel, 2020; Gel and Serra, 2017) in R (version 4.0.2) (R Core Team, 2020).

4.2.6 Analysis of identified variants:

After identifying the significantly differentiated variants in the *control-early* and *control-late* comparisons, I segregated locations common to both *early* and *late* populations which have diverged from the *controls*. Thus, I ended up with three exclusive groups of SNPs, *early* exclusive, *late* exclusive, and *early-late* overlap. I was mostly interested in the *early* and *late* exclusive group of variants and proceeded to annotate them using SnpEff (version 4.3t) variant effect predictor (Cingolani et al., 2012). SNPs were categorized based on estimation of their putative impact/deleteriousness (low, moderate, high, and modifier) and their effect type-description of consequence (see results section). Gene ontology (GO) analysis and pathway enrichment analysis was carried out with the identified genes from SnpEff using g:Profiler (Reimand et al., 2016). Further manual categorization of different GO terms was done. To identify statistical enrichment of overlap between significantly differentiated SNPs harboring genes among different sets (Tajima's *D* comparison, pairwise FST comparison, and CMH tests), I used the SuperExactTest package (Wang et al., 2022) in R (version 4.0.2) (R Core Team, 2020). Protein-protein interaction networks were constructed using STRING (https://string-db.org/) (Szklarczyk et al., 2021), and later analyzed and visualized using StringApp (Doncheva et al., 2019) in Cytoscape 3.7.2 (Shannon et al., 2003). Tajima's *D*, π, θ_w, KEGG enrichment analysis, and GO enrichment analysis results were plotted with ggplot2 package (Wickham, 2016; Wickham et al., 2020) in R (version 4.0.2) (R Core Team, 2020).

4.3 Results:

I performed a pooled sequencing of all 12 populations and analyzed variants (SNPs) to identify genomic regions under differential selection in *early* and *late* populations, compared to *control* populations. The main goal of this study was to reliably identify SNPs that were significantly different in terms of allele frequency and identify the genes where such SNPs were enriched, thus establishing genes and pathways which have been under selection in divergent chronotype populations.

4.3.1 Identification of SNPs and estimating population genomic parameters:

I identified a total of 1.37 million SNPs spanned over 17477 genes which were common among all four replicates of *early* populations, and 1.43 million SNPs spanned over 17515 genes in *late* populations when compared to the *control* populations, satisfying my quality control thresholds in the five major chromosomal arms (2L, 2R, 3L, 3R, and X), and these SNPs were used to prepare the SYNC files for further analysis (*Fig. 4.1A & B*). I ignored SNPs that passed the strict quality control in some but not all four replicates as I wanted to ignore allele frequency changes due to forces (e.g., genetic drift) other than selection, and only when a particular SNP passed quality control in all four replicates, it was considered for further analyses. These SNPs were distributed all over the five major chromosomal arms and did not show any discernible pattern of distribution or presence of bias in SNP calling (*Fig. 4.1A & B*). However, the SNP density patterns across chromosomal arms were similar in the case of both *early* and *late* populations, and in both populations, the X chromosome had a lower number of SNPs compared to the other arms (*Fig. 4.1A & B*). There were many more variants identified in the mpileup files compared to the 1.37 and 1.43 million SNPs used for *early* and *late* populations, respectively, but while using the mpileup files for different population genomic parameter estimations (π, D, and θ_w), I kept similar quality control of SNPs, so that similar number of SNPs are used for these estimations.

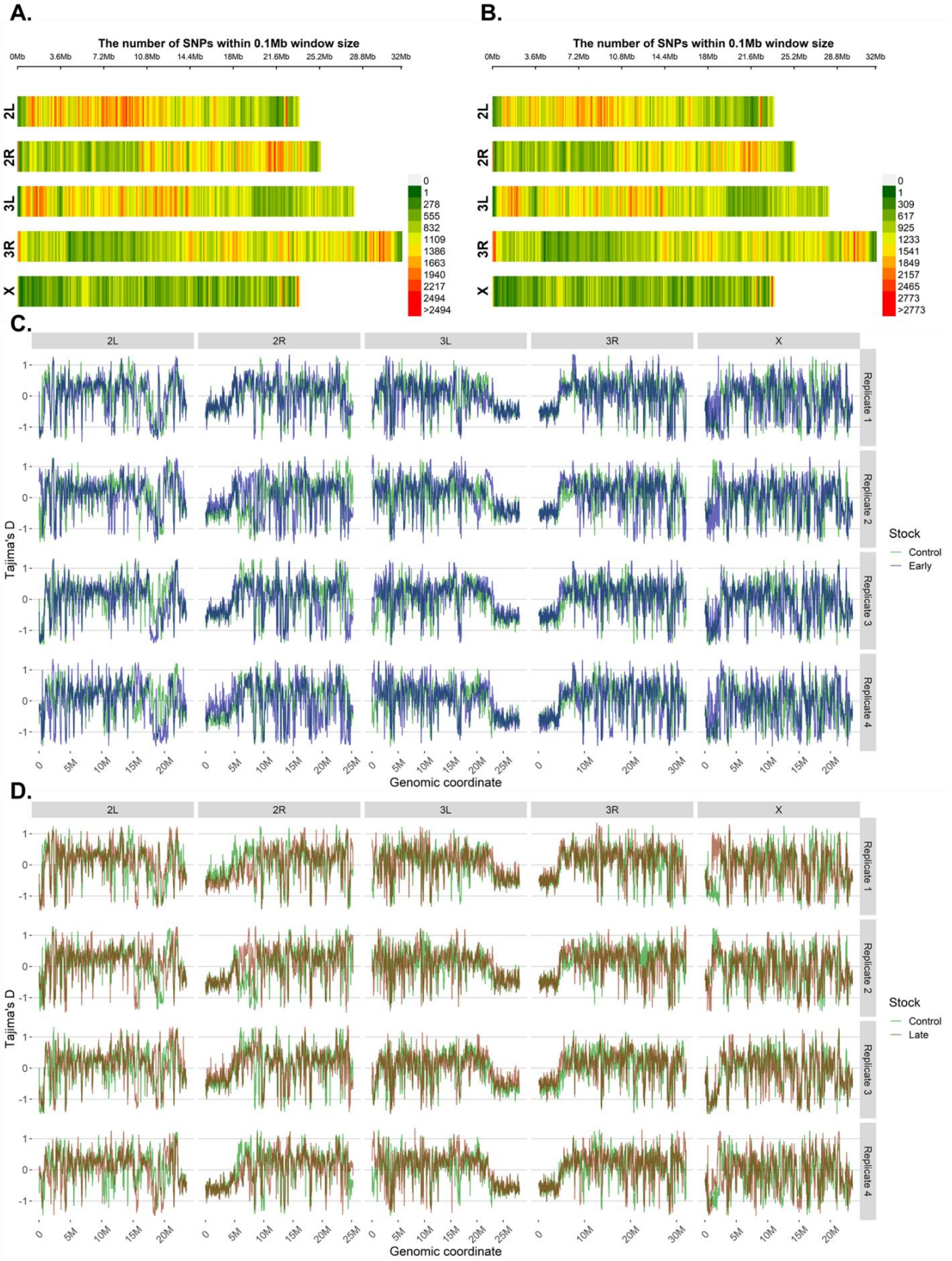

A.
The number of SNPs within 0.1Mb window size
0Mb
3.6Mb
7.2Mb
10.8Mb
14.4Mb
18Mb
21.6Mb
25.2Mb
28.8Mb
32Mb
2L
2R
3L
3R
X
0
1
278
555
832
1109
1386
1663
1940
2217
2494
>2494
B.
The number of SNPs within 0.1Mb window size
0Mb
3.6Mb
7.2Mb
10.8Mb
14.4Mb
18Mb
21.6Mb
25.2Mb
28.8Mb
32Mb
2L
2R
3L
3R
X
0
1
309
617
925
1233
1541
1849
2157
2465
2773
>2773
C.
2L
2R
3L
3R
X
Replicate 1
Replicate 2
Replicate 3
Replicate 4
Tajima's D
1
0
-1
Stock
Control
Early
0
5M
10M
15M
20M
25M
30M
Genomic coordinate
D.
2L
2R
3L
3R
X
Replicate 1
Replicate 2
Replicate 3
Replicate 4
Tajima's D
1
0
-1
Stock
Control
Late
0
5M
10M
15M
20M
25M
30M
Genomic coordinate

Figure 4.1: Description of genome-wide SNPs and Tajima's D values. **(A & B)** SNP density (passing quality control in all four replicate populations) along different chromosomal arms of *early* **(A)** and *late* **(B)** populations. These 1.37 and 1.43 million SNPs have been used for all downstream analyses. In both cases, a similar number of SNPs have been identified in *control* population but have not been shown. **(C & D)** Tajima's *D* value across 50Kb windows with 10Kb steps along different chromosomal arms (axes labels on top) and four replicate populations (axes labels on right) of *control*, *early* **(C)**, and *late* **(D)** populations. 159 regions in *early* and 87 regions in *late* populations showed significantly lower negative *D* values compared to *control* population and were identified to be under positive selection. Median genome-wide *D* values for the populations are – *control*: 0.121, *early*: 0.114, and *late*: 0.138. Line colors: blue – *early*, green – *control*, red – *late*. In **(A & B)** the SNP densities were calculated in 0.1Mb non-overlapping windows in five major chromosomal arms; the color guide is on the right.

I calculated Tajima's *D* (*D*) for 50Kb windows with a step size of 10Kb to identify loci under selection in *early*, *control*, and *late* populations (*Fig. 4.1C & D*). The goal of this analysis was to identify genomic regions where *early* and *late* populations have significantly lower negative *D* values compared to the *control* population, as those regions are most likely to be under significantly more positive selection. All populations showed overall high levels of heterozygosity and standing genetic variations (measured by *D* and π), as expected from large, outbreeding, random mating populations such as these, but overall *D* values were higher than were previously reported for different natural populations (Croze et al., 2017; Fabian et al., 2012). Median genome-wide *D* value of the respective populations were – *control*: 0.121 (range: -1.333 to 1.233), *early*: 0.114 (range: -1.393 to 1.192), and *late*: 0.138 (range: -1.39 to 1.17). I identified 159 regions of 10Kb intervals (total 1.59 Mb region) where *early* populations have significantly lower *D* values than the *control* populations, showing significantly stronger positive selection in these regions in *early* populations than that of *control* populations (*Fig. 4.1C*). Similarly, I identified 87 regions of 10Kb intervals (total 0.87 Mb) in *late* populations, where they have undergone significantly higher positive selection than *control* populations (*Fig. 4.1D*). In the *early* populations, these regions identified to be under positive selection harbor 296 genes (including pseudogenes, protein-coding genes, miRNA, snoRNA, lncRNA, tRNA, and antisense lncRNA genes). Some noteworthy genes are – *nord* (blue light induced gene; (Hall et al., 2018)), *Arrestin 1* (phototransduction; (Damulewicz et al., 2019)), *nervy* (regulates PDF expression; (Duvall and

Taghert, 2012)), *Thor* (contributes to circadian functions; (Nagoshi et al., 2010)), *reduced ocelli* (ocelli and photoreceptor formation; (Caldwell et al., 2007; Vosshall and Young, 1995)), *pickpocket 29* (neurotransmission; (Hill et al., 2017)), *mAChR-A* (regulates light regulate calcium response in LNVs; (Qin et al., 2019)), *spaghetti* (regulator of DBT; (Means et al., 2015)), *Leucyl-tRNA synthetase* (blue light induced gene; (Hall et al., 2018)), *dunce* (mutants show augmented light-induced phase delay and shortened period; (Levine et al., 1994)), *ninaD* (rhodopsin biosynthesis; (Gu et al., 2004)), and *timeless* (core circadian clock gene) (summarized in *Table 4.1*). In the *late* populations, the regions identified to be under positive selection harbor 112 genes. Among these, some directly or indirectly related to circadian rhythms are – *Blimp-1* (pupation timing, fat body timer; (Akagi et al., 2016)), *Rootletin* (sensory perception of sound and touch; (Chen et al., 2015)), *Tsc1* (keeps circadian period close to 24 hours and maintains robust circadian rhythms; (Zheng and Sehgal, 2010)), *bruchpilot* (adult locomotory behavior; (Wagh et al., 2006)), and *spaghetti* (regulator of DBT; (Means et al., 2015)) (summarized in *Table 4.1*). I also calculated π (average pairwise observed heterozygosity) in 50Kb windows with a step size of 10Kb along all five major chromosomal arms of all populations (*Appendix Fig. A4.9A & B*). Median π values were 0.00262 (range: 0-0.01; SD: 0.0018) for *control*, 0.00242 (range: 0-0.12; SD: 0.001772) for *early*, and 0.002641 (range: 0-0.011; SD: 0.001798) for *late* populations, reaffirming the high amount of heterozygosity and standing genetic variation within each population.

Table 4.1: Genes with direct or non-direct circadian functions with significantly lower negative Tajima's D value in *early* and *late* populations, compared to *control* population.

Population	Gene name	Function	Reference
early	*nord*	blue light induced gene	Hall et al., 2018
	Arrestin 1	phototransduction	Damulewicz et al., 2019
	nervy	regulates PDF expression	Duvall and Taghert, 2012
	Thor	Contributes to circadian functions	Nagoshi et al., 2010
	reduced ocelli	ocelli and photoreceptor formation	Caldwell et al., 2007; Vosshall and Young, 1995
	pickpocket 29	neurotransmission	Hill et al., 2017
	mAChR-A	regulates light regulate calcium response in LNVs	Qin et al., 2019
	spaghetti	regulator of DBT	Means et al., 2015
	Leucyl-tRNA synthetase	blue light induced gene	Hall et al., 2018
	dunce	mutants show augmented light-induced phase delay and shortened period	Levine et al., 1994
	ninaD	rhodopsin biosynthesis	Gu et al., 2004
	timeless	core circadian clock gene	
late	*Blimp-1*	pupation timing, fat body timer	Akagi et al., 2016
	Rootletin	sensory perception of sound and touch	Chen et al., 2015
	Tsc1	keeps circadian period close to 24 hours and maintains robust circadian rhythms	Zheng and Sehgal, 2010
	spaghetti	regulator of DBT	Means et al., 2015

4.3.2 Identification and characterization of significant SNPs by quantifying differential allele frequency changes between populations:

Next, I investigated the patterns of genetic differentiation in terms of allele frequency among our populations. The goal of this investigation was to identify SNPs that were different in terms of their allele frequencies in *early* and *late* populations compared to the *control* population. To this end, I calculated pairwise F_{ST} between all SNPs in *control-early* and *control-late* groups. In four replicates, median autosomal F_{ST} values in *early* population varied from 0.015 to 0.024 (average 0.0474), and in X chromosome, F_{ST} varied from 0.017 to 0.023 (average 0.055). In case of *late* populations, median autosomal F_{ST} values among four replicates varied from 0.013 to 0.016 (average 0.036), and in X chromosome it varied from 0.013 to 0.017 (average 0.041) (*Fig. 4.2A*). Owing to its smaller effective population size, the X chromosome has higher average F_{ST} values than the autosomes, as shown previously (Hutter et al., 2007; Kapun et al., 2020), and that is reflected in our populations. Overall differentiation between populations (*early* and *late*, both compared to *control*) varied widely, reaching a range of 0 to 1 in replicate 2 of *early* populations, while in the other three replicates, it varied from 0 to 0.95, 0.88, and 0.93. Whereas, in *late* populations, in four replicates, F_{ST} values varied from 0 to 0.67, 0.78, 0.64, and 0.83 (*Fig. 4.2B*). Next, I focused on loci where F_{ST} values in all four replicates show high values. I achieved this by calculating zF_{ST} values (a Z-transformation on the pairwise F_{ST} values of *control-early* and *control-late* comparisons) and using a 95-percentile threshold for each replicate, and choosing only the same loci where zF_{ST} values pass the threshold in all four replicates (*Fig. 4.2A & B*). I identified 686 such SNPs in *early* populations ($maxF_{ST}$ 1, $minF_{ST}$ 0.15, $maxzF_{ST}$ 14.45, $minzF_{ST}$ 1.98) and 79 SNPs in *late* populations ($maxF_{ST}$ 0.65, $minF_{ST}$ 0.12, $maxzF_{ST}$ 10.12, $minzF_{ST}$ 2.06) where allele frequencies have changed significantly from *control* populations in all four replicates. Though there is no consensus about cutoffs of F_{ST} values for significant differentiation, Hartl and Clark (Hartl et al., 1997) proposed a simple classification system where F_{ST} <0.05 : "little genetic differentiation", 0.05-0.15 : "moderate genetic differentiation", 0.15-0.25 : "means great genetic differentiation", and >0.25 : "very great genetic differentiation". While I adhere to our zF_{ST} based 95-percentile thresholds, it is important to note most of my identified significant SNPs have F_{ST} values greater than 0.15 (*early* average 0.361; range: 0.15 to 1, *late* average 0.229; range: 0.12 to 0.65), which can be categorized as "great to very great" genetic differentiation. Among the 686 SNPs identified in the *early* populations spanning over 127 genes, SnpEff predicted 112 to have an estimated "low" effect of putative impact/deleteriousness, 15 "moderate" effect SNPs, and 559 "modifier" effect SNPs (summarized in *Appendix Table A4.1*). 293 SNPs were categorized as "upstream_gene_variant", 104 were "synonymous_variant", 98 were "intron_variant", 71 were "downstream_gene_variant", and only 15 were categorized as "missense_variant – moderate effect" as per effect type description of consequence (*Fig. 4.2C & 4.3A$_{inset}$* and summarized in *Appendix Table A4.1*). These SNPs were present in genes like *kirre* (sleep traits; (Harbison et al., 2017)), *vrille* (core clock component; (Panda et al., 2002)), *notch* (compound eye development; (Kahali et al., 2009)), *timeless* (core clock component), *sgg* (modulates periodicity of the circadian clock; (Martinek et al., 2001)), *Egfr* (consolidation and maintenance of sleep; (Foltenyi et al., 2007)), *Ir64a* (odorant receptor; (Ai et al., 2013)), *nompC* (response to auditory stimulus and startle response; (Boyd-Gibbins et al., 2021; Zhang et al., 2013)), *dpr8* (synapse organization; (Carrillo et al., 2015)), *Nlg4* (sleep, expressed in clock neurons; (Li et al., 2013)), *Fbxl4* (negative regulation of circadian sleep/wake cycle; (Li et al., 2017)), *bruchpilot* (adult locomotory behavior; (Wagh et al., 2006)), *Basigin* (photoreceptor cell morphogenesis; (Munro et al., 2010)), *CCAP-R* (neuropeptide signaling pathway, ecdysis; (Cazzamali et al., 2003)), *Eip93F* (adult determinant during fly morphogenesis; FlyBase), *Tyramine β hydroxylase* (sleep; (Crocker and Sehgal, 2008)), and *Thor* (contributes to circadian functions; (Nagoshi et al., 2010)) (*Fig. 4.3A and* summarized in *Table 4.2*). Among the 79 SNPs identified in the *late* populations spanning over 47 genes, SnpEff predicted 9 to have an estimated "low" effect of putative impact/deleteriousness, 2 "moderate" effect SNPs, and 68 "modifier" effect SNPs (summarized in *Appendix Table A4.2*). 18 SNPs were categorized as "upstream_gene_variant", 9 were "synonymous_variant", 19 were "intron_variant", 20 were "downstream_gene_variant", and only 2 were categorized as "missense_variant – moderate effect" as per effect type description of consequence (*Fig. 4.2D & 4.3B$_{inset}$* and summarized in *Appendix Table A4.2*). These SNPs were present in genes like *Heterochromatin protein 5* (gene silencing by heterochromatin; FlyBase), *unplugged* (brain development; (Hirth et al., 2003)), *Cyp4p1, Cyp4p2, Cyp4p3* (metabolism of insect hormones;

FlyBase), *Netrin-A* (compound eye development; (Akin and Zipursky, 2016)), *Rootletin* (sensory perception of sound and touch; (Chen et al., 2015)), and *Sirt4* (determination of adult lifespan; (Wood et al., 2018)) (*Fig. 4.3B* and summarized in *Table 4.2*).

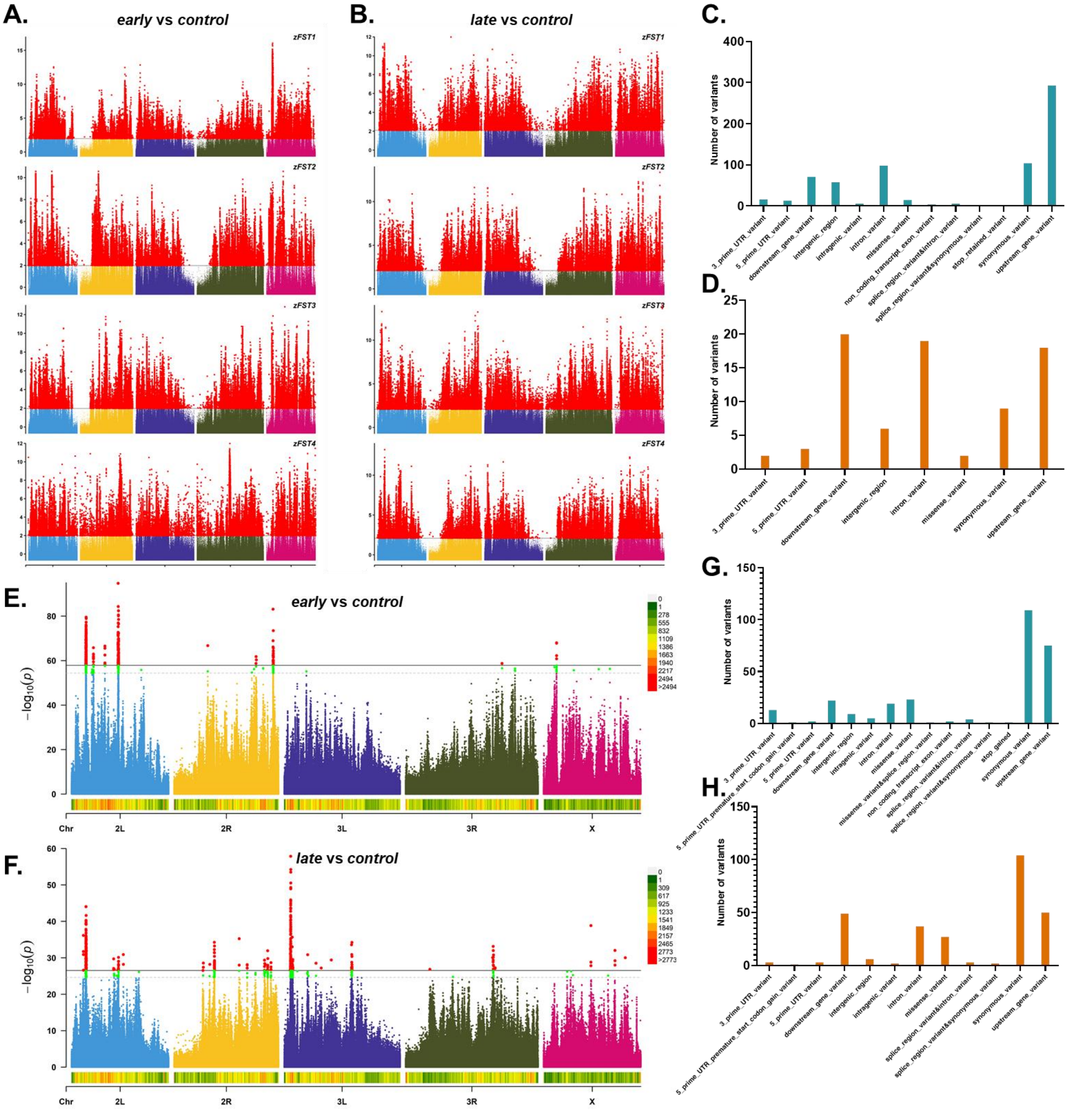


Figure 4.2: Results from F_{ST} and CMH test based allele frequency change. (A & B) zF_{ST} values plotted as Manhattan plots along different chromosomal arms and in four different replicates ($zF_{ST}1$, $zF_{ST}2$ and so on) from

control-early **(A)** and *control-late* **(B)** pairwise F_{ST} comparisons. **(C & D)** Effect type prediction of consequences of SNPs identified to be significantly different in *control-early* **(C)** and *control-late* **(D)** pairwise F_{ST} comparisons. **(E & F)** -log10(*p*) values plotted as Manhattan plots along different chromosomal arms from *control-early* **(E)** and *control-late* **(F)** pairwise CMH tests. Threshold for *early* was derived from permutation simulations ($CMHp_{corrected} = 1.396817^{-58}$) and for *late*, adjusted ($CMHp_{corrected} = 3.046445^{-27}$) to detect similar number of SNPs as of *control-early* comparison. **(G & H)** Effect type prediction of consequences of SNPs identified to be significantly different in *control-early* **(G)** and *control-late* **(H)** pairwise CMH tests. The dashed and solid lines in **(E & F)** depict top 333 (red points – derived from permutation simulations; 1000 permutations, 5% threshold) and 551 (green points – derived from permutation simulations; 1000 permutations, 10% threshold for *control-early* comparison and extended to *control-late* comparison) SNPs in each comparison. In **(E & F)** the SNP densities calculated from 0.1Mb non-overlapping windows were plotted at the bottom of each plot and the color guide is on the right.

There was no overlap between the loci identified in *early* and *late* populations. In *early* populations, a few significantly enriched gene ontology (GO) terms were “biological regulation”, “regulation of signaling”, “regulation of signal transduction”, “nervous system processes’, “instar larval or pupal development”, “metamorphosis” (*Appendix Fig. A4.11A*). In *late* populations, a few significantly enriched GO terms were “localization”, “primary metabolic process”, “response to stimulus”, “transport” (*Appendix Fig. A4.11B*). Next, to understand if specific biological pathways have been under differential selection in *early* and *late* populations, I performed a pathway enrichment analysis using g:Profiler (Reimand et al., 2016). Some pathways (KEGG terms) significantly enriched in *early* populations were “metabolic pathways”, “circadian rhythm - fly”, “apoptosis - fly”, “notch signaling pathway”, “longevity regulating pathway – multiple species”, “spliceosome”, “drug metabolism – other enzymes”, whereas in *late* populations, some significantly enriched pathways were “metabolic pathways”, “nicotinate and nicotinamide metabolism”, “longevity regulating pathway – multiple species”, “FoxO signaling pathway”, “phagosome”, “lysosome” (*Fig. 4.3E & F*).

Table 4.2: Genes with direct or non-direct circadian functions enriched in SNPs with significantly higher zF_{ST} values in *early* and *late* populations, compared to *control* population from pairwise F_{ST} comparisons.

Population	Gene name	Function	Reference
early	*kirre*	sleep traits	Harbison et al., 2017
	vrille	core clock component	Panda et al., 2002
	notch	compound eye development	Kahali et al., 2009
	timeless	core clock component	
	Egfr	consolidation and maintenance of sleep	Foltenyi et al., 2007
	Ir64a	odorant receptor	Ai et al., 2013
	nompC	response to auditory stimulus and startle response	Boyd-Gibbins et al., 2021; Zhang et al., 2013
	dpr8	synapse organization	Carrillo et al., 2015
	Nlg4	sleep, expressed in clock neurons	Li et al., 2013
	Fbxl4	negative regulation of circadian sleep/wake cycle	Li et al., 2017
	bruchpilot	adult locomotory behavior	Wagh et al., 2006
	Basigin	photoreceptor cell morphogenesis	Munro et al., 2010
	CCAP-R	neuropeptide signaling pathway, ecdysis	Cazzamali et al., 2003
	Eip93F	adult determinant during fly morphogenesis	FlyBase
	Tyramine β hydroxylase	sleep	Crocker and Sehgal, 2008
	Thor	contributes to circadian functions	Nagoshi et al., 2010
	sgg	regulates period of circadian clock	Martinek et al., 2001
late	*Heterochromatin protein 5*	gene silencing by heterochromatin	FlyBase
	unplugged	brain development	Hirth et al., 2003
	Cyp4p1, Cyp4p2, Cyp4p3	metabolism of insect hormones	FlyBase
	Netrin-A	compound eye development	Akin and Zipursky, 2016
	Rootletin	sensory perception of sound and touch	Chen et al., 2015
	Sirt4	determination of adult lifespan	Wood et al., 2018

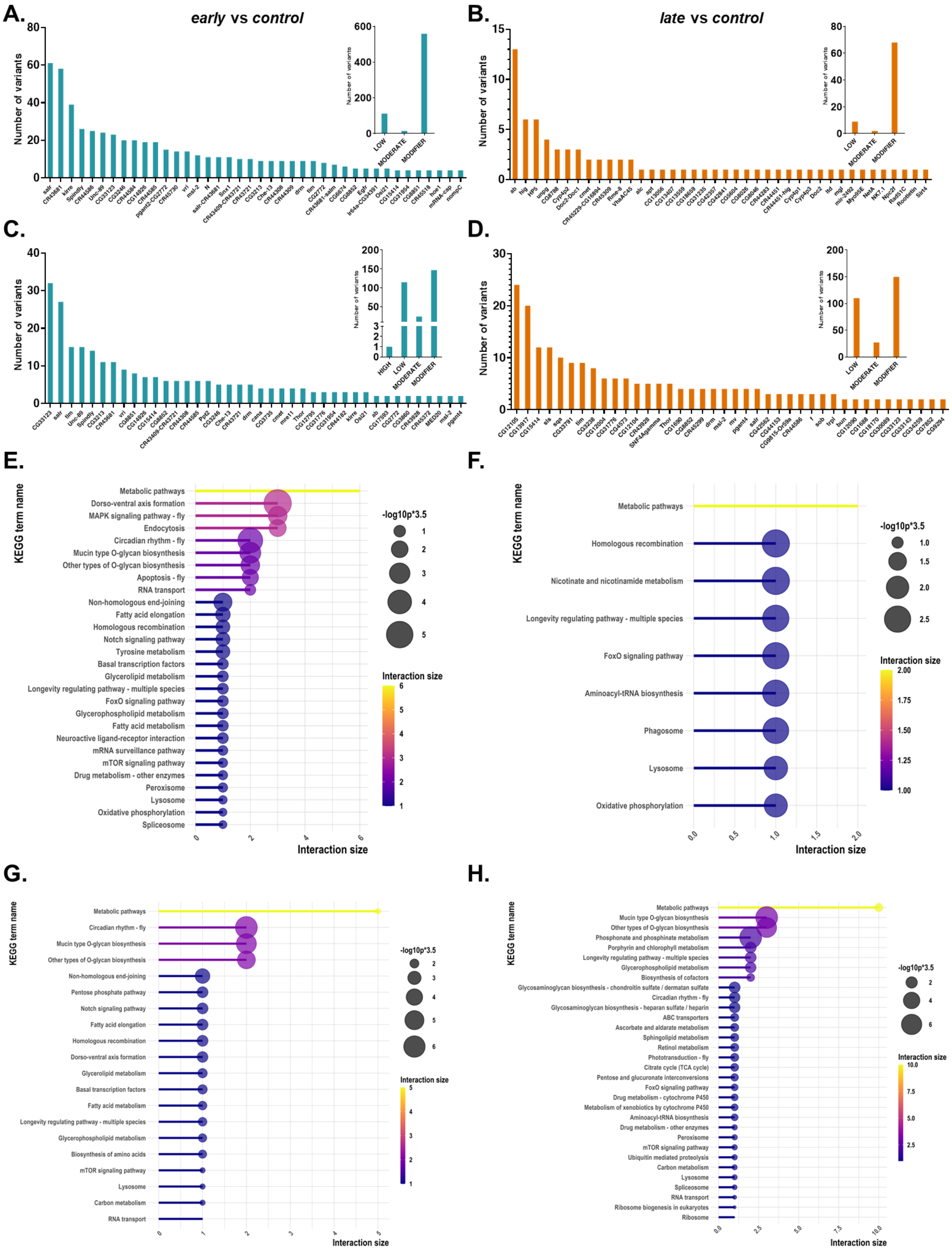

A.
early vs control
Number of variants
80
60
40
20
0
salr
CR43681
kirre
Spindly
CR44586
Unc-89
CG33123
CG3246
CR44584
CG14926
CR44585
pgant2-CG2772
CR45730
vri
msl-2
N
salr-CR43681
Snx1
CR43409-CR43721
CR43721
CG3213
Che-13
CR44308
CR44309
drm
tim
CG2772
CR43681-salm
CG6674
CG8852
Egfr
Ir64a-CG34391
Osi21
CG15414
CG31954
CG8851
CR45518
hoe1
mRNA-cap
nompC
Number of variants
600
400
200
0
LOW
MODERATE
MODIFIER
B.
late vs control
Number of variants
15
10
5
0
ab
hig
HP5
unpg
CG8788
Cyp4p2
Doc2-Doc1
cmet
CR45229-CG16894
CR45309
Rme-8
VhaAC45
alc
apt
CG13056
CG13407
CG13559
CG18659
CG31235
CG42357
CG42841
CG5604
CG8026
CG8046
CR44283
CR44451
CR44451-hig
Cyp4p1
Cyp4p3
Doc2
ltd
mgl
mir-2492
Myo95E
NetA
NK7.1
Npc2f
Rad51C
Rootletin
Sirt4
Number of variants
80
60
40
20
0
LOW
MODERATE
MODIFIER
C.
Number of variants
40
30
20
10
0
CG33123
salr
tim
Unc-89
Spindly
CG3213
CR43681
vri
CG8851
CG14926
CG15414
CG8852
CR43409-CR43721
CR44308
CR44585
Ppt2
CG3246
Che-13
CR43721
drm
cana
CG3735
cmet
mre11
Thor
CG12795
CG31776
CG31954
CR44182
kirre
Osi21
ab
CG17593
CG2772
CG3860
CR43928
CR45372
MED20
msl-2
pgant4
Number of variants
200
150
100
50
3
2
1
0
HIGH
LOW
MODERATE
MODIFIER
D.
Number of variants
30
20
10
0
CG12105
CG13917
CG15414
sls
sqa
CG33791
tim
CG3238
CG12004
CG31776
CG4573
CG12104
CR43928
SNF4Agamma
Thor
CG1690
CG8852
CR45299
drm
msl-2
mv
pgant4
salr
CG42562
CG44153
CG9815-Or59a
CR44586
f
sob
trpl
bun
CG12090
CG1688
CG18170
CG30069
CG33123
CG33143
CG34259
CG7852
CG9294
Number of variants
200
150
100
50
0
LOW
MODERATE
MODIFIER
E.
KEGG term name
Metabolic pathways
Dorso-ventral axis formation
MAPK signaling pathway - fly
Endocytosis
Circadian rhythm - fly
Mucin type O-glycan biosynthesis
Other types of O-glycan biosynthesis
Apoptosis - fly
RNA transport
Non-homologous end-joining
Fatty acid elongation
Homologous recombination
Notch signaling pathway
Tyrosine metabolism
Basal transcription factors
Glycerolipid metabolism
Longevity regulating pathway - multiple species
FoxO signaling pathway
Glycerophospholipid metabolism
Fatty acid metabolism
Neuroactive ligand-receptor interaction
mRNA surveillance pathway
mTOR signaling pathway
Drug metabolism - other enzymes
Peroxisome
Lysosome
Oxidative phosphorylation
Spliceosome
0
2
4
6
Interaction size
-log10p*3.5
1
2
3
4
5
Interaction size
6
5
4
3
2
1
F.
KEGG term name
Metabolic pathways
Homologous recombination
Nicotinate and nicotinamide metabolism
Longevity regulating pathway - multiple species
FoxO signaling pathway
Aminoacyl-tRNA biosynthesis
Phagosome
Lysosome
Oxidative phosphorylation
0.0
0.5
1.0
1.5
2.0
Interaction size
-log10p*3.5
1.0
1.5
2.0
2.5
Interaction size
2.00
1.75
1.50
1.25
1.00
G.
KEGG term name
Metabolic pathways
Circadian rhythm - fly
Mucin type O-glycan biosynthesis
Other types of O-glycan biosynthesis
Non-homologous end-joining
Pentose phosphate pathway
Notch signaling pathway
Fatty acid elongation
Homologous recombination
Dorso-ventral axis formation
Glycerolipid metabolism
Basal transcription factors
Fatty acid metabolism
Longevity regulating pathway - multiple species
Glycerophospholipid metabolism
Biosynthesis of amino acids
mTOR signaling pathway
Lysosome
Carbon metabolism
RNA transport
0
1
2
3
4
5
Interaction size
-log10p*3.5
2
3
4
5
6
Interaction size
5
4
3
2
1
H.
KEGG term name
Metabolic pathways
Mucin type O-glycan biosynthesis
Other types of O-glycan biosynthesis
Phosphonate and phosphinate metabolism
Porphyrin and chlorophyll metabolism
Longevity regulating pathway - multiple species
Glycerophospholipid metabolism
Biosynthesis of cofactors
Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate
Circadian rhythm - fly
Glycosaminoglycan biosynthesis - heparan sulfate / heparin
ABC transporters
Ascorbate and aldarate metabolism
Sphingolipid metabolism
Retinol metabolism
Phototransduction - fly
Citrate cycle (TCA cycle)
Pentose and glucuronate interconversions
FoxO signaling pathway
Drug metabolism - cytochrome P450
Metabolism of xenobiotics by cytochrome P450
Aminoacyl-tRNA biosynthesis
Drug metabolism - other enzymes
Peroxisome
mTOR signaling pathway
Ubiquitin mediated proteolysis
Carbon metabolism
Lysosome
Spliceosome
RNA transport
Ribosome biogenesis in eukaryotes
Ribosome
0.0
2.5
5.0
7.5
10.0
Interaction size
-log10p*3.5
2
4
6
Interaction size
10.0
7.5
5.0
2.5

Figure 4.3: Results from CMH test based allele frequency change. **(A & B)** Number of variants in genes (top 40) and effect of putative impact/deleteriousness (insets) in *control-early* **(A)** and *control-late* **(B)** pairwise F_{ST} comparisons. **(C & D)** Number of variants in genes (top 40) and effect of putative impact/deleteriousness (insets) in *control-early* **(C)** and *control-late* **(D)** pairwise CMH tests. **(E & F)** Enrichment analysis of KEGG pathways of genes from *control-early* (top 40 - **E**) and *control-late* **(F)** pairwise F_{ST} comparisons. **(G & H)** Enrichment analysis of KEGG pathways of genes from *control-early* (top 40 - **G**) and *control-late* (top 40 - **H**) pairwise CMH tests. The length of the horizontal lines in the lollipop charts **(E, F, G & H)** are the interaction sizes as derived from the KEGG pathway analysis with g:Profiler and the size of the bubbles depict -log10(*p*) values from the enrichment analysis, scaled by a constant of 3.5 for representation purposes, and the color codes also depict the same.

I used the CMH test to identify loci where allele frequency has concurrently changed among all four replicate populations of *early* and *late* populations compared to that of *control* populations. In *early* and *late* populations, I found 287 exclusive SNPs significantly different from *control* populations using the thresholds described in materials and methods (*Fig. 4.2E & F*). I also identified 46 mutual SNPs in *early* and *late* populations, which have significantly different allele frequencies compared to *control* populations. In the *early* populations, these 287 SNPs span over 65 genes, and some noticeable ones are – *timeless* (core clock component), *Leucyl-tRNA synthetase* (blue light induced gene; (Hall et al., 2018)), *vrille* (core clock component; (Panda et al., 2002)), *Thor* (contributes to circadian functions; (Nagoshi et al., 2010)), *Toll-7* (innate immune response; (Chowdhury et al., 2019)) (*Fig. 4.3C* and summarized in *Table 4.3*). Among these SNPs, SnpEff predicted 115 to have an estimated "low" effect of putative impact/deleteriousness, 24 "moderate" effect SNPs, 147 "modifier" effect, and 1 "high" effect SNPs (summarized in *Appendix Table A4.3*). 75 SNPs were categorized as "upstream_gene_variant", 109 were "synonymous_variant", 19 were "intron_variant", 22 were "downstream_gene_variant", and only 23 were categorized as "missense_variant – moderate effect" and 1 was categorized as "stop_gained – High effect" as per effect type description of consequence (*Fig. 4.2G & 4.3C*$_{inset}$ and summarized in *Appendix Table A4.4*). In the *late* populations, these 287 SNPs span over 107 genes and some mentionable ones are - *spaghetti-squash activator* (required for starvation induced autophagy; (Tang et al., 2011)), CG31776 (nsSNPs present in dark adapted flies; (Izutsu et al., 2012)), *timeless* (core clock component), *Thor* (contributes to circadian functions; (Nagoshi et al., 2010)), *SNF4Aγ* (lipid metabolism, autophagy and response to starvation; (Johnson et al., 2010; Lippai et al., 2008)), *Or59a* (detection of chemical stimulus involved in sensory perception of smell; (Kreher et al., 2005)), *trpl* (attenuated light response of TIM; (Yang et al., 1998)), *ETHR* (activation of ecdysis motor program; (Mark et al., 2021)), CG4329 (sensory perception of sound; (Senthilan et al., 2012)), CG7879 (regulation of alternative mRNA splicing via spliceosome; (Park et al., 2004)), *Ir62a* (detection of chemical stimulus; (Benton et al., 2009)), *hfp* (alternative mRNA splicing, via spliceosome; (Van Buskirk and Schüpbach, 2002)) (*Fig. 4.3D and* summarized in *Table 4.3*). Among these SNPs, SnpEff predicted 110 to have an estimated "low" effect of putative impact/deleteriousness, 27 "moderate" effect SNPs, and 150 "modifier" effect SNPs. 50 SNPs were categorized as "upstream_gene_variant", 104 were "synonymous_variant", 37 were "intron_variant", 49 were "downstream_gene_variant", and only 27 were categorized as "missense_variant – moderate effect" as per effect type description of consequence (*Fig. 4.2H & 4.3D*$_{inset}$ and summarized in *Appendix Table A4.4*). Among the 46 SNPs spanning over 10 genes, which were common between *early* and *late* populations (*early*-*late* overlap), SnpEff predicted 31 to have an estimated "low" effect of putative impact/deleteriousness, 4 "moderate" effect SNPs, and 11 "modifier" effect SNPs. 4 SNPs were categorized as "upstream_gene_variant", 31 were "synonymous_variant", 4 were "intron_variant", 1 was "downstream_gene_variant", and only 4 were categorized as "missense_variant – moderate effect" as per the effect-type description of consequence (summarized in *Appendix Table A4.4*). *timeless* (core clock component) is one of the major genes which accumulated 14 SNPs common to both *early* and *late* populations. In *early* populations, a few significantly enriched gene ontology (GO) terms were "transcytosis", "RNA polymerase III type 2 promoter sequence-specific DNA binding", "Mre11 complex", "transcription factor binding", and a few enriched pathways were "metabolic pathways", "circadian rhythm - fly", "notch signaling pathway", "dorso-ventral axis formation", "longevity regulating pathway – multiple species", "lysosome" (*Fig. 4.3G & Appendix Fig. A4.11C*). In *late* populations, some enriched GO terms were "miRNA loading onto RISC involved in gene silencing by miRNA", "chemoattractant activity", "mRNA splice site selection", "sensory perception of sound", "sensory perception of mechanical stimulus", and some enriched pathways are "metabolic pathways", "longevity regulating pathway – multiple species", "circadian rhythm - fly", "retinol metabolism", "phototransduction - fly", "FoxO signaling pathway", "citrate cycle (TCA cycle)", "ubiquitin mediated proteolysis", "spliceosome", "drug metabolism – other enzymes" (*Fig. 4.3H & Appendix Fig. A4.11D*).

Table 4.3: Genes with direct or non-direct circadian functions enriched in significant SNPs in *early* and *late* populations, compared to *control* population from CMH tests.

Population	Gene name	Function	Reference
early	*timeless*	core clock component	
	vrille	core clock component	Panda et al., 2002
	Leucyl-tRNA synthetase	blue light induced gene	Hall et al., 2018
	Thor	contributes to circadian functions	Nagoshi et al., 2010
	Toll-7	innate immune response	Chowdhury et al., 2019
late	*spaghetti-squash activator*	required for starvation induced autophagy	Tang et al., 2011
	CG31776	nsSNPs present in dark adapted flies	Izutsu et al., 2012
	timeless	core clock component	
	Thor	contributes to circadian functions	Nagoshi et al., 2010
	SNF4Aγ	lipid metabolism, autophagy and response to starvation	Johnson et al., 2010
	Or59a	detection of chemical stimulus involved in sensory perception of smell	Kreher et al., 2005
	trpl	attenuated light response of TIM	Yang et al., 1998
	ETHR	activation of ecdysis motor program	Mark et al., 2021
	CG4329	sensory perception of sound	Senthilan et al., 2012
	CG7879	regulation of alternative mRNA splicing via spliceosome	Park et al., 2004
	Ir62a	detection of chemical stimulus	Benton et al., 2009
	hfp	alternative mRNA splicing, via spliceosome	Van Buskirk and Schüpbach, 2002

4.3.3 Signatures of selective sweeps in populations:

Pool-hmm estimates allele frequencies and detects selective sweeps using pooled-sequencing data from a single sample using the hidden Markov model (HMM) method from *Boitard et al., 2012* (Boitard et al., 2012). In *Pool-hmm*, each polymorphic in the genome can have three hidden states: “Neutral”, “Intermediate”, and “Selection”, which are associated with different patterns of allele frequencies. Based on the observed allele frequency spectrum (AFS) calculated from the pooled-sequencing data, *Pool-hmm* predicts the most likely hidden state at each site, where windows of sites with the hidden state of “Selection” can be considered as a sweep signal. I identified multiple sweep windows in all four replicates of *early*, *control*, and *late* populations. In four replicates of *control* populations, a total of 12.86Mb, 11.66Mb, 11.74Mb, and 13.32Mb regions in five chromosomal arms were identified to be under selective sweep (*Appendix Fig. A4.15A*). Similarly, for *early* populations, these values were 12.99Mb, 13.59Mb, 12.8Mb, 10.79Mb, and for *late* populations – 11.18Mb, 10.88Mb, 10.34Mb, 11.7Mb respectively (*Appendix Fig. A4.15B & C*). When I calculated sweep regions common among all four replicates, I saw a drastic reduction in the number of regions under sweep in all populations. In *control* populations, the regions common among all four replicates indicate a total of 2.42Mb region under selective sweep among all five major chromosomal arms (2L: 0.14Mb, 2R: 0.6Mb, 3L: 0.44Mb, 3R: 0.16Mb, and X: 1.1Mb) and containing 43 genes (*Fig. 4.4B & E*). In *early* populations, these sweep regions were of 1.02Mb (2L: 0.29Mb, 2R: 0.19Mb, 3L: 0.22Mb, 3R: 0.19Mb, and X: 0.47Mb) containing 79 genes, and in *late* populations, these regions covered a total of 2.9Mb region (2L: 0.07Mb, 2R: 0.39Mb, 3L: 0.31Mb, 3R: 0.41Mb, and X: 1.73Mb) containing 234 genes (*Fig. 4.4A & D*). Next, I compared the regions under selection between all four replicates of *control* population with those of *early* population and derived *early* population exclusive sweep regions. These regions were even smaller, a total of 0.41Mb spanning over all five major chromosomal arms (2L: 0.02Mb, 2R: 0.12Mb, 3L: 0.1Mb, 3R: 0.04Mb, and X: 0.14Mb) containing 20 genes (*Fig. 4.4C*). Similarly, the *late* population exclusive sweep regions were of total 1.34Mb in length (2L: 0.04Mb, 2R: 0.3Mb, 3L: 0.15Mb, 3R: 0.09Mb, and X: 0.76Mb) containing 123 genes (*Fig. 4.4F*). Pool-hmm results showed that all replicates of all three populations show significant signatures of selective sweeps all over their major chromosomal arms (~1/10th of their genome in each replicate), though when I compared among different replicates of a population, the regions under sweep common to all replicates were relatively small (1.02Mb in early, 2.42Mb in control, and 2.9Mb in late populations). Considering each replicate are phenocopies of these populations, these regions common to all four replicates of each population can be most associated with their respective phenotypes. Moreover, most of the sweep regions also show negative D values and low π values, showing they most probably are true sweep regions (*Appendix Table A4.7 & A4.8*).

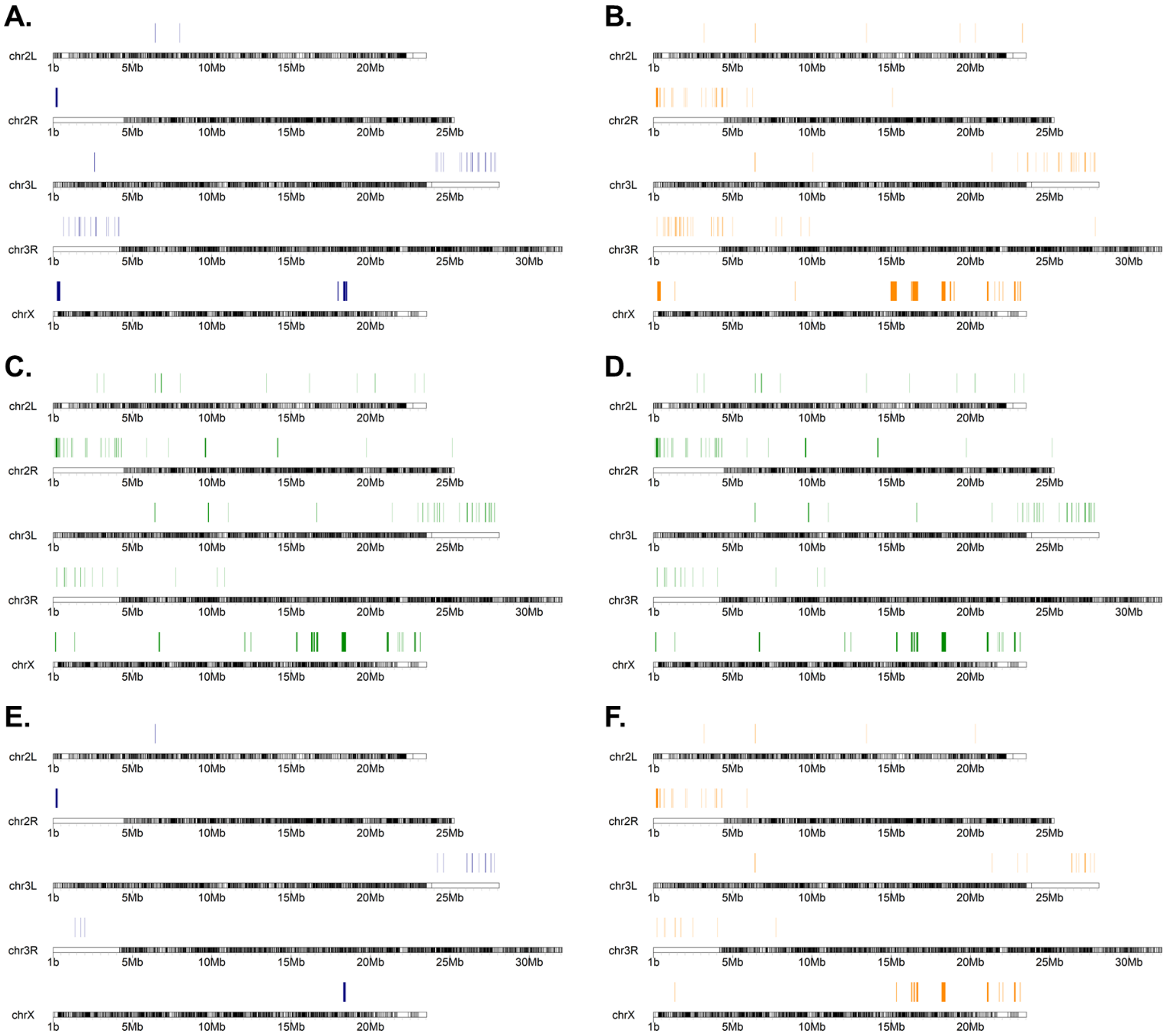


Figure 4.4: Results from selective sweep analysis using *Pool-hmm*. (A, B, D, & E) Sweep windows common among all four replicates of *control* (green – **B** & **E**), *early* (blue - **A**), and *late* (red - **D**) populations. **(C & F)**. Sweep windows exclusive to *early* (blue - **C**) and *late* (red - **F**) populations when compared to those of *control* populations. In *control* populations, 2.42Mb region was under selective sweep among all four replicates, and 1.02Mb and 2.9Mb regions under *early* and *late* populations respectively. When I removed the regions under sweep in *control* populations **(B & E)** from those of *early* **(A)** and *late* **(D)** populations, the resulting set contained 0.41Mb region in *early* population **(C)** and 1.34Mb region in *late* population **(F)**, which were exclusive to these populations.

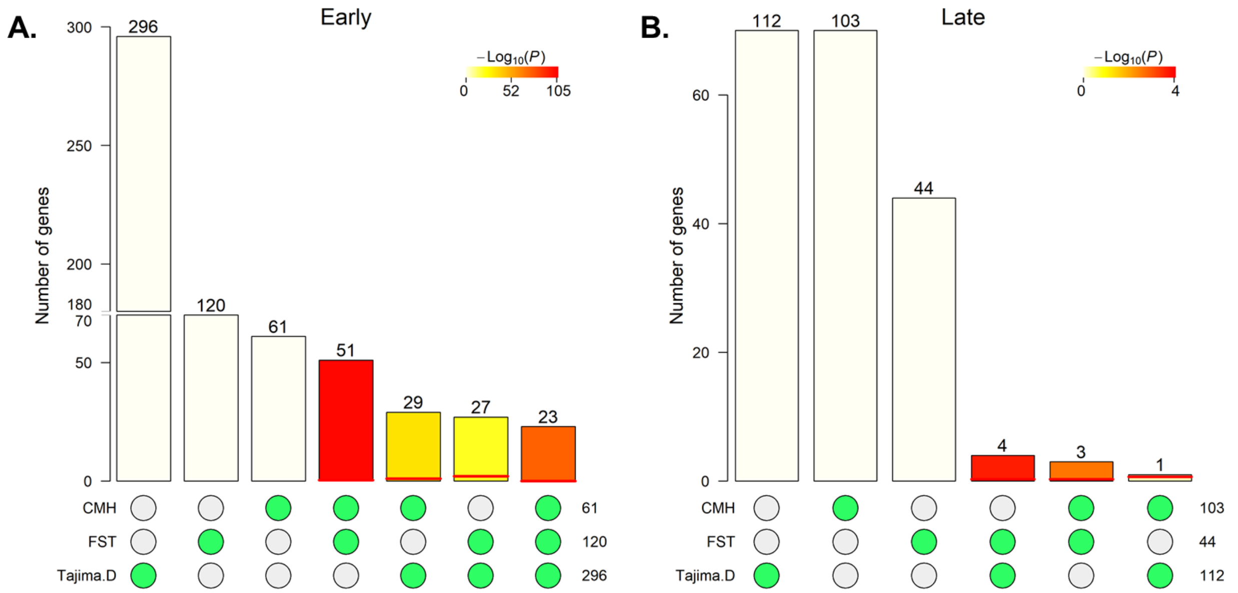


Figure 4.5: SuperExactTest for overlap between sets. (A & B) Number of genes in each comparison set (**A** – *control-early*, **B** – *control-late*) for diffesrent methods (comparison of Tajima's D in windows, pairwise F_{ST} comparison, and CMH tests). Overlap of genes enriched in differentiated SNPs among different methods were high for *early* population (CMH+F_{ST}:51 – *SuperExactTest*$_{pvalue}$ = 1.57-105, CMH+Tajima.D:29 – *SuperExactTest*$_{pvalue}$ = 1.45-35, F_{ST}+Tajima.D:27 – *SuperExactTest*$_{pvalue}$ = 5.96-23, CMH+F_{ST}+Tajima.D:23 – *SuperExactTest*$_{pvalue}$ = 5.41-76) compared to *late* population (CMH+F_ST:3 – *SuperExactTest*$_{pvalue}$ = 0.002, CMH+Tajima.D:1 – *SuperExactTest*$_{pvalue}$ = 0.048, F_{ST}+Tajima.D:4 – *SuperExactTest*$_{pvalue}$ = 0.0002, CMH+F_{ST}+Tajima.D:0 – *SuperExactTest*$_{pvalue}$ = 1).

4.4 Discussion:

In this study, I compared genomes of populations of *Drosophila melanogaster* which evolved significantly divergent chronotypes from their controls, using a long-term experimental evolution experiment for about ~15 years and ~260 generations. These populations are unique in terms of their strict maintenance regime, insights provided on the evolution of circadian clock properties with chronotypes, and population-level genetically independent replicates. All these characteristics make them ideal candidates for population genomic studies investigating genome-wide changes in evolved divergent chronotypes, in this case, early and late eclosion rhythm chronotypes. As expected, I observed all populations maintain a high level of standing genetic variation (average D value > 0; high average/median π value compared to other studies (Yukilevich et al., 2010)), and there were signatures of selection present in the form of selective sweeps (using *Pool-hmm*) in all 12 samples (four replicates each of *early*, *control* and *late* populations) or regions of significantly low *D* value in *early* and *late* populations compared to *control* populations (*Fig. 4.1C & D, 4.4, Appendix Fig. A4.9 & A.415*). Though signatures of selection were present in all populations, including *control*, there were a large number of loci that were exclusive to *early* and *late* populations compared to *control*. When I investigated allele frequency changes of different SNPs in *early* and *late* populations compared to the *control* population, I found a high number of SNPs where allele frequency had changed (*Fig. 4.1, 4.2 & 4.3*). I detected significantly different SNPs through different methods (F_{ST} and CMH), and analyzed each list of SNPs separately as different methods of detection of allele frequency changes have their own merits and demerits (*Fig. 4.2 & 4.3*) (Kofler et al., 2011b; Vlachos et al., 2019). Using these different methods, I identified different sets of genes harboring significantly differentiated SNPs in *early* and *late* populations compared to the *control* population (*Fig. 4.1, 4.2 & 4.3, Fig. 4.5A & B*). I also identified several enriched pathways plausibly under differential selection in these two different populations.

4.4.1 Photosensitivity, sleep, and clock related changes in *early* populations:

In the *early* populations, genes with significantly differentiated SNPs included light-mediated (induced), photoreceptor regulation, and phototransduction genes like *nord, Arrestin 1, reduced ocelli, ninaD, notch, Basigin,* and *Leucyl-tRNA synthetase* (summarized in *Table 4.1, 4.2 & 4.3*) (Caldwell et al., 2007; Damulewicz et al., 2019; Gu et al., 2004; Hall et al., 2018; Kahali et al., 2009; Munro et al., 2010). Previous studies from our lab on these populations have shown *early* populations may have evolved heightened light sensitivity and light-mediated positive masking response, suggesting that there may be photosensitivity differences of the circadian clock between *early* and *late* populations (Ghosh et al., 2021; Kumar et al., 2007; Vaze et al., 2012a). Most of the SNPs present in these genes are in regulatory regions, suggesting that regulatory changes in multiple genes related to light input pathways to the circadian clock can modulate chronotype in *Drosophila melanogaster* (*Fig. 4.2C & G*). I also identified multiple core circadian clock genes (*timeless, vrille, sgg*) and genes affecting circadian functions in Drosophila (*nervy*, *Thor*, *spaghetti*, and *Fbxl4*) harboring many significantly differentiated SNPs in *early* populations (*Fig. 4.2A & E*) (Duvall and Taghert, 2012; Hermann-Luibl et al., 2014; Li et al., 2017; Means et al., 2015; Nagoshi et al., 2010). These genes can modulate circadian clock function directly or indirectly, and fine-tuning of circadian phase may be achieved by a change in their protein structure-function as predicted by SnpEff (Cingolani et al., 2018). Interestingly, I also found multiple sleep regulating genes having multiple SNPs present in them in *early* populations – *kirre*, *Egfr*, *Tyramine β hydroxylase*, and *Nlg4* (*Fig. 4.2A & E*) (Crocker and Sehgal, 2008; Dreyer et al., 2019; Foltenyi et al., 2007; Harbison et al., 2017; Li et al., 2013). As sleep and circadian clock are tightly correlated processes, I speculated that along with divergent chronotypes, there would be differences in sleep among the *early* and *late* populations, which may be governed by differences in the mentioned genes. Indeed, I found *early* and *late* populations have significantly lower nighttime sleep amount and quality compared to the *control* populations under different light intensities and shorter daylengths (see chapter 3). Directly correlated to the adult eclosion rhythm phenotypes of these populations, I found changes in genes (genes accumulating more SNPs) regulating the ecdysis behavior, like *CCAP-R* and *Eip93F* (Cazzamali et al., 2003; FlyBase). *CCAP-R* is directly involved in the modulation of circadian control of the adult eclosion behavior, as shown by abnormal timing of eclosion in populations of flies lacking *CCAP* neurons (Park et al., 2003). *Eip93F* has a tightly regulated temporal expression pattern induced by the ecdysone hormone – only during the early stages of Drosophila metamorphosis, and *Eip93F* mutants die during the early stages of pupal development (Baehrecke and Thummel, 1995; Lee et al., 2000). There were also different genes having significant allele frequency changes related to other biological functions – odorant reception (*Ir64a*), response to auditory stimulus and startle response (*nompC*), synapse organization (*dpr8*), neurotransmission (*pickpocket 29*), and adult locomotory behavior (*bruchpilot*) (Ai et al., 2013; Boyd-Gibbins et al., 2021; Carrillo et al., 2015; Hill et al., 2017; Wagh et al., 2006). In the *early* populations, I observed putative biological pathways under selection to be – "circadian rhythm – fly", "notch signaling pathway", "spliceosome", etc., and some significantly enriched GO terms were – "RNA polymerase III type 2 promoter sequence-specific DNA binding", "transcription factor binding", "nervous system processes", "instar larval or pupal development", "metamorphosis", etc. Many of these pathways can affect development time, morphogenesis (particularly eclosion, and possibly timing of adult eclosion), expression of circadian genes at specific times directly and indirectly (*Fig. 4.2E & G, Appendix Fig. A4.11A & C*). For example, the notch signaling pathway has been implicated to be involved in the regulation of sensitivity to sleep loss, which directly correlates with our *early* population showing less nighttime sleep amount and quality under different light intensities and shorter daylengths (see chapter 3) (Seugnet et al., 2011). All these changes were specific to the *early* populations. When I grouped genes under selection from different analysis sets, I observed significant enrichment of the protein-protein interaction network than expected (*p* value = 0.0291) for genes from all three sets (Tajima's *D* based comparison, pairwise F_{ST} comparison, and CMH test) with 314 nodes and 347 edges in the network (*Appendix Fig. A4.14A*). When I only used the genes from the allele frequency-based tests (pairwise F_{ST} comparison and CMH test), I did not see significant enrichment of the interaction network (*p* value = 0.272), but I observed 101 nodes and 41 edges in the network (*Appendix Fig. A4.9A, A4.16A*). There were significant overlaps between genes enriched in significantly differentiated SNPs from different methods – CMH+F_{ST}: 51 (*SuperExactTest*$_{pvalue}$ = 1.57^{-105}), CMH+Tajima.D: 29 (*SuperExactTest*$_{pvalue}$ = 1.45^{-35}), F_{ST}+Tajima.D: 27 (*SuperExactTest*$_{pvalue}$ = 5.96^{-23}), CMH+F_{ST}+Tajima.D: 23 (*SuperExactTest*$_{pvalue}$ = 5.41^{-76}) (*Fig. 4.5A*).

4.4.2 Clock, mRNA splicing, and ecdysis related changes in *late* populations:

In the *late* populations, genes with significantly differentiated SNPs included core circadian clock genes or genes affecting circadian functions – *timeless, Thor*, *Blimp-1*, and *spaghetti* (summarized in *Table 4.1, 4.2 & 4.3*)(Akagi et al., 2016; Means et al., 2015; Nagoshi et al., 2010). Though I also found multiple significantly differentiated SNPs in *timeless* in *early* populations, there were many SNPs in *timeless* exclusive to *late* population at very different regions of the *timeless* locus (Summarized in *Appendix Table A4.5 & A4.S6*). There were multiple genes in *late* populations related to brain development (*unplugged*), compound eye development (*Netrin-A*), autophagy regulation (*spaghetti-squash activator* and *SNF4Aγ*), etc. (Akin and Zipursky, 2016; Hill et al., 2017; Hirth et al., 2003; Johnson et al., 2010; Lippai et al., 2008; Tang et al., 2011). I also found SNPs in *trpl*, which is responsible for the attenuated light response of TIM (Yang et al., 1998). More importantly, I identified several SNPs in genes responsible for the regulation of alternative mRNA splicing via spliceosome (CG7879), alternative mRNA splicing, via spliceosome (*hfp*), and identified GO term “mRNA splice site selection” and enriched pathway “spliceosome” – all of which indicate significant changes in the mRNA splicing machinery which has been proposed to regulate period and phases of the circadian clock of *Drosophila melanogaster* directly or indirectly, along with temperature sensitivity of the circadian clock (*Fig. 4.3F & H, Appendix Fig. A4.11B & D*) (Van Buskirk and Schüpbach, 2002; Evantal et al., 2018; Majercak et al., 2004; Park et al., 2004; Shakhmantsir et al., 2018). Indeed, previous studies from the lab had shown significantly more phase lability and temperature sensitivity and longer free-running period in the *late* population compared to *control* and *early* populations under temperature cycles in both eclosion and locomotor activity rhythms (Abhilash et al., 2019, 2020; Nikhil et al., 2016b). The *late* population also had SNPs in genes involved in activation of ecdysis motor program (*ETHR*), detection of chemical stimulus involved in sensory perception of smell (*Or59A*), sensory perception of sound and touch (*Rootletin*), determination of adult lifespan (*Sirt4*), along with enriched GO terms like "response to stimulus", "primary metabolic process", "phagosome", "lysosome", "sensory perception of sound", "sensory perception of mechanical stimulus", "miRNA loading onto RISC involved in gene silencing by miRNA", and enriched pathways like "circadian rhythm - fly", "retinol metabolism", "phototransduction - fly", "ubiquitin mediated proteolysis", "metabolic pathways", "nicotinate and nicotinamide metabolism", "longevity regulating pathway – multiple species", "FoxO signaling pathway" etc. (*Fig. 4.3F & H, Appendix Fig. A4.11B & D*) (Chen et al., 2015; Kreher et al., 2005; Mark et al., 2021; Wood et al., 2018). Previously, it was observed that virgin female flies of the *late* population have lower lifespan compared to the *early* population (Nikhil et al., 2016a), which in retrospect, may be governed by changes in longevity regulating pathways and determinant of adult lifespan (*Sirt4*). Ecdysis Triggering Hormone Receptor (ETHR) knockdown delays the ecdysis behavior, while overexpression of ETHR can accelerate ecdysis – this suggests that SNPs, if present within ETHR, deleteriously affect ETHR signaling, may lead to delayed eclosion as observed in the *late* population (Kim et al., 2015). Previous work has shown that FoxO signaling has non cell-autonomous effect on the central circadian clock function, and as *foxo* is majorly expressed in the fat body, it suggests that changes in genes expressed in peripheral tissues can also affect circadian behavior in the *late* population (Zheng et al., 2007). When I grouped genes under selection from different analysis sets, I observed significant enrichment of protein-protein interaction network than expected (*p* value = 0.00032) for genes from all three sets (Tajima's *D* based comparison, pairwise F_{ST} comparison, and CMH test) with 214 nodes and 131 edges in the network (*Appendix Fig. A4.14B*). The protein-protein interaction networks were larger in *early* population compared to the *late* population, both in terms of number of nodes and edges – this suggests that in *early* population, the genes that harbor most significantly differentiated SNPs are more biologically connected (more physical interactions among them) than in *late* population. When I only used the genes from the allele frequency-based tests (pairwise F_{ST} comparison and CMH test), I again observed significant enrichment of the interaction network (*p* value = 0.00376), and I observed 130 nodes and 40 edges in the network (*Appendix Fig. A4.9B, A4.16B*). There were significant overlaps between genes enriched in significantly differentiated SNPs from different methods – CMH+F_{ST}: 3 ($SuperExactTest_{pvalue}$ = 0.002), CMH+Tajima.D: 1 ($SuperExactTest_{pvalue}$ = 0.048), F_{ST}+Tajima.D: 4 ($SuperExactTest_{pvalue}$ = 0.0002), CMH+F_{ST}+Tajima.D: 0 ($SuperExactTest_{pvalue}$ = 1), but these overlaps were much smaller when compared to those of the *early* population (*Fig. 4.5B*).

There were also a few genes harboring SNPs different from *control* populations in both *early* and *late* populations (*early-late* overlap set from CMH test). Among these, most noteworthy is the *timeless* gene which has 14 SNPs common to both *early* and *late* populations, along with SNPs exclusive to both populations. This suggests that *timeless* may have evolved as a major player regulating phase divergence in our populations, along with *vrille* in *early* populations. Apart from these core clock genes, various clock output genes and clock regulatory genes showed significant SNPs exclusively in *early* and *late* populations. The majority of the SNPs I detected are in regulatory regions, which correlates well with chronotypes being highly quantitative and polygenic traits. Most of the SNPs I identified may contribute a very small part to the final phase divergence among our populations. As I observed temperature-dependent phase lability is higher in the *late* populations, I expect this phenomenon to be regulated by temperature mediated differential splicing of *timeless* or *period*, for which I have some supporting data from this study in terms of multiple significant SNPs in *timeless* or different spliceosome complex genes (Anduaga et al., 2019; Majercak et al., 1999, 2004; Shakhmantsir et al., 2018). All the significant SNPs and genes I identified in this study have been selected based on their presence in all four replicate populations, which largely excludes the possibility of them being the result of genetic drift and purely rising because of the selection pressure acting on the standing genetic variation. The large number of significant SNPs being present in non-clock genes and the diverse array of non-circadian phenotypes evolving differentially among populations across all four replicates suggest that chronotype divergence in populations does not necessarily stem from changes in clock genes but can be achieved by multiple small-effect changes in different genes in a population and selection on chronotypes may result in correlated changes in other phenotypes (e.g., sleep, lifespan, fecundity, non-clock photosensitivity, etc.). Our populations were derived from natural populations; as we have successfully maintained large outbreeding populations with large standing genetic variation, it is expected that whatever correlation we see among traits or loci will most likely be a natural genetic correlation among these traits and loci, and I believe the current study fills in a large gap in this aspect. I did not observe any high effect polymorphism (barring one "stop_gained" variant in the *early*-exclusive CMH list), and a majority of the polymorphisms were of "low" and "modifier" effects, and very few of "moderate" effect (*Summarized in Appendix Table A4.1-A4.4*). This suggests that when selection for chronotype divergence acts on natural standing genetic variations, multiple small effect variants get differentially selected in different chronotypes, and the abundant high effect genes/polymorphisms observed to be affecting chronotypes in different organisms most probably stem from high inbreeding and random fixation of loci in the organisms studied. Overall, I hope this study will serve as an important database of information and enable the identification of putative targets for further work investigating the mechanisms by which chronotype divergence may occur in *Drosophila melanogaster*.

Chapter 5

Screening of the DrosDel collection for differential phasing of eclosion rhythm

5.1 Introduction:

The length of a day, defined as the time it takes the earth to rotate once around its own axis completely, has not always been ~24 hours as we see today. According to fossil evidence, one of the first organisms to appear on ancient earth were cyanobacteria, 3.5 billion years ago (Dodd et al., 2017), when the length of a day was ~8 hours (Turcotte et al., 1977); the day length was ~21 hours when the eukaryotic cells emerged 1.7 billion years ago (Meckien, 2014). Around 4 million years ago, the first human ancestors arose, when the day length was already very close to 24 hours long. Circadian clocks, as we know, are shaped by evolutionary forces (i.e., natural selection, genetic drift etc.), and one of the major selection pressures is the day length on earth (~24) hours as most organisms on earth have a circadian clock with free running period close to 24 hours. Even cyanobacteria, which are hypothesized to be one of the first organisms to appear on earth, have a circadian clock with close to 24-hour periodicity but have been postulated to have a circadian clock with much shorter periodicity (close to 8 hours) when they first evolved (Turcotte et al., 1977). The circadian clocks in organisms evolved to match the periodicity of the revolution of the earth, conceivably because the rotation of earth gave rise to the most prominent and reliable cycling environmental cue – the light-dark cycle, which is repeated every ~24 hours and gives faithful signals to the circadian clock of organisms to synchronize/entrain to this period. As circadian clocks are endogenous, and their outputs can be affected by both genetic and environmental factors, the clock outputs can show a wide range of differences in a population. These differences are generally not reflected in the free running period, as most organisms are under cyclic environments. The differences in clock output of different rhythmic behaviors or physiological processes majorly express themselves as differences in the rhythmic phenomenon's waveform and phase. Waveforms are one of the most central and critical dimensions in circadian biology

(Gorman et al., 2017). Waveforms and phases of a circadian rhythm can vary depending on the environmental condition in an individual and can also vary among different individuals in a population due to their inherent genetic differences (Gorman et al., 2017; Roenneberg et al., 2007). Different individuals in a population adopt a specific temporal relationship to the zeitgeber (phase of entrainment – Ψ_{ent}). These variations in Ψ_{ent}, if inherited, give rise to chronotypes in a population (Roenneberg et al., 2007). Chronotypes have been reviewed in detail in chapters 2 and 3, and the genetic basis of chronotype variation also has been described in detail in chapter 4.

Genetically tractable model systems, such as *Drosophila melanogaster*, are valuable resources for understanding and discovering basic principles in biology conserved throughout evolution. One main reason behind the success of Drosophila in basic science research is the ability to carry out genetic screens for the factors playing roles in different biological processes and then characterize how these factors affect those processes. Genome-wide screens for various phenotypes have been done in Drosophila facilitated by the identification of genes that modify/govern certain phenotypes (St Johnston, 2002). A very useful tool for such approaches is characterized sets of chromosomal deletions that help to rapidly screen most of the genome at a somewhat lower resolution (Ryder et al., 2004). One of the most famous such set of deletions is the “core deficiency kit” – a set of 220 stocks of Drosophila that have deletions covering ~85% of the euchromatic genome (Bloomington Drosophila Stock Center 2003). The disadvantage of using this traditional “core deficiency kit” is – they are genetically heterogenous, and the deletions are not molecularly mapped. Various other sets of molecular mapped deficiency lines have been described, like the Exelixis collection (Parks et al., 2004; Thibault et al., 2004). However, these other lines have much smaller deficiencies in size, and screening numerous lines for particular phenotypes is not always possible to achieve in a genome-wide screen. Heterozygous deletion screens have proved to be useful in assessing a number of dominant effects of potential target genes responsible for a specific phenotype (St Johnston, 2002). The DrosDel deficiency core kit describes 209 deficiency lines created using FLP-mediated recombination between FRT sites, a method originally described by Golic and Golic, and created by Ryder and colleagues (Golic and Golic, 1996; Ryder et al., 2004, 2007). Different deficiencies in this collection can cover up to 65% of the *Drosophila melanogaster* genome (release 5.1). I made use of the DrosDel collection for their large deletion sizes and isogenized background (w^{1118}). These deficiencies were constructed in isogenic background (w^{1118}) and do not contain any additional mutations, which reduces the likelihood of epistatic interactions between mutations carried on the deficiency chromosome, and not the deficiency itself. Previously there has been one study focusing on the effects of different deficiencies from the DrosDel kit on the free-running period of the locomotor activity rhythms of *Drosophila melanogaster*; more specifically, they investigated the effects of deficiencies affecting developmental time from a previous screening on the free-running period (Takahashi et al., 2013).

In this chapter, I aimed to identify deficiencies, and the genes within, which affect the chronotype of *Drosophila melanogaster* eclosion rhythm. The chronotypes were measured by comparing the mean phases of emergence of the deficiency lines with their controls as described in the *materials and methods* section. Among the 114 lines I screened, I found 10 of them to be of significantly different chronotypes compared to their controls.

A.

Stock Id		Del seg	Stock Id		Del Seg	Stock Id		Del Seg	Stock Id		Del Seg	Stock Id		Del Seg
X			2L			2R			3L			3R		
Bloomington	Kyoto		Bloomington	Kyoto		Bloomington	Kyoto		Bloomington	Kyoto		Bloomington	Kyoto	
ED6443	150167	1B14-1E1	ED62	150002	21D1-21E2	ED1552	150410	42A11-42C7	ED4079	150419	61A5-61B1	ED5071	150448	82A1-82E4
ED6521	150168	1E3-1F4	ED105	150060	21E2-22A1	ED1673	150123	42E4-43D3	ED4177	150229	61C1-61E2	ED5147	150143	82E8-83A1
ED6558	NA		ED7762	150200	22A6-22D3	ED1725	150026	43E4-44B5	ED4287	150422	62B4-62E5	ED5197	150452	83B7-83D2
ED6574	150169	2E1-3A2	ED136	150376	22F4-23A3	ED1770	150510	44D8-45B4	ED4341	150321	63F6-64B9	ED7665	150259	84B4-84E11
ED411	150291	3A3-3A8	ED4559	150140	23C4-23F6	ED2098	150029	47A7-47C6	ED210	150380	64B9-64C13	ED5296	150241	84F6-85C3
ED6630	150343	3B1-3C5	ED250	150004	24F4-25A7	ED2219	150031	47D6-48B6	ED211	150276	65A9-65B4	ED5429	***150525***	85D19-85F8
ED6712	150172	3D3-3F1	ED385	150010	26B1-26D7	ED2308	150033	49D3-49E7	ED4408	150426	66A22-66C5	ED5518	150153	86C7-86E13
ED6716	150173	3F2-4B3	ED441	NA		ED2354	150034	50E6-51B1	ED4421	150230	66D12-67B3	ED5559	150243	86E11-87B11
ED6727	150344	4B6-4D5	ED489	150296	27E4-28B1	ED2436	***150035***	51F11-52D11	ED4457	150428	67E2-68A7	ED5612	***150159***	87C7-87F6
ED6829	150175	5C7-5F3	ED12527	150364	28C4-28D3	ED3181	150514	53C9-53F10	ED4475	NA		ED5642	NA	
ED6906	150178	7A3-7B2	ED647	150016	29E1-29F5	ED3683	150039	55C2-56C4	ED4486	150432	69C4-69F6	ED5705	150244	88E12-89A5
ED6957	150345	8B6-8C13	ED690	150091	30B3-30E4	ED3791	150417	57B1-57D4	ED4502	***150433***	70A3-70C10	ED5780	150045	89E11-90C1
ED6991	150181	8F9-9B4	ED8142	150202	31E1-32A4	ED3943	150129	57F10-58D7	ED217	150072	70F4-71E1	ED5815	150464	90F4-91B8
ED7005	150346	9B1-9D3	ED775	150397	33B8-34A3	ED4071	150418	60C8-60E8	ED223	150280	73A1-73B5	ED5938	150247	91D4-92A11
ED7161	150184	11A1-11B14	ED784	150399	34A4-34B6	ED1874	NA		ED4685	150436	73D5-74E2	ED6025	150466	92A11-92E2
ED7170	150186	11B15-11E8	ED793	150400	34E4-35B4	ED3728	150127	56D10-56E2	ED224	***150281***	75B1-75C6	ED6058	150251	93D4-93F6
ED7225	150257	12C4-12E8	ED1054	150306	35B10-35D4				ED229	NA		ED6085	***150049***	93F14-94B5
ED7294	150351	13B1-13C3	ED1161	150312	36A10-36C9				ED4978	150442	78D5-79A2	ED6096	150467	94B5-94E7
ED7331	150191	13C3-13F1	ED1196	150111	36E6-37B1				ED230	150073	79C2-80A4	ED6187	150051	95D10-96A7
ED7364	150258	14A8-14A6	ED1202	150315	37B1-37C5				ED4288	***150423***	63A6-63B7	ED6232	150052	96F10-97D2
ED7374	150193	15A1-15E3	ED1303	150407	37E5-38C6				ED4543	***150434***	70C6-70F4	ED6265	150471	97E2-98A7
ED447	150294	17C1-17F1	ED1378	150118	38F1-39D2				ED4858	150232	76D3-77C1	ED6316	150475	99A5-99C1
ED7635	150197	19A2-19C1	ED1473	***150023***	39B4-40A5				ED4786	150439	75F7-76A5	ED6361	150478	100C7-100E3
ED7664	150198	19F1-19F6	ED334	***150007***	25F2-26B2				ED4483	150431	69A5-69D3	ED5644	150163	88A4-88C9
ED7067	150347	10B8-10C10	ED548	150084	28E1-28E9				ED4470	150429	68A6-68E1	ED6220	150253	96A7-96C3
ED7147	150183	10D7-11A1	ED578	150086	28F1-29A3							ED6346	150477	100A5-100B1
			ED729	150095	31B1-31D7									

B.

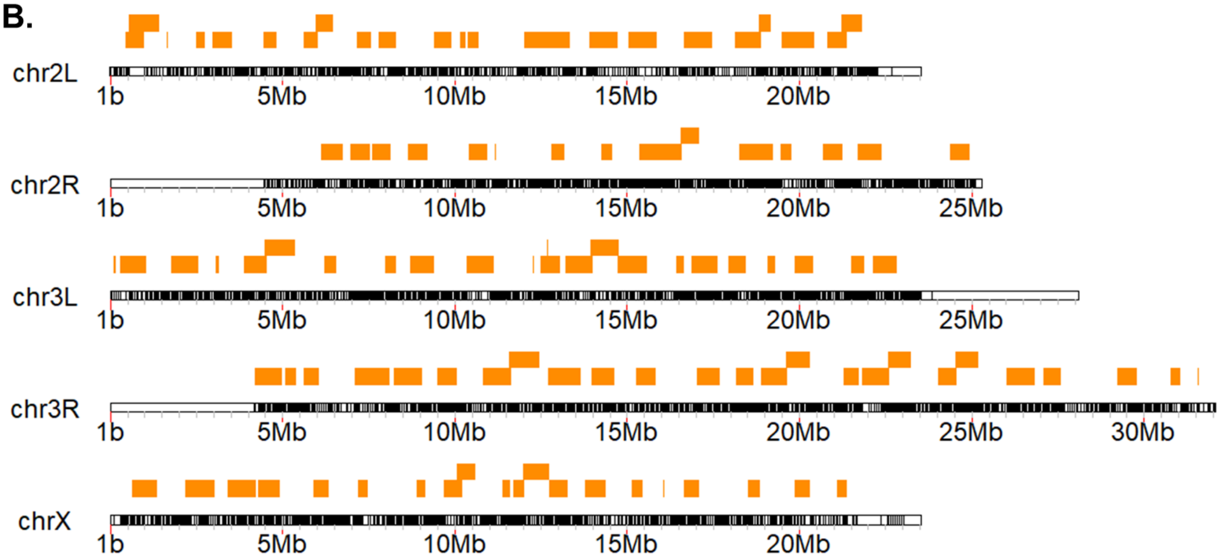


Figure 5.1: Description of deficiency lines chosen for screening for advanced/delayed phase of eclosion rhythm. A. Identifiers of flies chosen for screening from Bloomington and Kyoto stock centers (BDSC and Kyoto-DGRC), the deleted segment, and the chromosome the deficiencies are present. Italicized and bold identifiers are deemed as lines with advanced/delayed phases compared to their background control, as indicated by the results of my study. **B.** Visualization of spans of the deleted segments in different deficiency lines screened across different chromosomal arms.

5.2 Materials and methods:

5.2.1 Deficiency line procurement and fly husbandry:

In a first attempt, I chose a total of 114 deficiency lines from the DrosDel core kit keeping the major criteria of the deletions being broad and non-overlapping. These 114 deletions cover ~55% of the euchromatin (converted to *Drosophila melanogaster* genome release 6 – *Fig. 5.1B*). These 114 deficiency lines (tabulated in *Fig. 5.1A*) were procured from Kyoto Drosophila Genome Resource Center (DGRC – https://www.dgrc.kit.ac.jp/) in batches from 2017-2020. The genomic regions deleted in these lines varied widely – from 0.003Mb to 1.24Mb. The median deficiency size was 0.53Mb. All lines are here onwards referred to in this chapter by the last three digits of their Kyoto DGRC identifier (tabulated in *Fig. 5.1A*). Further details of these lines can be found at https://drosdel.org.uk/. These lines were screened for their mean phase and consolidation of the emergence rhythm compared to their background genetic control (line 534). Each line was cultured in multiple vials, and a total of 700-800 flies were collected in cages (described in chapter 2). Eggs were collected one day after providing the cage populations with yeast paste, and ~100 eggs were put into 6-10 vials (depending on the number of eggs obtained). Most of the lines did not have sufficient egg output, so they were cultured for months in 20-30 vials simultaneously to achieve a greater number of flies.

5.2.2 Eclosion assay, analyses, and statistics:

Eclosion assay was carried out for at least 2-3 days, as mentioned in chapter 2, under ~70 lux LD12:12 and 25±0.5°C. The 2-hour resolution emergence data were first averaged over cycles for a vial and then averaged over vials for a line. I analyzed this data to detect rhythmicity and entrainment. Detection of rhythmicity was achieved through JTK-cycle (Hughes et al., 2010) employed in MetaCycle2d (Wu et al., 2016), keeping type-I error rate as 5%. Criteria for entrainment was $T_{observed}$ must equal to $T_{environment}$ ($T_{observed}$ is the observed period of the eclosion rhythm, and $T_{environment}$ is the duration of the light/ dark cycle – in this case, 24 hours). $T_{observed}$ was calculated with JTK-cycle employed in MetaCycle2d with percentage eclosion data for each vial. Vials showing a "JTK_pvalue" less than 0.05 were considered to be rhythmic, and among the rhythmic vials, the ones with "JTK_period" within a range of ±1 hour of $T_{environment}$ were considered to be entrained. This method of identifying entrainment does not exclude the possibility of detecting any masking responses. The emergence data was then converted into frequencies (as percentage emergence throughout 24 hours) and polar coordinates in Microsoft Excel 365 using custom macro considering ZT00 as 0°/0 radian. I derived the phase of the center of mass of the rhythm (Ψ_{CoM}) by calculating mean phase of emergence (θ in polar coordinate system) and mean consolidation of emergence (R in polar coordinate system – also, the normalized amplitude of the rhythm) as Ψ_{CoM} captures all characteristics of the emergence waveform and has been used previously to describe features of phase divergence in the eclosion rhythm (Abhilash et al., 2019; Ghosh et al., 2021). I pooled data for the background genetic control (line 534) over multiple cycles, vials, and experiments and derived ±3SD of θ and R values for the Ψ_{CoM} of line 534. This ±3SD of θ and R for the Ψ_{CoM} values served as a predefined empirical threshold (confidence cone and confidence donut; *Fig. 5.3*) for choosing deficiency lines (Barde and Barde, 2012; Grafarend, 2007). For each line, parameters of Ψ_{CoM} were calculated, and a custom R (version 4.0.2) code was used to identify "hits" whose θ and R values lay outside the predefined threshold (R Core Team, 2020), and Ψ_{CoM} was deemed to be advanced/delayed compared to line 534. The "hits" were at least assayed two times independently in different experiments, and data was pooled. Deficiency spans along chromosomal arms were plotted using the karyoploteR package (Gel, 2020) in R (version 4.0.2), polar plots were plotted using the Plotly graphing library (Inc., 2015) in R (version 4.0.2), and emergence profiles were plotted using GraphPad Prism 8 (R Core Team, 2020). For the "hits", I reassigned the deleted segments to *Drosophila melanogaster* reference genome (release 6) using FlyBase FeatureMapper and identified genes present in those regions. Further gene ontology (GO) analysis and KEGG pathway enrichment analyses of those identified genes were done using g:Profiler (Reimand et al., 2016). Network analysis of protein-protein interaction was carried out in STRING (https://string-db.org/) (Szklarczyk et al., 2021). Identified genes in the 10 "hits" are presented in *Appendix table A5.1*.

5.3 Results:

The goal of this study was to identify additional genes responsible for fine-tuning of the phase of emergence rhythm in *Drosophila melanogaster*. As these lines have large genomic regions deleted in their chromosomes, if any deletion covers any genes that are essential for the generation of the eclosion rhythm itself, the line would be arrhythmic. So, overall, I expected to detect deletions in genes involved in the regulation of phases of the eclosion rhythm directly or indirectly under an LD12:12 cycle but do not affect the rhythmicity or entrainment themselves.

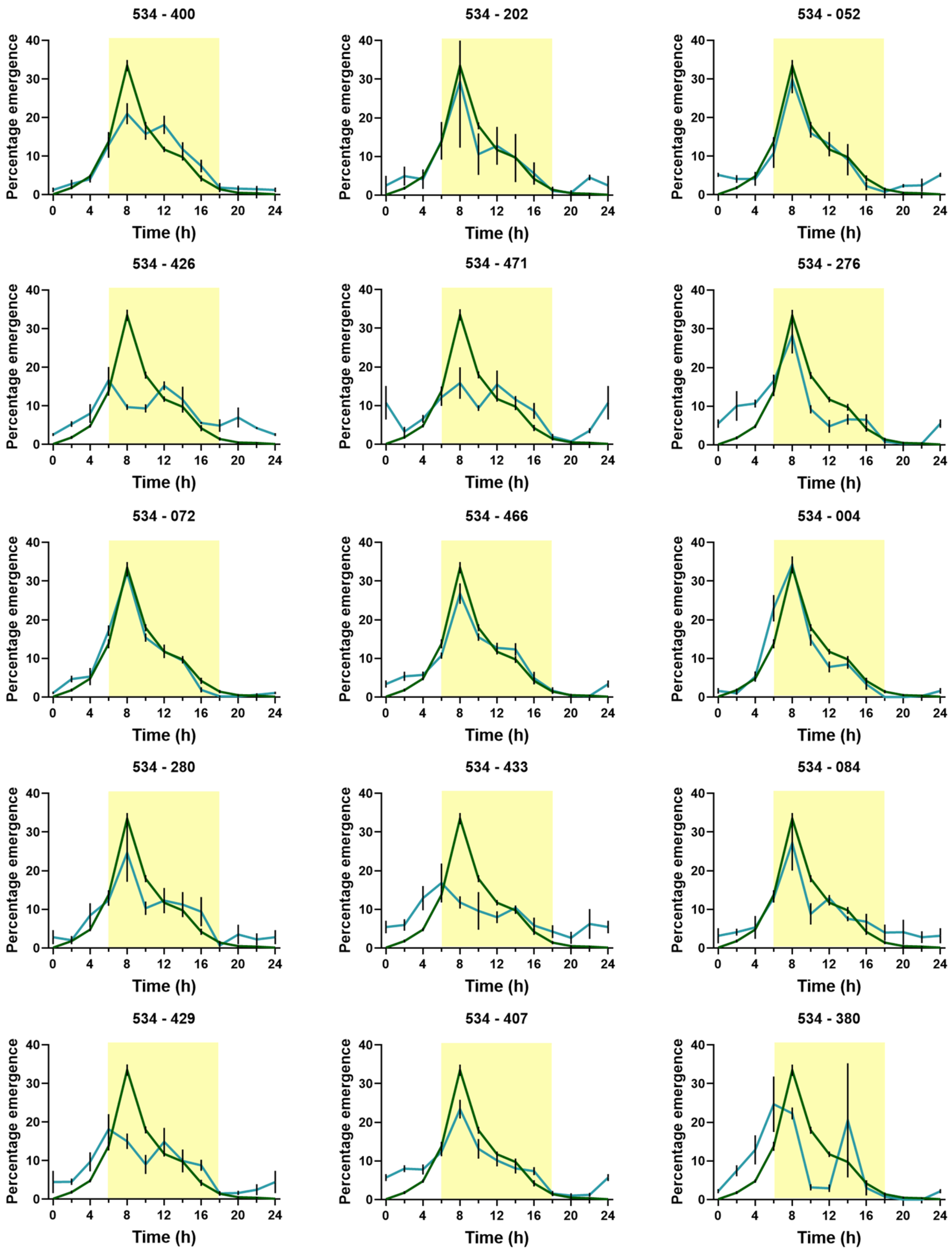


Figure 5.2: Average emergence profile of representative rhythmic deficiency lines. Among 40 rhythmic deficiency lines, 15 are plotted for representation purposes. The dark green line is the background control and blue

lines are individual deficiency lines. The title of each plot depicts the identifiers of strains – the first is the background control identifier (534–same in all), and the second, after the "-" is the deficiency line identifier. Error bars are ±SEM. Yellow shaded region depicts 12 hours of light.

5.3.1 Detection of rhythmicity and entrainment in deficiency lines:

Among the 114 deficiency lines chosen for the eclosion assay, I found only 40 showed significant rhythmicity and entrainment to an LD12:12 cycle. A few (~10) more lines showed significant rhythmicity using the threshold mentioned above in JTK-cycle analysis but were deemed to be not entrained as their period values did not match the environmental period value according to the predefined threshold. I present the results for a total of 25 of these lines, which were rhythmic and entrained for the eclosion rhythm in this chapter (*Fig. 5.2 & 5.5*). Figure 5.2 depicts 15 lines (400, 202, 052, 426, 471, 276, 072, 466, 004, 280, 433, 084, 429, 407, and 380) which were rhythmic and entrained but not deemed to have advanced/delayed Ψ_{CoM} compared to their background genetic control – line 534. Some of these lines have a high amplitude, and unimodal emergence rhythm (lines 202, 052, 276, 072, 466, 004, 084 – *Fig. 5.2*). Some other lines (lines 400, 426, 471, 433, 429, 380 – *Fig. 5.2*) showed low amplitude, often bimodal rhythm, reducing their normalized amplitude (R) and giving rise to a mean phase of emergence (θ) close to that of their control (line 534). Overall, the difference in the emergence waveform varied widely from their control.

5.3.2 Identification of "hits" – lines showing significantly advanced/delayed phase of emergence:

With the strict predefined thresholds, I identified 10 lines (lines 023, 525, 007, 281, 433, 423, 035, 159, 434, and 049; *Fig. 5.4, 5.5 & 5.6*) which had significantly more advanced or delayed mean phase of emergence (Ψ_{CoM}) compared to their background genetic control (line 534). Among these lines, 281 (θ: 5.38, R: 0.54), 035 (θ: 22.54, R: 0.75), 049 (θ: 32.11, R: 0.74), and 434 (θ:

21.76, R: 0.4) were shown to have significantly more advanced mean phase of emergence (Ψ_{CoM}) than line 534 (θ: 48.9, R: 0.64; $3SD_{\theta}$: 10, $3SD_{R}$: 0.094) (*Fig. 5.4 & 5.5*). Lines 423 (θ: 81.9, R: 0.4), 159 (θ: 61.7, R: 0.78), 525 (θ: 193.86, R: 0.37), 007 (θ: 65.8, R: 0.52), 433 (θ: 101.8, R: 0.41), and 023 (θ: 94.32, R: 0.5) were shown to have significantly more delayed Ψ_{CoM} compared to that of line 534 (*Fig. 5.4 & 5.5*).

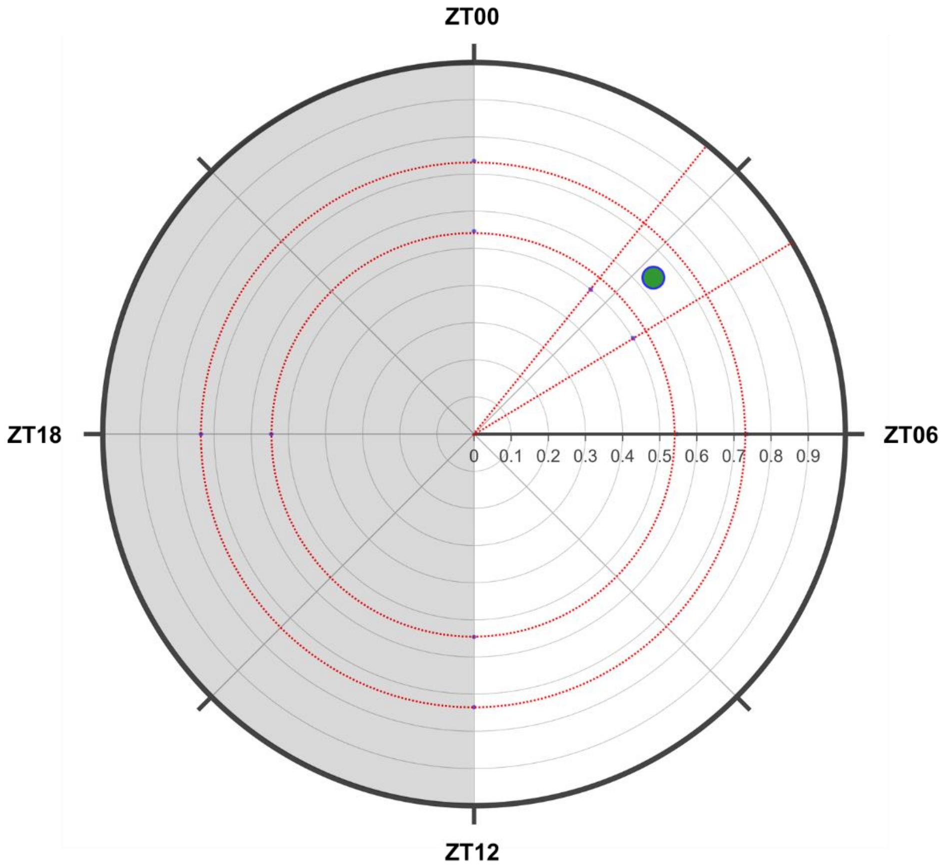


Figure 5.3: Calculation of thresholds for screening. The green filled circle depicts the phase of the center of mass (Ψ_{CoM}) of the background genetic control flies. The distance from the center depicts the consolidation of emergence (R value in polar coordinate system), and the angle it makes with the ZT00 (0° in polar coordinate system) mark depicts the mean phase angle of emergence (θ in polar coordinate system). The Ψ_{CoM} value for the background genetic control flies was calculated over multiple vials, days, and experiments. I chose a ±3SD threshold for both θ (red dotted lines from the origin – forming the confidence cone) and R (red dotted circles around the origin – forming the confidence donut) for screening for advanced/delayed phase of emergence. Dark shaded region depicts 12 hours of darkness.

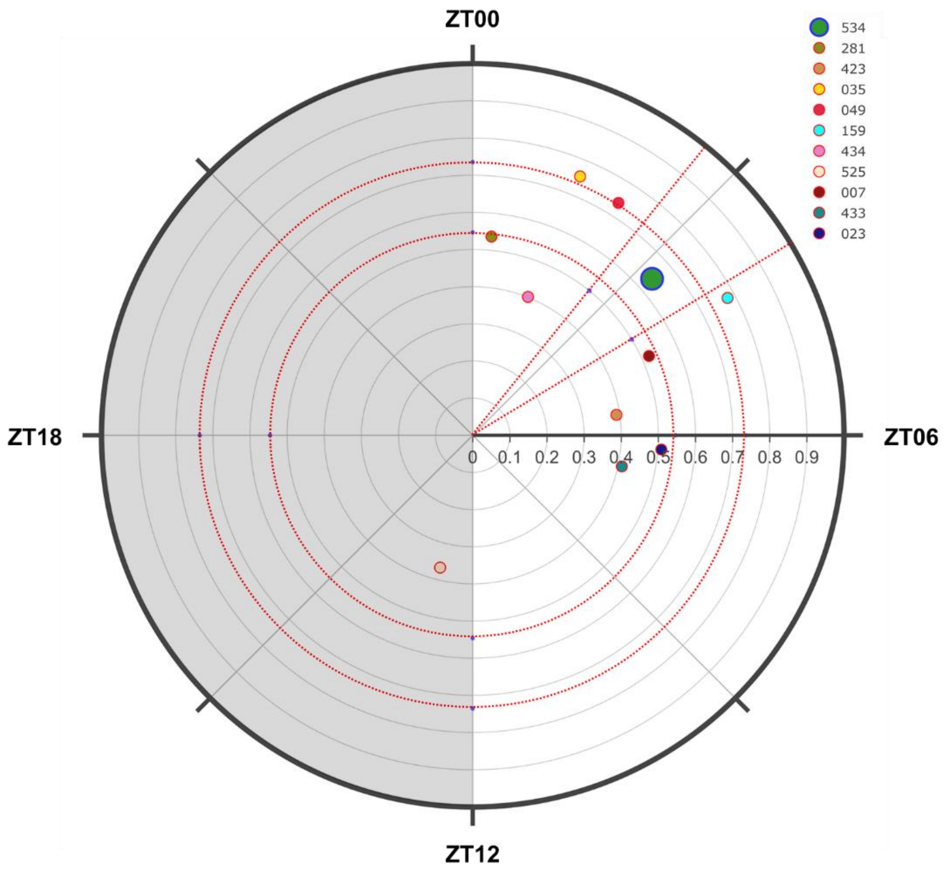


Figure 5.4: Identification of phase advanced/delayed lines based on set thresholds. 10 lines were identified to have advanced (4 lines) and delayed (6 lines) Ψ_{CoM} based on the thresholds set beforehand. The legend guide on the right depicts the line identifiers matching with the color codes used.

5.3.3 Description of the early chronotype (advanced Ψ_{CoM}) lines:

Line 281 showed the peak of their emergence at the same time as line 534 did, but their emergence started much before the lights-ON signal (>30% emergence before 6 hours of lights-ON), and thus can be categorized as a true early chronotype (*Fig. 5.4 & 5.5*). Line 035 showed an advancement of the whole emergence profile along with their peak of emergence, which happens before lights-ON (*Fig. 5.4 & 5.5*). Their consolidation of emergence is also significantly higher than their controls, driven by quick dissipation of emergence after lights-ON (*Fig. 5.4 & 5.5*). This line can also be considered a true early chronotype. Line 049 showed their peak of emergence at the same time the controls did, but a larger percentage of them emerged before lights-ON and emergence stopped before controls did, contributing to an advanced Ψ_{CoM} compared to the controls (*Fig. 5.4 & 5.5*). Though the differences were marginal but still statistically significant, I consider this line as a moderately early chronotype. Line 434 showed considerably high emergence at night, more during late night (*Fig. 5.4 & 5.5*). Their peak of emergence occurs at the same time as of controls, but emergence drops immediately after the peak occurs, and there is very low emergence throughout the day, thus giving rise to a significantly advanced Ψ_{CoM} compared to the controls (*Fig. 5.4 & 5.5*). This line can also be considered a true early chronotype.

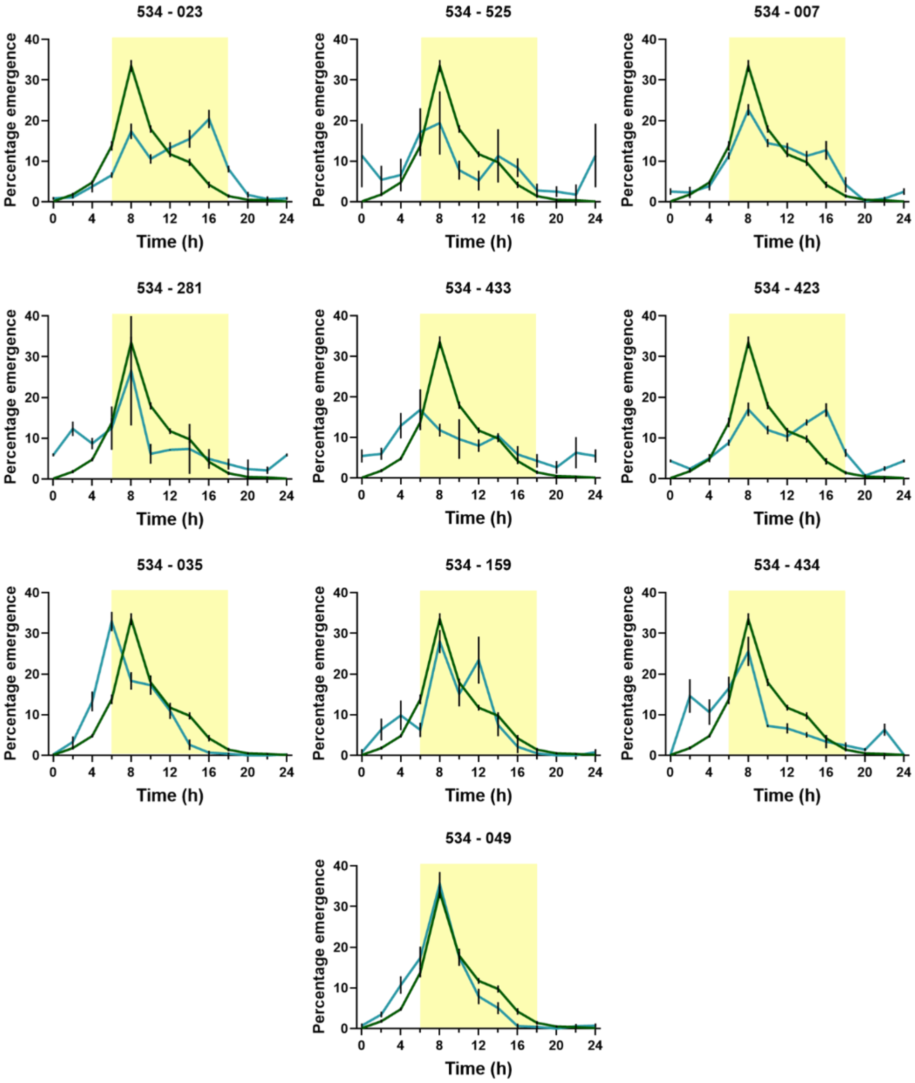


Figure 5.5: Average emergence profiles of identified phase advanced/delayed lines along with their background genetic control flies. The dark green line is the background control and blue lines are individual deficiency lines. The title of each subplot depicts the identifiers of strains – the first is the background control identifier (534–same in all), and the second, after the "-" is the deficiency line identifier. Error bars are ±SEM.

5.3.4 Description of the late chronotype (delayed Ψ_{CoM}) lines:

Line 159 showed their peak of emergence at the same time as of controls, but they had substantially higher emergence later in the day (~ZT6), contributing to their Ψ_{CoM} being significantly delayed compared to the controls (*Fig. 5.4 & 5.5*). Restricting the majority of their emergence between ZT2-6 also helps them achieve a higher consolidation of emergence (*Fig. 5.4 & 5.5*). I considered line 159 to be of true late chronotype as they showed higher emergence in the latter part of the day. Line 007 showed high emergence throughout the day, and considerably higher emergence in the latter part of the day, though their phase of peak emergence coincides with the phase of peak emergence of the controls (*Fig. 5.4 & 5.5*). Distributed emergence throughout the day leads to significantly lower consolidation in this line, and higher emergence at the later part of the day compared to controls leads to significantly delayed mean phase of emergence (*Fig. 5.4 & 5.5*). Thus, line 007 can also be considered a true late chronotype. Line 423 showed a bimodal emergence pattern, having the majority of the emergence at the later part of the day and low amplitude of the emergence rhythm (*Fig. 5.4 & 5.5*). This line had significantly delayed phase of emergence and significantly lower consolidation of emergence compared to the controls, and thus, it can be considered to be of true late chronotype. Line 023 also showed a bimodal emergence pattern, but the peak at ZT10 was higher than that in the morning (ZT2) (*Fig. 5.4 & 5.5*). This line had most of the emergence occurring at the end of the day and thus showed significantly delayed phase of emergence, and because their emergence was distributed throughout the day, they had significantly lower consolidation of emergence (*Fig. 5.4 & 5.5*). I considered line 023 to be of true late chronotype. Line 433 had a low amplitude eclosion rhythm, and emergence was spread throughout the whole day, where high emergence was observed before lights-ON, during late night (*Fig. 5.4 & 5.5*). Due to their high emergence during nighttime and owing to the fact that mean phase of emergence calculation takes place taking ZT00 as origin, line 433 showed a significantly delayed phase of emergence and very low consolidation of emergence compared to their controls (*Fig. 5.4 & 5.5*). As the delayed phase of emergence is not immediately evident from their emergence profile and can as well be an artifact due to the way the mean phase of emergence calculation was done, I consider line 433 to be of moderately late chronotype. Line 525 showed a bimodal emergence profile, showing high emergence before lights-ON, immediately after lights-ON, and also just before lights-OFF (*Fig. 5.4 & 5.5*). Interestingly, this line also showed high emergence at midnight (ZT18) (*Fig. 5.4 & 5.5*). This high midnight emergence, coupled with high emergence at the end of the day and before lights-ON, leads to their extremely delayed phase of emergence (mean phase of emergence lying in the dark part of the day) and low consolidation of emergence compared to the controls (*Fig. 5.4 & 5.5*). I considered this line to be of paradoxical late chronotype.

5.3.5 Identification of genes affected by the deficiencies and their characterization:

I identified ~595 unique genes present in the genomic segments deleted in the 10 “hits” (*Fig. 5.6*). When I did a GO enrichment analysis, I found these genes were significantly enriched (FDR corrected *p* value = 0.05) in biological processes like nucleosome and chromatin assembly, chromosome and chromatin organization, chromatin and gene silencing, fatty acid elongation, larval development, negative regulation of gene expression, DNA-templated transcription, etc. In the KEGG pathway enrichment analysis, though there was not any significant enrichment of any pathway, affected pathways included spliceosome, mitophagy, RNA degradation, mTOR signaling pathway, glycolysis/gluconeogenesis, etc. The genes also showed significantly more interactions than expected (p value = 0.0148) in a protein-protein interaction network analysis. The high number of edges in the network (553) and higher average node degree (2.26) suggest that the selected genes show high degree of interaction among themselves.

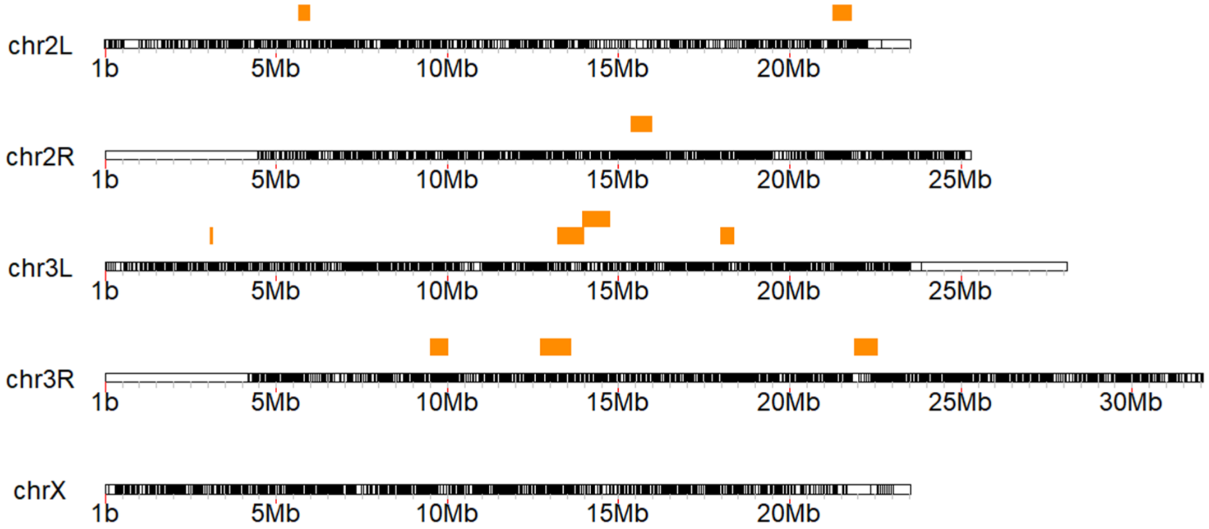


Figure 5.6: Visualization of spans of the deleted segments identified phase advanced/delayed lines. The identified deficiency lines are present over four major chromosomal arms (2L, 2R, 3L, and 3R). They span over ~595 unique genes, and these genes are majorly involved in nucleosome and chromatin assembly, chromatin and gene silencing, ecdysone induced genes, etc.

5.4 Discussion:

In this chapter, I have screened broad, non-overlapping, heterozygous deficiencies in different chromosomal arms generated by Ryder and colleagues (Ryder et al., 2004, 2007) for chronotype variation. The DrosDel kit has proved to be an indispensable resource for Drosophila genetics research over the years and helped many discoveries related to – small molecule discovery for autophagy enhancers (Sarkar et al., 2007), apoptosis protection (Ravikumar et al., 2006), activation of early zygotic genome (Liang et al., 2008), manipulation of genes in the nervous system of flies (Venken et al., 2011), mechanism for treatment of Huntington’s disease (Sarkar et al., 2008), gene disruption projects (Bellen et al., 2011), database curation (Lyne et al., 2007), creation of other deletion strategies (Bateman et al., 2006), study of the *Minute* loci (Marygold et al., 2007), study of aggression and toxicity in Huntington's disease (Ravikumar et al., 2008), identification of sleep promoting factors (Koh et al., 2008), heterochromatin formation (Rudolph et al., 2007), GPCR characterization (Lear et al., 2005), identification of sleep arousal and dopamine arousal pathways (Pfeiffenberger and Allada, 2012), sleep and immunity functions (Toda et al., 2019), PDF neurons promoting sleep (Chung et al., 2009), etc. To the best of my knowledge, there has been only one study till now using the DrosDel kit directly screening for circadian phenotypes (free running period) with respect to developmental time (Takahashi et al., 2013).

I have successfully screened 114 deficiencies spread over four major chromosomal arms (two autosomes and the X chromosome – *Fig. 5.1A & B*) and identified 10 deficiencies which significantly alter the mean phase of emergence (Ψ_{CoM}) of the eclosion rhythm compared to their background genetic control (line 534 – an isogenized w^{1118} line; *Fig. 5.3 & 4*). All these deficiencies were heterozygous deficiencies, allowing me to identify the dominant effects of genes on this phenotype. Among the 10 lines identified, the largest deletion was of 0.93Mb (line 159 – *Df(3R)ED5612*, 130 genes), and the smallest deletion was of 0.08Mb (line 423 – *Df(3L)ED4288*, 19 genes) (*Fig. 5.4, 5.5 & 5.6*). The average deletion size of the 10 selected lines was 0.58Mb (*Fig. 5.6*). 4 lines were deemed to be of early chronotype according to the predefined thresholds – lines 281 (*Df(3L)ED224*), 035 (*Df(2R)ED2436*), 049 (*Df(3R)ED6085*), and 434 (*Df(3L)ED4543*) (*Fig. 5.4 & 5.5*). 6 lines were deemed to be of late chronotype – lines 423 (*Df(3L)ED4288*), 159 (*Df(3R)ED5612*), 525 (*Df(3R)ED5429*), 007 (*Df(2L)ED334*), 433 (*Df(3L)ED4502*), and 023 (*Df(2L)ED1473*) (*Fig. 5.4 & 5.5*). The early chronotype lines had genes deleted which are related to positive regulation of execution phase of apoptosis, eye pigmentation, detection and response to stimulus, response to light stimulus, rhodopsin biosynthesis, detection and response to chemical stimulus, indirect involvement with circadian rhythm, locomotory behavior, learning and memory, startle response, etc. according to GO analysis. Whereas late chronotype lines had genes deleted which are related to ion transport, peptide transport, endocytosis, response to light stimulus, phototaxis, gene expression, RNA splicing, mRNA splicing via spliceosome, 4 genes known to directly affect circadian rhythms (*Ras85D*, *Dh44*, *Ac3*, *Kdm5*, *timeout*), molting cycle, 2 genes known to directly affect sleep (*CalR*, *Men*), aggressive behavior, defense response to bacteria and fungi, chromosome and nucleosome organization, histone modifications, mating and reproduction, mTOR signaling pathway, etc. according to GO and pathway analyses.

One gene, *Eip93F*, was common between this screen and the pooled-sequencing analysis from chapter 4 for early chronotype. *Eip93F* encodes a DNA-binding protein which is important as an adult determinant during fly metamorphosis (FlyBase), and its plausible role in the emergence rhythm has been discussed in chapter 4. While two genes, *Ugt37A2* and *Ugt37A3* were common between both approaches in late chronotype. These two neighboring genes are predicted to enable UDP-glycosyltransferase activity and are expressed in the adult head. Human orthologues of these genes are implicated in several diseases, like alcoholic pancreatitis, bilirubin metabolic disorder, etc. A close relative and neighbor of these genes, *Ugt35B1*, shows a strong circadian rhythm in flies (Pogue-Geile et al., 2006); however, *Ugt37A2* and *Ugt37A3* do not have any known direct or indirect role in the fly circadian system. Further studies will be needed to establish their roles in the phasing of the emergence rhythm. A piece of cautionary advice will be to not directly correlate the results from the pooled-sequencing study (see chapter 4) and the results from this screening. The study in chapter 4 was performed using large populations with high standing genetic variation, whereas the screening here was done with inbred, isogenized lines containing heterozygous chromosomal deficiencies. Though I identified multiple genes harboring significantly differentiated SNPs in different chronotypes (see chapter 4), right now, the effects of those SNPs on the structure-function of the gene products are highly uncertain (as the analyses used only predictive algorithms), while the deficiency lines screened in this chapter contains confirmed heterozygous deletions of genomic regions.

I categorized different chronotypes as "true" early (lines 281, 035, and 434)/late (lines 159, 007, 423, and 023), "moderately" early (line 049)/late (line 433), and "paradoxical" late (line 525) chronotypes. Though the phase of the Center of Mass (Ψ_{CoM}) is a strong estimator of the average emergence waveform, bimodality, multiple peaks, gate-width of emergence, etc. can affect and cloud inferences drawn solely on the basis of Ψ_{CoM}. Line 433 is an example of this (*Fig. 5.5*). In this line, the emergence appears to be advanced compared to control, as shown by high emergence before lights-ON, but by virtue of how Ψ_{CoM} is calculated (ZT00 is 0°/0 radian) and large gate-width of emergence, line 433 shows a significantly delayed phase of emergence compared to control (*Fig. 5.4 & 5.5*). An example of "paradoxical" late chronotype can be seen in line 525, which showed high emergence at midnight, this high midnight emergence, coupled with the high emergence at late night gives rise to the extremely delayed mean phase of emergence, which may be an artifact from the analysis (*Fig. 5.4 & 5.5*). Nevertheless, as this line showed considerably higher emergence at the later part of the day, I still classify it as a late chronotype.

In conclusion, in this chapter, I have screened 114 lines from the DrosDel core kit (Ryder et al., 2004, 2007) and identified 10 lines that were of early and late chronotypes compared to their controls. These deletions were molecularly mapped and covered ~595 genes throughout two autosomes (Chromosomes 2 and 3). However, I chose some of the broadest and non-overlapping deficiencies from this collection to begin the screening procedure. Due to various limitations imposed during the past 2 years, it has not been possible to carry out finer mapping of the phenotypes using more restricted deletions. It is expected in the future, narrower deletions within the identified deficiencies will be screened from the DrosDel second version and Exelixis kits (Parks et al., 2004; Ryder et al., 2007; Thibault et al., 2004). I expect this screening procedure, when completed, will identify more specific genes controlling the phase of emergence in *Drosophila melanogaster* and be an invaluable resource to the fly chronobiology community.

Chapter 6

Summary, conclusions, and future directions

In this chapter, I will summarize all my results so far (chapters 2-5) and attempt to make a few general comments and discuss avenues for further studies. Though the discussion section of each chapter has future experiments in reference to that chapter, here I will note more specific experiments.

a) In the second chapter, I examined the positive masking response to the lights-ON signal in the *early*, *control*, and *late* populations, in order to understand the difference in Ψ_{ent} among populations and the contribution of non-clock mechanisms in this difference. I hypothesized that our selection protocol has inadvertently resulted in selection for masking, a non-clock phenomenon, in the *early* population due to the temporal placement of our selection window (which includes the lights-ON transition). I designed experiments to delineate the contributions of enhanced masking to light and that of the circadian clock in regulating/mediating enhanced emergence in the morning window of *early* population. Using a series of experiments comprising phase-shift protocols, LD-DD transition, and *T*-cycle experiments, I found that our *early* populations have evolved positive masking, and their apparent entrained phases are largely contributed by masking. Through skeleton *T*-cycle experiments, I found that in addition to the evolution of greater masking, our *early* populations have also evolved advanced Ψ_{ent}. Furthermore, this study systematically outlined experimental approaches to examine the relative contributions of clock versus non-clock control of an entrained behavior. Although it has previously been suggested that masking may confer an adaptive advantage to organisms, here, I provided experimental evidence for the evolution of masking as a means of phasing that can complement clock control of an entrained behavior.

b) Using the experimental paradigms reported in the third chapter, I first established that high light intensity photoperiods can separate the circadian clock controlled peak and the masking peak to lights-ON signal efficiently. Next, I showed that *early* and *control* populations both showed higher masking to the lights-ON signal under long photoperiods compared to the *late* populations, but only *early* populations showed a trend of higher masking with increasing day length, while *control* populations did not show that. I speculated that the masking response under different photoperiods is dependent on the day length and limitation of total activity duration (α). This is shown under short photoperiods where the activity starts much before lights-ON and does not show any masking response at all. I also investigated sleep in our populations under low and high light intensity LD cycles. I showed that higher light intensity during the daytime reduces night sleep in divergent chronotypes compared to the *control* populations. Using high light intensity photoperiods, I also showed that under short day lengths, night sleep amount and quality are adversely affected only in the divergent chronotype populations, compared to the *control* populations. Also, the long photoperiod data suggested that these adverse effects on nighttime sleep in divergent chronotypes can be overcome by using artificial light later in the day, as, under long daylengths, there is no difference in nighttime sleep in divergent chronotypes compared to the *control* population.

c) In the next chapter, chapter four, I analyzed ~100X pooled-sequencing data from all populations (total 12 – *early*, *control*, and *late*, and their four replicates each). The goal of this study was to understand the genetic architecture of these divergent chronotype populations and identify loci that have changed significantly in terms of allele frequencies in either *early* or *late* populations, compared to the *control* populations. Another major goal was to identify signatures of selection in their genomes and to dissect regions where either *early* or *late* populations have undergone more positive selection than the *control* populations. I successfully achieved both goals – identifying genomic regions under more positive selection and identifying loci where divergent chronotypes have undergone significantly more allele frequency change compared to the *control* population. I concluded that *early* and *late* chronotypes are brought about by changes at the genome level, and these changes are at very different loci in both. Overall, I established a database of putative variations associated with divergent chronotypes, which are plausibly natural genetic correlations, and I hope that this work will pave the way to investigate novel targets regulating phase divergence in Drosophila.

d) In order to further understand the genetic basis of chronotype divergence, I took another approach to discover loci affecting chronotype divergence in *Drosophila melanogaster* by screening several lines from the DrosDel collection for differential phasing of the eclosion rhythm. I chose 114 non-overlapping deletions covering ~55% of the euchromatin. Among the lines assayed for eclosion rhythms, only 40 could be entrained to a LD12:12 regime. By using a stringent cutoff of ±3 SD of the phase angle of the mean phase of emergence and normalized amplitude (mean vector length in polar coordinates) of the control background, I found only 10 lines spanning over four major chromosomal arms (2L, 2R, 3L, and 3R) to be different from their background genetic control (early and late chronotype). These 10 lines cover a total of 595 genes and are enriched in genes involved in eye pigmentation, detection and response to stimulus, response to light stimulus, rhodopsin biosynthesis, indirect involvement with circadian rhythm, locomotory behavior, learning and memory, startle response, ion transport, peptide transport, phototaxis, gene expression, RNA splicing, mRNA splicing via spliceosome, 4 genes known to directly affect circadian rhythms, molting cycle, 2 genes known to directly affect sleep, aggressive behavior, defense response to bacteria and fungi, chromosome and nucleosome organization, histone modifications, mating and reproduction, mTOR signaling pathway, etc. I proposed that these genes may act as fine regulators of phases of the eclosion rhythm in *Drosophila melanogaster*, but not have a large role to play in the generation of the eclosion rhythm *per se*.

As discussed in chapters 2 and 3, masking seems to be one major influence on the phase of emergence, at least under simplistic environments with sharp light/temperature transitions. Masking also affects the shape of the waveform and the allowed time duration when flies are allowed to emerge or show activity (called “gate-width” of emergence/activity). Masking and its influence on circadian clock governed behaviors has been historically neglected by circadian researchers, probably fueled by the stunning variety of masking responses observed in different organisms and different behavioral and physiological rhythms, and the lack of molecular/neuronal mechanisms explaining masking. As a field, chronobiology needs greater efforts towards understanding the masking phenomenon, its relationship/regulation by the circadian clock, and the influence masking may exert on the circadian clock. In all probability, the phase of entrainment for different biological rhythms observed in nature or in laboratory, is governed by both masking and the circadian clock. The interaction between them or the hierarchy in which they operate on a particular rhythmic behavior or physiological process is needed to be understood. While identifying or planning interventions with light (e.g., light therapy), it is important to keep in mind that light or any other zeitgeber may also have a masking effect, which, if ignored, may lead to wrong interpretation of the results. This is particularly important in the case of divergent chronotypes, as I have shown in chapters 2 and 3. Just by virtue of differences in masking response, one organism/individual can assume a different phase than others under different environmental conditions, particularly, under more naturalistic conditions, where a multitude of zeitgebers is present. In the context of results from chapter 2, I think the evolution of masking may serve to be advantageous in certain ecological contexts, such that it can help organisms regulate the timing of behavorial and physiological processes with respect to specific zeitgeber(s), which may be more "important" and/or "reliable" depending on the ecological and temporal niche of the organism. The fact that masking can evolve just as circadian clocks do, and has adaptive value, makes it even more important to study and understand.

The following two chapters of my thesis, chapters 4 and 5, deal with the genetic basis of chronotype divergence in fruit flies. Here I showed that there are different loci that have evolved differently in divergent chronotypes, and plausibly, different loci in their genome are responsible for giving rise to different chronotypes. I identified many genes where allele frequencies have significantly changed in *early* and *late* populations compared to the *control* population, and genomic regions undergoing more positive selection in the divergent chronotypes. As chronotypes are highly quantitative traits, and as I do not identify many loci predicted to have large effects on the gene product's structure-function, it leads me to believe that selection acts on the standing genetic variation in a population and the interaction of multiple loci give rise to different chronotypes in the population. When researchers identify large effect mutations in inbred lines affecting chronotypes, those are most likely to be non-existing in nature and do not reflect a natural genetic correlation between the loci and the phenotype. My studies on these large populations of fruit flies with high standing genetic variations give me confidence in the natural genetic correlations between the identified loci and the phenotypes. The screening described in chapter 5 illustrates the power of *Drosophila melanogaster* as a powerhouse for genetics. Using mapped heterozygous chromosomal deficiencies, I identified 595 genes whose heterozygous deletion leads to a significantly advanced or delayed phase of emergence. I speculate that these genes are important for fine-tuning the Ψ_{ent} of the eclosion rhythm in Drosophila.

To advance the propositions of this thesis, many future avenues may be pursued. I list a few specific experiments/directions here:

1) A very important next step will be to characterize the neuronal and molecular basis of the masking phenomenon, particularly masking to light cues. As mentioned in chapter 2, eclosion hormone expressing neurons control this masking response to certain extent. How the strength of masking is determined may involve stronger/weaker coupling between light sensitive modalities and these neurons. In the case of eclosion rhythm, the circadian/non-circadian photosensitivity of the neuronal connections to prothoracic gland (Prothoracicotrophic Hormone neurons – PTTH neurons; containing Short Neuropeptide F – sNPF) itself may play some role and must be investigated. For the locomotor activity rhythm, it is necessary to observe the molecular clock in the major groups of clock cells under different photoperiods and different light intensities. It is imperative to find out the parametric effect of light on the masking phenomenon. Along this line, experiments including tracking the masking peaks under different light intensity skeleton photoperiods of different pulse duration will prove to be valuable. Masking responses may be elicited by different brain structures and involve a different group of molecules for different behaviors, so it is important to investigate the neuroanatomical and molecular basis of masking in different rhythms separately. The molecules or cells important for masking in the eclosion rhythm may be related to cells expressing metamorphosis related genes and

light sensitive genes, while in case of the locomotor activity rhythms, the cells responsible may be the ones related to locomotion directly, and again, cells expressing light sensitive genes.

2) The preliminary sleep differences reported in my thesis are, in my opinion, just the tip of the iceberg of what we could learn about the evolution of sleep in divergent chronotypes from our populations. The differences in sleep parameters under high light intensity and different photoperiods illustrate similarities with sleep aberrations in divergent chronotype humans. It will be very important to study if the differences in the divergent chronotypes are due to their different circadian locomotor activity patterns or due to changes in their sleep homeostat. Along these lines, sleep induction and sleep disruption experiments with our populations will shed light. Also, all my experiments were done with virgin male flies; these experiments should be repeated with virgin females, mated males, and mated female flies.
3) The SNPs and genes I associated with divergent chronotypes should serve as a repertoire of information on natural alleles which are genetically correlated with divergent chronotypes in *Drosophila melanogaster*. I have only used the SNPs which were common to all four replicates, whereas there are significantly different alleles in all replicates separately, which were not investigated further. These SNPs may have been byproducts of genetic drift, but still may lead to divergent chronotypes and are certainly strongly correlated with the chronotype divergence observed in our populations and should be characterized further. Recreating specific high effect and significant SNPs in an isogenized background using advanced genetic engineering techniques may reveal the roles of high effect SNPs/genes in creating divergent chronotypes.

4) The screening performed in chapter 5 yielded 10 lines with large chunk of the genome deleted in them. All genes covered in these deletions will certainly not be involved in chronotype divergence. Further narrower deletion screening with molecularly mapped lines will lead to better identification of causal genes related to chronotype divergence and should be conducted.

The generalizability of the conclusions of my thesis will depend on the reproduction of similar experiments with different organisms and validity of the hypotheses while investigating further into mechanisms behind these observations. However, I expect my conclusions regarding masking and sleep in divergent chronotypes to broadly hold up and the analysis of the genetic elements to be at least reproducible in other insect species. It will even not be surprising if orthologs of certain genes and SNPs therein turn out to be important for similar phenotypes in other organisms, even humans. I hope readers of this thesis will take away a few semi-philosophical messages – “Do not ignore the component of masking in your study of rhythms”, “Genetic variations affecting phenotypes are very different in case of populations than individuals and will give insights that may surprise you”, and last but not the least “No matter what data you get, negative or positive, you try to explain them as best as you can”.

Appendix 1

VANESSA – Shiny apps for accelerated time-series analysis and visualization of Drosophila circadian rhythm and sleep data

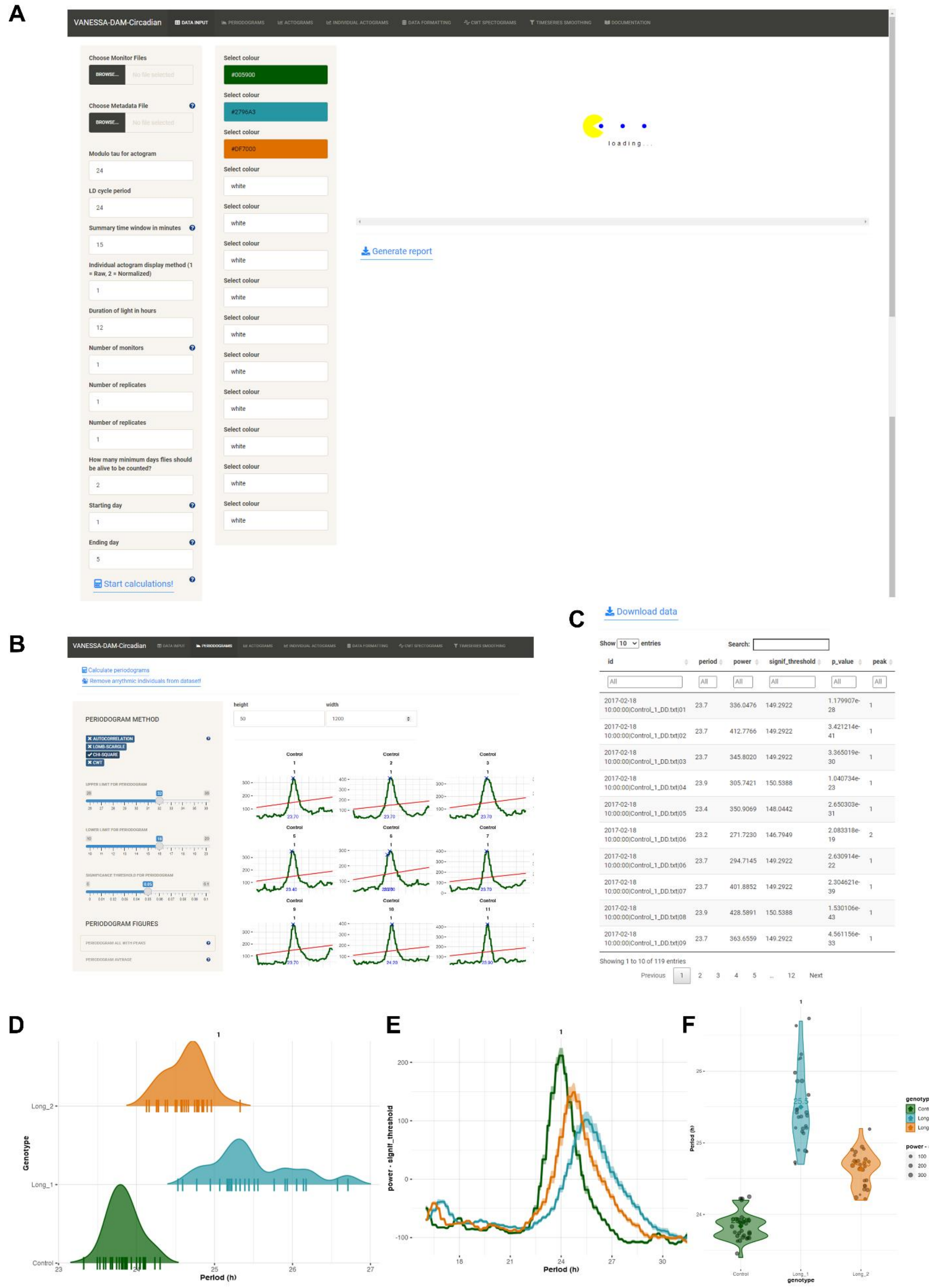
This section has been accepted for publication in Journal of Biological Rhythms as “VANESSA – Shiny apps for accelerated time-series analysis and visualization of Drosophila circadian rhythm and sleep data, Ghosh and Sheeba, 2022, JBR”.

Timeseries analysis involves the evaluation of sequential data obtained over time – data collected at specific intervals, where time is an independent variable. Circadian rhythm and sleep researchers are mostly focused on events occurring at a *circa-dian* (~approximately a day) scale. Since the proliferation of circadian rhythm research in the late 1970's, multiple methods have been developed to decipher patterns in data collected to study rhythms in different organisms. Mostly these researchers wish to estimate periodicity, phase, and robustness of the rhythms; whereas sleep researchers are interested in sleep architecture, sleep parameter estimation, and latency to sleep. Rhythm parameters such as presence or absence of rhythmicity, period, and amplitude tell us about the robustness of the rhythm and whether a periodicity of circadian time scales is present or not. Sleep parameters such as total sleep time, bout numbers, bout lengths, latency tell us the amount, quality of sleep, and overall sleep architecture. Calculation of these parameters enable us to compare individuals under different conditions, genotypes, treatment groups etc., and let us uncover underlying mechanisms.

Over the past five decades, *Drosophila melanogaster* has emerged as a widely used model to study circadian rhythms and sleep. The *Drosophila* Activity Monitor (DAM) systems from Trikinetics (https://www.trikinetics.com/) are the most often used systems for automated recording of *Drosophila* locomotor activity data for circadian rhythm and sleep research. Over the years various free and paid tools have emerged to analyze DAM system outputs, some notable ones with Graphical User Interface (GUI) are - ClockLab (*ClockLab | Actimetrics*), El Temps (*el temps principal*), ShinyR-DAM (Cichewicz and Hirsh, 2018), ActogramJ (Schmid et al., 2011), RhythmicAlly (Abhilash and Sheeba, 2019) etc.

There are fewer publicly available open-source tools for sleep analysis, one of the major tools being pySolo (Gilestro and Cirelli, 2009), a python-based program with a GUI and another tool – ShinyR-DAM, an R-based tool with a GUI and a webserver. However, these tools each have their own set of limitations, including, but not limited to – analysis of only one single DAM monitor file at a time preventing quick comparisons between monitors or genotypes/treatments/conditions/sex during analysis and visualization, inability to calculate and plot replicates separately, unavailability of multiple widely used periodogram methods within one tool, lack of access to reproducible minimal codes for publication, high probability of human errors due to manual preparation of input files without metadata, inability to produce publication-quality figures with customization, unavailability of automatic period-power detection from periodograms, lack of features to normalize activity counts of individuals to facilitate comparison among individuals, among *identifiers* (genotypes/treatments/conditions/sex etc.), and among experiments, inability to provide estimates of important sleep features separately for light and dark phase of the day such as latency, details of sleep bouts etc.

To alleviate these shortcomings in available open-source tools, I developed VANESSA (Visualization and ANalysis of timE SerieS dAta), an R-based set of tools to explore and analyze timeseries data primarily from DAM systems. Here I present the first two tools in VANESSA – (a) VANESSA-DAM-CRA (for circadian rhythm analysis) and (b) VANESSA-DAM-SA (for sleep analysis). Both of these apps have easy-to use GUIs, written completely in R (R Core Team, 2021) and GUIs deployed with Shiny (Winston et al., 2019). Both the apps are available as R packages from GitHub and are also deployed on shinyapps.io server to be used directly from a browser with an internet connection. VANESSA makes use of core R functions, packages from the TidyVerse, and some packages from the rethomics framework (Geissmann et al., 2019) – *behavr*, *damr*, *ggetho*, *zeitgebr* and *sleepr*. All other packages used are listed separately. A brief description of functionalities of VANESSA-DAM-CRA and VANESSA-DAM-SA follows:


A
VANESSA-DAM-Circadian
DATA INPUT
PERIODOGRAMS
ACTOGRAMS
INDIVIDUAL ACTOGRAMS
DATA FORMATTING
CWT SPECTOGRAMS
TIMESERIES SMOOTHING
DOCUMENTATION
BROWSE...
loading...
Generate report
Start calculations!
B
Calculate periodograms
Remove arrythmic individuals from dataset
PERIODOGRAM METHOD
AUTOCORRELATION
LOMB-SCARGLE
CHI-SQUARE
CWT
UPPER LIMIT FOR PERIODOGRAM
LOWER LIMIT FOR PERIODOGRAM
SIGNIFICANCE THRESHOLD FOR PERIODOGRAM
PERIODOGRAM FIGURES
PERIODOGRAM ALL WITH PEAKS
PERIODOGRAM AVERAGE
Control
C
Download data
D
Genotype
Long_2
Long_1
Control
Period (h)
E
power - signif_threshold
Period (h)
F
Period (h)
genotype
Control
Long_1
Long_2
power

Choose Monitor Files: No file selected

Choose Metadata File: No file selected

Modulo tau for actogram: 24

LD cycle period: 24

Summary time window in minutes: 15

Individual actogram display method (1 = Raw, 2 = Normalized): 1

Duration of light in hours: 12

Number of monitors: 1

Number of replicates: 1

Number of replicates: 1

How many minimum days flies should be alive to be counted?: 2

Starting day: 1

Ending day: 5

Select colour: #005900

Select colour: #2796A3

Select colour: #DF7000

Select colour: white

Select colour: white

Select colour: white

Select colour: white

Select colour: white

Select colour: white

Select colour: white

Select colour: white

Select colour: white

Select colour: white

height: 50

width: 1200

Show 10 entries

Search:

id	period	power	signif_threshold	p_value	peak
2017-02-18 10:00:00|Control_1_DD.txt|01	23.7	336.0476	149.2922	1.179907e-28	1
2017-02-18 10:00:00|Control_1_DD.txt|02	23.7	412.7766	149.2922	3.421214e-41	1
2017-02-18 10:00:00|Control_1_DD.txt|03	23.7	345.8020	149.2922	3.365019e-30	1
2017-02-18 10:00:00|Control_1_DD.txt|04	23.9	305.7421	150.5388	1.040734e-23	1
2017-02-18 10:00:00|Control_1_DD.txt|05	23.4	350.9069	148.0442	2.650303e-31	1
2017-02-18 10:00:00|Control_1_DD.txt|06	23.2	271.7230	146.7949	2.083318e-19	2
2017-02-18 10:00:00|Control_1_DD.txt|06	23.7	294.7145	149.2922	2.630914e-22	1
2017-02-18 10:00:00|Control_1_DD.txt|07	23.7	401.8852	149.2922	2.304621e-39	1
2017-02-18 10:00:00|Control_1_DD.txt|08	23.9	428.5891	150.5388	1.530106e-43	1
2017-02-18 10:00:00|Control_1_DD.txt|09	23.7	363.6559	149.2922	4.561156e-33	1

Showing 1 to 10 of 119 entries

Previous 1 2 3 4 5 … 12 Next

Figure A1.1: General description and period-power calculation in VANESSA-DAM. (A) The home tab of VANESSA-DAM – the Data Input tab. This tab has different parameters that can be changed according to the user's analysis needs – modulo-Tao, the LD cycle period in the experiment, summary time window, light duration, total number of monitors being analyzed, replicate declaration, subset data to specific days and indicate for how many days flies must be alive to be included in calculations and plots. This tab can also be used to customize colors for different *identifiers*. (B) Examples of periodogram calculations using Chi-square method from the "Periodogram" tab. Users can easily define upper and lower limits of periods to be scanned and also define significance threshold for the periodogram method. (C) Example of a preview of the tabulated data file that can be downloaded with all period and power values for all individuals for all significant peaks in periodogram. (D) Example of a density plot of period values of individuals from different identifiers. All individual period values are plotted as short horizontal lines - "|" on the x-axis. (E) Average periodogram of all individuals from different identifiers. (F) Violin plot of period values of all individuals from different identifiers are plotted along with the mean period value of that identifier embedded in the plot.

VANESSA-DAM common features:

The only requirements of both the apps are the creation of a metadata file, containing details of monitors, *identifiers* (genotypes/treatments/conditions/sex etc.), replicate information (if any used, e.g., Genotype A, B, C; Treatment 1, 2, 3), experiment start and end date time, and the monitor files scanned from the DAMScan program (Trikintetics^{TM}) (*Fig. 1 A & B*). The information in the metadata file can serve as experimental records and is used to determine zeitgeber on/off time for analyses and plots, and for assigning *identifiers* for each channel of the monitors. The metadata file can be prepared manually in any spreadsheet program or created using the apps themselves. The data is then curated (data for dead flies removed – *see Documentation tab of the app*), data pertaining to the days specified by the user is extracted, day-wise normalization of activity counts is done, and sleep bouts are identified based on 5 minutes of immobility (Hendricks et al., 2000; Shaw et al., 2000). A detailed user guide is available as Supplementary Material (*Supplementary methods 1 & 2*) and from https://github.com/orijitghosh/VANESSA-DAM. All plots made by the app can be resized, recolored and re-binned (binning of data over time) anytime dynamically.

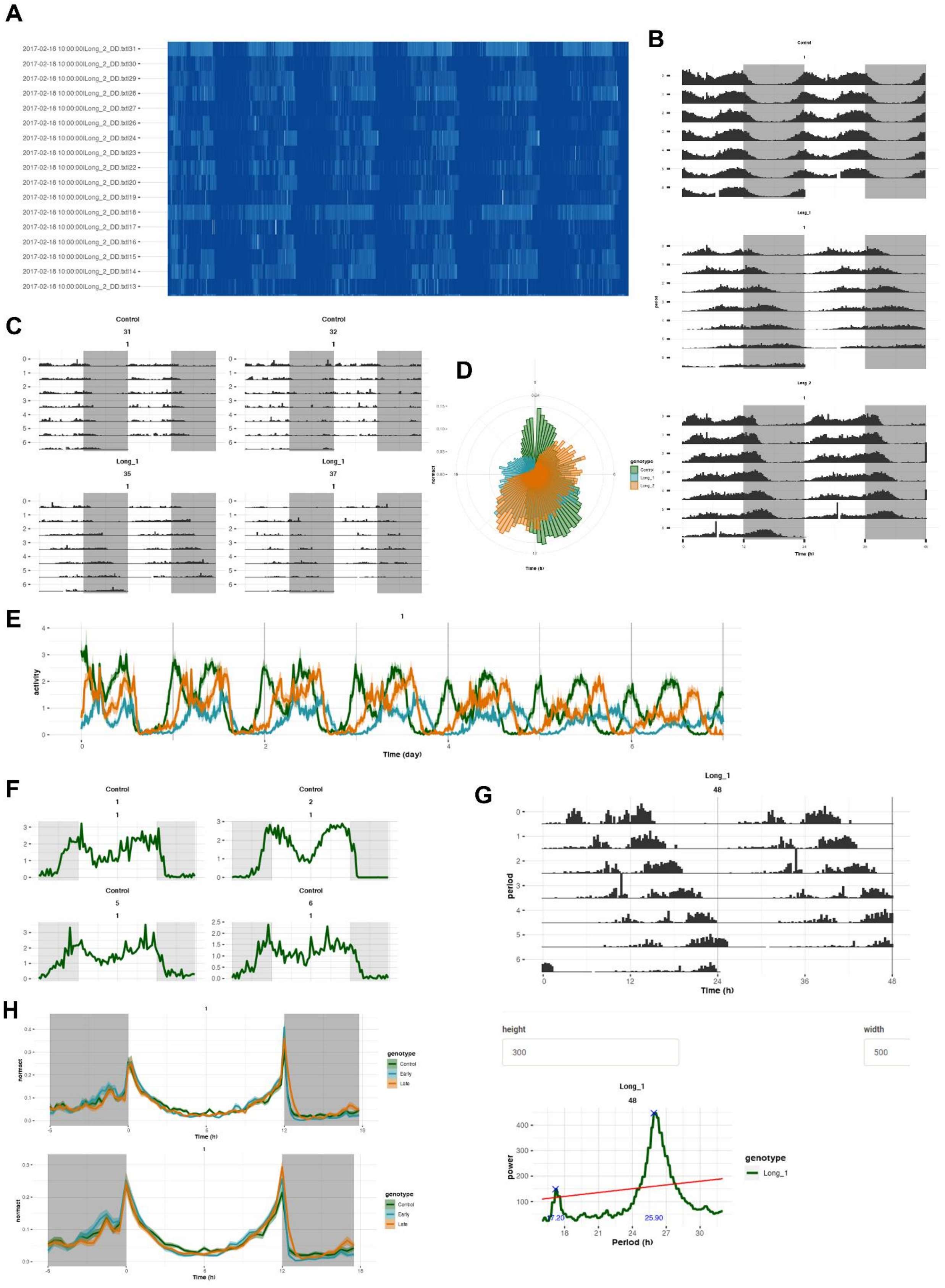

A
2017-02-18 10:00:00|Long_2_DD.txt|31
2017-02-18 10:00:00|Long_2_DD.txt|30
2017-02-18 10:00:00|Long_2_DD.txt|29
2017-02-18 10:00:00|Long_2_DD.txt|28
2017-02-18 10:00:00|Long_2_DD.txt|27
2017-02-18 10:00:00|Long_2_DD.txt|26
2017-02-18 10:00:00|Long_2_DD.txt|24
2017-02-18 10:00:00|Long_2_DD.txt|23
2017-02-18 10:00:00|Long_2_DD.txt|22
2017-02-18 10:00:00|Long_2_DD.txt|20
2017-02-18 10:00:00|Long_2_DD.txt|19
2017-02-18 10:00:00|Long_2_DD.txt|18
2017-02-18 10:00:00|Long_2_DD.txt|17
2017-02-18 10:00:00|Long_2_DD.txt|16
2017-02-18 10:00:00|Long_2_DD.txt|15
2017-02-18 10:00:00|Long_2_DD.txt|14
2017-02-18 10:00:00|Long_2_DD.txt|13
B
Control
Long_1
Long_2
period
Time (h)
C
Control
31
Control
32
Long_1
35
Long_1
37
D
normact
genotype
Control
Long_1
Long_2
Time (h)
E
activity
Time (day)
F
Control
G
Long_1
48
period
Time (h)
height
300
width
500
Long_1
48
power
17.20
25.90
Period (h)
genotype
Long_1
H
normact
Time (h)
genotype
Control
Early
Late

Figure A1.2: Actograms and activity profile visualization in VANESSA-DAM-CRA. (A) Activity profiles plotted as an ethogram where darkest blue color depicts highest activity and complete white indicates zero activity. Ethograms are an easy way to look at data from all individuals at a glance and notice any abnormal or unexpected patterns. (B) Batch actograms of different *identifiers*. The light-dark shading can be quickly altered by changing the "Duration of light in hours" parameters in the "Data input" tab. (C) Actograms of all individuals of one *identifier* can be visualized together. (D) Average activity profiles can also be visualized in a polar scale. (E) Day-wise activity profiles for all *identifiers* can be plotted averaged across individuals. (F) Activity profiles of each individual averaged across days can also be visualized separately. (G) A separate tab, "Individual actograms" can help visualize and inspect each actogram individually for the purpose of choosing representative actograms along with their periodograms. (H) Average activity profile of different *identifiers* can be plotted first averaged across days for each individual, then averaged across individuals. The top panel is plotted in 15 minutes bin and the bottom panel is plotted in 30 minutes bin. This binning can be changed quickly by changing the values of the "Summary time window in minutes" parameter in the "Data input" tab.

VANESSA-DAM-CRA specific features:

Periodograms:

Four popular periodogram methods have been incorporated in the app – (i) autocorrelation, (ii) chi-square (greedy chi-square method, improved method over the classical chi-square method), (iii) Lomb-Scargle, (iv) continuous wavelet transformation (CWT) (*Fig. 1 C & D*). Users can choose upper limit, lower limit and significance threshold for the periodogram analysis for all methods (*Fig. 1 C*). Arrhythmic individuals can be removed from the dataset after determining periodicity so that they do not affect further analysis and visualization of average profiles, average periodograms and batch actograms. All period and power values of each individual can be downloaded as csv files where different peaks of periodograms for all individuals will be saved (*Fig. 1 D*). Four useful visualizations are provided for period data – (i) all individual periodograms can be plotted captioned with monitor and channel number, *identifier* and replicate number with user specified colors of the particular *identifiers*; The highest peak of the periodograms will be identified and printed in these plots, (ii) periodograms can be averaged over individuals of an *identifier* and plotted, (iii) violin plots of all *identifiers* can be plotted along with individual data points and the mean of each *identifier* will be calculated and printed on the plots, (iv) a density plot of periods of *identifiers* can be plotted along with individual period values on the *x*-axis (*Fig. 1 E-G*).

Actograms and average profiles:

Activity counts can be visualized as heatmap "*ethograms*" for raw data and curated data (*Fig. 2 A*). All double-plotted actograms for all individuals can be visualized either together enabling quick and clear visualization of individuals that exhibit abnormal/unique and thus interesting patterns of activity, or separately, one-by-one along with their periodograms displayed below them enabling ease of picking representative actograms (*Fig. 2 B & C*). Double-plotted batch actograms are also available for each *identifier* separately (*Fig. 2 G*). Average profiles for individuals and all individuals of an *identifier* can be visualized either as a timeseries of consecutive days or averaging across days (*Fig. 2 E*). Average profiles of activity can also be visualized via a circular plot for each *identifier* (*Fig. 2 D*). All these plots can be plotted either with raw activity counts or with normalized activity (*Fig. 2 F & H*).

CWT spectrogram and timeseries smoothing:

CWT spectrograms can be visualized for individuals over several days (*Supplementary methods 2 – step 18*). This function does not require the metadata file, instead the input is the monitor file itself. Wavelet powers for different periods scanned can also be visualized. The timeseries smoothing function operates as a separate module and takes one single monitor file as input (*Supplementary methods 2 – step 19*). Users can upload a monitor file, change the binning of the data, use two popular methods of timeseries smoothening (lowpass Butterworth filter and kernel smoothing). For Butterworth filter, users can change "*filter order or generic filter model*" and "*critical frequencies of the filter*" and immediately check the effect of the binning and filtering on average profile of individuals of the monitor and on the individual profiles and download the smoothened data with applied parameters. Similarly, for kernel smoothing, users can specify the "*kernel smoothing bandwidth*".

VANESSA-DAM-SA specific features:

Sleep profiles:

Sleep bouts of each individual can be visualized as heatmaps across days. The other methods of sleep bout visualization are – (i) plots of sleep bouts across days for each individual (ii) sleep bouts can be averaged across chosen days for each individual and plotted, (iii) sleep bouts of individuals of an *identifier* can be averaged over chosen days and plotted, (iv) sleep bouts can be first averaged across days for each individual and then averaged across individuals of an identifier and plotted in a linear manner or in a circular scale (*Fig. 3 A-G*).

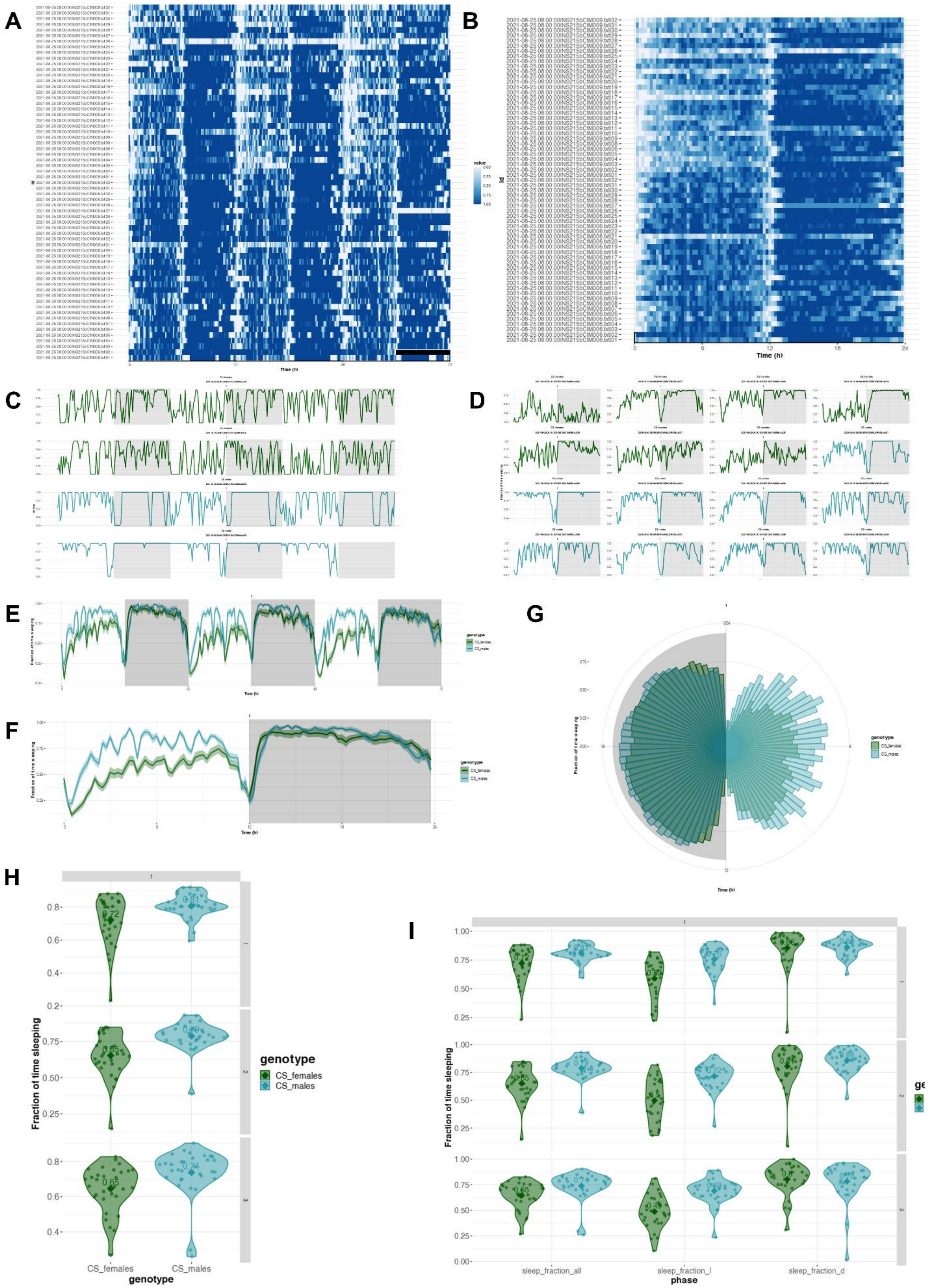

A
B
Id
value
0.00
0.25
0.50
0.75
1.00
Time (h)
0
6
12
18
24
C
D
E
F
G
genotype
CS_females
CS_males
Fraction of time sleeping
H
1
2
3
0.72
0.81
0.65
0.79
0.74
genotype
CS_females
CS_males
I
sleep_fraction_all
sleep_fraction_l
sleep_fraction_d
phase

Figure A1.3: Somnograms and sleep fraction analyses in VANESSA-DAM-SA. (A) Somnogram of all individuals across different days selected can be visualized where clearest white color depicts zero sleep and darkest blue denotes maximum sleep. (B) Somnogram of all individuals averaged across selected days can be visualized where clearest white color depicts zero sleep and darkest blue denotes maximum sleep. (C) Day-wise sleep profiles of each individual from different *identifiers* can be plotted separately. (D) Sleep profiles of each individual averaged across selected days can be plotted. (E) Day-wise sleep profiles of each *identifier* averaged across all individuals. (F) Sleep profiles of each *identifier* first averaged across individuals then averaged across selected days. (G) Sleep profiles of each *identifier* first averaged across individuals then averaged across selected days in polar scale. (H) Fraction of time sleeping for each individual during the entire day for different *identifiers* as a violin plot for each day separately. (I) Fraction of time sleeping in whole day (*sleep_fraction_all*), only in the light part of the day (*sleep_fraction_l*) and only in the dark part of the day (*sleep_fraction_d*) of each individual of different *identifiers* as a violin plot for each day separately.

Sleep fractions, bout, and latency analysis:

Sleep analysis results such as sleep levels, bout details, latency data can be downloaded as csv files for all individuals separately for light and dark phases. Total sleep fractions and total sleep times of different *identifiers* for chosen days can be plotted as violin plots. Sleep fractions and total times in the light and dark part of the day can also be plotted for different *identifiers* as violin plots. Bout analysis can be visualized in different ways – (i) sleep and awake bouts can be plotted for individuals of an *identifier*, (ii) violin plot of number of sleep and awake bouts for an *identifier*, (iii) violin plot of mean sleep and awake bout lengths for an *identifier*, (iv) violin plot of sleep latency for an *identifier*, (v) for all measurements, light and dark phase values can be plotted as separate violin plots (*Fig. 3 H & I, Fig. 4 A-K*).

Reproducible codes:

Both apps have capability of generating HTML files with all parameters used by the user for analysis in the session and codes to reproduce the same exact analyses to generate the tables and plots (*Fig. 5 A-C*). Advanced users can tweak different codes for analyses to suit their needs and generate custom plots. For timeseries smoothing, a HTML report file is generated with all parameters used for reproducing similar smoothing and binning of data.

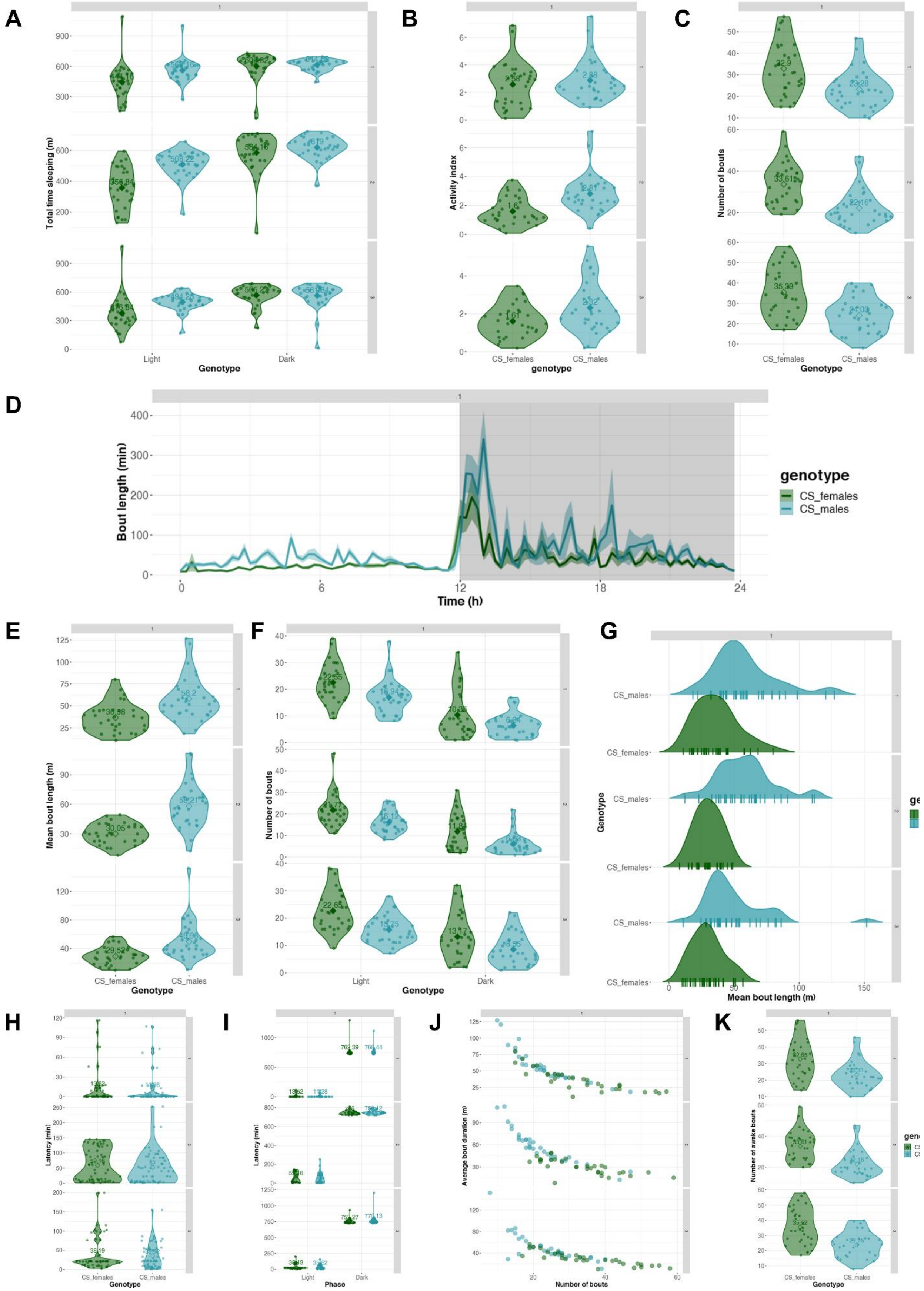
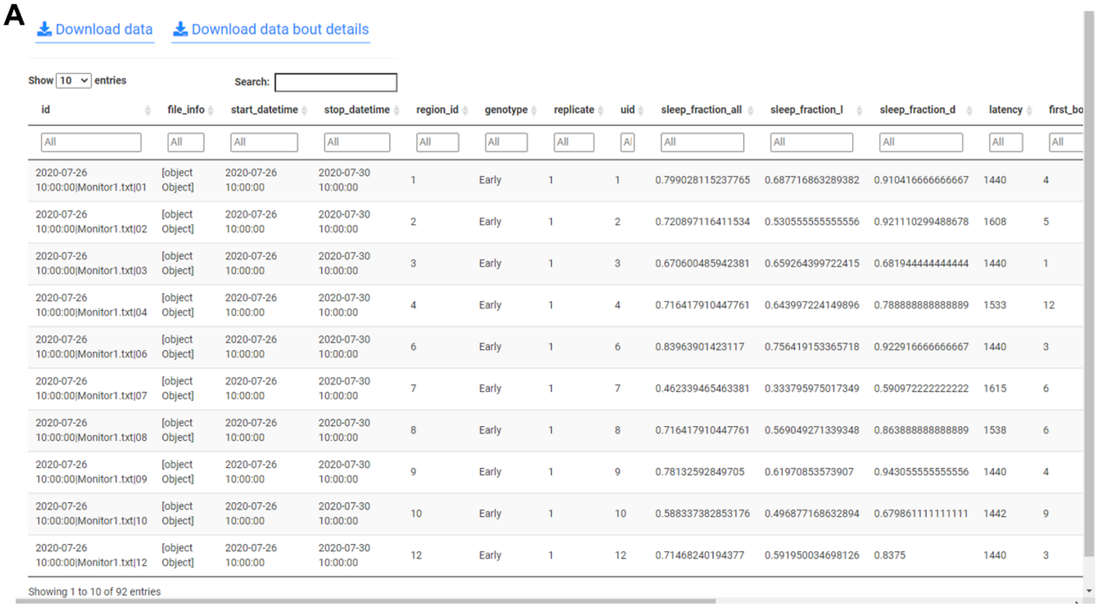

A
Total time sleeping (m)
Light
Dark
Genotype
B
Activity index
CS_females
CS_males
genotype
C
Number of bouts
Genotype
D
Bout length (min)
Time (h)
genotype
CS_females
CS_males
E
Mean bout length (m)
Genotype
F
Number of bouts
Light
Dark
Genotype
G
Genotype
Mean bout length (m)
H
Latency (min)
Genotype
I
Latency (min)
Phase
J
Average bout duration (m)
Number of bouts
K
Number of awake bouts
Genotype

Figure A1.4: Sleep time, bout and latency analyses in VANESSA-DAM-SA. (A) Total time sleeping in the light part of the day (Light) and in the dark part of the day (Dark) of each individual of different *identifiers* as a violin plot for each day separately. (B) Activity index (activity count per waking minute) of each individual of different *identifiers* as a violin plot for each day separately. (C) Violin plot of total number of bouts of all individuals of each *identifier* for each day separately. (D) Starting time of bout and bout lengths in minutes are plotted averaged across all individuals of each *identifier*. (E) Violin plot of mean bout length in minutes of all individuals of each *identifier* for each day separately. (F) Violin plot of number of bouts in dark (D) and light (L) parts of the day of all individuals of each *identifier* for each day separately. (G) Density plot of mean bout lengths in minutes for each *identifier* and mean bout lengths for each individual plotted along the *x*-axis for each day separately. (H) Violin plot of latency to first bout in minutes of all individuals of each *identifier* for each day separately. (I) Violin plot of latency to first bout in minutes in dark (D) and light (L) parts of the day of all individuals of each *identifier* for each day separately. (J) Plot of average bout duration in minutes vs number of bouts for all individuals colored by *identifiers*. (K) Violin plot of total number of awake bouts of all individuals of each *identifier* for each day separately. *Note: Latency calculated as the time taken to the first sleep bout from ZT00.*

The codes for these apps are hosted on GitHub (https://github.com/orijitghosh/VANESSA-DAM) and freely available under an MIT license along with toy/demo data. The apps are also hosted on a server and can directly be used from a browser with internet connection (https://cryptodice.shinyapps.io/vanessa-dam-cra/ and https://cryptodice.shinyapps.io/vanessa-dam-sa/). Additionally, VANESSA is available from GitHub as R packages and the instructions and usage are on the VANESSA-DAM GitHub page. VANESSA is also easily extendable to analyze locomotor activity data or any other numerical rhythmic behavioral data from other organisms recorded with systems other than DAM, and I will try our best to provide support for converting user data to DAM compatible files. Such an example is available in the Supplementary materials (*Supplementary figure 1*) where mouse locomotor activity data recorded with Trikinetics software was easily converted to VANESSA compatible files easily with a few lines of code. I plan to host an online webserver for converting other format data to VANESSA compatible files, as and when I get requests from users. For now, only mouse data conversion is in place as a function from the vanessadamcra package.

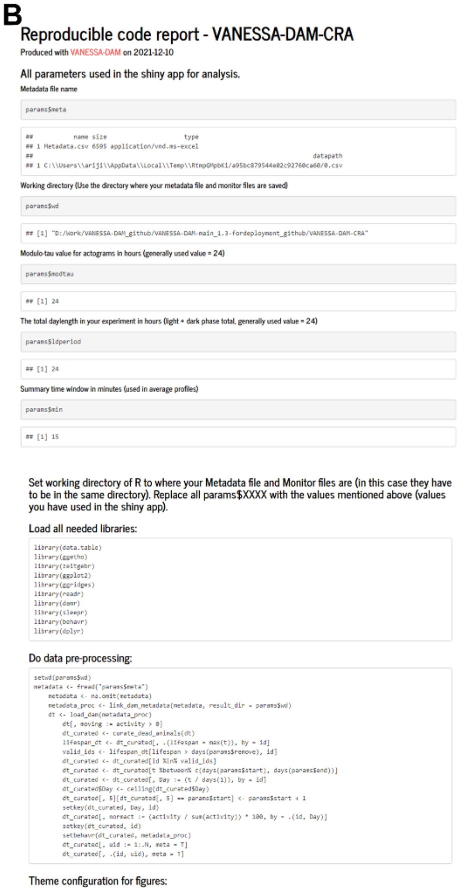
A



id	file_info	start_datetime	stop_datetime	region_id	genotype	replicate	uid	sleep_fraction_all	sleep_fraction_l	sleep_fraction_d	latency	first_bo
All	All	All	All	All	All	All	Al	All	All	All	All	All
2020-07-26 10:00:00|Monitor1.txt|01	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	1	Early	1	1	0.799028115237765	0.687716863289382	0.910416666666667	1440	4
2020-07-26 10:00:00|Monitor1.txt|02	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	2	Early	1	2	0.720897116411534	0.530555555555556	0.921110299488678	1608	5
2020-07-26 10:00:00|Monitor1.txt|03	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	3	Early	1	3	0.670600485942381	0.659264399722415	0.681944444444444	1440	1
2020-07-26 10:00:00|Monitor1.txt|04	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	4	Early	1	4	0.716417910447761	0.643997224149896	0.788888888888889	1533	12
2020-07-26 10:00:00|Monitor1.txt|06	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	6	Early	1	6	0.83963901423117	0.756419153365718	0.922916666666667	1440	3
2020-07-26 10:00:00|Monitor1.txt|07	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	7	Early	1	7	0.462339465463381	0.333795975017349	0.590972222222222	1615	6
2020-07-26 10:00:00|Monitor1.txt|08	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	8	Early	1	8	0.716417910447761	0.569049271339348	0.863888888888889	1538	6
2020-07-26 10:00:00|Monitor1.txt|09	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	9	Early	1	9	0.78132592849705	0.61970853573907	0.943055555555556	1440	4
2020-07-26 10:00:00|Monitor1.txt|10	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	10	Early	1	10	0.588337382853176	0.496877168632894	0.679861111111111	1442	9
2020-07-26 10:00:00|Monitor1.txt|12	[object Object]	2020-07-26 10:00:00	2020-07-30 10:00:00	12	Early	1	12	0.71468240194377	0.591950034698126	0.8375	1440	3

Showing 1 to 10 of 92 entries

B



Reproducible code report - VANESSA-DAM-CRA

Produced with VANESSA-DAM on 2021-12-10

All parameters used in the shiny app for analysis.

Metadata file name

```
params$meta
```

```
##          name size                     type
## 1 Metadata.csv 6595 application/vnd.ms-excel
##                                                                                 datapath
## 1 C:\\Users\\ariji\\AppData\\Local\\Temp\\RtmpGMpbKi/a95bc879544e02c92760ca60/0.csv
```

Working directory (Use the directory where your metadata file and monitor files are saved)

```
params$wd
```

```
## [1] "D:/Work/VANESSA-DAM_github/VANESSA-DAM-main_1.3-fordeployment_github/VANESSA-DAM-CRA"
```

Modulo-tau value for actograms in hours (generally used value = 24)

```
params$modtau
```

```
## [1] 24
```

The total daylength in your experiment in hours (light + dark phase total, generally used value = 24)

```
params$ldperiod
```

```
## [1] 24
```

Summary time window in minutes (used in average profiles)

```
params$min
```

```
## [1] 15
```

Set working directory of R to where your Metadata file and Monitor files are (in this case they have to be in the same directory). Replace all params$XXXX with the values mentioned above (values you have used in the shiny app).

Load all needed libraries:

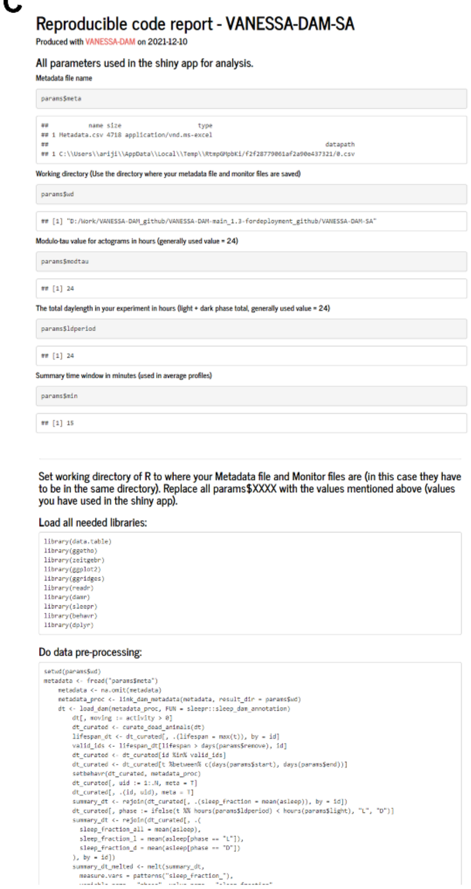
```
library(data.table)
library(ggetho)
library(zeitgebr)
library(ggplot2)
library(ggridges)
library(readr)
library(damr)
library(sleepr)
library(behavr)
library(dplyr)
```

Do data pre-processing:

```
setwd(params$wd)
metadata <- fread("params$meta")
    metadata <- na.omit(metadata)
    metadata_proc <- link_dam_metadata(metadata, result_dir = params$wd)
    dt <- load_dam(metadata_proc)
        dt[, moving := activity > 0]
        dt_curated <- curate_dead_animals(dt)
        lifespan_dt <- dt_curated[, .(lifespan = max(t)), by = id]
        valid_ids <- lifespan_dt[lifespan > days(params$remove), id]
        dt_curated <- dt_curated[id %in% valid_ids]
        dt_curated <- dt_curated[t %between% c(days(params$start), days(params$end))]
        dt_curated <- dt_curated[, Day := (t / days(1)), by = id]
        dt_curated$Day <- ceiling(dt_curated$Day)
        dt_curated[, 5][dt_curated[, 5] == params$start] <- params$start + 1
        setkey(dt_curated, Day, id)
        dt_curated[, normact := (activity / sum(activity)) * 100, by = .(id, Day)]
        setkey(dt_curated, id)
        setbehavr(dt_curated, metadata_proc)
        dt_curated[, uid := 1:.N, meta = T]
        dt_curated[, .(id, uid), meta = T]
```

Theme configuration for figures:

C



Reproducible code report - VANESSA-DAM-SA

Produced with VANESSA-DAM on 2021-12-10

All parameters used in the shiny app for analysis.

Metadata file name

```
params$meta
```

```
##          name size                     type
## 1 Metadata.csv 4718 application/vnd.ms-excel
##                                                                                 datapath
## 1 C:\\Users\\ariji\\AppData\\Local\\Temp\\RtmpGMpbKi/f2f28779061af2a90e437321/0.csv
```

Working directory (Use the directory where your metadata file and monitor files are saved)

```
params$wd
```

```
## [1] "D:/Work/VANESSA-DAM_github/VANESSA-DAM-main_1.3-fordeployment_github/VANESSA-DAM-SA"
```

Modulo-tau value for actograms in hours (generally used value = 24)

```
params$modtau
```

```
## [1] 24
```

The total daylength in your experiment in hours (light + dark phase total, generally used value = 24)

```
params$ldperiod
```

```
## [1] 24
```

Summary time window in minutes (used in average profiles)

```
params$min
```

```
## [1] 15
```

Set working directory of R to where your Metadata file and Monitor files are (in this case they have to be in the same directory). Replace all params$XXXX with the values mentioned above (values you have used in the shiny app).

Load all needed libraries:

```
library(data.table)
library(ggetho)
library(zeitgebr)
library(ggplot2)
library(ggridges)
library(readr)
library(damr)
library(sleepr)
library(behavr)
library(dplyr)
```

Do data pre-processing:

```
setwd(params$wd)
metadata <- fread("params$meta")
    metadata <- na.omit(metadata)
    metadata_proc <- link_dam_metadata(metadata, result_dir = params$wd)
    dt <- load_dam(metadata_proc, FUN = sleepr::sleep_dam_annotation)
        dt[, moving := activity > 0]
        dt_curated <- curate_dead_animals(dt)
        lifespan_dt <- dt_curated[, .(lifespan = max(t)), by = id]
        valid_ids <- lifespan_dt[lifespan > days(params$remove), id]
        dt_curated <- dt_curated[id %in% valid_ids]
        dt_curated <- dt_curated[t %between% c(days(params$start), days(params$end))]
        setbehavr(dt_curated, metadata_proc)
        dt_curated[, uid := 1:.N, meta = T]
        dt_curated[, .(id, uid), meta = T]
        summary_dt <- rejoin(dt_curated[, .(sleep_fraction = mean(asleep)), by = id])
        dt_curated[, phase := ifelse(t %% hours(params$ldperiod) < hours(params$light), "L", "D")]
        summary_dt <- rejoin(dt_curated[, .(
          sleep_fraction_all = mean(asleep),
          sleep_fraction_l = mean(asleep[phase == "L"]),
          sleep_fraction_d = mean(asleep[phase == "D"])
        ), by = id])
        summary_dt_melted <- melt(summary_dt,
          measure.vars = patterns("sleep_fraction_"),
```

Figure A1.5: Examples of downloadable sleep analysis data and reproducible code reports from VANESSA-DAM. (A) Example of the tabulated sleep analysis data that can be downloaded from VANESSA-DAM-SA. (B) Snippets of reproducible code report with all used parameters for analysis to reproduce exact same analysis and plots from VANESSA-DAM-CRA. (C) Snippets of reproducible code report with all used parameters for analysis to reproduce exact same analysis and plots from VANESSA-DAM-SA.

VANESSA is easily extendable to analysis of ultradian or infradian rhythms, and with a few modifications, can be used for analyzing rhythms other than locomotor activity. I welcome feature requests, suggestions, bug reports and collaborations via GitHub or email. I have provided tips regarding interpretation of different plotting options, proper usage of different periodogram methods with appropriate parameters in the VANESSA-DAM wiki (https://github.com/orijitghosh/VANESSA-DAM/wiki/Good-practices/). I believe VANESSA apps will be extremely useful to fly circadian rhythm and sleep researchers, especially those working with high throughput behavioral screenings where analyses need to be automated, and reproducibility of analysis methods is critical. Future updates will include but not be limited to – calculation of the angle-doubled center of mass for bimodal actograms, support of generating metadata file for more than 12 monitor files, in-built statistical tests, subjective and objective phase marking, support for more data acquisition systems for more organisms, sleep analysis in specific time windows etc.

Appendix 2

Supplementary information for Chapter 2

Statistically significant fixed factors or interaction terms are *italicized,* and fixed factors or interaction terms used for calculating Tukey's HSD are in bold in all ANOVA tables.

Appendix Table A2.1: ANOVA table summarizing the effect of Selection on the difference in percentage emergence between pre- and post-masking days in the window specified in Figure 3 (Dashed rectangle).

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*144.09*	*1*	*144.09*	*3*	*0.36*	*391.05*	*0.00*
Selection	***Fixed***	***248.81***	***2***	***124.40***	***6***	***0.62***	***200.30***	***0.00***
Block	Random	1.10	3	0.36	6	0.62	0.59	0.64
Selection* Block	Random	3.72	6	0.62	0	0		

Appendix Table A2.2: ANOVA table summarizing the effect of Selection on the difference in percentage emergence between pre- and post-masking days in the window specified in Figure 3 (Solid rectangle).

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*1660.22*	*1*	*1660.22*	*3*	*6.69*	*248.04*	*0.00*
Selection	***Fixed***	***1485.42***	***2***	***742.71***	***6***	***16.07***	***46.20***	***0.00***
Block	Random	20.08	3	6.69	6	16.07	0.41	0.74
Selection* Block	Random	96.45	6	16.07	0	0.00		

Appendix Table A2.3: ANOVA table summarizing the effect of Selection on the difference of percentage emergence between pre-masking and post-masking days from ZT0-0.5.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*1123.80*	*1*	*1123.80*	*3*	*8.28*	*135.65*	*0.00*
Selection	***Fixed***	***684.15***	***2***	***342.07***	***6***	***6.17***	***55.43***	***0.00***
Block	Random	24.85	3	8.28	6	6.17	1.34	0.34
Selection* Block	Random	37.02	6	6.17	0	0.00		

Appendix Table A2.4: ANOVA table summarizing the effects of Selection, full *T*-cycle and their interactions on Ψ_{CoM}.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*311448.8*	*1*	*311448.8*	*3*	*57.97*	*5371.89*	*0.00*

Selection	*Fixed*	*57207.8*	*2*	*28603.9*	*6*	*87.12*	*328.31*	*0.00*
T-cycle	*Fixed*	*49298.1*	*2*	*24649.0*	*6*	*37.50*	*657.14*	*0*
Block	Random	173.9	3	58.0	4.68	87.01	0.66	0.60
Selection*T-cycle	***Fixed***	***20203.8***	***4***	***5051.0***	***12***	***37.61***	***134.28***	***0***
Selection*Block	Random	522.7	6	87.1	12	37.61	2.31	0.10
T-cycle*Block	Random	225.1	6	37.5	12	37.61	0.99	0.46
Selection*T-cycle*Block	Random	451.4	12	37.6	0.00	0.00		

Appendix Table A2.5: ANOVA table summarizing the effects of Selection, full *T*-cycle and their interactions on Ψ_{Peak}.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*959.24*	*1*	*959.24*	*3*	*0.50*	*1898.00*	*0.00*
Selection	*Fixed*	*283.42*	*2*	*141.71*	*6*	*1.01*	*140.20*	*0.00*
T-cycle	*Fixed*	*201.88*	*2*	*100.94*	*6*	*0.31*	*321.95*	*0.00*
Block	Random	1.51	3	0.50	3.41	0.83	0.60	0.65
Selection*T-cycle	***Fixed***	***141.80***	***4***	***35.45***	***12***	***0.48***	***73.03***	***0***
Selection*Block	Random	6.06	6	1.01	12	0.48	2.08	0.13
T-cycle*Block	Random	1.88	6	0.31	12	0.48	0.64	0.69
Selection*T-cycle*Block	Random	5.82	12	0.48	0.00	0		

Appendix Table A2.6: ANOVA table summarizing the effects of Selection, full *T*-cycle and their interactions on R.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*13.95*	*1*	*13.95*	*3*	*0.00*	*24096.85*	*0.00*
Selection	*Fixed*	*0.37*	*2*	*0.18*	*6*	*0.00*	*67.34*	*0.00*
T-cycle	Fixed	0.00	2	0.00	6	0.00	1.35	0.32
Block	Random	0.00	3	0.00	8.79	0.00	0.14	0.93
Selection*T-cycle	***Fixed***	***0.01***	***4***	***0.00***	***12***	***0.00***	***8.03***	***0.00***
Selection*Block	Random	0.01	6	0.00	12	0.00	4.71	0.01
T-cycle*Block	Random	0.01	6	0.00	12	0.00	3.32	0.03
Selection*T-cycle*Block	Random	0.00	12	0.00	0.00	0		

Appendix Table A2.7: ANOVA table summarizing the effect of Selection on SSD calculated between T22 skeleton and full *T*-cycle profiles.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*7532872*	*1*	*7532872*	*3*	*9269.89*	*812.61*	*0.00*
Selection	***Fixed***	***7823911***	***2***	***3911955***	***6***	***38013.65***	***102.90***	***0.00***
Block	Random	27810	3	9270	6	38013.65	0.24	0.86
Selection*Block	Random	228082	6	38014	0	0.00		

Appendix Table A2.8: ANOVA table summarizing the effect of Selection on SSD calculated between T24 skeleton and full *T*-cycle profiles.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*2484756*	*1*	*2484756*	*3*	*13279.91*	*187.10*	*0.00*
Selection	**Fixed**	**356923**	**2**	**178461**	**6**	**38904.85**	**4.58**	**0.06**
Block	Random	39840	3	13280	6	38904.85	0.34	0.79
Selection*Block	Random	233429	6	38905	0	0.00		

Appendix Table A2.9: ANOVA table summarizing the effect of Selection on SSD calculated between T26 skeleton and full *T*-cycle profiles.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*2137588*	*1*	*2137588*	*3*	*7183.08*	*297.58*	*0.00*
Selection	***Fixed***	***973610***	***2***	***486805***	***6***	***3074.88***	***158.31***	***0.00***
Block	Random	21549	3	7183	6	3074.88	2.33	0.17
Selection*Block	Random	18449	6	3075	0	0.00		

Appendix Table A2.10: ANOVA table summarizing the effects of Selection, skeleton *T*-cycle and their interactions on R.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*6.36*	*1*	*6.36*	*3*	*0.00*	*1997.57*	*0.00*
Selection	*Fixed*	*0.13*	*2*	*0.06*	*6*	*0.00*	*45.72*	*0.00*
T-cycle	*Fixed*	*0.05*	*2*	*0.02*	*6*	*0.00*	*8.73*	*0.01*
Block	Random	0.00	3	0.00	1.72	0.00	1.48	0.44
Selection*T-cycle	***Fixed***	***0.24***	***4***	***0.06***	***12***	***0.00***	***23.70***	***0.00***

Selection*Block	Random	0.00	6	0.00	12	0.00	0.56	0.75
T-cycle*Block	Random	0.01	6	0.00	12	0.00	1.27	0.33
Selection*T-cycle*Block	Random	6.36	1	6.36	3	0.00	1997.57	0.00

Appendix Table A2.11: ANOVA table summarizing the effects of Selection, skeleton *T*-cycle and their interactions on Ψ_{CoM}.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*358228.4*	*1*	*358228.4*	*3*	*263.84*	*1357.71*	*0.00*
Selection	*Fixed*	*34817.6*	*2*	*17408.8*	*6*	*340.88*	*51.07*	*0.00*
T-cycle	*Fixed*	*93361.1*	*2*	*46680.5*	*6*	*106.99*	*436.30*	*0*
Block	Random	791.5	3	263.8	4.63	323.29	0.81	0.54
Selection*T-cycle	***Fixed***	***3718.3***	***4***	***929.6***	***12***	***124.57***	***7.46***	***0.00***
Selection*Block	Random	2045.3	6	340.9	12	124.57	2.73	0.06
T-cycle*Block	Random	641.9	6	107.0	12	124.57	0.85	0.55
Selection*T-cycle*Block	Random	1494.9	12	124.6	0.00	0.00		

Appendix Table A2.12: ANOVA table summarizing the effects of Selection, skeleton *T*-cycle and their interactions on Ψ_{Peak}.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*1336.30*	*1*	*1336.30*	*3*	*0.34*	*3838.33*	*0.00*
Selection	*Fixed*	*140.84*	*2*	*70.42*	*6*	*2.66*	*26.42*	*0.00*
T-cycle	*Fixed*	*190.58*	*2*	*95.29*	*6*	*2.85*	*33.37*	*0.00*
Block	Random	1.04	3	0.34	2.91	2.99	0.11	0.94
Selection*T-cycle	**Fixed**	**23.16**	**4**	**5.79**	**12**	**2.52**	**2.29**	**0.11**
Selection*Block	Random	15.98	6	2.66	12	2.52	1.05	0.43
T-cycle*Block	Random	17.12	6	2.85	12	2.52	1.13	0.40
Selection*T-cycle*Block	Random	30.29	12	2.52	0.00	0		

Appendix Table A2.13: Ψ_{CoM}, Ψ_{Peak}, R, SSD values of *early*, *control* and *late* populations under full or skeleton *T*-cycles and between them.

		Full *T*-cycles			**Skeleton *T*-cycles**			**SSD values**			
		T22	**T24**	**T26**	**T22**	**T24**	**T26**		**T22**	**T24**	**T26**
early	**Ψ_{CoM}** *±8.07* *±14.7*	53.33	49.37	39.99	114.27	52.38	8.63	***early***	1934.22 *±211.5*	698.93 *±213.9*	824.88 *±60.15*

	Ψ_{Peak} *±0.91* *±2.1*	1.95	2.05	2	5.6	4.8	-0.25
	R *±0.04* *±0.06*	0.75	0.76	0.74	0.60	0.51	0.29
control	Ψ_{CoM} *±8.07* *±14.7*	141.69	83.26	35.56	176.08	108.23	37.79
	Ψ_{Peak} *±0.91* *±2.1*	9.65	2.8	1.55	9.5	8.48	2.55
	R *±0.04* *±0.06*	0.61	0.58	0.60	0.39	0.52	0.44
late	Ψ_{CoM} *±8.07* *±14.7*	219.70	146.94	67.23	210.21	105.92	84.23
	Ψ_{Peak} *±0.91* *±2.1*	12.66	10.48	3.31	9.6	8.15	6.4
	R *±0.04* *±0.06*	0.48	0.47	0.56	0.23	0.37	0.39

control	218.94 *±211.5*	334.61 *±213.9*	220.79 *±60.15*
late	223.74 *±211.5*	331.58 *±213.9*	220.49 *±60.15*

Note: All phase values are phase relationships with lights-ON. Italicized numbers are ±95% CI used. For each phase marker, the top CI is for full *T*-cycles and the bottom CI is for skeleton *T*-cycles.

Appendix Table A2.14: ANOVA table summarizing the effect of *T*-cycle on percentage emergence at 2 hours prior to lights-ON under three full *T*-cycles in early flies.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*1426.55*	*1*	*1426.55*	*3*	*8.24*	*173.11*	*0.00*
T-cycle	**Fixed**	**32.38**	**2**	**16.19**	**6**	**4.02**	**4.02**	**0.07**
Block	Random	24.72	3	8.24	6	4.02	2.04	0.20
T-cycle*Block	Random	24.12	6	4.02	0	0		

Appendix Table A2.15: ANOVA table summarizing the effect of *T*-cycle on percentage emergence at 4 hours prior to lights-ON under three full *T*-cycles in early flies.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*85.29*	*1*	*85.29*	*3*	*2.08*	*40.95*	*0.00*
T-cycle	***Fixed***	***67.70***	***2***	***33.85***	***6***	***2.23***	***15.11***	***0.00***
Block	Random	6.24	3	2.08	6	2.23	0.92	0.48
T-cycle*Block	Random	13.43	6	2.23	0	0		

Appendix Table A2.16: ANOVA table summarizing the effect of *T*-cycle on percentage emergence at 6 hours prior to lights-ON under three full *T*-cycles in early flies.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*26.25*	*1*	*26.25*	*3*	*0.26*	*100.06*	*0.00*
T-cycle	***Fixed***	***27.85***	***2***	***13.92***	***6***	***0.51***	***27.24***	***0.00***

Block	Random	0.78	3	0.26	6	0.51	0.51	0.68
T-cycle*Block	Random	3.06	6	0.51	0	0		

Appendix Table A2.17: ANOVA table summarizing the effect of Selection on percentage of entrainment under T20.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	Fixed	8066.67	1	8066.66	0	0.00		
Selection	***Fixed***	***10433.33***	***2***	***5216.66***	***2***	***50.00***	***104.33***	***0.00***
Block	Random	0.00	1	0.00	2	50.00	0.00	1.00
Selection*Block	Random	100.00	2	50.00	0	0.00		

Appendix Table A2.18: ANOVA table summarizing the effect of Selection on percentage of entrainment under T28.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*36816.67*	*1*	*36816.67*	*1.00*	*150.00*	*245.44*	*0.04*
Selection	***Fixed***	***2433.33***	***2***	***1216.67***	***2.00***	***50.00***	***24.33***	***0.03***
Block	Random	150.00	1	150.00	2.00	50.00	3.00	0.22
Selection*Block	Random	100.00	2	50.00	0	0.00		

Appendix Table A2.19: ANOVA table summarizing the effects of Selection, regimes and their interactions on total activity within 15 minutes of lights-ON.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*7025.74*	*1*	*7025.74*	*3*	*74.86*	*93.85*	*0.00*
Selection	Fixed	30.53	2	15.26	6	64.89	0.23	0.79
Block	Random	224.58	3	74.86	4.32	58.80	1.27	0.38
Regime	*Fixed*	*377.53*	*2*	*188.76*	*6*	*17.44*	*10.82*	*0.01*
Selection*Block	Random	389.37	6	64.89	12	23.52	2.75	0.06
Selection*Regime	**Fixed**	**118.70**	**4**	**29.67**	**12**	**23.52**	**1.26**	**0.33**
Block*Regime	Random	104.64	6	17.44	12	23.52	0.74	0.62
Selection*Block*Regime	Random	282.35	12	23.53	0.00	0.00		

Appendix Table A2.20: ANOVA table summarizing the effect of Selection on difference in percentage emergence between Assay day 1 and 2 in 6 hour window under temperature cycle experiment (warm temperature ON 6 hour advanced).

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*1368.85*	*1*	*1368.85*	*3*	*43.03*	*31.80*	*0.01*
Selection	***Fixed***	***687.93***	***2***	***343.96***	***6***	***12.85***	***26.76***	***0.00***
Block	Random	129.10	3	43.03	6	12.85	3.34	0.09
Selection*Block	Random	77.11	6	12.85	0	0.00		

Appendix Table A2.21: ANOVA table summarizing the effect of Selection on difference in HTISE score under temperature cycle experiment (warm temperature ON 6 hour advanced).

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	Fixed	4.95	1	4.95	3	0.50	9.76	0.05
Selection	***Fixed***	***9.58***	***2***	***4.79***	***6***	***0.50***	***9.51***	***0.01***
Block	Random	1.52	3	0.50	6	0.50	1.00	0.45
Selection*Block	Random	3.02	6	0.50	0	0		

Appendix 3

Supplementary information for Chapter 3

Statistically significant fixed factors or interaction terms are *italicized,* and fixed factors or interaction terms used for calculating Tukey's HSD are in bold in all ANOVA tables.

Appendix Table A3.1: ANOVA table summarizing the effects of Selection, regimes and their interactions on total activity within 15 minutes of lights-ON.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*7025.74*	*1*	*7025.74*	*3.00*	*74.86*	*93.85*	*0.00*
Selection	Fixed	30.53	2	15.26	6.00	64.89	0.23	0.79
Block	Random	224.58	3	74.86	4.32	58.80	1.27	0.38
Regime	*Fixed*	*377.53*	*2*	*188.76*	*6.00*	*17.44*	*10.82*	*0.01*
Selection*Block	Random	389.37	6	64.89	12.00	23.52	2.75	0.06
Selection*Regime	**Fixed**	**118.70**	**4**	**29.67**	**12.00**	**23.52**	**1.26**	**0.33**
Block*Regime	Random	104.64	6	17.44	12.00	23.52	0.74	0.62
Selection*Block*Regime	Random	282.35	12	23.53	0.00	0.00		

Appendix Table A3.2: ANOVA table summarizing the effects of Selection on normalized activity within 5 minutes of lights-ON under 70 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*87.64*	*1*	*87.64*	*3*	*0.12*	*697.32*	*0.00*
Selection	**Fixed**	**0.35**	**2**	**0.17**	**6**	**0.04**	**3.92**	**0.08**
Block	Random	0.37	3	0.12	6	0.04	2.79	0.13
Selection*Block	Random	0.26	6	0.04	0	0		

Appendix Table A3.3: ANOVA table summarizing the effects of Selection on normalized activity within 10 minutes of lights-ON under 70 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*396.53*	*1*	*396.53*	*3*	*0.89*	*441.59*	*0.00*
Selection	**Fixed**	**1.41**	**2**	**0.70**	**6**	**0.46**	**1.51**	**0.29**
Block	Random	2.69	3	0.89	6	0.46	1.91	0.22
Selection*Block	Random	2.80	6	0.46	0	0		

Appendix Table A3.4: ANOVA table summarizing the effects of Selection on normalized activity within 15 minutes of lights-ON under 70 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*857.98*	*1*	*857.98*	*3*	*1.77*	*483.86*	*0.00*
Selection	**Fixed**	**2.63**	**2**	**1.31**	**6**	**1.29**	**1.02**	**0.41**
Block	Random	5.31	3	1.77	6	1.29	1.37	0.33

Selection*Block	Random	7.75	6	1.29	0	0		

Appendix Table A3.5: ANOVA table summarizing the effects of Selection on normalized activity within 5 minutes of lights-ON under 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*90.56*	*1*	*90.56*	*3*	*0.30*	*296.25*	*0.00*
Selection	***Fixed***	***1.05***	***2***	***0.52***	***6***	***0.03***	***13.92***	***0.00***
Block	*Random*	*0.91*	*3*	*0.30*	*6*	*0.03*	*8.03*	*0.01*
Selection*Block	Random	0.22	6	0.03	0	0		

Appendix Table A3.6: ANOVA table summarizing the effects of Selection on normalized activity within 10 minutes of lights-ON under 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*309.92*	*1*	*309.92*	*3*	*0.95*	*325.71*	*0.00*
Selection	***Fixed***	***2.55***	***2***	***1.27***	***6***	***0.18***	***6.99***	***0.02***
Block	*Random*	*2.85*	*3*	*0.95*	*6*	*0.18*	*5.21*	*0.04*
Selection*Block	Random	1.09	6	0.18	0	0		

Appendix Table A3.7: ANOVA table summarizing the effects of Selection on normalized activity within 15 minutes of lights-ON under 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*594.16*	*1*	*594.16*	*3*	*1.82*	*326.43*	*0.00*
Selection	**Fixed**	**4.01**	**2**	**2.00**	**6**	**0.42**	**4.75**	**0.05**
Block	Random	5.46	3	1.82	6	0.42	4.30	0.06
*Selection*Block*	*Random*	*2.53*	*6*	*0.42*	*0*	*0*		

Appendix Table A3.8: ANOVA table summarizing the effects of Selection on raw activity within 5 minutes of lights-ON under 70 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*4775.06*	*1*	*4775.06*	*3*	*6.15*	*775.40*	*0.00*
Selection	**Fixed**	**0.83**	**2**	**0.41**	**6**	**18.02**	**0.02**	**0.97**
Block	Random	18.47	3	6.15	6	18.02	0.34	0.79
Selection*Block	Random	108.12	6	18.02	0	0		

Appendix Table A3.9: ANOVA table summarizing the effects of Selection on raw activity within 10 minutes of lights-ON under 70 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*21357.56*	*1*	*21357.56*	*3*	*33.01*	*646.90*	*0.00*
Selection	**Fixed**	**0.46**	**2**	**0.23**	**6**	**71.32**	**0.00**	**0.99**
Block	Random	99.05	3	33.02	6	71.32	0.46	0.71
Selection*Block	Random	427.92	6	71.32	0	0		

Appendix Table A3.10: ANOVA table summarizing the effects of Selection on raw activity within 15 minutes of lights-ON under 70 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*46696.31*	*1*	*46696.31*	*3*	*72.14*	*647.29*	*0.00*
Selection	**Fixed**	**3.23**	**2**	**1.62**	**6**	**137.13**	**0.01**	**0.98**
Block	Random	216.42	3	72.14	6	137.13	0.52	0.68
Selection*Block	Random	822.83	6	137.14	0	0		

Appendix Table A3.11: ANOVA table summarizing the effects of Selection on raw activity within 5 minutes of lights-ON under 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*11850.51*	*1*	*11850.51*	*3*	*7.16*	*1654.15*	*0.00*
Selection	**Fixed**	**40.1**	**2**	**20.05**	**6**	**20.81**	**0.96**	**0.43**
Block	Random	21.49	3	7.16	6	20.81	0.34	0.79
Selection*Block	Random	124.87	6	20.81	0	0		

Appendix Table A3.12: ANOVA table summarizing the effects of Selection on raw activity within 10 minutes of lights-ON under 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*41192.39*	*1*	*41192.39*	*3*	*26.44*	*1557.66*	*0.00*
Selection	**Fixed**	**59.17**	**2**	**29.58**	**6**	**70.75**	**0.41**	**0.67**
Block	Random	79.33	3	26.44	6	70.75	0.37	0.77
Selection*Block	Random	424.56	6	70.76	0	0		

Appendix Table A3.13: ANOVA table summarizing the effects of Selection on raw activity within 15 minutes of lights-ON under 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*79873.17*	*1*	*79873.17*	*3*	*61.37*	*1301.33*	*0.00*
Selection	**Fixed**	**60.15**	**2**	**30.08**	**6**	**124.97**	**0.24**	**0.79**
Block	Random	184.13	3	61.38	6	124.97	0.49	0.70
Selection*Block	Random	749.83	6	124.97	0	0		

Appendix Table A3.14: ANOVA table summarizing the effects of Selection on normalized activity within 5 minutes of lights-ON under 500 lux LD04:20.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*16.51*	*1*	*16.51*	*3*	*0.08*	*199.11*	*0.00*
Selection	**Fixed**	**0.16**	**2**	**0.08**	**6**	**0.10**	**0.79**	**0.49**
Block	Random	0.24	3	0.08	6	0.10	0.79	0.54
Selection*Block	Random	0.62	6	0.10	0	0		

Appendix Table A3.15: ANOVA table summarizing the effects of Selection on normalized activity within 10 minutes of lights-ON under 500 lux LD04:20.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*52.91*	*1*	*52.91*	*3*	*0.39*	*132.48*	*0.00*
Selection	**Fixed**	**0.75**	**2**	**0.37**	**6**	**0.23**	**1.62**	**0.27**
Block	Random	1.19	3	0.39	6	0.23	1.72	0.26
Selection*Block	Random	1.39	6	0.23	0	0		

Appendix Table A3.16: ANOVA table summarizing the effects of Selection on normalized activity within 15 minutes of lights-ON under 500 lux LD04:20.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*105.71*	*1*	*105.71*	*3*	*0.79*	*133.13*	*0.00*
Selection	**Fixed**	**2.09**	**2**	**1.04**	**6**	**0.33**	**3.13**	**0.11**
Block	Random	2.38	3	0.79	6	0.33	2.38	0.16
Selection*Block	Random	1.99	6	0.33	0	0		

Appendix Table A3.17: ANOVA table summarizing the effects of Selection on normalized activity within 5 minutes of lights-ON under 500 lux LD08:16.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*17.25*	*1*	*17.25*	*3*	*0.07*	*242.46*	*0.00*

Selection	**Fixed**	**0.04**	**2**	**0.02**	**6**	**0.07**	**0.30**	**0.75**
Block	Random	0.21	3	0.07	6	0.07	0.93	0.48
Selection*Block	Random	0.45	6	0.07	0	0		

Appendix Table A3.18: ANOVA table summarizing the effects of Selection on normalized activity within 10 minutes of lights-ON under 500 lux LD08:16.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*62.50*	*1*	*62.50*	*3*	*0.19*	*317.52*	*0.00*
Selection	**Fixed**	**0.13**	**2**	**0.06**	**6**	**0.27**	**0.25**	**0.78**
Block	Random	0.59	3	0.19	6	0.27	0.71	0.57
Selection*Block	Random	1.64	6	0.27	0	0		

Appendix Table A3.19: ANOVA table summarizing the effects of Selection on normalized activity within 15 minutes of lights-ON under 500 lux LD08:16.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*122.76*	*1*	*122.76*	*3*	*0.31*	*393.59*	*0.00*
Selection	**Fixed**	**0.24**	**2**	**0.12**	**6**	**0.57**	**0.21**	**0.81**
Block	Random	0.93	3	0.31	6	0.57	0.53	0.67
Selection*Block	Random	3.47	6	0.57	0	0		

Appendix Table A3.20: ANOVA table summarizing the effects of Selection on raw activity within 5 minutes of lights-ON under 500 lux LD04:20.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*939.34*	*1*	*939.34*	*3*	*8.85*	*106.06*	*0.00*
Selection	**Fixed**	**0.00**	**2**	**0.00**	**6**	**3.84**	**0.00**	**0.99**
Block	Random	26.56	3	8.85	6	3.84	2.30	0.17
Selection*Block	Random	23.04	6	3.84	0	0		

Appendix Table A3.21: ANOVA table summarizing the effects of Selection on raw activity within 10 minutes of lights-ON under 500 lux LD04:20.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*2917.51*	*1*	*2917.51*	*3*	*31.81*	*91.71*	*0.00*
Selection	**Fixed**	**5.12**	**2**	**2.56**	**6**	**9.77**	**0.26**	**0.77**
Block	Random	95.43	3	31.81	6	9.77	3.25	0.10
Selection*Block	Random	58.65	6	9.77	0	0		

Appendix Table A3.22: ANOVA table summarizing the effects of Selection on raw activity within 15 minutes of lights-ON under 500 lux LD04:20.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*5657.65*	*1*	*5657.65*	*3*	*55.95*	*101.10*	*0.00*
Selection	**Fixed**	**26.50**	**2**	**13.25**	**6**	**16.48**	**0.80**	**0.49**
Block	Random	167.87	3	55.95	6	16.48	3.39	0.09
Selection*Block	Random	98.92	6	16.48	0	0		

Appendix Table A3.23: ANOVA table summarizing the effects of Selection on raw activity within 5 minutes of lights-ON under 500 lux LD08:16.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*1065.30*	*1*	*1065.30*	*3*	*3.82*	*278.84*	*0.00*
Selection	**Fixed**	**0.18**	**2**	**0.09**	**6**	**4.44**	**0.02**	**0.97**
Block	Random	11.46	3	3.82	6	4.44	0.85	0.51
Selection*Block	Random	26.66	6	4.44	0	0		

Appendix Table A3.24: ANOVA table summarizing the effects of Selection on raw activity within 10 minutes of lights-ON under 500 lux LD08:16.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*3964.59*	*1*	*3964.59*	*3*	*12.66*	*313.01*	*0.00*
Selection	**Fixed**	**6.39**	**2**	**3.19**	**6**	**17.70**	**0.18**	**0.83**
Block	Random	37.99	3	12.66	6	17.70	0.71	0.57
Selection*Block	Random	106.23	6	17.70	0	0		

Appendix Table A3.25: ANOVA table summarizing the effects of Selection on raw activity within 15 minutes of lights-ON under 500 lux LD08:16.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*7950.03*	*1*	*7950.03*	*3*	*24.82*	*320.22*	*0.00*
Selection	**Fixed**	**19.04**	**2**	**9.52**	**6**	**35.55**	**0.26**	**0.77**
Block	Random	74.47	3	24.82	6	35.55	0.69	0.58
Selection*Block	Random	213.31	6	35.55	0	0		

Appendix Table A3.26: ANOVA table summarizing the effects of Selection on normalized activity within 5 minutes of lights-ON under 500 lux LD16:08.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*54.17*	*1*	*54.17*	*3*	*0.04*	*1236.40*	*0.00*
Selection	***Fixed***	***2.18***	***2***	***1.09***	***6***	***0.07***	***15.21***	***0.00***
Block	Random	0.13	3	0.04	6	0.07	0.60	0.63
Selection*Block	Random	0.43	6	0.07	0	0		

Appendix Table A3.27: ANOVA table summarizing the effects of Selection on normalized activity within 10 minutes of lights-ON under 500 lux LD16:08.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*215.99*	*1*	*215.99*	*3*	*0.05*	*3973.91*	*0.00*
Selection	***Fixed***	***7.27***	***2***	***3.63***	***6***	***0.17***	***21.25***	***0.00***
Block	Random	0.16	3	0.05	6	0.17	0.31	0.81
Selection*Block	Random	1.02	6	0.17	0	0		

Appendix Table A3.28: ANOVA table summarizing the effects of Selection on normalized activity within 15 minutes of lights-ON under 500 lux LD16:08.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*423.49*	*1*	*423.49*	*3*	*0.09*	*4514.89*	*0.00*
Selection	***Fixed***	***11.29***	***2***	***5.64***	***6***	***0.34***	***16.19***	***0.00***
Block	Random	0.28	3	0.09	6	0.34	0.26	0.84
Selection*Block	Random	2.09	6	0.34	0	0		

Appendix Table A3.29: ANOVA table summarizing the effects of Selection on normalized activity within 5 minutes of lights-ON under 500 lux LD20:04.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*24.94*	*1*	*24.94*	*3*	*0.07*	*328.80*	*0.00*
Selection	***Fixed***	***1.88***	***2***	***0.94***	***6***	***0.10***	***9.03***	***0.01***
Block	Random	0.22	3	0.07	6	0.10	0.72	0.57
Selection*Block	Random	0.62	6	0.10	0	0		

Appendix Table A3.30: ANOVA table summarizing the effects of Selection on normalized activity within 10 minutes of lights-ON under 500 lux LD20:04.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*92.08*	*1*	*92.08*	*3*	*0.08*	*1066.34*	*0.00*
Selection	***Fixed***	***5.46***	***2***	***2.73***	***6***	***0.07***	***34.71***	***0.00***
Block	Random	0.25	3	0.08	6	0.07	1.09	0.42
Selection*Block	Random	0.47	6	0.07	0	0		

Appendix Table A3.31: ANOVA table summarizing the effects of Selection on normalized activity within 15 minutes of lights-ON under 500 lux LD20:04.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*178.46*	*1*	*178.46*	*3*	*0.17*	*1009.71*	*0.00*
Selection	***Fixed***	***8.25***	***2***	***4.12***	***6***	***0.22***	***18.70***	***0.00***
Block	Random	0.53	3	0.17	6	0.22	0.80	0.53
Selection*Block	Random	1.32	6	0.22	0	0		

Appendix Table A3.32: ANOVA table summarizing the effects of Selection on raw activity within 5 minutes of lights-ON under 500 lux LD16:08.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*3462.38*	*1*	*3462.38*	*3*	*19.97*	*173.34*	*0.00*
Selection	***Fixed***	***187.81***	***2***	***93.90***	***6***	***5.95***	***15.77***	***0.00***
Block	Random	59.92	3	19.97	6	5.95	3.35	0.09
Selection*Block	Random	35.71	6	5.95	0	0		

Appendix Table A3.33: ANOVA table summarizing the effects of Selection on raw activity within 10 minutes of lights-ON under 500 lux LD16:08.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*13985.46*	*1*	*13985.46*	*3*	*77.44*	*180.59*	*0.00*
Selection	***Fixed***	***642.33***	***2***	***321.17***	***6***	***24.09***	***13.32***	***0.00***
Block	Random	232.33	3	77.44	6	24.09	3.21	0.10
Selection*Block	Random	144.6	6	24.1	0	0		

Appendix Table A3.34: ANOVA table summarizing the effects of Selection on raw activity within 15 minutes of lights-ON under 500 lux LD16:08.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p

Intercept	*Fixed*	*27214.47*	*1*	*27214.47*	*3*	*142.31*	*191.22*	*0.00*
Selection	***Fixed***	***1057.72***	***2***	***528.86***	***6***	***47.52***	***11.12***	***0.00***
Block	Random	426.95	3	142.32	6	47.52	2.99	0.11
Selection*Block	Random	285.16	6	47.53	0	0		

Appendix Table A3.35: ANOVA table summarizing the effects of Selection on raw activity within 5 minutes of lights-ON under 500 lux LD20:04.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*571.39*	*1*	*571.39*	*3*	*3.27*	*174.24*	*0.00*
Selection	***Fixed***	***65.20***	***2***	***32.60***	***6***	***5.29***	***6.15***	***0.03***
Block	Random	9.83	3	3.27	6	5.29	0.61	0.62
Selection*Block	Random	31.77	6	5.29	0	0		

Appendix Table A3.36: ANOVA table summarizing the effects of Selection on raw activity within 10 minutes of lights-ON under 500 lux LD20:04.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*2802.09*	*1*	*2802.09*	*3*	*15.23*	*183.96*	*0.00*
Selection	***Fixed***	***436.75***	***2***	***218.37***	***6***	***21.54***	***10.13***	***0.01***
Block	Random	45.69	3	15.23	6	21.54	0.70	0.58
Selection*Block	Random	129.25	6	21.54	0	0		

Appendix Table A3.37: ANOVA table summarizing the effects of Selection on raw activity within 15 minutes of lights-ON under 500 lux LD20:04.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*5218.36*	*1*	*5218.36*	*3*	*45.47*	*114.75*	*0.00*
Selection	***Fixed***	***509.41***	***2***	***254.70***	***6***	***43.78***	***5.81***	***0.03***
Block	Random	136.42	3	45.47	6	43.78	1.03	0.44
Selection*Block	Random	262.72	6	43.78	0	0		

Appendix Table A3.38: ANOVA table summarizing the effects of Selection and Regime on number of sleep bouts in daytime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*5830.60*	*1*	*5830.60*	*3*	*3.81*	*1527.69*	*0.00*

Selection	*Fixed*	*27.10*	*2*	*13.55*	*6*	*1.84*	*7.35*	*0.02*
Block	Random	11.45	3	3.81	0.00	0.00	785.61	
Regime	Fixed	2.30	1	2.30	3	3.68	0.62	0.48
Selection*Block	Random	11.06	6	1.84	6	5.51	0.33	0.89
Selection*Regime	**Fixed**	**6.57**	**2**	**3.28**	**6**	**5.51**	**0.59**	**0.58**
Block*Regime	Random	11.04	3	3.68	6	5.51	0.66	0.60
Selection*Block*Regime	Random	33.11	6	5.51	0	0		

Appendix Table A3.39: ANOVA table summarizing the effects of Selection and Regime on mean sleep bout length in daytime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*24239.42*	*1*	*24239.42*	*3*	*6.79*	*3566.96*	*0.00*
Selection	Fixed	237.05	2	118.52	6	36.69	3.23	0.11
Block	Random	20.39	3	6.8	0			
Regime	*Fixed*	*235.61*	*1*	*235.61*	*3*	*2.22*	*105.68*	*0.00*
Selection*Block	Random	220.18	6	36.7	6	47.83	0.76	0.62
Selection*Regime	**Fixed**	**69.34**	**2**	**34.67**	**6**	**47.83**	**0.72**	**0.52**
Block*Regime	Random	6.69	3	2.23	6	47.83	0.04	0.98
Selection*Block*Regime	Random	287.02	6	47.84	0	0		

Appendix Table A3.40: ANOVA table summarizing the effects of Selection and Regime on total sleep in daytime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*3185953*	*1*	*3185953*	*3*	*2301.60*	*1384.23*	*0.00*
Selection	Fixed	7459	2	3730	6	3719.56	1.00	0.42
Block	Random	6905	3	2302	5.16	3837.70	0.6	0.64
Regime	Fixed	4286	1	4286	3	1083.82	3.95	0.14
Selection*Block	Random	22317	6	3720	6	965.68	3.85	0.06
Selection*Regime	**Fixed**	**761**	**2**	**381**	**6**	**965.68**	**0.39**	**0.69**
Block*Regime	Random	3251	3	1084	6	965.68	1.12	0.41
Selection*Block*Regime	Random	5794	6	966	0	0		

Appendix Table A3.41: ANOVA table summarizing the effects of Selection and Regime on latency to sleep in daytime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*187725.4*	*1*	*187725.4*	*3*	*356.96*	*525.89*	*0.00*
Selection	Fixed	664.4	2	332.2	6	665.15	0.49	0.63
Block	Random	1070.9	3	357	5.90	675.74	0.52	0.67
Regime	Fixed	17.1	1	17.1	3	88.55	0.19	0.68
*Selection*Block*	*Random*	*3990.9*	*6*	*665.2*	*6*	*77.96*	*8.53*	*0.00*
Selection*Regime	**Fixed**	**74**	**2**	**37**	**6**	**77.96**	**0.47**	**0.64**
Block*Regime	Random	265.7	3	88.6	6	77.96	1.13	0.40
Selection*Block*Regime	Random	467.8	6	78	0	0		

Appendix Table A3.42: ANOVA table summarizing the effects of Selection and Regime on number of sleep bouts in nighttime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*4947.93*	*1*	*4947.93*	*3*	*3.30*	*1495.24*	*0.00*
Selection	*Fixed*	*72.16*	*2*	*36.08*	*6*	*2.47*	*14.60*	*0.00*
Block	Random	9.92	3	3.30	0.17	0.56	5.81	0.70
Regime	*Fixed*	*81.60*	*1*	*81.60*	*3*	*0.29*	*273.71*	*0.00*
Selection*Block	Random	14.82	6	2.47	6	2.19	1.12	0.44
Selection*Regime	**Fixed**	**11.51**	**2**	**5.76**	**6**	**2.19**	**2.61**	**0.15**
Block*Regime	Random	0.89	3	0.29	6	2.19	0.13	0.93
Selection*Block*Regime	Random	13.19	6	2.19	0	0		

Appendix Table A3.43: ANOVA table summarizing the effects of Selection and Regime on mean sleep bout length in nighttime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*96580.32*	*1*	*96580.32*	*3*	*203.20*	*475.28*	*0.00*
Selection	*Fixed*	*4220.8*	*2*	*2110.4*	*6*	*231.57*	*9.11*	*0.01*
Block	Random	609.62	3	203.21	0.08	37.21	5.46	0.81
Regime	*Fixed*	*8633.29*	*1*	*8633.29*	*3*	*5.85*	*1474.56*	*0.00*
Selection*Block	Random	1389.47	6	231.58	6	200.21	1.15	0.43
Selection*Regime	**Fixed**	**8.28**	**2**	**4.14**	**6**	**200.21**	**0.02**	**0.97**
Block*Regime	Random	17.56	3	5.85	6	200.21	0.02	0.99
Selection*Block*Regime	Random	1201.31	6	200.22	0	0		

Appendix Table A3.44: ANOVA table summarizing the effects of Selection and Regime on total sleep in nighttime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*6648430*	*1*	*6648430*	*3*	*5617.98*	*1183.41*	*0.00*
Selection	*Fixed*	*22125*	*2*	*11063*	*6*	*1964.79*	*5.63*	*0.04*
Block	Random	16854	3	5618	7.99	2409.81	2.33	0.15
Regime	*Fixed*	*61339*	*1*	*61339*	*3*	*497.34*	*123.33*	*0.00*
*Selection*Block*	*Random*	*11789*	*6*	*1965*	*6*	*52.32*	*37.55*	*0.00*
Selection*Regime	***Fixed***	***4498***	***2***	***2249***	***6***	***52.32***	***42.98***	***0.00***
*Block*Regime*	*Random*	*1492*	*3*	*497*	*6*	*52.32*	*9.50*	*0.01*
Selection*Block*Regime	Random	314	6	52	0	0		

Appendix Table A3.45: ANOVA table summarizing the effects of Selection and Regime on latency to sleep in nighttime under 70 lux and 500 lux LD12:12.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*16254.8*	*1*	*16254.8*	*3*	*83.11*	*195.58*	*0.00*
Selection	Fixed	56.06	2	28.03	6	26.79	1.04	0.40
Block	Random	249.33	3	83.11	2.76	38.61	2.15	0.28
Regime	Fixed	303.3	1	303.3	3	32.34	9.37	0.05
Selection*Block	Random	160.79	6	26.8	6	20.53	1.30	0.37
Selection*Regime	Fixed	99.1	2	49.55	6	20.53	2.41	0.17
Block*Regime	Random	97.04	3	32.35	6	20.53	1.57	0.29
Selection*Block*Regime	Random	123.19	6	20.53	0	0		

Appendix Table A3.46: ANOVA table summarizing the effects of Selection and Regime on number of sleep bouts in daytime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*18104.37*	*1*	*18104.37*	*3*	*4.65*	*3892.44*	*0.00*
Selection	Fixed	1.05	2	0.52	6	6.78	0.07	0.92
Block	Random	13.95	3	4.65	10.17	10.38	0.44	0.72
Regime	*Fixed*	*4006.11*	*4*	*1001.53*	*12*	*5.73*	*174.66*	*0*
*Selection*Block*	*Random*	*40.68*	*6*	*6.78*	*24*	*2.13*	*3.17*	*0.01*
Selection*Regime	***Fixed***	***46.65***	***8***	***5.83***	***24***	***2.13***	***2.73***	***0.02***
*Block*Regime*	*Random*	*68.81*	*12*	*5.73*	*24*	*2.13*	*2.68*	*0.01*

Selection*Block*Regime	Random	51.21	24	2.13	0	0		

Appendix Table A3.47: ANOVA table summarizing the effects of Selection and Regime on mean sleep bout length in daytime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*68515.93*	*1*	*68515.93*	*3*	*35.53*	*1928.09*	*0.00*
Selection	Fixed	229.33	2	114.66	6	114.78	0.99	0.42
Block	Random	106.61	3	35.54	5.17	107.85	0.32	0.80
Regime	*Fixed*	*965.73*	*4*	*241.43*	*12*	*17.75*	*13.59*	*0.00*
*Selection*Block*	*Random*	*688.69*	*6*	*114.78*	*24*	*24.68*	*4.65*	*0.00*
Selection*Regime	**Fixed**	**203.63**	**8**	**25.45**	**24**	**24.68**	**1.03**	**0.44**
Block*Regime	Random	213.1	12	17.76	24	24.68	0.72	0.71
Selection*Block*Regime	Random	592.34	24	24.68	0	0		

Appendix Table A3.48: ANOVA table summarizing the effects of Selection and Regime on total sleep in daytime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*12024765*	*1*	*12024765*	*3*	*8478.47*	*1418.27*	*0.00*
Selection	Fixed	15909	2	7954	6	4229.62	1.88	0.23
Block	Random	25435	3	8478	6.49	4637.85	1.82	0.23
Regime	*Fixed*	*3802188*	*4*	*950547*	*12*	*1746.36*	*544.29*	*0*
*Selection*Block*	*Random*	*25378*	*6*	*4230*	*24*	*1338.13*	*3.16*	*0.01*
Selection*Regime	**Fixed**	**23585**	**8**	**2948**	**24**	**1338.13**	**2.20**	**0.06**
Block*Regime	Random	20956	12	1746	24	1338.13	1.30	0.27
Selection*Block*Regime	Random	32115	24	1338	0	0		

Appendix Table A3.49: ANOVA table summarizing the effects of Selection and Regime on latency to sleep in daytime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*138163.3*	*1*	*138163.3*	*3*	*559.22*	*247.06*	*0.00*
Selection	Fixed	471.8	2	235.9	6	285.52	0.82	0.48
Block	Random	1677.7	3	559.2	9.20	374.18	1.49	0.27
Regime	*Fixed*	*31058*	*4*	*7764.5*	*12*	*135.78*	*57.18*	*0*

Selection*Block	_Random_	_1713.2_	_6_	_285.5_	_24_	_47.12_	_6.05_	_0.00_
___Selection*Regime___	___Fixed___	___1962.2___	___8___	___245.3___	___24___	___47.12___	___5.20___	___0.00___
Block*Regime	_Random_	_1629.4_	_12_	_135.8_	_24_	_47.12_	_2.88_	_0.01_
Selection*Block*Regime	Random	1131	24	47.1	0	0		

Appendix Table A3.50: ANOVA table summarizing the effects of Selection and Regime on number of sleep bouts in nighttime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	_Fixed_	_16169.46_	_1_	_16169.46_	_3_	_7.17_	_2253.51_	_0.00_
Selection	_Fixed_	_165.07_	_2_	_82.53_	_6_	_9.58_	_8.61_	_0.01_
Block	Random	21.53	3	7.18	9.26	12.63	0.56	0.64
Regime	_Fixed_	_3562.76_	_4_	_890.69_	_12_	_4.67_	_190.39_	_0_
Selection*Block	_Random_	_57.51_	_6_	_9.58_	_24_	_1.62_	_5.89_	_0.00_
___Selection*Regime___	___Fixed___	___102.24___	___8___	___12.78___	___24___	___1.62___	___7.86___	___0.00___
Block*Regime	_Random_	_56.14_	_12_	_4.68_	_24_	_1.62_	_2.87_	_0.01_
Selection*Block*Regime	Random	39.01	24	1.63	0	0		

Appendix Table A3.51: ANOVA table summarizing the effects of Selection and Regime on mean sleep bout length in nighttime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	_Fixed_	_126611.5_	_1_	_126611.5_	_3_	_86.76_	_1459.24_	_0.00_
Selection	Fixed	1854.6	2	927.3	6	279.51	3.31	0.10
Block	Random	260.3	3	86.8	7.49	333.59	0.26	0.85
Regime	_Fixed_	_4789.5_	_4_	_1197.4_	_12_	_136.07_	_8.79_	_0.00_
Selection*Block	_Random_	_1677.1_	_6_	_279.5_	_24_	_81.99_	_3.40_	_0.01_
___Selection*Regime___	___Fixed___	___1865.1___	___8___	___233.1___	___24___	___81.99___	___2.84___	___0.02___
Block*Regime	Random	1632.9	12	136.1	24	81.99	1.66	0.14
Selection*Block*Regime	Random	1967.9	24	82	0	0		

Appendix Table A3.52: ANOVA table summarizing the effects of Selection and Regime on total sleep in nighttime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	_Fixed_	_14707088_	_1_	_14707088_	_3_	_1332.74_	_11035.19_	_0.00_

Selection	*Fixed*	*44055*	*2*	*22028*	*6*	*1126.14*	*19.56*	*0.00*
Block	Random	3998	3	1333	8.83	1685.20	0.79	0.52
Regime	*Fixed*	*2270207*	*4*	*567552*	*12*	*1086.61*	*522.31*	*0*
Selection*Block	Random	6757	6	1126	24	527.54	2.13	0.08
Selection*Regime	***Fixed***	***24840***	***8***	***3105***	***24***	***527.54***	***5.89***	***0.00***
Block*Regime	Random	13039	12	1087	24	527.54	2.06	0.06
Selection*Block*Regime	Random	12661	24	528	0	0		

Appendix Table A3.53: ANOVA table summarizing the effects of Selection and Regime on latency to sleep in nighttime under 500 lux photoperiods.

	Effect (F/R)	SS	Degrees of Freedom	MS	Error df	Error MS	F	p
Intercept	*Fixed*	*126686.2*	*1*	*126686.2*	*3*	*438.26*	*289.06*	*0.00*
Selection	Fixed	596	2	298	6	492.18	0.60	0.57
Block	Random	1314.8	3	438.3	5.92	501.28	0.87	0.50
Regime	*Fixed*	*39299.9*	*4*	*9825*	*12*	*129.72*	*75.73*	*0*
*Selection*Block*	*Random*	*2953.1*	*6*	*492.2*	*24*	*120.61*	*4.08*	*0.00*
Selection*Regime	**Fixed**	**1610.4**	**8**	**201.3**	**24**	**120.61**	**1.66**	**0.15**
Block*Regime	Random	1556.7	12	129.7	24	120.61	1.07	0.42
Selection*Block*Regime	Random	2894.8	24	120.6	0	0		

Appendix Table A3.54: Summary of sleep parameters under 500 lux LD08:16 second set of experiments. Mean length of sleep bouts, Total sleep time, and Latency to sleep are in minutes. Stock labels: E – *early*, C – *control*, and L – *late*. 1, 2, 3, 4 are replicate populations. Latency is calculated from ZT00.

Stock	**Phase**	**Number of sleep bouts**	**Mean length of sleep bouts**	**Total sleep time**	**Latency to sleep**
E1	*light*	9.06	66.59	346.16	47.13
E2	*light*	12.81	39.47	314.77	47.99
E3	*light*	8.65	69.94	350.18	56.73
E4	*light*	14.34	27.84	304.73	39.82
C1	*light*	12.91	35.41	317.76	45.37
C2	*light*	13.09	37.17	321.04	36.84
C3	*light*	9.21	76.13	386.85	43.17
C4	*light*	10.37	67.42	364.82	56.57
L1	*light*	11.15	57.09	369.04	37.89
L2	*light*	12.47	41.17	345.81	29.17
L3	*light*	12.58	29.96	287.06	67.48

L4	*light*	8.75	88.47	365.40	43.86
E1	*dark*	14.07	73.63	684.05	572.72
E2	*dark*	15.92	65.10	697.44	569.52
E3	*dark*	15.55	64.19	777.18	542.28
E4	*dark*	20.04	42.76	700.24	544.15
C1	*dark*	17.51	53.61	678.23	551.22
C2	*dark*	14.25	60.93	658.17	567.52
C3	*dark*	12.66	100.40	750.66	556.55
C4	*dark*	13.21	98.25	815.72	535.14
L1	*dark*	15.14	64.52	660.87	592.42
L2	*dark*	14.93	58.16	634.70	592.70
L3	*dark*	13.33	81.64	735.23	552.24
L4	*dark*	11.85	92.34	747.75	561.51

Appendix Table A3.55: Summary of sleep parameters under 500 lux LD04:20 second set of experiments. Mean length of sleep bouts, Total sleep time, and Latency to sleep are in minutes. Stock labels: E – *early*, C – *control*, and L – *late*. 1, 2, 3, 4 are replicate populations. Latency is calculated from ZT00.

Stock	**Phase**	**Number of sleep bouts**	**Mean length of sleep bouts**	**Total sleep time**	**Latency to sleep**
E1	*light*	9.06	68.63	169.71	39.34
E2	*light*	12.81	39.26	153.69	45.01
E3	*light*	8.65	65.40	185.8	22.06
E4	*light*	14.34	33.00	155.56	37.53
C1	*light*	12.91	54.64	164.6	40.44
C2	*light*	13.09	37.71	162.5	24.57
C3	*light*	9.21	126.15	245.88	19.18
C4	*light*	10.37	89.66	202.32	27.20
L1	*light*	11.15	47.01	172.44	33.42
L2	*light*	12.47	55.42	172.01	19.96
L3	*light*	12.58	48.16	152.12	42.11
L4	*light*	8.75	57.94	166.86	27.67
E1	*dark*	14.07	64.58	857.75	282.13
E2	*dark*	15.92	47.69	845	286.39
E3	*dark*	15.55	43.87	868.06	277.63
E4	*dark*	20.04	49.56	910.89	300.43
C1	*dark*	17.51	50.48	926.82	269.78
C2	*dark*	14.25	53.63	833.28	295.43
C3	*dark*	12.66	50.30	869.34	324.5
C4	*dark*	13.21	85.98	951.58	274.61
L1	*dark*	15.14	42.95	902.77	271.02

L2	*dark*	14.93	49.57	807.22	320.16
L3	*dark*	13.33	74.93	884.67	308.93
L4	*dark*	11.85	62.76	909	265.26

Appendix Table A3.56: Summary of sleep parameters under 500 lux LD16:08 second set of experiments. Mean length of sleep bouts, Total sleep time, and Latency to sleep are in minutes. Stock labels: E – *early*, C – *control*, and L – *late*. 1, 2, 3, 4 are replicate populations. Latency is calculated from ZT00.

Stock	**Phase**	**Number of sleep bouts**	**Mean length of sleep bouts**	**Total sleep time**	**Latency to sleep**
E1	*light*	9.06	56.17	573.94	103.83
E2	*light*	12.81	36.59	542.45	98.82
E3	*light*	8.65	49.11	644.61	53.18
E4	*light*	14.34	31.52	512.28	70.66
C1	*light*	12.91	53.36	569.68	84.4
C2	*light*	13.09	41.92	533.21	93.54
C3	*light*	9.21	73.01	586.69	71.87
C4	*light*	10.37	38.96	535.82	66.09
L1	*light*	11.15	50.11	523.56	77.90
L2	*light*	12.47	54.41	601.82	89.28
L3	*light*	12.58	68.15	609.63	74.91
L4	*light*	8.75	51.32	567	50.11
E1	*dark*	14.07	126.77	365.14	1016.75
E2	*dark*	15.92	107.42	356.60	1013.21
E3	*dark*	15.55	102.32	407.37	993.87
E4	*dark*	20.04	82.01	377.02	983.23
C1	*dark*	17.51	118.99	380.94	1003.71
C2	*dark*	14.25	145.00	410.02	976.79
C3	*dark*	12.66	165.28	425	985.10
C4	*dark*	13.21	111.73	392.42	999.23
L1	*dark*	15.14	96.06	399.28	983.42
L2	*dark*	14.93	67.45	360.64	1009.37
L3	*dark*	13.33	193.27	419.07	989.53
L4	*dark*	11.85	137.77	434.39	977.81

Appendix Table A3.57: Summary of sleep parameters under 500 lux LD20:04 second set of experiments. Mean length of sleep bouts, Total sleep time, and Latency to sleep are in minutes. Stock labels: E – *early*, C – *control*, and L – *late*. 1, 2, 3, 4 are replicate populations. Latency is calculated from ZT00.

Stock	**Phase**	**Number of sleep bouts**	**Mean length of sleep bouts**	**Total sleep time**	**Latency to sleep**
E1	*light*	9.06	30.70	601.19	61.48

E2	*light*	12.81	43.48	623.87	78.44
E3	*light*	8.65	50.60	813.55	44.88
E4	*light*	14.34	31.13	635.07	64.70
C1	*light*	12.91	36.46	706.17	65.65
C2	*light*	13.09	33.36	699.14	55.84
C3	*light*	9.21	26.45	682.28	47.79
C4	*light*	10.37	35.10	705.45	54.93
L1	*light*	11.15	30.87	632.79	77.51
L2	*light*	12.47	28.89	687.97	46.42
L3	*light*	12.58	25.98	675.99	73.80
L4	*light*	8.75	33.29	674.68	50.91
E1	*dark*	14.07	60.64	182.92	1221.53
E2	*dark*	15.92	64.61	187.5	1221.64
E3	*dark*	15.55	140.28	221.11	1219.13
E4	*dark*	20.04	72.23	189.81	1225.10
C1	*dark*	17.51	99.20	197.78	1226.14
C2	*dark*	14.25	111.11	208.78	1221.59
C3	*dark*	12.66	82.76	202.47	1222.30
C4	*dark*	13.21	102.22	208.57	1216.10
L1	*dark*	15.14	65.89	204.29	1221.98
L2	*dark*	14.93	83.63	213.64	1219.81
L3	*dark*	13.33	91.05	207.01	1228
L4	*dark*	11.85	85.27	204.32	1220.39

Appendix Table A3.58: Summary of sleep parameters under 500 lux LD12:12 second set of experiments. Mean length of sleep bouts, Total sleep time, and Latency to sleep are in minutes. Stock labels: E – *early*, C – *control*, and L – *late*. 1, 2, 3, 4 are replicate populations. Latency is calculated from ZT00.

Stock	**Phase**	**Number of sleep bouts**	**Mean length of sleep bouts**	**Total sleep time**	**Latency to sleep**
E1	*light*	9.06	49.90	387.69	65.11
E2	*light*	12.81	37.69	365.86	103.67
E3	*light*	8.65	53.07	479.86	66.40
E4	*light*	14.34	41.19	407.45	70.12
C1	*light*	12.91	33.89	399.24	66.28
C2	*light*	13.09	38.79	406.38	72.39
C3	*light*	9.21	71.20	434.58	73.33
C4	*light*	10.37	46.15	423.08	62.01
L1	*light*	11.15	61.17	391.19	74.34
L2	*light*	12.47	44.73	389.23	93.08
L3	*light*	12.58	56.95	433.26	71.56

L4	*light*	8.75	59.75	415.60	56.94
E1	*dark*	14.07	65.51	498.18	755.24
E2	*dark*	15.92	35.81	461.65	761.60
E3	*dark*	15.55	57.98	584.79	743.12
E4	*dark*	20.04	38.64	491.92	746.69
C1	*dark*	17.51	63.58	497.69	750.61
C2	*dark*	14.25	71.73	514.21	759.16
C3	*dark*	12.66	99.24	579.07	746.33
C4	*dark*	13.21	64.59	518.29	748.14
L1	*dark*	15.14	49.75	473.22	754.54
L2	*dark*	14.93	43.47	506.15	743.87
L3	*dark*	13.33	71.36	552.56	749.57
L4	*dark*	11.85	66.39	549.56	752.24

Appendix 4

Supplementary information for Chapter 4

Appendix Table A4.1: SNP characterization of significant SNPs exclusively in *early* populations from pairwise F_{ST} comparisons.

LOW	**112**
splice_region_variant&intron_variant	6
splice_region_variant&synonymous_variant	1
stop_retained_variant	1
synonymous_variant	104
MODERATE	**15**
missense_variant	15
MODIFIER	**559**
3_prime_UTR_variant	16
5_prime_UTR_variant	13
downstream_gene_variant	71
intergenic_region	58
intragenic_variant	6
intron_variant	98
non_coding_transcript_exon_variant	4
upstream_gene_variant	293
Grand Total	**686**

Appendix Table A4.2: SNP characterization of significant SNPs exclusively in *late* populations from pairwise F_{ST} comparisons.

LOW	**9**
synonymous_variant	9
MODERATE	**2**
missense_variant	2
MODIFIER	**68**
3_prime_UTR_variant	2
5_prime_UTR_variant	3
downstream_gene_variant	20
intergenic_region	6
intron_variant	19
upstream_gene_variant	18
Grand Total	**79**

Appendix Table A4.3: SNP characterization of significant SNPs exclusively in *early* populations from CMH tests.

HIGH	**1**
stop_gained	1
LOW	**115**
5_prime_UTR_premature_start_codon_gain_variant	1

splice_region_variant&intron_variant	4
splice_region_variant&synonymous_variant	1
synonymous_variant	109
MODERATE	**24**
missense_variant	23
missense_variant&splice_region_variant	1
MODIFIER	**147**
3_prime_UTR_variant	13
5_prime_UTR_variant	2
downstream_gene_variant	22
intergenic_region	9
intragenic_variant	5
intron_variant	19
non_coding_transcript_exon_variant	2
upstream_gene_variant	75
Grand Total	**287**

Appendix Table A4.4: SNP characterization of significant SNPs exclusively in *late* populations from CMH tests.

LOW	**110**
5_prime_UTR_premature_start_codon_gain_variant	1
splice_region_variant&intron_variant	3
splice_region_variant&synonymous_variant	2
synonymous_variant	104
MODERATE	**27**
missense_variant	27
MODIFIER	**150**
3_prime_UTR_variant	3
5_prime_UTR_variant	3
downstream_gene_variant	49
intergenic_region	6
intragenic_variant	2
intron_variant	37
upstream_gene_variant	50
Grand Total	**287**

Appendix Table A4.5: SNP characterization of significant SNPs in *timeless* and *vrille* exclusively in *early* populations from CMH tests.

Type of variant	**Category of variant**	**Gene name**
3_prime_UTR_variant	MODIFIER	tim
3_prime_UTR_variant	MODIFIER	tim

synonymous_variant	LOW	tim
missense_variant&splice_region_variant	MODERATE	tim
intron_variant	MODIFIER	tim
intron_variant	MODIFIER	tim
intron_variant	MODIFIER	tim
intron_variant	MODIFIER	tim
intron_variant	MODIFIER	tim
synonymous_variant	LOW	tim
synonymous_variant	LOW	tim
synonymous_variant	LOW	tim
5_prime_UTR_variant	MODIFIER	tim
upstream_gene_variant	MODIFIER	tim
upstream_gene_variant	MODIFIER	tim

Type of variant	**Category of variant**	**Gene name**
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri
upstream_gene_variant	MODIFIER	vri

Appendix Table A4.6: SNP characterization of significant SNPs in *timeless* exclusively in *late* populations from CMH tests.

Type of variant	**Category of variant**	**Gene name**
intron_variant	MODIFIER	tim
splice_region_variant&synonymous_variant	LOW	tim
synonymous_variant	LOW	tim
synonymous_variant	LOW	tim
intron_variant	MODIFIER	tim
synonymous_variant	LOW	tim
synonymous_variant	LOW	tim
synonymous_variant	LOW	tim
synonymous_variant	LOW	tim

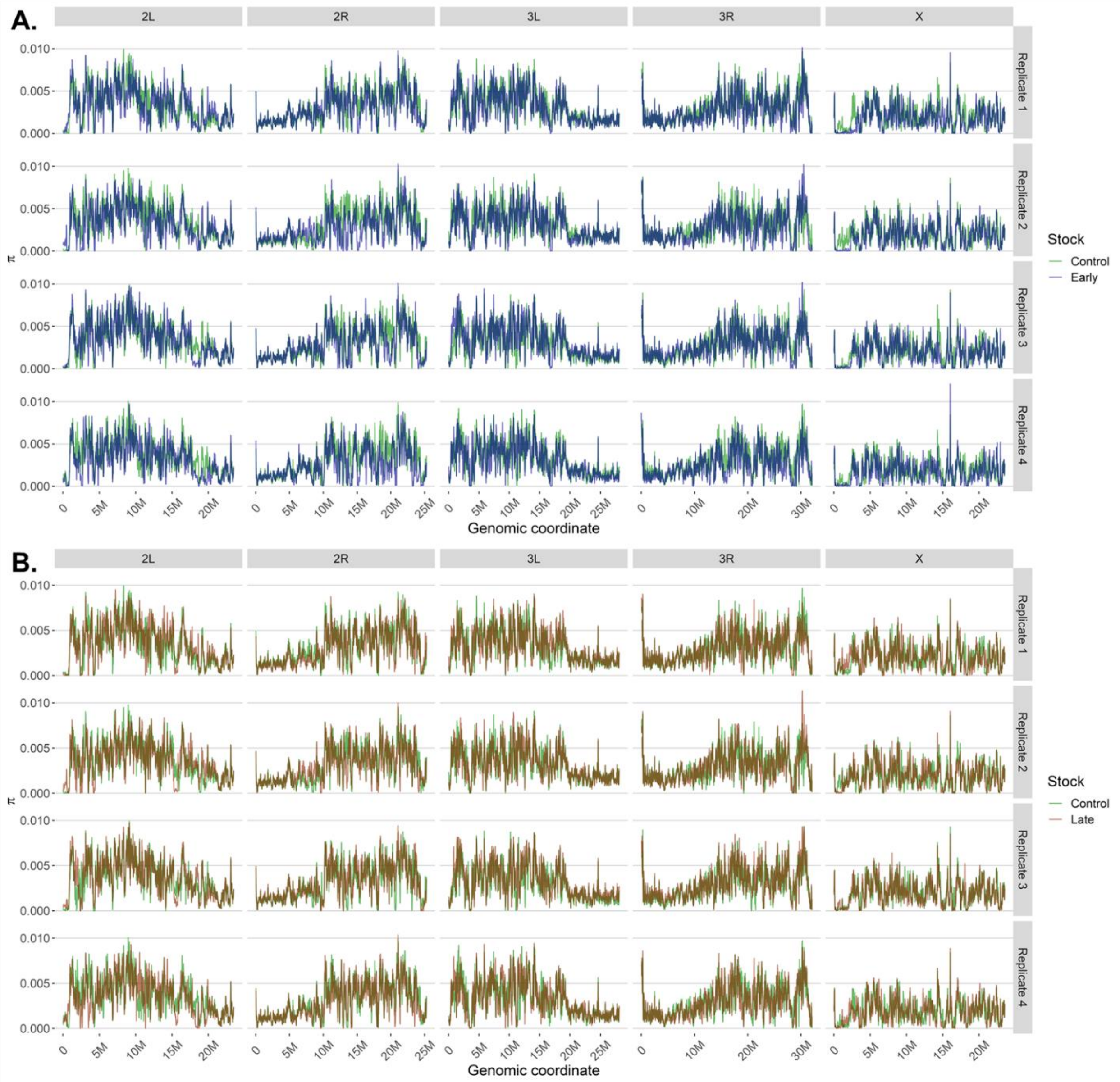
Appendix Table A4.7: Average Tajima's *D* values in selective sweep windows detected in all four replicates of the *early* population using *Pool-hmm*.

Sweep regions	**Average *D* Value**
2L:6427152-6443453	0.075
2R:165886-241481	-0.778
2R:243490-286145	-0.616
3L:24218954-24224496	-0.356
3L:24595121-24601240	-0.495
3L:26072652-26089293	-0.802
3L:26387768-26413463	-0.710
3L:26836519-26841797	-0.569
3L:27212412-27222493	-0.536
3L:27237197-27252886	-0.571
3L:27565708-27571919	-0.587
3L:27604816-27609696	-0.664
3L:27796463-27801898	-0.679
3R:1378942-1389707	-0.446
3R:1731842-1750040	-0.698
3R:1985987-1994833	-0.435
X:18288670-18427409	-0.879

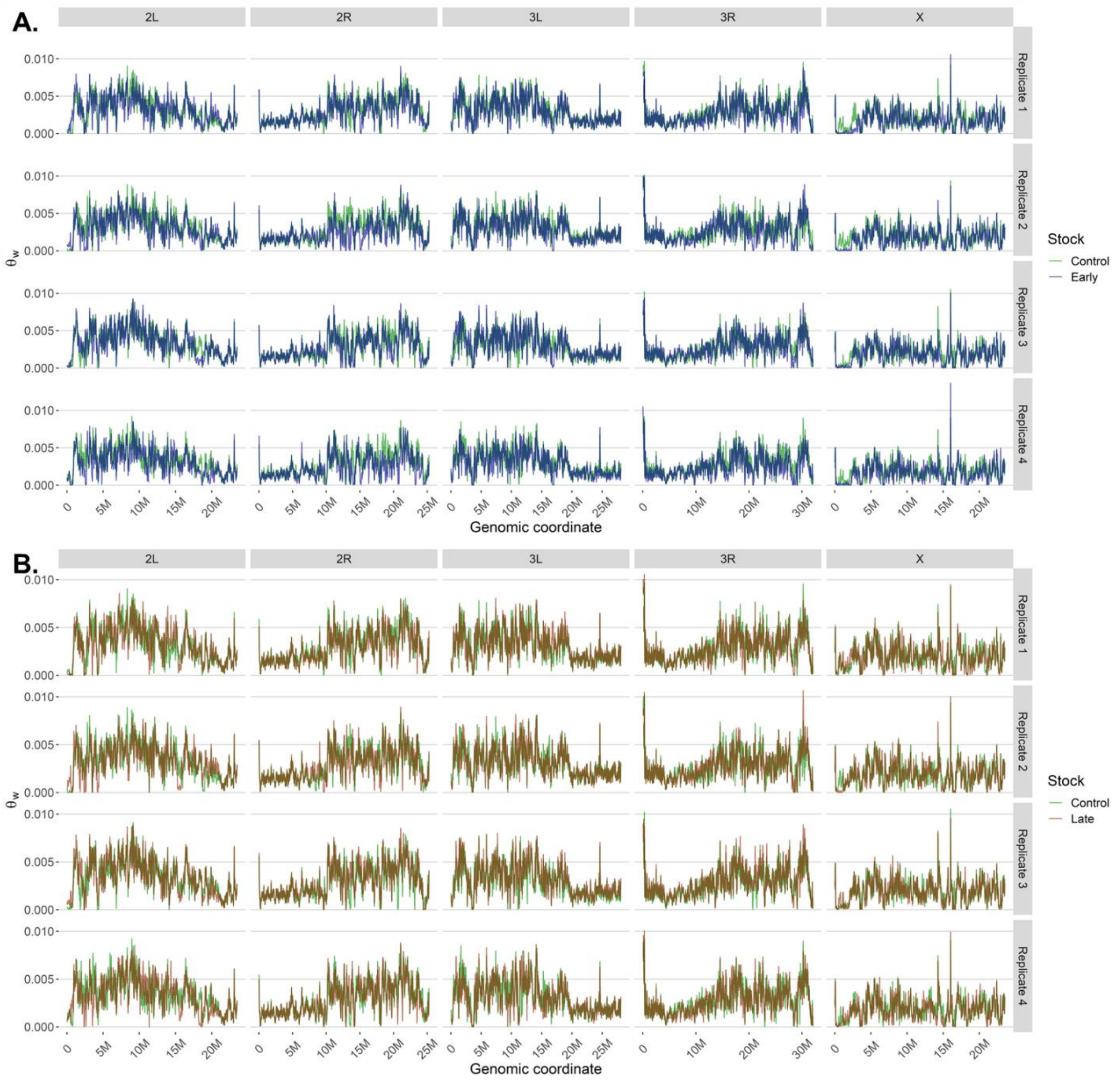
Appendix Table A4.8: Average Tajima's *D* values in selective sweep windows detected in all four replicates of the *late* population using *Pool-hmm*.

Sweep regions	Average *D* Value
2L:3203669-3209864	0.425
2L:6427152-6444980	0.200
2L:13449665-13456771	0.593
2L:20215222-20218183	-0.008
2L:20303232-20312906	-0.126
2R:167066-286594	-0.721
2R:406655-430990	-0.667
2R:663098-668522	-0.297
2R:699235-703517	-0.286
2R:1161290-1177433	-0.581
2R:1234382-1245021	-0.557
2R:2019474-2030305	-0.461
2R:2110623-2117977	-0.353
2R:3039829-3050166	-0.573
2R:3304209-3313133	-0.365
2R:3881154-3890194	-0.457
2R:3960674-3999918	-0.696
2R:4300402-4307442	-0.250
2R:4310363-4320760	-0.616
2R:5902477-5909622	0.359
3L:6405972-6431765	0.424
3L:21365975-21378683	-0.056
3L:22990230-22996709	0.061
3L:23573960-23578445	-0.389
3L:26382695-26413457	-0.584
3L:26674894-26682887	-0.600
3L:26836552-26844539	-0.411
3L:27212412-27222542	-0.490
3L:27225036-27252861	-0.545
3L:27565631-27571919	-0.608
3L:27814609-27827044	-0.514
3R:220082-229674	-0.464
3R:685326-691751	-0.648
3R:734298-744666	-0.479
3R:1369553-1375549	-0.380
3R:1378915-1389707	-0.430
3R:1729975-1744282	-0.751
3R:2477862-2488599	-0.373
3R:4045480-4057938	-0.428
3R:7731336-7736692	0.383
X:1353135-1367904	-0.397
X:15321393-15348579	-1.010
X:16270099-16324145	-0.732
X:16340838-16356930	-0.212
X:16417169-16485959	-1.071
X:16596275-16696614	-0.639
X:18182443-18417500	-0.763
X:21039100-21137905	-0.241
X:21806666-21821848	-0.366
X:22022951-22032923	-0.398
X:22060069-22066818	-0.378
X:22757082-22841415	-0.682
X:23120341-23150558	-0.635

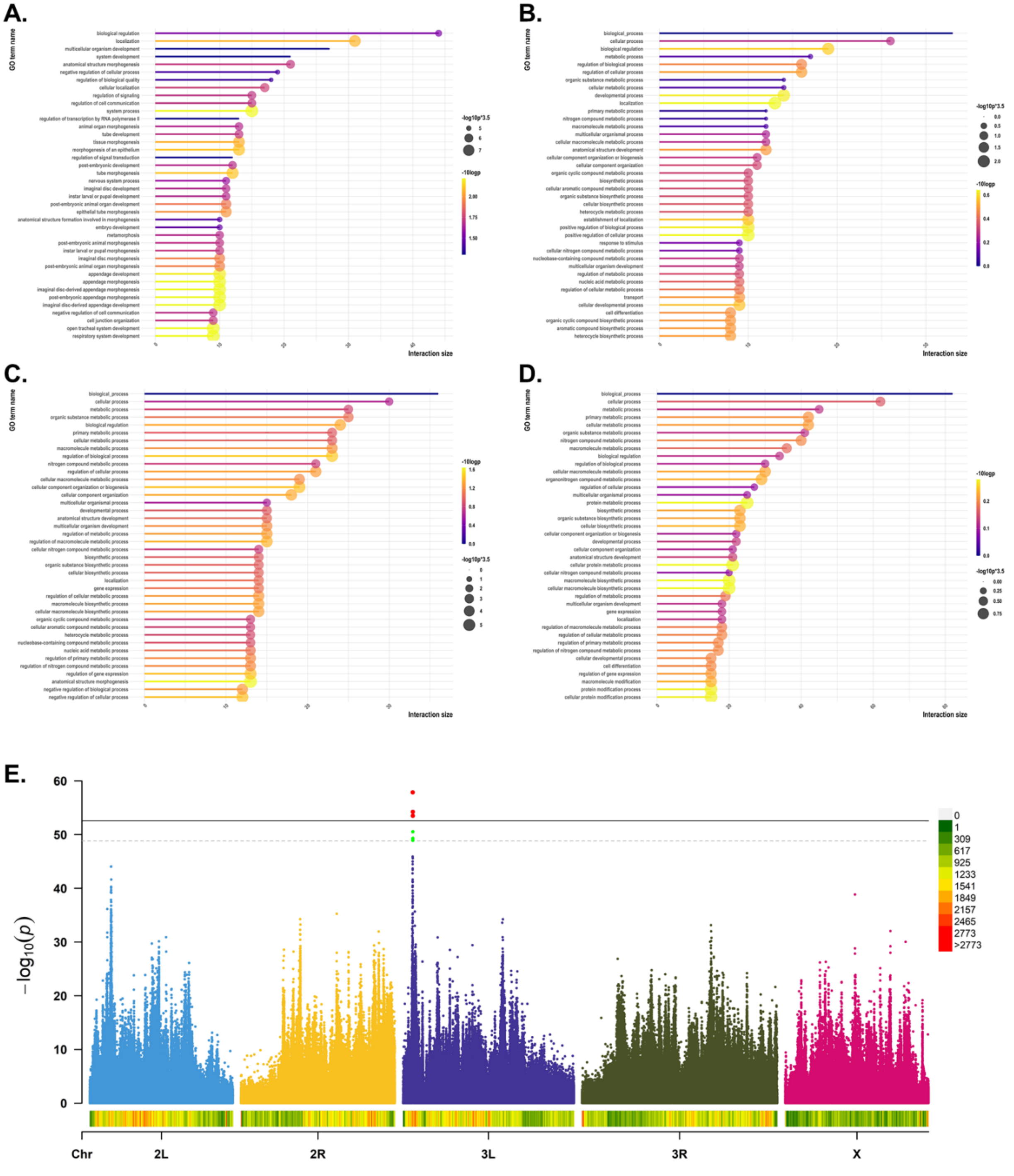
Appendix Figure A4.9: Description of genome-wide π values. (**A & B**) π values across 50Kb windows with 10Kb steps along different chromosomal arms (axes labels on top) and four replicate populations (axes labels on right) of *control*, *early* (**A**), and *late* (**B**) populations. Median genome-wide π values for the populations are – *control*: 0.00262, *early*: 0.00242, and *late*: 0.002641. Line colors: blue – *early*, green – *control*, red – *late*.



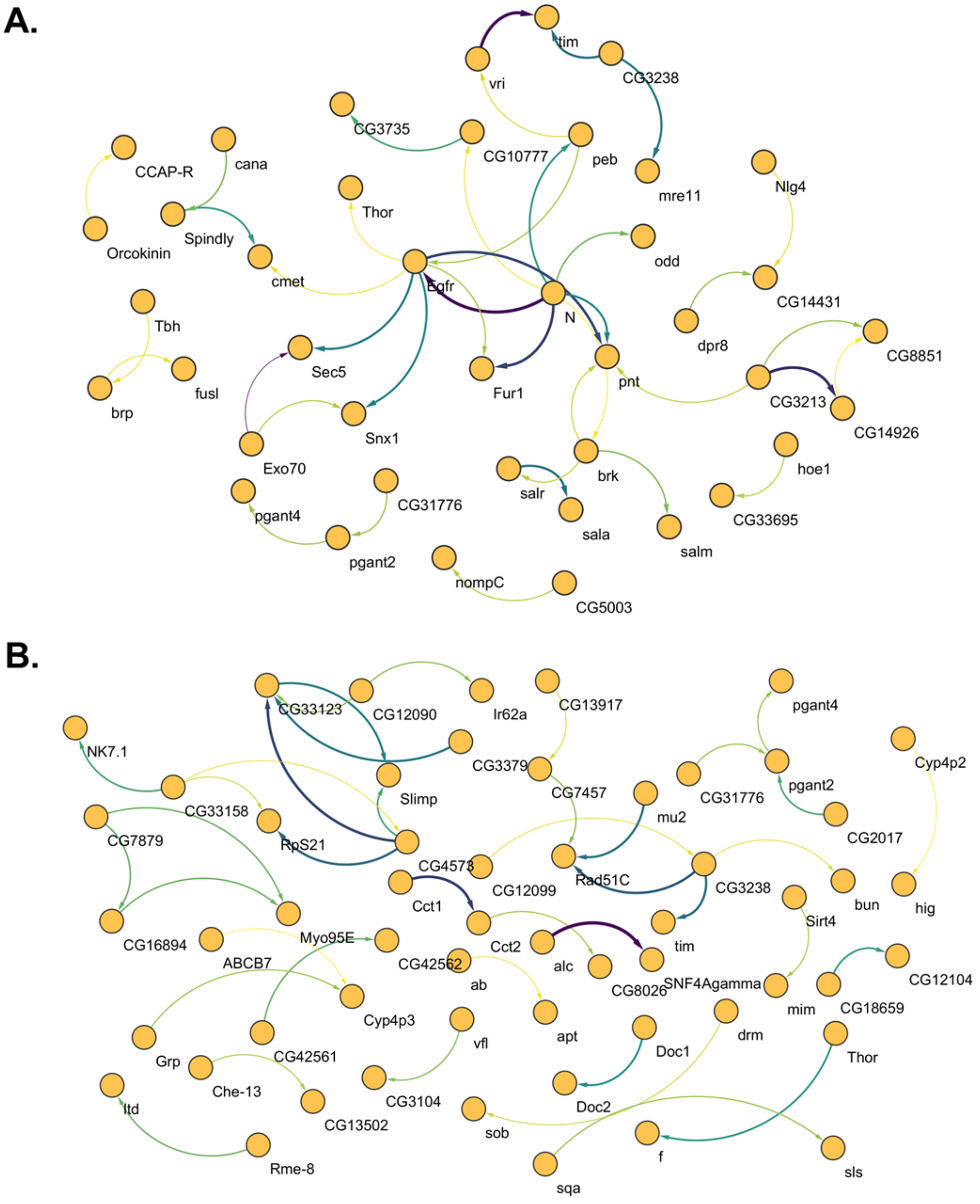
Appendix Figure A4.10: Description of genome-wide θ_w values. (**A & B**) θ_w values across 50Kb windows with 10Kb steps along different chromosomal arms (axes labels on top) and four replicate populations (axes labels on right) of *control*, *early* (**A**), and *late* (**B**) populations. Median genome-wide θ_w values for the populations are – *control*: 0.002553, *early*: 0.002356, and *late*: 0.002578. Line colors: blue – *early*, green – *control*, red – *late*.



Appendix Figure A4.11: Gene ontology analysis and *control-late* CMH test. (**A & B**) Enrichment analysis of GO terms of genes from *control-early* (top 40 - **A**) and *control-late* (top 40 - **B**) pairwise F_{ST} comparisons. (**C & D**) Enrichment analysis of GO terms of genes from *control-early* (top 40 - **C**) and *control-late* (top 40 - **D**) pairwise CMH tests. (**E**) -log10(*p*) values plotted as Manhattan plots along different chromosomal arms from *control-late* pairwise CMH tests. Threshold for *late* was derived from permutation simulations ($CMH_{pcorrected}$ = 4.556171-55). The length of the horizontal lines in the lollipop charts (**A, B, C & D**) are the interaction sizes as derived from the GO analysis with g:Profiler and the size of the bubbles depict -log10(*p*) values from the enrichment analysis, scaled by a constant of 3.5 for representation purposes, and the color codes also depict the same. In (**E**) the SNP densities calculated from 0.1Mb non-overlapping windows were plotted at the bottom of the plot and the color guide is on the right.

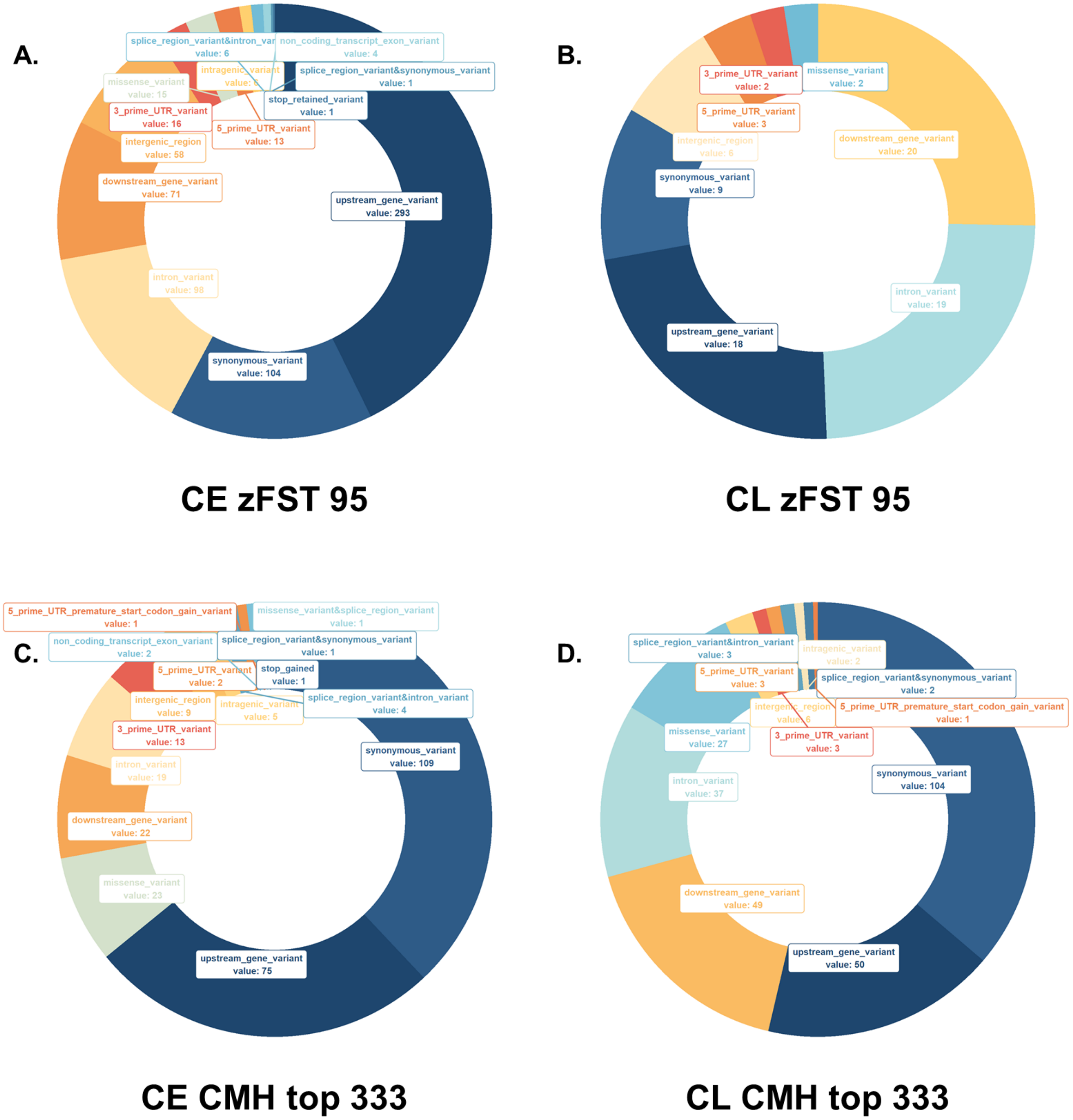


Appendix Figure A4.12: Interaction network for F$_{ST}$+CMH genes. (**A & B**) protein-protein interaction network constructed with all genes from allele frequency bases tests (F$_{ST}$ and CMH) for *control-early* (**A**) and *control-late* (**B**) comparisons. Each node represents a protein coding gene, edges are colored and weighted in a continuous scale based on STRING combined score of that particular interaction. Total node numbers were 101 and 130, and total edge numbers were 41 and 40 for *early* and *late* populations respectively. Singletons (nodes without edges) were removed for ease of visualization.


A.
tim
vri
CG3238
CG3735
CG10777
peb
CCAP-R
cana
mre11
Nlg4
Thor
Spindly
Orcokinin
odd
cmet
Egfr
CG14431
N
Tbh
dpr8
CG8851
Sec5
pnt
Fur1
CG3213
fusl
brp
Snx1
CG14926
hoe1
Exo70
brk
salr
CG31776
pgant4
sala
CG33695
salm
pgant2
nompC
CG5003
B.
CG33123
CG12090
Ir62a
CG13917
pgant4
NK7.1
CG3379
Cyp4p2
Slimp
CG7457
CG31776
pgant2
CG33158
mu2
CG7879
RpS21
CG2017
CG4573
Rad51C
CG3238
Cct1
CG12099
bun
hig
Sirt4
CG16894
Myo95E
Cct2
tim
ABCB7
CG42562
alc
ab
CG8026
SNF4Agamma
CG12104
mim
CG18659
Cyp4p3
apt
drm
vfl
Doc1
Thor
Grp
CG42561
Che-13
CG3104
Doc2
ltd
CG13502
sob
f
sls
Rme-8
sqa

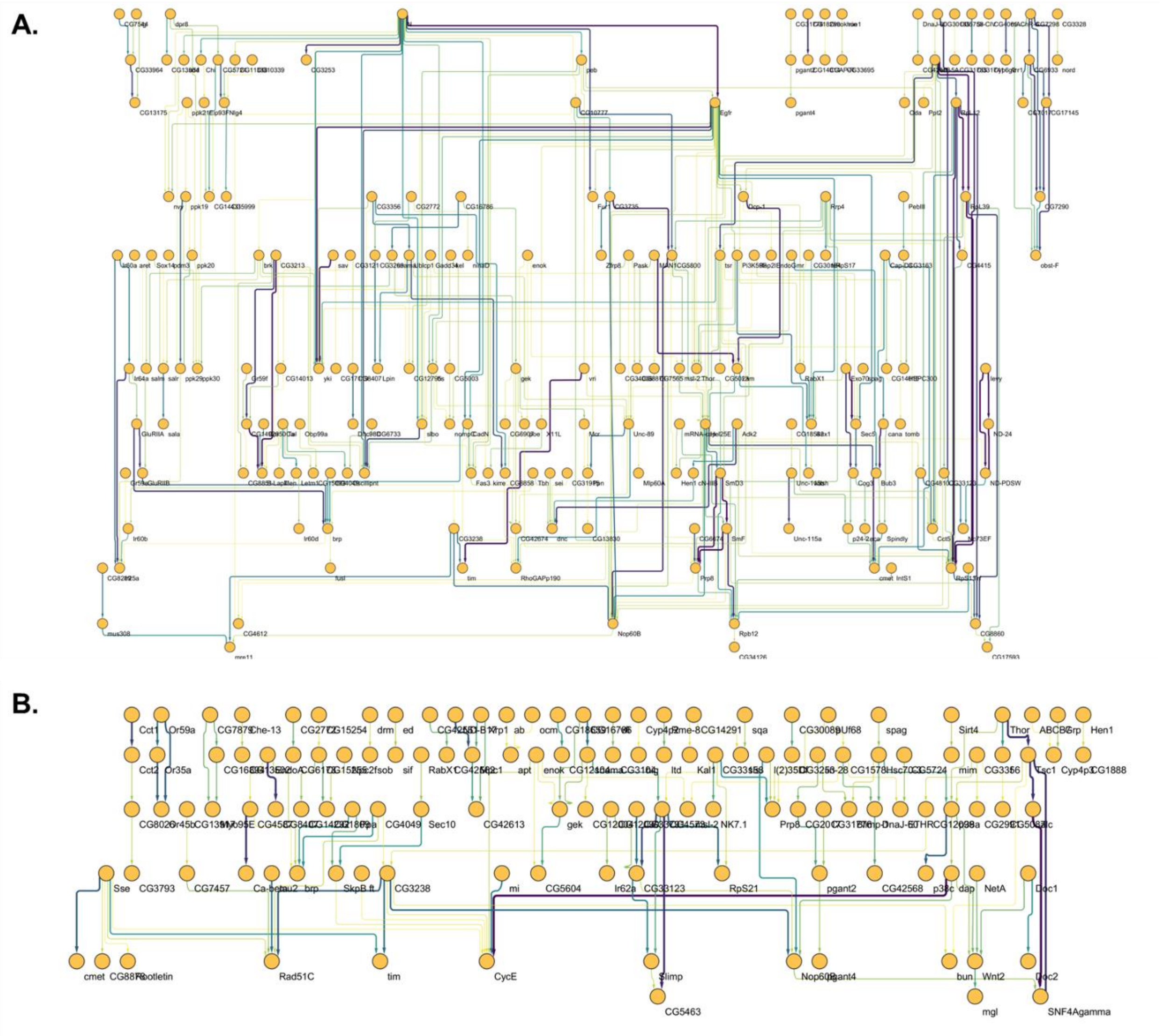
Appendix Figure A4.13: Type of variants based on their effect type-description in F$_{ST}$ and CMH test based comparisons. (**A & B**) Composition of different type of variants significantly different in terms of allele frequency in pairwise F$_{ST}$ comparison of *control-early* (**A**) and *control-late* (**B**). (**C & D**)

Composition of different type of variants significantly different in terms of allele frequency in pairwise CMH tests of *control-early* (**C**) and *control-late* (**D**).

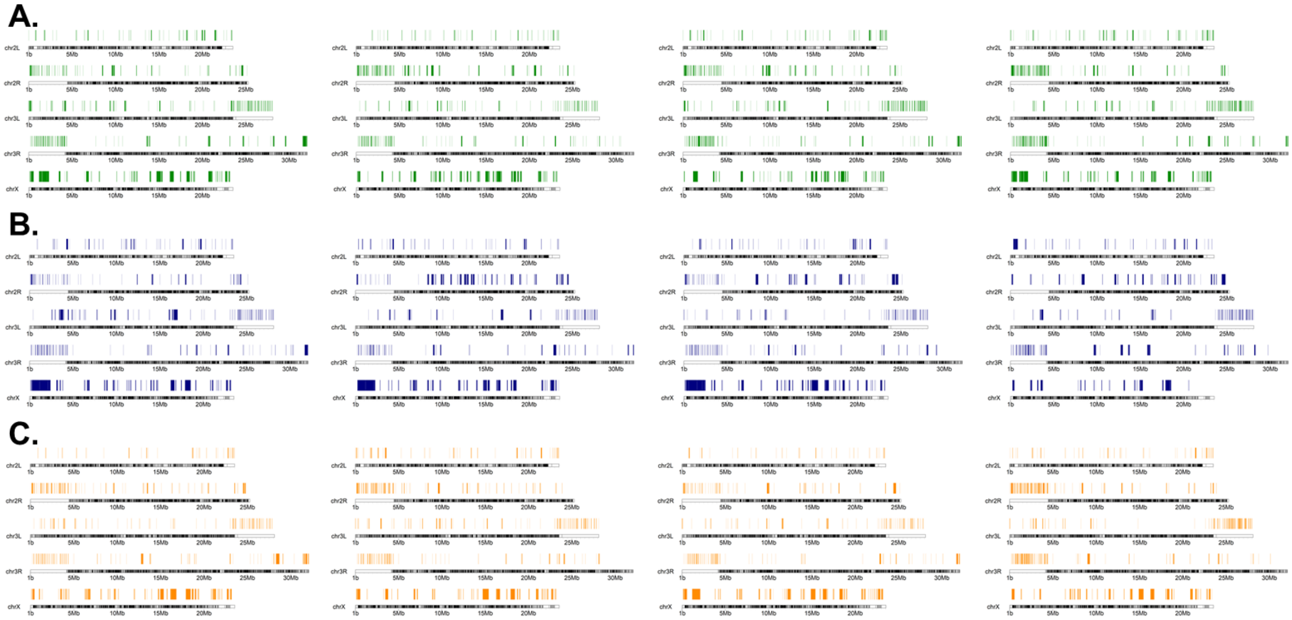


Appendix Figure A4.14: Interaction network for Tajima.D+F$_{ST}$+CMH genes. (**A & B**) protein-protein interaction network constructed with all genes from all tests (Comparison of Tajima's *D* values, pairwise

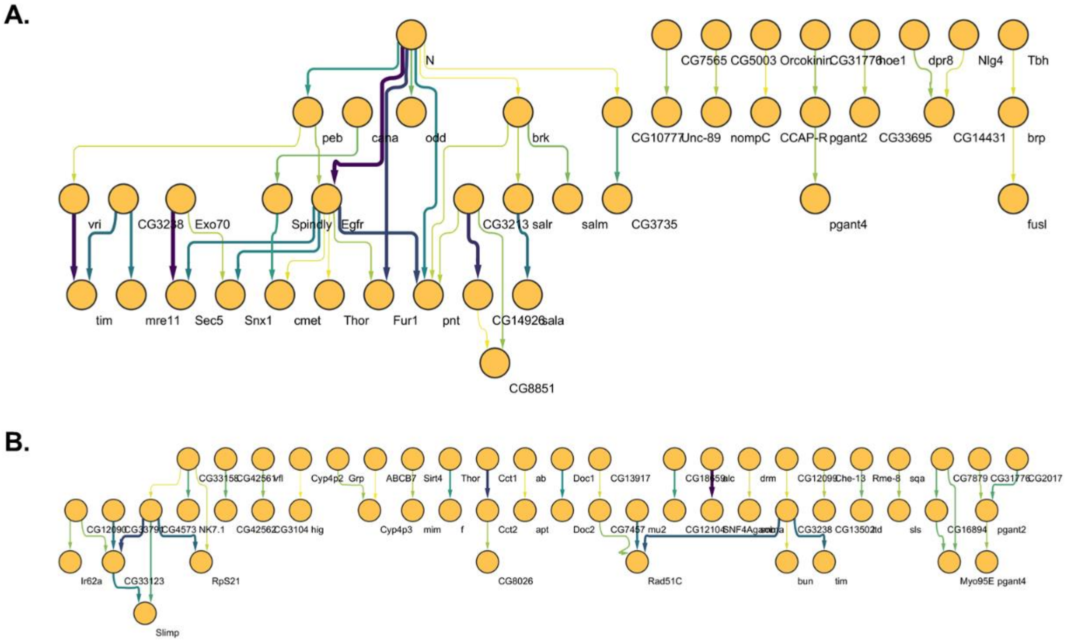
F$_{ST}$ comparison, and CMH test) for *control-early* (**A**) and *control-late* (**B**) comparisons. Each node represents a protein coding gene, edges are colored and weighted in a continuous scale based on STRING combined score of that particular interaction. Total node numbers were 314 and 214, and total edge numbers were 347 and 131 for *early* and *late* populations respectively. Singletons (nodes without edges) were removed for ease of visualization.


A.
dpr8
CG33964
CG3253
CG13175
Egfr
pgant4
CG3356
CG2772
Pebll
CG7290
enok
obst-F
CG14013
yki
Lpin
gek
vri
RabX1
GluRIIA
sala
Obp99a
stbo
Unc-89
Adk2
ND-24
Fas3
kirre
Tbh
sei
Mlp60A
Hen1
Unc-115b
Cog3
Bub3
ND-PDSW
Ir60b
Ir60d
brp
CG3238
CG42674
CG13830
SmF
Unc-115a
Spindly
tim
RhoGAPp190
Prp8
cmet
lrtS1
mus308
CG4612
Nop60B
Rpb12
CG8860
B.
Cct1
Or59a
Che-13
CG15254
drm
ed
ab
ocm
CG14291
sqa
spag
Sirt4
Thor
Hen1
Cct2
Or35a
sif
apt
enok
ltd
Kal1
mim
Cyp4p3
CG1888
CG8026
CG4049
Sec10
CG42613
gek
NK7.1
Prp8
Sse
CG3793
CG7457
brp
SkpB
CG3238
mi
CG5604
Ir62a
CG33123
RpS21
pgant2
CG42568
dap
NetA
Doc1
cmet
Rad51C
tim
CycE
Slimp
pgant4
bun
Wnt2
Doc2
CG5463
mgl
SNF4Agamma

Appendix Figure A4.15: Results from selective sweep analysis using Pool-hmm. (**A, B, & C**) Spans of selective sweep windows along different chromosomal arms and four replicates of *control* (green - **A**), *early* (blue - **B**), and *late* (red - **C**) populations as detected by *Pool-hmm*.



Appendix Figure A4.16: Interaction network for F_{ST}+CMH genes. (**A & B**) protein-protein interaction network constructed with all genes from allele frequency bases tests (F_{ST} and CMH) for *control-early* (**A**) and *control-late* (**B**) comparisons. Each node represents a protein coding gene, edges are colored and weighted in a continuous scale based on STRING combined score of that particular interaction. Total node numbers were 101 and 130, and total edge numbers were 41 and 40 for *early* and *late* populations respectively. Singletons (nodes without edges) were removed for ease of visualization. Same as *Appendix Figure A4.16c & D*, but in orthogonal arrangement.


A.
N
CG7565
CG5003
Orcokinin
CG31776
hoe1
dpr8
Nlg4
Tbh
peb
cana
odd
brk
CG10777
Unc-89
nompC
CCAP-R
pgant2
CG33695
CG14431
brp
vri
CG3238
Exo70
Spindly
Egfr
CG3213
salr
salm
CG3735
pgant4
fusl
tim
mre11
Sec5
Snx1
cmet
Thor
Fur1
pnt
CG14926
sala
CG8851
B.
CG33158
CG42561
vfl
Cyp4p2
Grp
ABCB7
Sirt4
Thor
Cct1
ab
Doc1
CG13917
CG18659
alc
drm
CG12099
Che-13
Rme-8
sqa
CG7879
CG31776
CG2017
CG12090
CG33791
CG4573
NK7.1
CG42562
CG3104
hig
Cyp4p3
mim
f
Cct2
apt
Doc2
CG7457
mu2
CG12104
SNF4Agamma
CG3238
CG13502
td
sls
CG16894
pgant2
Ir62a
CG33123
RpS21
CG8026
Rad51C
bun
tim
Myo95E
pgant4
Slimp

Appendix 5

Supplementary information for Chapter 5

Appendix Table A5.1: Identified "hits" and the genes the deletions contain.

150281	150423	150035	150049	150159	150434	150525	150007	150433	150023
Eip75B	*CG32486*	*CR43276*	*asRNA:CR45046*	*Men*	*CG32137*	*rump*	*lncRNA:CR44821*	*caps*	*Hr39*
lncRNA:CR45921	*PAN3*	*CR43275*	*CG31465*	*lncRNA:CR44231*	*Meics*	*Ras85D*	*Hsp60C*	*CG32119*	*l(2)k14505*
snoRNA:Me28S-A30	*snRNA:U5:63BC*	*CR43105*	*fit*	*Ugt37A2*	*ssp2*	*Rlb1*	*CG12511*	*tRNA:Asp-GTC-1-10*	*CG8671*
lncRNA:CR45922	*Cht7*	*CR43186*	*CG17819*	*Ugt37A3*	*Nxf3*	*mRpL47*	*lncRNA:CR44819*	*tRNA:CR32123:Psi*	*mir-4974*
CG32192	*tRNA:Cys-GCA-1-1*	*dup*	*dnd*	*lncRNA:CR46238*	*CG13738*	*Kdm3*	*lncRNA:CR44820*	*lncRNA:CR43146*	*Cyp6t2Psi*
CG42393	*tRNA:Cys-GCA-1-2*	*SRPK*	*CG6678*	*beat-Vc*	*Hsc70-1*	*CG8176*	*CG7236*	*lncRNA:CR43913*	*Mondo*
CG44006	*tRNA:Cys-GCA-1-3*	*Stacl*	*CG43844*	*lncRNA:CR44233*	*CG17364*	*mir-2283*	*CG14007*	*lncRNA:CR45752*	*Gr39a*
CG44005	*tRNA:Lys-CTT-1-12*	*lncRNA:CR45020*	*Qsox3*	*lncRNA:CR44232*	*CG17362*	*by*	*asRNA:CR44818*	*lncRNA:CR45753*	*crc*
CG44004	*tRNA:Lys-CTT-1-13*	*Mtk*	*Qsox2*	*CG31345*	*CG9040*	*mura*	*CG34011*	*lncRNA:CR43912*	*CG46314*
CR45391	*tRNA:Cys-GCA-1-4*	*CG30472*	*Eip93F*	*lncRNA:CR46217*	*26-29-p*	*lncRNA:CR42549*	*CG11149*	*Sfp70A4*	*lncRNA:CR44785*
lncRNA:CR45392	*CG45066*	*CG34188*	*Theg*	*beat-Va*	*CG17361*	*RnpS1*	*sip1*	*lncRNA:CR45253*	*lncRNA:CR42696*
lncRNA:CR43253	*CG45067*	*CG8180*	*Idh3b*	*lncRNA:CR45587*	*CG17359*	*CG9386*	*CG11030*	*lncRNA:CR45255*	*dimm*
lncRNA:CR32194	*promL*	*CG12964*	*Mitofilin*	*lncRNA:CR44234*	*Nprl3*	*CG8199*	*CG14006*	*lncRNA:CR45254*	*Tsp39D*
CG34253	*tRNA:Met-CAT-2-1*	*Ir52a*	*mRpL35*	*CG10126*	*upSET*	*AP-1mu*	*CG11147*	*CG42481*	*dtr*
CG13698	*CG11537*	*Ir52b*	*CG6028*	*lncRNA:CR46218*	*Ptip*	*MBD-like*	*CG11029*	*CG43147*	*Gr39b*
mRpS26	*lncRNA:CR45819*	*Ir52c*	*Cchl*	*asRNA:CR44235*	*endos*	*CG9393*	*obst-E*	*SP*	*CG8665*
Polr3D	*Alg2*	*Ir52d*	*ND-42*	*d-cup*	*CG6650*	*Vps45*	*CG9171*	*lncRNA:CR43911*	*nrv3*
mus304	*Usp5*	*Pgant1*	*CG13409*	*CR33929*	*asRNA:CR42871*	*CG16789*	*asRNA:CR31912*	*lncRNA:CR44557*	*His-Psi:CR31616*
CG32195	*BtbVII*	*Khc-73*	*CG45099*	*CG10909*	*CG6661*	*CG16790*	*CG31913*	*lncRNA:CR44555*	*His-Psi:CR31615*
CG7341		*CG30471*	*CG6015*	*beat-Vb*	*Hsc70Cb*	*CG9396*	*CG14005*	*lncRNA:CR44558*	*His-Psi:CR31754*
CG42853		*CG30467*	*pit*	*lncRNA:CR44236*	*Neurl4*	*CG9399*	*CG7239*	*CG14113*	*His1:CG33801*
Cyp312a1		*CG8187*	*Fadd*	*grsm*	*CG6833*	*Kap-alpha3*	*CG11034*	*lncRNA:CR44559*	*His2B:CG33910*
lncRNA:CR44669		*asRNA:CR45927*	*scaRNA:MeU5-G38*	*Spc25*	*CG13484*	*Son*	*lncRNA:CR44575*	*CG17687*	*His-Psi:CR33802*
geko		*Vha36-1*	*how*	*Cyp304a1*	*asRNA:CR46067*	*CG8301*	*lncRNA:CR44576*	*Nplp2*	*His4:CG33909*
CG7330		*eIF2Bgamma*	*CG13408*	*CG14384*	*Frl*	*CG33654*	*CG43307*	*CG14111*	*His3:CG33803*
CG13699		*CG8192*	*BomT3*	*CG7381*	*Pex1*	*P58IPK*	*COX6CL*	*SNCF*	*His1:CG31617*
hid		*CG12963*	*BomBc3*	*CG7091*	*btl*	*bocks*	*TrissinR*	*asRNA:CR45886*	*His2B:CG17949*
lncRNA:CR45234		*CG8195*	*CG13407*	*Paip2*	*CG8100*	*CG8312*	*rau*	*CG14107*	*His2A:CG31618*
CG7320		*Flo1*	*Ir94a*	*CG31342*	*Fbp1*	*CG9427*	*CG44574*	*ImpL1*	*sncRNA:430a*
lncRNA:CR45935		*CG30466*	*Ir94b*	*CG14383*	*Sox21a*	*CG8319*	*Sfp26Ac*	*CG14110*	*His4:CG31611*

lncRNA:CR45937		*Cdk5*	*Ir94c*	*yellow-f*	*Sox21b*	*Calr*	*CG43185*	*CG10171*	*His3:CG31613*
lncRNA:CR45936		*CG8204*	*CG42390*	*yellow-f2*	*lncRNA:CR45887*	*sisRNA:CR46357*	*CG9029*	*Poc1*	*His1:CG33804*
CheA75a		*CG42524*	*SKIP*	*CG7488*	*D*	*SpdS*	*lncRNA:CR44577*	*sens*	*His2B:CG33908*
CG5103		*mir-278*	*lncRNA:CR45224*	*CG17327*	*nan*	*asRNA:CR45053*	*lncRNA:CR44578*	*CG10222*	*His-Psi:CR33805*
lncRNA:CR43306		*fus*	*Gld2*	*CG44194*	*nuf*	*Scm*	*Acp26Ab*	*flr*	*His4:CG33907*
CG13700		*asRNA:CR45143*	*mir-1010*	*CG7518*	*saturn*	*Dh44*	*Acp26Aa*	*CG32121*	*His3:CG33806*
grim		*CG8207*	*CR43696*	*CG8031*	*CG7768*	*Fst*	*lncRNA:CR44579*	*CG33263*	*His1:CG33807*
lncRNA:CR45975		*Vha14-1*	*CR43697*	*CG11656*	*CG7924*	*Dhc1*	*CG9021*	*CG14106*	*His2B:CG33906*
asRNA:CR45974		*CG30091*	*CG7084*	*CtBP*	*CG34244*	*CR43441*	*lncRNA:CR43808*	*CG14105*	*His2A:CG33808*
lncRNA:CR46353		*asRNA:CR43429*	*CG34377*	*CG46280*	*CG7906*	*Nepl12*	*bchs*	*CG10713*	*sncRNA:430b*
rpr		*Impbeta11*	*asRNA:CR46261*	*CG46281*	*lncRNA:CR45399*	*p23*	*CG14000*	*tRNA:Val-AAC-2-3*	*His4:CG33905*
		CG30082	*CG7080*	*l(3)87Df*	*lncRNA:CR44561*	*nmdyn-D7*	*CG9016*	*tRNA:Val-AAC-2-4*	*His3:CG33809*
		Cep89	*CG33721*	*ry*	*lncRNA:CR44560*	*CG9444*	*asRNA:CR43926*	*CG10154*	*His1:CG33810*
		CG30090	*CG13862*	*CG11668*	*fz*	*eca*	*dsf*	*CG10725*	*His2B:CG33904*
		asRNA:CR44372	*CG5391*	*snk*	*lncRNA:CR45025*	*CG18542*	*lncRNA:CR45288*	*CG10140*	*His-Psi:CR33811*
		CG30088	*CG5388*	*CG11670*	*CG13482*	*Unc-115b*	*lncRNA:CR45289*	*CG14109*	*His4:CG33903*
		CG30087	*asRNA:CR46054*	*CG45122*	*CG13481*	*p24-2*	*Sfp26Ad*	*cmb*	*His3:CG33812*
		CG33460	*CG5386*	*Hsc70-2*	*CG43120*	*CG32939*	*Gpdh1*	*JMJD7*	*His1:CG33813*
		CG33461	*rdhB*	*CG31157*	*CG3868*	*Unc-115a*	*lncRNA:CR44986*	*CG10738*	*His2B:CG33902*
		CG33462	*Sar1*	*CG7966*	*stwl*	*trbd*	*CG9044*	*CG10116*	*His2A:CG33814*
		CG46433	*JMJD6*	*pic*	*lncRNA:CR45233*	*dmt*	*CG13999*	*CG10089*	*sncRNA:430c*
		CG42662	*Muted*	*sim*	*CG3919*	*scaRNA:PsiU1-6*	*CG13998*	*stv*	*His4:CG33901*
		CG30080	*CG7071*	*lncRNA:CR44967*	*bbg*	*hyd*	*Vm26Ab*	*Abp1*	*His3:CG33815*
		CG30083	*CG5382*	*CG43063*	*lncRNA:CR46032*	*FoxP*	*Vm26Ac*	*Tgi*	*His1:CG33816*
		CG30089	*Polr3F*	*timeout*	*lncRNA:CR46031*	*alphaTub85E*	*Vm26Aa*	*Spt20*	*His2B:CG33900*
		Zasp52	*PyK*	*CG34308*	*CG9592*	*side-VII*	*psd*	*Vps36*	*His2A:CG33817*
		tun	*CG7069*	*lncRNA:CR46019*	*CG4613*	*Pnn*	*CG13992*	*Liprin-beta*	*sncRNA:430d*
		CG33465	*CG18596*	*2mit*	*CG43246*	*CG34409*	*Ucp4C*	*CG10710*	*His4:CG33899*
		Poxn	*CG34149*	*lncRNA:CR45109*	*Mpcp2*	*CG31415*	*Ucp4B*	*bru3*	*His3:CG33818*
		CG8249	*CG43342*	*CG8138*	*Gbs-70E*	*CG12948*	*asRNA:CR43465*	*lncRNA:CR45825*	*His1:CG33819*
		CG12970	*Gpdh3*	*CG8508*	*asRNA:CR44843*	*PpD3*	*chic*	*lncRNA:CR45178*	*His2B:CG33898*
		Gpo1	*CAH8*	*CG14380*	*CG34039*	*Tti1*	*eIF4A*	*lncRNA:CR45120*	*His2A:CG33820*
		Rif1	*CG7059*	*CG8141*	*Lk*	*Rpt3R*	*ifc*	*mir-289*	*sncRNA:430e*

		sli	*CG13857*	*CG8483*	*CG42758*	*Alg12*	*Kdm5*	*CG43184*	*His4:CG33897*
		asRNA:CR46471	*tRNA:Ser-GCT-2-3*	*CG8476*	*HGTX*	*Mpi*		*CG8757*	*His3:CG33821*
		CG33463	*tRNA:Ser-GCT-2-4*	*Ace*	*shd*	*Snap24*		*asRNA:CR46266*	*His1:CG33822*
		lncRNA:CR44373	*tRNA:Ser-GCT-1-1*	*CG11686*	*CG9628*	*CG8478*		*CG8750*	*His2B:CG33896*
		bdg	*tRNA:Ser-GCT-2-5*	*Ravus*	*RecQ5*	*MED6*		*Tsp68C*	*His2A:CG33823*
		Diap2	*CG13856*	*Su(var)3-7*	*dlp*	*Naa80*		*Hml*	*sncRNA:430f*
		bug	*CG13855*	*TBC1D5*	*lncRNA:CR45888*	*CG9471*		*CG8745*	*His4:CG33895*
		Mlf	*CG13850*	*CG8630*	*CG43121*	*Whamy*		*dysc*	*His3:CG33824*
		CG8299	*lqfR*	*CG15888*	*ome*	*topi*		*CG13737*	*His1:CG33825*
		Cyp4aa1	*Nop56*	*apn*	*tRNA:iMet-CAT-1-4*	*RpS29*		*Rgl*	*His2B:CG33894*
		COX6AL	*mats*	*Osi22*	*tRNA:iMet-CAT-1-5*	*snoRNA:Psi18S-1275*		*CG8833*	*His2A:CG33826*
		mir-4919	*pinta*	*wntD*	*CG4914*	*asRNA:CR31514*		*DCTN1-p150*	*sncRNA:430g*
		Strn-Mlck	*CG13847*	*CG8773*	*CG5048*	*CG12947*		*CG32137*	*His4:CG33893*
		CG8366	*CG12499*	*CG8774*	*gdrd*	*MtnA*		*Meics*	*His3:CG33827*
		lncRNA:CR45021	*CG34288*	*CG32473*	*CG13474*	*CG8500*		*ssp2*	*His1:CG33828*
		CG8314	*CG34376*	*CG43208*	*CG17177*	*CG12945*		*Nxf3*	*His2B:CG33892*
		Pex11ab	*Rpn7*	*CG44142*	*CG13476*	*CG8507*		*CG13738*	*His2A:CG33829*
		CG8320	*AP-2mu*	*CR46407*	*CG13473*	*CG8516*		*Hsc70-1*	*sncRNA:430h*
		ATPCL	*CG7054*	*CG45080*	*CG13471*	*asRNA:CR46144*		*CG17364*	*His4:CG33891*
			Pebp1	*PK2-R2*	*CG42507*	*CG9467*		*CG17362*	*His3:CG33830*
			CG5377	*PK2-R1*	*Trl*	*CG8526*		*CG9040*	*His1:CG33831*
			Nrx-1	*mthl12*		*FBXO11*			*His2B:CG33890*
			mir-4952	*poly*		*asRNA:CR44035*			*His2A:CG33832*
			Pfdn5	*Dic1*		*CG8534*			*sncRNA:430i*
			CG5376	*CheA87a*		*eloF*			*His4:CG33889*
			CCT1	*Lip3*		*CG16904*			*His3:CG33833*
			tHMG1	*CG34309*		*CG9459*			*His1:CG33834*
			tHMG2	*CG9813*		*CG9458*			*His2B:CG33888*
			Octbeta1R	*lncRNA:CR44138*		*CG42857*			*His2A:CG33835*
			CG5346	*CG8870*		*lncRNA:CR45790*			*sncRNA:430j*
			CG33099	*mRpS21*		*CG34302*			*His4:CG33887*
			CG33093	*Droj2*		*lncRNA:CR42858*			*His3:CG33836*
			CG5326	*CG9799*		*Teh1*			*His1:CG33837*

			lncRNA:CR45646	*CCHa2*		*Glut4EF*			*His2B:CG33886*
			lncRNA:CR45647	*CG14374*		*CG46467*			*His2A:CG33838*
			AdipoR	*CG14377*		*lncRNA:CR45029*			*sncRNA:430k*
			asRNA:CR44062	*GILT1*		*Art4*			*His4:CG33885*
			bond	*Dpm3*		*mir-9371*			*His3:CG33839*
			sit	*yellow-e3*		*Gr85a*			*His1:CG33840*
			CG33110	*yellow-e2*		*Spn85F*			*His2B:CG33884*
			CSN6	*yellow-e*		*CG5359*			*His2A:CG33841*
			Dph5	*Ir87a*		*Mical*			*sncRNA:430l*
			CG33107	*lncRNA:CR45591*		*CG31407*			*His4:CG33883*
			CG6937	*Act87E*		*CG3909*			*His3:CG33842*
			btn	*yrt*		*CG11722*			*His1:CG33843*
				lncRNA:CR42756		*mtTFB2*			*His2B:CG33882*
				lncRNA:CR45679		*CG12811*			*His2A:CG33844*
				side-IV					*sncRNA:430m*
				asRNA:CR17025					*His4:CG33881*
				asRNA:CR46354					*His3:CG33845*
				mir-252					*His1:CG33846*
				lncRNA:CR45914					*His2B:CG33880*
				lncRNA:CR45589					*His2A:CG33847*
				CG12538					*sncRNA:430n*
				CG46457					*His4:CG33879*
				lncRNA:TS26					*His3:CG33848*
				CG31337					*His1:CG33849*
				lncRNA:CR43848					*His2B:CG33878*
				CG14370					*His2A:CG33850*
				CG14369					*sncRNA:430o*
				CG9759					*His4:CG33877*
				CG9757					*His3:CG33851*
				CG9269					*His1:CG33852*
				lncRNA:CR45680					*His2B:CG33876*
				lncRNA:CR45594					*His2A:CG33853*
				CG10841					*sncRNA:430p*
				sqd					*His4:CG33875*

									His3:CG33854
									His1:CG33855
									His2B:CG33874
									His2A:CG33856
									sncRNA:430q
									His4:CG33873
									His3:CG33857
									His1:CG33858
									His2B:CG33872
									His2A:CG33859
									sncRNA:430r
									His4:CG33871
									His3:CG33860
									His1:CG33861
									His2B:CG33870
									His2A:CG33862
									sncRNA:430s
									His4:CG33869
									His3:CG33863
									His1:CG33864
									His2B:CG33868
									His2A:CG33865
									sncRNA:430t
									His-Psi:CR33867
									His3:CG33866
									His-Psi:CR31614
									Lamp1
									CR42545
									nompB
									CR42546
									CG2201
									lncRNA:CR43148
									Df31
									lncRNA:CR44993

									Ac3
									lncRNA:CR43144
									Cul2
									CG2225
									asRNA:CR46066
									eEF2
									sncRNA:254
									nolo
									step
									CG1416
									CG31612
									CG11630
									lncRNA:CR44269
									lncRNA:CR44786

References

Abhilash, L. (2020). Circadian organization and mechanisms of entrainment in populations of drosophila melanogaster selected for divergent timing of eclosion. Jawaharlal Nehru Centre for Advanced Scientific Research.

Abhilash, L., and Sharma, V.K. (2016). On the relevance of using laboratory selection to study the adaptive value of circadian clocks. Physiol. Entomol. 41, 293-306.

Abhilash, L., and Sharma, V.K. (2020). Mechanisms of photic entrainment of activity/rest rhythms in populations of Drosophila selected for divergent timing of eclosion. Chronobiol. Int. 37, 469–484.

Abhilash, L., Ghosh, A., and Sheeba, V. (2019). Selection for timing of eclosion results in co-evolution of temperature responsiveness in Drosophila melanogaster. J. Biol. Rhythms 34, 1–14.

Abhilash, L., Kalliyil, A., and Sheeba, V. (2020). Responses of activity rhythms to temperature cues evolve in Drosophila populations selected for divergent timing of eclosion. J. Exp. Biol. 223.

Aguiar, G.F., da Silva, H.P., and Marques, N. (1991). Patterns of daily allocation of sleep periods: a case study in an Amazonian riverine community. Chronobiologia 18, 9–19.

Ai, M., Blais, S., Park, J.Y., Min, S., Neubert, T.A., and Suh, G.S.B. (2013). Ionotropic glutamate receptors IR64a and IR8a form a functional odorant receptor complex in Vivo in Drosophila. J. Neurosci. 33, 10741–10749.

Akagi, K., Sarhan, M., Sultan, A.R.S., Nishida, H., Koie, A., Nakayama, T., and Ueda, H. (2016). A biological timer in the fat body comprising blimp-1, βftz-f1 and shade regulates pupation timing in Drosophila melanogaster. Dev. 143, 2410–2416.

Akin, O., and Zipursky, S.L. (2016). Frazzled promotes growth cone attachment at the source of a Netrin gradient in the Drosophila visual system. Elife 5.

Allebrandt, K. V., Amin, N., Müller-Myhsok, B., Esko, T., Teder-Laving, M., Azevedo, R.V.D.M., Hayward, C., Van Mill, J., Vogelzangs, N., Green, E.W., et al. (2013). A KATP channel gene effect on sleep duration: From genome-wide association studies to function in Drosophila. Mol. Psychiatry 18, 122–132.

An, H., Zhu, Z., Zhou, C., Geng, P., Xu, H., Wang, H., Chen, R., Qu, X., Qian, H., Gao, Y., et al. (2014). Chronotype and a PERIOD3 variable number tandem repeat polymorphism in Han Chinese pilots. Int. J. Clin. Exp. Med. 7, 3770–3776.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics.

Anduaga, A.M., Evanta, N., Patop, I.L., Bartok, O., Weiss, R., and Kadener, S. (2019). Thermosensitive alternative splicing senses and mediates 2 temperature adaptation in drosophila. Elife 8.

Archer, S.N., Robilliard, D.L., Skene, D.J., Smits, M., Williams, A., Arendt, J., and Von Schantz, M. (2003). A length polymorphism in the circadian clock gene Per3 is linked to delayed sleep phase syndrome and extreme diurnal preference. Sleep 26, 413–415.

Aschoff, J. (1955). Tagesperiodik bei Mäusestämmen unter konstanten Umgebungsbedingungen. Pflüger’s Arch. Für Die Gesamte Physiol. Des Menschen Und Der Tiere 262, 51–59.

Aschoff, J. (1960). Exogenous and endogenous components in circadian rhythms. Cold Spring Harb. Symp. Quant. Biol. 25, 11–28.

Aschoff, J. (1965). Circadian rhythms in man. Science 148, 1427–1432.

Aschoff, J., and von Goetz, C. (1988). Masking of circadian activity rhythms in male golden hamsters by the presence of females. Behav. Ecol. Sociobiol. 22, 409–412.

Aschoff, J., and Honma, K.I. (1999). Masking and parametric effects of high-frequency light-dark cycles. Jpn. J. Physiol. 49, 11–18.

Aschoff, J., and Meyer-Lohmann, J. (1954). Angeborene 24-stunden-periodik beim kücken. Pflüger's Arch. Für Die Gesamte Physiol. Des Menschen Und Der Tiere 260, 170–176.

Aschoff, J., and Pohl, H. (1978). Phase relations between a circadian rhythm and its zeitgeber within the range of entrainment. Naturwissenschaften 65, 80–84.

Aschoff, J., and Wever, R. (1962). Über Phasenbeziehungen zwischen biologischer Tagesperiodik und Zeitgeberperiodik. Z. Vgl. Physiol. 46, 115–128.

Auger, R.R., Burgess, H.J., Emens, J.S., Deriy, L. V., Thomas, S.M., and Sharkey, K.M. (2015). Clinical practice guideline for the treatment of intrinsic circadian rhythm sleep-wake disorders: Advanced Sleep-Wake Phase Disorder (ASWPD), Delayed Sleep-Wake Phase Disorder (DSWPD), Non-24-Hour Sleep-Wake Rhythm Disorder (N24SWD), and Irregular Sleep-W. J. Clin. Sleep Med. 11, 1199–1236.

Baehrecke, E.H., and Thummel, C.S. (1995). The Drosophila E93 Gene from the 93F Early Puff Displays Stage- and Tissue-Specific Regulation by 20-Hydroxyecdysone. Dev. Biol. 171, 85–97.

Baker, J.D., McNabb, S.L., and Truman, J.W. (1999). The hormonal coordination of behavior and physiology at adult ecdysis in Drosophila melanogaster. J. Exp. Biol. 202, 3037–3048.

Barclay, N.L., Eley, T.C., Mill, J., Wong, C.C.Y., Zavos, H.M.S., Archer, S.N., and Gregory, A.M. (2011). Sleep quality and diurnal preference in a sample of young adults: Associations with 5HTTLPR, PER3, and CLOCK 3111. Am. J. Med. Genet. Part B Neuropsychiatr. Genet. 156, 681–690.

Barde, P., and Barde, M. (2012). What to use to express the variability of data: Standard deviation or standard error of mean? Perspect. Clin. Res. 3, 113.

Barghi, N., Tobler, R., Nolte, V., Jakšić, A.M., Mallard, F., Otte, K.A., Dolezal, M., Taus, T., Kofler, R., and Schlötterer, C. (2019). Genetic redundancy fuels polygenic adaptation in Drosophila. PLoS Biol. 17, e3000128.

Bateman, J.R., Lee, A.M., and Wu, C.T. (2006). Site-specific transformation of Drosophila via φC31 integrase-mediated cassette exchange. Genetics 173, 769–777.

Beling, I. (1929). Über das zeitgedächtnis der bienen. Z. Vgl. Physiol. 9, 259–338.

Bellen, H.J., Levis, R.W., He, Y., Carlson, J.W., Evans-Holm, M., Bae, E., Kim, J., Metaxakis, A., Savakis, C., Schulze, K.L., et al. (2011). The Drosophila gene disruption project: Progress using transposons with distinctive site specificities. Genetics 188, 731–743.

Benton, R., Vannice, K.S., Gomez-Diaz, C., and Vosshall, L.B. (2009). Variant Ionotropic Glutamate Receptors as Chemosensory Receptors in Drosophila. Cell 136, 149–162.

Binkley, S., Mosher, K., and Reilly, K.B. (1983). Circadian rhythms in house sparrows: Lighting ad lib. Physiol. Behav. 31, 829–837.

Biswas, S., and Akey, J.M. (2006). Genomic insights into positive selection. Trends Genet. 22, 437–446.

Bittman, E.L. (2020). Entrainment Is NOT Synchronization: An Important Distinction and Its Implications. J. Biol. Rhythms 36, 196-199.

Bloch, G., Barnes, B.M., Gerkema, M.P., and Helm, B. (2013). Animal activity around the clock with no overt circadian rhythms: Patterns, mechanisms and adaptive value. Proc. R. Soc. B Biol. Sci. 280, 20130019.

Blume, C., Garbazza, C., and Spitschan, M. (2019). Effects of light on human circadian rhythms, sleep and mood. Somnologie 23, 147–156.

Boitard, S., Schlötterer, C., Nolte, V., Pandey, R.V., and Futschik, A. (2012). Detecting selective sweeps from pooled next-generation sequencing samples. Mol. Biol. Evol. 29, 2177–2186.

Boitard, S., Kofler, R., Françoise, P., Robelin, D., Schlötterer, C., and Futschik, A. (2013). Pool-hmm: A Python program for estimating the allele frequency spectrum and detecting selective sweeps from next generation sequencing of pooled samples. Mol. Ecol. Resour. 13, 337–340.

Bonduriansky, R., and Day, T. (2009). Nongenetic Inheritance and Its Evolutionary Implications. Annu. Rev. Ecol. Evol. Syst. 40, 103–125.

Bordyugov, G., Abraham, U., Granada, A.E., Rose, P., Imkeller, K., Kramer, A., and Herzel, H. (2015). Tuning the phase of circadian entrainment. J. R. Soc. Interface 12, 20150282.

Boyd-Gibbins, N., Tardieu, C.H., Blunskyte, M., Kirkwood, N., Somers, J., and Albert, J.T. (2021). Turnover and activity-dependent transcriptional control of NompC in the Drosophila ear. IScience 24, 102486.

Brady, J. (1987). Crcadian rhythms - endogenous or exogenous? J. Comp. Physiol. A 161, 711–714.

Brown, F.A., Frank A. (1970). HYPOTHESIS OF ENVIRONMENTAL TIMING OF THE CLOCK. In The Biological Clock, (Academic Press), pp. 13–59.

Bullock, B. (2019). An interdisciplinary perspective on the association between chronotype and well-being. Yale J. Biol. Med. 92, 359–364.

Bunning, E. (1930). Uber die tagesperiodischen Bewegungen der Primarblatter von Phaseolus multiflorus. II. Die Bewegungen beim Thermokonstanz. Ber Deut Bot Ges 48, 227–252.

Burke, M.K., Dunham, J.P., Shahrestani, P., Thornton, K.R., Rose, M.R., and Long, A.D. (2010). Genome-wide analysis of a long-term evolution experiment with Drosophila. Nature 467, 587–590.

Burke, M.K., King, E.G., Shahrestani, P., Rose, M.R., and Long, A.D. (2014). Genome-wide association study of extreme longevity in Drosophila melanogaster. Genome Biol. Evol. 6, 1–11.

Burke, M.K., Barter, T.T., Cabral, L.G., Kezos, J.N., Phillips, M.A., Rutledge, G.A., Phung, K.H., Chen, R.H., Nguyen, H.D., Mueller, L.D., et al. (2016). Rapid divergence and convergence of life-history in experimentally evolved Drosophila melanogaster. Evolution 70, 2085–2098.

Van Buskirk, C., and Schüpbach, T. (2002). half pint regulates alternative splice site selection in Drosophila. Dev. Cell 2, 343–353.

Caldwell, J.C., Fineberg, S.K., and Eberl, D.F. (2007). Reduced ocelli encodes the leucine rich repeat protein pray for elves in Drosophila melanogaster. Fly (Austin). 1, 146–152.

Carpen, J.D., Archer, S.N., Skene, D.J., Smits, M., and Von Schantz, M. (2005a). A single-nucleotide polymorphism in the 5′-untranslated region of the hPER2 gene is associated with diurnal preference. J. Sleep Res. 14, 293–297.

Carpen, J.D., Von Schantz, M., Smits, M., Skene, D.J., and Archer, S.N. (2006a). A silent polymorphism in the PER1 gene associates with extreme diurnal preference in humans. J. Hum. Genet. 51, 1122–1125.

Carrillo, R.A., Özkan, E., Menon, K.P., Nagarkar-Jaiswal, S., Lee, P.T., Jeon, M., Birnbaum, M.E., Bellen, H.J., Garcia, K.C., and Zinn, K. (2015). Control of Synaptic Connectivity by a Network of Drosophila IgSF Cell Surface Proteins. Cell 163, 1770–1782.

Carskadon, M.A., Acebo, C., and Jenni, O.G. (2004). Regulation of adolescent sleep: Implications for behavior. In Annals of the New York Academy of Sciences, (John Wiley & Sons, Ltd), pp. 276–291.

Cassidy, J.J., Jha, A.R., Posadas, D.M., Giri, R., Venken, K.J.T., Ji, J., Jiang, H., Bellen, H.J., White, K.P., and Carthew, R.W. (2013). MiR-9a minimizes the phenotypic impact of genomic diversity by buffering a transcription factor. Cell 155, 1556-1567.

Cazzamali, G., Hauser, F., Kobberup, S., Williamson, M., and Grimmelikhuijzen, C.J.P. (2003). Molecular identification of a Drosophila G protein-coupled receptor specific for crustacean cardioactive peptide. Biochem. Biophys. Res. Commun. 303, 146–152.

Chandrashekaran, M.K. (1967). Studies on phase-shifts in endogenous rhythms - II. The dual effect of light on the entrainment of the eclosion rhythm in Drosophila pseudoobscura. Z. Vgl. Physiol. 56, 163–170.

Chandrashekaran, M.K., and Loher, W. (1969). The effect of light intensity on the circadian rhythms of eclosion in Drosophila pseudoobscura. Z. Vgl. Physiol. 62, 337–347.

Chen, J. V., Kao, L.R., Jana, S.C., Sivan-Loukianova, E., Mendonça, S., Cabrera, O.A., Singh, P., Cabernard, C., Eberl, D.F., Bettencourt-Dias, M., et al. (2015). Rootletin organizes the ciliary rootlet to achieve neuron sensory function in Drosophila. J. Cell Biol. 211, 435–453.

Chowdhury, M., Li, C.F., He, Z., Lu, Y., Liu, X.S., Wang, Y.F., Tony Ip, Y., Strand, M.R., and Yu, X.Q. (2019). Toll family members bind multiple Spätzle proteins and activate antimicrobial peptide gene expression in Drosophila. J. Biol. Chem. 294, 10172–10181.

Chung, B.Y., Kilman, V.L., Keath, J.R., Pitman, J.L., and Allada, R. (2009). The GABAA Receptor RDL Acts in Peptidergic PDF Neurons to Promote Sleep in Drosophila. Curr. Biol. 19, 386–390.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80-92.

Cingolani, P., Cunningham, F., McLaren, W., and Wang, K. (2018). Variant annotations in VCF format. January.

Clayton, D.L., and Paietta, J. V. (1972). Selection for circadian eclosion time in Drosophila melanogaster. Science (80-.). 178, 994–995.

Cormack, R.M., Hartl, D.L., and Clark, A.G. (1990). Principles of Population Genetics. Biometrics 46, 546.

Crocker, A., and Sehgal, A. (2008). Octopamine regulates sleep in Drosophila through protein kinase A-dependent mechanisms. J. Neurosci. 28, 9377–9385.

Croze, M., Wollstein, A., Božičević, V., Živković, D., Stephan, W., and Hutter, S. (2017). A genome-wide scan for genes under balancing selection in Drosophila melanogaster. BMC Evol. Biol. 17, 1–12.

Cutler, D.J., and Jensen, J.D. (2010). To Pool, or Not to Pool? Genetics 186, 41–43.

Daan, S. (2000). Colin Pittendrigh, Jurgen Aschoff, and the natural entrainment of circadian systems. J. Biol. Rhythms 15, 195–207.

Daan, S. (2010). A history of chronobiological concepts. In Protein Reviews, (Springer, New York, NY), pp. 1–35.

Daan, S., and Aschoff, J. (2001). The Entrainment of Circadian Systems. (Springer, Boston, MA), pp. 7–43.

Daan, S., and Pittendrigh, C.S. (1976). A functional analysis of circadian pacemakers in nocturnal rodents III. Heavy water and constant light: Homeostasis of frequency? J. Comp. Physiol. A 106, 267–290.

Damulewicz, M., Świątek, M., Łoboda, A., Dulak, J., Bilska, B., Przewłocki, R., and Pyza, E. (2019). Daily regulation of phototransduction, circadian clock, DNA repair, and immune gene expression by heme oxygenase in the retina of drosophila. Genes (Basel) 10, 6.

Davis, F.C., and Mannion, J. (1988). Entrainment of hamster pup circadian rhythms by prenatal melatonin injections to the mother. Am. J. Physiol. Integr. Comp. Physiol. 255, R439--R448.

De Mairan, J.J.. (1729). Observation Botanique.

Dodd, M.S., Papineau, D., Grenne, T., Slack, J.F., Rittner, M., Pirajno, F., O'Neil, J., and Little, C.T.S. (2017). Evidence for early life in Earth's oldest hydrothermal vent precipitates. Nature 543, 60–64.

Dominoni, D.M., Helm, B., Lehmann, M., Dowse, H.B., and Partecke, J. (2013). Clocks for the city: Circadian differences between forest and city songbirds. Proc. R. Soc. B Biol. Sci. 280, 20130593.

Doncheva, N.T., Morris, J.H., Gorodkin, J., and Jensen, L.J. (2019). Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. J. Proteome Res. 18, 623–632.

Dowle, M., and Srinivasan, A. (2020). data.table: Extension of `data.frame`.

Drake, C.L., Belcher, R., Howard, R., Roth, T., Levin, A.M., and Gumenyuk, V. (2015). Length polymorphism in the Period 3 gene is associated with sleepiness and maladaptive circadian phase in night-shift workers. J. Sleep Res. 24, 254–261.

Dreyer, A.P., Martin, M.M., Fulgham, C. V., Jabr, D.A., Bai, L., Beshel, J., and Cavanaugh, D.J. (2019). A circadian output center controlling feeding: Fasting rhythms in Drosophila. PLoS Genet. 15, e1008478.

du Monceau M. (1758). La physique des arbres : où il est traité de l'anatomie des plantes et de l'économie végétale: pour servir d'introduction au traité complet des bois & des forests: avec une dissertation sur l'utilité des méthodes de botanique ; & une explication d (A Paris :H.L. Guerin & L.F. Delatour,).

Duffy, J.F., and Wright, K.P. (2005). Entrainment of the human circadian system by light. J. Biol. Rhythms 20, 326–338.

Duffy, J.F., Rimmer, D.W., and Czeisler, C.A. (2001). Association of intrinsic circadian period with morningness-eveningness, usual wake time, and circadian phase. Behav. Neurosci. 115, 895–899.

Dunster, G.P., de la Iglesia, L., Ben-Hamo, M., Nave, C., Fleischer, J.G., Panda, S., and de la Iglesia, H.O. (2018). Sleepmore in Seattle: Later school start times are associated with more sleep and better performance in high school students. Sci. Adv. 4, 6200–6212.

Duvall, L.B., and Taghert, P.H. (2012). The circadian neuropeptide pdf signals preferentially through a specific adenylate cyclase isoform ac3 in m pacemakers of drosophila. PLoS Biol. 10, e1001337.

Ebisawa, T., Uchiyama, M., Kajimura, N., Mishima, K., Kamei, Y., Katoh, M., Watanabe, T., Sekimoto, M., Shibui, K., Kim, K., et al. (2001). Association of structural polymorphisms in the human period3 gene with delayed sleep phase syndrome. EMBO Rep. 2, 342–346.

Eelderink-Chen, Z., Olmedo, M., Bosman, J., and Merrow, M. (2015). Using Circadian Entrainment to Find Cryptic Clocks. Methods Enzymol. 551, 73–93.

Eggleston, W.B., Schlitz, D.M.J., and Engels, W.R. (1988). P-M hybrid dysgenesis does not mobilize other transposable element families in D. melanogaster. Nature 331, 368–370.

Ehret, C.F. (1974). The sense of time: evidence for its molecular basis in the eukaryotic gene-action system. Adv. Biol. Med. Phys. 15, 47–77.

Emery, I.F., Noveral, J.M., Jamison, C.F., and Siwicki, K.K. (1997). Rhythms of Drosophila period gene expression in culture. Proc. Natl. Acad. Sci. U. S. A. 94, 4092–4096.

Emery, N., Markosian, N., and Sullivan, M. (2020). Time. In The {Stanford} Encyclopedia of Philosophy, E.N. Zalta, ed. (Metaphysics Research Lab, Stanford University), p.

Engelmann, W. (1969). Phase shifting of eclosion in Drosophila pseudoobscura as a function of the energy of the light pulse. Z. Vgl. Physiol. 64, 111–117.

Evantal, N., Anduaga, A.M., Bartok, O., Patop, I.L., Weiss, R., Evanta, N., Patop, I.L., Bartok, O., Weiss, R., and Kadener, S. (2018). Thermosensitive alternative splicing senses and mediates 2 temperature adaptation in drosophila. Elife 8, e44642.

Fabian, D.K., Kapun, M., Nolte, V., Kofler, R., Schmidt, P.S., Schlötterer, C., and Flatt, T. (2012). Genome-wide patterns of latitudinal differentiation among populations of Drosophila melanogaster from North America. Mol. Ecol. 21, 4748–4769.

Foltenyi, K., Greenspan, R.J., and Newport, J.W. (2007). Activation of EGFR and ERK by rhomboid signaling regulates the consolidation and maintenance of sleep in Drosophila. Nat. Neurosci. 10, 1160–1167.

Fraekkel, G. (1935). Observations and Experiments on the Blow-fly (Calliphora erythrocephala) during the First Day after Emergence. Proc. Zool. Soc. London 105, 893–904.

Frías-Lasserre, D., S, A.L., and Villagra, C.A. (2019). Differences in larval emergence chronotypes for sympatric Rhagoletis brncici Frías and Rhagoletis conversa (Bréthes) (Diptera, Tephritidae). Rev. Bras. Entomol. 63, 195–198.

Frisch, K. v (1950). Die sonne als kompaßim leben der bienen. Experientia 6, 210–221.

Fry, F.E.J. (1947). Effects of the environment on animal activity. Publ. Ontario Fish. Res. Lab. 55, 1–62.

Futschik, A., and Schlötterer, C. (2010). The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. Genetics 186, 207–218.

Gamble, F.W., and Keeble, F.W. (1900). Memoirs: hippolyte varians: a study in colour-change. J. Cell Sci. 2, 589–698.

Gao, Q., Sheng, J., Qin, S., and Zhang, L. (2019). Chronotypes and affective disorders: A clock for mood? Brain Sci. Adv. 5, 145–160.

García-Allegue, R., Lax, P., Madariaga, A.M., and Madrid, J.A. (1999). Locomotor and feeding activity rhythms in a light-entrained diurnal rodent, Octodon degus. Am. J. Physiol. - Regul. Integr. Comp. Physiol. 277, R523-R531.

Garland, T., and Rose, M.R. (2009). Experimental evolution: Concepts, methods, and applications of selection experiments (University of California Press).

Gel, B. (2020). karyoploteR: Plot customizable linear genomes displaying arbitrary data.

Gel, B., and Serra, E. (2017). karyoploteR : an R / Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 33, 3088–3090.

Ghosh, A., Sharma, P., Dansana, S., and Sheeba, V. (2021). Evidence for Co-Evolution of Masking With Circadian Phase in Drosophila Melanogaster. J. Biol. Rhythms 36, 254–270.

Ghosh, A., and Sheeba, V. (2022). VANESSA – Shiny apps for accelerated time-series analysis and visualization of Drosophila circadian rhythm and sleep data. J. Biol. Rhythms (in press)

Giannotti, F., Cortesi, F., Sebastiani, T., and Ottaviano, S. (2002). Circadian preference, sleep and daytime behaviour in adolescence. J. Sleep Res. 11, 191–199.

Godoy-Herrera, R., Bustamante, M., Campos, P., and Cancino, J.L. (1997). The development of larval behaviours in sympatric chilean populations of drosophila melanogaster and drosophila simulans. Behaviour 134, 105–125.

Goldin, A.P., Sigman, M., Braier, G., Golombek, D.A., and Leone, M.J. (2020). Interplay of chronotype and school timing predicts school performance. Nat. Hum. Behav. 4, 387–396.

Gorman, M.R., Harrison, E.M., and Evans, J.A. (2017). Circadian waveform and its significance for clock organization and plasticity. In Biological Timekeeping: Clocks, Rhythms and Behaviour, (Springer, New Delhi), pp. 59–79.

Gottlieb, D.J., O’Connor, G.T., Wilk, J.B., DJ, G., GT, O., and JB, W. (2007). Genome-wide association of sleep and circadian phenotypes. BMC med. Gen., 8, 1-8.

Grafarend, E.W. (2007). Linear and nonlinear models: fixed effects, random effects, and mixed models. Choice Rev. Online 44, 44-2750-44–2750.

Granada, A.E., Bordyugov, G., Kramer, A., and Herzel, H. (2013). Human Chronotypes from a Theoretical Perspective. PLoS One 8, e59464.

Graves, J.L.L., Hertweck, K.L.L., Phillips, M.A.A., Han, M.V. V., Cabral, L.G.G., Barter, T.T.T., Greer, L.F.F., Burke, M.K.K., Mueller, L.D.D., Rose, M.R.R., et al. (2017). Genomics of parallel experimental evolution in drosophila. Mol. Biol. Evol. 34, 831–842.

Gu, G., Yang, J., Mitchell, K.A., and O’Tousa, J.E. (2004). Drosophila NinaB and NinaD Act Outside of Retina to Produce Rhodopsin Chromophore. J. Biol. Chem. 279, 18608–18613.

Halberg, F. (1959). Physiologic 24-hour periodicity; general and procedural considerations with reference to the adrenal cycle. Z. Vitamin-, Horm. Fermentforsch 10, 225.

Hall, H., Ma, J., Shekhar, S., Leon-Salas, W.D., and Weake, V.M. (2018). Blue light induces a neuroprotective gene expression program in Drosophila photoreceptors. BMC Neurosci. 19, 1-18.

Hamblen-Coyle, M.J., Wheeler, D.A., Rutila, J.E., Rosbash, M., and Hall, J.C. (1992). Behavior of period-altered circadian rhythm mutants of Drosophila in light: Dark cycles (Diptera: Drosophilidae). J. Insect Behav. 5, 417–446.

Harbison, S.T., Serrano Negron, Y.L., Hansen, N.F., and Lobell, A.S. (2017). Selection for long and short sleep duration in Drosophila melanogaster reveals the complex genetic network underlying natural variation in sleep. PLoS Genet. 13, e1007098.

Hardin, P.E. (2011). Molecular genetic analysis of circadian timekeeping in Drosophila (Elsevier Inc.).

Harker, J.E. (1965). The Effect of Photoperiod on the Developmental Rate of Drosophila Pupae. J. Exp. Biol. 43, 411–423.

Hartl, D.L., Clark, A.G., and Clark, A.G. (1997). Principles of population genetics (Sinauer associates Sunderland, MA).

Hastings, J.W., and Sweeney, B.M. (1957). ON THE MECHANISM OF TEMPERATURE INDEPENDENCE IN A BIOLOGICAL CLOCK. Proc. Natl. Acad. Sci. 43, 804–811.

Helm, B., and Visser, M.E. (2010). Heritable circadian period length in a wild bird population. Proc. R. Soc. B Biol. Sci. 277, 3335–3342.

Hendricks, J.C., Finn, S.M., Panckeri, K.A., Chavkin, J., Williams, J.A., Sehgal, A., and Pack, A.I. (2000). Rest in Drosophila Is a Sleep-like State. Neuron 25, 129–138.

Hermann-Luibl, C., Yoshii, T., Senthilan, P.R., Dircksen, H., and Helfrich-Förster, C. (2014). The ion transport peptide is a new functional clock neuropeptide in the fruit fly drosophila melanogaster. J. Neurosci. 34, 9522–9536.

Hill, A., Zheng, X., Li, X., McKinney, R., Dickman, D., and Ben-Shahar, Y. (2017). The Drosophila postsynaptic DEG/ENaC channel ppk29 contributes to excitatory neurotransmission. J. Neurosci. 37, 3171–3180.

Hirth, F., Kammermeier, L., Frei, E., Walldorf, U., Noll, M., and Reichert, H. (2003). An urbilaterian origin of the tripartite brain: Developmental genetic insights from Drosophila. Development 130, 2365–2373.

Horn, M., Mitesser, O., Hovestadt, T., Yoshii, T., Rieger, D., and Helfrich-Förster, C. (2019). The Circadian Clock Improves Fitness in the Fruit Fly, Drosophila melanogaster. Front. Physiol. 10, 1374.

Horne, J.A., and Ostberg, O. (1976). A self assessment questionnaire to determine Morningness Eveningness in human circadian rhythms. Int. J. Chronobiol. 4, 97–110.

Hsu, P.K., Ptáček, L.J., and Fu, Y.H. (2015). Genetics of human sleep behavioral phenotypes. In Methods in Enzymology, (Academic Press), pp. 309–324.

Hu, Y., Shmygelska, A., Tran, D., Eriksson, N., Tung, J.Y., and Hinds, D.A. (2016). GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. Nat. Commun. 7, 1-9.

Hughes, M.E., Hogenesch, J.B., and Kornacker, K. (2010). JTK-CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. J. Biol. Rhythms 25, 372–380.

Hur, Y.M., Bouchard, T.J., and Lykken, D.T. (1998). Genetic and environmental influence on morningness-eveningness2. Pers. Individ. Dif. 25, 917–925.

Hutter, S., Li, H., Beisswanger, S., De Lorenzo, D., and Stephan, W. (2007). Distinctly different sex ratios in African and European populations of Drosophila melanogaster inferred from chromosomewide single nucleotide polymorphism data. Genetics 177, 469–480.

Ikeda, K., Daimon, T., Sezutsu, H., Udaka, H., and Numata, H. (2019). Involvement of the clock gene period in the circadian rhythm of the silkmoth Bombyx mori. J. Biol. Rhythms 34, 283–292.

Inc., P.T. (2015). Collaborative data science.

Infante-Rivard, C., Dumont, M., and Montplaisir, J. (1989). Sleep disorder symptoms among nurses and nursing aides. Int. Arch. Occup. Environ. Health 61, 353–358.

Iwase, T., Kajimura, N., Uchiyama, M., Ebisawa, T., Yoshimura, K., Kamei, Y., Shibui, K., Kim, K., Kudo, Y., Katoh, M., et al. (2002). Mutation screening of the human Clock gene in circadian rhythm sleep disorders. Psychiatry Res. 109, 121–128.

Izutsu, M., Zhou, J., Sugiyama, Y., Nishimura, O., Aizu, T., Toyoda, A., Fujiyama, A., Agata, K., and Fuse, N. (2012). Genome features of “dark-fly”, a drosophila line reared long-term in a dark environment. PLoS One 7, e33288.

Jackson, F.R., Genova, G.K., Huang, Y., Kleyner, Y., Suh, J., Roberts, M.A., Sundram, V., and Akten, B. (2005). Genetic and biochemical strategies for identifying Drosophila genes that function in circadian control. Methods Enzymol. 393, 663–682.

Johnson, M.S. (1939). Effect of continuous light on periodic spontaneous activity of white-footed mice (Peromyscus). J. Exp. Zool. 82, 315–328.

Johnson, S.A., and Milner, M.J. (1987). The final stages of wing development in Drosophila melanogaster. Tissue Cell 19, 505–513.

Johnson, C.H., Elliott, J.A., and Foster, R. (2003). Entrainment of circadian programs. Chronobiol. Int. 20, 741–774.

Johnson, E.C., Kazgan, N., Bretz, C.A., Forsberg, L.J., Hector, C.E., Worthen, R.J., Onyenwoke, R., and Brenman, J.E. (2010). Altered metabolism and persistent starvation behaviors caused by reduced AMPK function in Drosophila. PLoS One 5, 1–11.

Jones, S.E., Tyrrell, J., Wood, A.R., Beaumont, R.N., Ruth, K.S., Tuke, M.A., Yaghootkar, H., Hu, Y., Teder-Laving, M., Hayward, C., et al. (2016). Genome-Wide Association Analyses in 128,266 Individuals Identifies New Morningness and Sleep Duration Loci. PLoS Genet. 12, e1006125.

Jones, S.E., Lane, J.M., Wood, A.R., van Hees, V.T., Tyrrell, J., Beaumont, R.N., Jeffries, A.R., Dashti, H.S., Hillsdon, M., Ruth, K.S., et al. (2019). Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. Nat. Commun. 10, 1-11.

Kahali, B., Bose, A., Karandikar, U., Bishop, C.P., and Bidwai, A.P. (2009). On the mechanism underlying the divergent retinal and bristle defects of M8* (E(spl)D) in drosophila. Genesis 47, 456–468.

Kalmbach, D.A., Schneider, L.D., Cheung, J., Bertrand, S.J., Kariharan, T., Pack, A.I., and Gehrman, P.R. (2017). Genetic basis of chronotype in humans: Insights from three landmark gwas. Sleep 40.

Kalmus, H. (1940). Diurnal rhythms in the axolotl larva and in drosophila. Nature 145, 72–73.

Kantermann, T., Juda, M., Merrow, M., and Roenneberg, T. (2007). The Human Circadian Clock's Seasonal Adjustment Is Disrupted by Daylight Saving Time. Curr. Biol. 17, 1996–2000.

Kapun, M., Barron, M.G., Staubach, F., Obbard, D.J., Axel, R., Vieira, J., Goubert, C., Rota-Stabelli, O., Kankare, M., Bogaerts-Marquez, M., 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.

Katzenberg, D., Young, T., Finn, L., Lin, L., King, D.P., Takahashi, J.S., and Mignot, E. (1998). A CLOCK polymorphism associated with human diurnal preference. Sleep 21, 569–576.

Kawecki, T.J., Lenski, R.E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M.C. (2012). Experimental evolution. Trends Ecol. Evol. 27, 547-560.

Kempinger, L., Dittmann, R., Rieger, D., and Helfrich-Förster, C. (2009). The nocturnal activity of fruit flies exposed to artificial moonlight is partly caused by direct light effects on the activity level that bypass the endogenous clock. Chronobiol. Int. 26, 151–166.

Kerkhof, G.A., and Van Dongen, H.P.A. (1996). Morning-type and evening-type individuals differ in the phase position of their endogenous circadian oscillator. Neurosci. Lett. 218, 153–156.

Kim, D.H., Han, M.R., Lee, G., Lee, S.S., Kim, Y.J., and Adams, M.E. (2015). Rescheduling Behavioral Subunits of a Fixed Action Pattern by Genetic Manipulation of Peptidergic Signaling. PLoS Genet. 11, e1005513.

Kivelä, L., Papadopoulos, M.R., and Antypa, N. (2018). Chronotype and Psychiatric Disorders. Curr. Sleep Med. Reports 4, 94–103.

Klei, L., Reitz, P., Miller, M., Wood, J., Maendel, S., Gross, D., Waldner, T., Eaton, J., Monk, T.H., and Nimgaonkar, V.L. (2005). Heritability of morningness-eveningness and self-report sleep measures in a family-based sample of 521 hutterites. Chronobiol. Int. 22, 1041–1054.

Kleinhoonte, A. (1929). Über die durch das Licht regulierten autonomen Bewegungen der Canavalia-Blätter (Société hollandaise des sciences).

Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., and Ding, L. (2009). VarScan: Variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25, 2283-2285.

Kofler, R., Orozco-terWengel, P., de Maio, N., Pandey, R.V., Nolte, V., Futschik, A., Kosiol, C., and Schlötterer, C. (2011a). Popoolation: A toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS One 6, e15925.

Kofler, R., Pandey, R.V., and Schlötterer, C. (2011b). PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics 27, 3435-3436.

Koh, K., Joiner, W.J., Wu, M.N., Yue, Z., Smith, C.J., and Sehgal, A. (2008). Identification of SLEEPLESS, a sleep-promoting factor. Science 321, 372–376.

Koren, D., Dumin, M., and Gozal, D. (2016). Role of sleep quality in the metabolic syndrome. Diabetes, Metab. Syndr. Obes. Targets Ther. 9, 281–310.

Koskenvuo, M., Hublin, C., Partinen, M., Heikkilä, K., and Kaprio, J. (2007). Heritability of diurnal type: A nationwide study of 8753 adult twin pairs. J. Sleep Res. 16, 156–162.

Kramer, G. (1950). Weitere Analyse der Faktoren, welche die Zugaktivität des gekäfigten Vogels orientieren. Naturwissenschaften 37, 377–378.

Kreher, S.A., Kwon, J.Y., and Carlson, J.R. (2005). The molecular basis of odor coding in the Drosophila larva. Neuron 46, 445–456.

Krüger, E., Mena, W., Lahr, E.C., Johnson, E.C., and Ewer, J. (2015). Genetic analysis of Eclosion hormone action during Drosophila larval ecdysis. Dev. 142, 4279–4287.

Kumar, S., Vaze, K.M., Kumar, D., and Sharma, V.K. (2006). Selection for early and late adult emergence alters the rate of pre-adult development in Drosophila melanogaster. BMC Dev. Biol. 6, 1-14.

Kumar, S., Kumar, D., Paranjpe, D. a, R, A.C., and Sharma, V.K. (2007). Selection on the timing of adult emergence results in altered circadian clocks in fruit flies Drosophila melanogaster. J. Exp. Biol. 210, 906–918.

Kunorozva, L., Stephenson, K.J., Rae, D.E., and Roden, L.C. (2012). Chronotype and PERIOD3 variable number tandem repeat polymorphism in individual sports athletes. Chronobiol. Int. 29, 1004–1010.

Lane, J.M., Vlasac, I., Anderson, S.G., Kyle, S.D., Dixon, W.G., Bechtold, D.A., Gill, S., Little, M.A., Luik, A., Loudon, A., et al. (2016). Genome-wide association analysis identifies novel loci for chronotype in 100,420 individuals from the UK Biobank. Nat. Commun. 7, 1–10.

Langmead, B., Salzberg, S.L., and Langmead (2013). Bowtie2. Nat. Methods 9, 357-359.

Lear, B.C., Merrill, C.E., Lin, J.M., Schroeder, A., Zhang, L., and Allada, R. (2005). A G Protein-coupled receptor, groom-of-PDF, is required for PDF neuron action in circadian behavior. Neuron 48, 221–227.

Lee, C.Y., Wendel, D.P., Reid, P., Lam, G., Thummel, C.S., and Baehrecke, E.H. (2000). E93 directs steroid-triggered programmed cell death in Drosophila. Mol. Cell 6, 433–443.

Lehmann, M., Spoelstra, K., Visser, M.E., and Helm, B. (2012). Effects of temperature on circadian clock and chronotype: An experimental study on a passerine bird. Chronobiol. Int. 29, 1062–1071.

Levine, J.D., Casey, C.I., Kalderon, D.D., and Jackson, F.R. (1994). Altered circadian pacemaker functions and cyclic AMP rhythms in the drosophila learning mutant dunce. Neuron 13, 967–974.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Li, Q., Li, Y., Wang, X., Qi, J., Jin, X., Tong, H., Zhou, Z., Zhang, Z.C., and Han, J. (2017). Fbxl4 Serves as a Clock Output Molecule that Regulates Sleep through Promotion of Rhythmic Degradation of the GABAA Receptor. Curr. Biol. 27, 3616-3625.e5.

Li, Y., Zhou, Z., Zhang, X., Tong, H., Li, P., Zhang, Z.C., Jia, Z., Xie, W., and Han, J. (2013). Drosophila neuroligin 4 regulates sleep through modulating GABA transmission. J. Neurosci. 33, 15545–15554.

Liang, H.L., Nien, C.Y., Liu, H.Y., Metzstein, M.M., Kirov, N., and Rushlow, C. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature 456, 400–403.

LiLin-Yin (2020). CMplot: Circle Manhattan Plot.

Lippai, M., Csikós, G., Maróy, P., Lukácsovich, T., Juhász, G., and Sass, M. (2008). SNF4Aγ, the Drosophila AMPK γ subunit is required for regulation of developmental and stress-induced autophagy. Autophagy 4, 476–486.

Lu, B., Liu, W., Guo, F., and Guo, A. (2008). Circadian modulation of light-induced locomotion responses in Drosophila melanogaster. Genes, Brain Behav. 7, 730–739.

Lu, W., Meng, Q.J., Tyler, N.J.C., Stokkan, K.A., and Loudon, A.S.I. (2010). A Circadian Clock Is Not Required in an Arctic Mammal. Curr. Biol. 20, 533–537.

Lyne, R., Smith, R., Rutherford, K., Wakeling, M., Varley, A., Guillier, F., Janssens, H., Ji, W., Mclaren, P., North, P., et al. (2007). FlyMine: An integrated database for Drosophila and Anopheles genomics. Genome Biol. 8, 1–16.

Majercak, J., Sidote, D., Hardin, P.E., and Edery, I. (1999). How a circadian clock adapts to seasonal decreases in temperature and day length. Neuron 24, 219–230.

Majercak, J., Chen, W.-F., and Edery, I. (2004). Splicing of the period Gene 3'-Terminal Intron Is Regulated by Light, Circadian Clock Factors, and Phospholipase C. Mol. Cell. Biol. 24, 3359–3372.

Manfredini, R., Fabbian, F., Cappadona, R., and Modesti, P.A. (2018). Daylight saving time, circadian rhythms, and cardiovascular health. Intern. Emerg. Med. 13, 641–646.

Mark, B., Bustos-González, L., Cascallares, G., Conejera, F., and Ewer, J. (2021). The circadian clock gates Drosophila adult emergence by controlling the timecourse of metamorphosis. Proc. Natl. Acad. Sci. U. S. A. 118.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal 17, 10-12.

Martin, J.S., Hébert, M., Ledoux, E., Gaudreault, M., and Laberge, L. (2012). Relationship of chronotype to sleep, Light exposure, and work-related fatigue in student workers. Chronobiol. Int. 29, 295–304.

Martinek, S., Inonog, S., Manoukian, A.S., and Young, M.W. (2001). A role for the segment polarity gene shaggy/GSK-3 in the Drosophila circadian clock. Cell 105, 769–779.

Marygold, S.J., Roote, J., Reuter, G., Lambertsson, A., Ashburner, M., Millburn, G.H., Harrison, P.M., Yu, Z., Kenmochi, N., Kaufman, T.C., et al. (2007). The ribosomal protein genes and Minute loci of Drosophila melanogaster. Genome Biol. 8, 1–26.

Mcmahon, D.P., and Hayward, A. (2016). Why grow up? A perspective on insect strategies to avoid metamorphosis. Ecol. Entomol. 41, 505–515.

McNabb, S.L., and Truman, J.W. (2008). Light and peptidergic eclosion hormone neurons stimulate a rapid eclosion response that masks circadian emergence in Drosophila. J. Exp. Biol. 211, 2263–2274.

McNabb, S.L., Baker, J.D., Agapite, J., Steller, H., Riddiford, L.M., and Truman, J.W. (1997). Disruption of a behavioral sequence by targeted death of peptidergic neurons in Drosophila. Neuron 19, 813–823.

Means, J.C., Venkatesan, A., Gerdes, B., Fan, J.-Y.Y., Bjes, E.S., and Price, J.L. (2015). Drosophila Spaghetti and Doubletime Link the Circadian Clock and Light to Caspases, Apoptosis and Tauopathy. PLoS Genet. 11, e1005171.

Mecacci, L., and Rocchetti, G. (1998). Morning and evening types: Stress-related personality aspects. Pers. Individ. Dif. 25, 537–542.

Meckien, R. (2014). When a day lasted only 4 hours.

Mishima, K., Tozawa, T., Satoh, K., Saitoh, H., and Mishima, Y. (2005). The 3111T/C polymorphism of hClock is associated with evening preference and delayed sleep timing in a Japanese population sample. Am. J. Med. Genet. - Neuropsychiatr. Genet. 133 B, 101–104.

Morioka, E., Matsumoto, A., and Ikeda, M. (2012). Neuronal influence on peripheral circadian oscillators in pupal Drosophila prothoracic glands. Nat. Commun. 3, 1–11.

Mrosovsky, N. (1994). In praise of masking: Behavioural responses of retinally degenerate mice to dim light. Chronobiol. Int. 11, 343–348.

Mrosovsky, N. (1999). Masking: History, definitions, and measurement. Chronobiol. Int. 16, 415–429.

Mueller, L.D., Phillips, M.A., Barter, T.T., Greenspan, Z.S., and Rose, M.R. (2018). Genome-Wide Mapping of Gene–Phenotype Relationships in Experimentally Evolved Populations. Mol. Biol. Evol. 35, 2085–2095.

Munro, M., Akkam, Y., and Curtin, K.D. (2010). Mutational analysis of Drosophila basigin function in the visual system. Gene 449, 50–58.

Myers, E.M., Yu, J., and Sehgal, A. (2003). Circadian control of eclosion: Interaction between a central and peripheral clock in Drosophila melanogaster. Curr. Biol. 13, 526–533.

Nagoshi, E., Sugino, K., Kula, E., Okazaki, E., Tachibana, T., Nelson, S., and Rosbash, M. (2010). Dissecting differential gene expression within the circadian neuronal circuit of Drosophila. Nat. Neurosci. 13, 60–68.

Nikhil, K., Vaze, K.M., and Sharma, V.K. (2015). Late emergence chronotypes of fruit flies Drosophila melanogaster exhibit higher accuracy of entrainment. Chronobiol. Int. 32, 1477–1485.

Nikhil, K., Ratna, K., and Sharma, V.K. (2016). Life-history traits of Drosophila melanogaster populations exhibiting early and late eclosion chronotypes. BMC Evol. Biol. 16, 46.

Nikhil, K.L., Goirik, G., Karatgi, R., and Sharma, V.K. (2014). Role of temperature in mediating morning and evening emergence chronotypes in fruit flies drosophila melanogaster. J. Biol. Rhythms 29, 427–441.

Nikhil, K.L., Abhilash, L., Sharma, V.K.K., Lakshman, A., and Sharma, V.K.K. (2016b). Molecular correlates of circadian clocks in fruit fly Drosophila melanogaster populations exhibiting early and late emergence chronotypes. J. Biol. Rhythms 31, 125–141.

Nikhil, K.L., Vaze, K.M., Karatgi, R., and Sharma, V.. (2016c). Circadian clock properties of fruit flies Drosophila melanogaster exhibiting early and late emergence chronotypes. Chronobiol. Int. 33, 22–38.

Ocampo-Garcés, A., Mena, W., Hernández, F., Cortés, N., and Palacios, A.G. (2006). Circadian chronotypes among wild-captured west Andean octodontids. Biol. Res. 39, 209–220.

Orozco-Terwengel, P., Kapun, M., Nolte, V., Kofler, R., Flatt, T., and Schlãtterer, C. (2012). Adaptation of Drosophila to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. Mol. Ecol. 21, 4931–4941.

Osland, T.M., Bjorvatn, B., Steen, V.M., and Pallesen, S. (2011). Association study of a variable-number tandem repeat polymorphism in the clock gene period3 and chronotype in norwegian university students. Chronobiol. Int. 28, 764–770.

Page, T.L. (1989). Masking in invertebrates. Chronobiol. Int. 6, 3-11.

Palaksha, and Shakunthala, V. (2014). Effect of different light regimes on eclosion rhythm of Drosophila agumbensis and Drosophila nagarholensis. Biol. Rhythm Res. 45, 219–227.

Panda, S., Hogenesch, J.B., and Kay, S.A. (2002). Circadian rhythms from flies to human. Nature 417, 329–335.

Park, J.H., Schroeder, A.J., Helfrich-Förster, C., Jackson, F.R., and Ewer, J. (2003). Targeted ablation of CCAP neuropeptide-containing neurons of Drosophila causes specific defects in execution and circadian timing of ecdysis behavior. Development 130, 2645–2656.

Park, J.W., Parisky, K., Celotto, A.M., Reenan, R.A., and Graveley, B.R. (2004). Identification of alternative splicing regulators by RNA interference in Drosophila. Proc. Natl. Acad. Sci. U. S. A. 101, 15974–15979.

Parks, A.L., Cook, K.R., Belvin, M., Dompe, N.A., Fawcett, R., Huppert, K., Tan, L.R., Winter, C.G., Bogart, K.P., Deal, J.E., et al. (2004). Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. Nat. Genet. 36, 288–292.

Parsons, M.J., Lester, K.J., Barclay, N.L., Archer, S.N., Nolan, P.M., Eley, T.C., and Gregory, A.M. (2014). Polymorphisms in the circadian expressed genes PER3 and ARNTL2 are associated with diurnal preference and GNβ3 with sleep measures. J. Sleep Res. 23, 595–604.

Pavlidis, T. (1967). A mathematical model for the light affected system in the drosophila eclosion rhythm. Bull. Math. Biophys. 29, 291–310.

Peabody, N.C., and White, B.H. (2013). Eclosion gates progression of the adult ecdysis sequence of Drosophila. J. Exp. Biol. 216, 4395–4402.

Pedrazzoli, M., Louzada, F.M., Pereira, D.S., Benedito-Silva, A.A., Lopez, A.R., Martynhak, B.J., Korczak, A.L., Koike, B.D.V., Barbosa, A.A., D’Almeida, V., et al. (2007). Clock polymorphisms and circadian rhythms phenotypes in a sample of the Brazilian population. Chronobiol. Int. 24, 1–8.

Pegoraro, M., Flavell, L., Menegazzi, P., Colombi, P., Dao, P., Helfrich-Förster, C., and Tauber, E. (2019). The genetic basis of diurnal preference in Drosophila melanogaster. 1–11.

Perea, C.S., Niño, C.L., López-León, S., Gutiérrez, R., Ojeda, D., Arboleda, H., Camargo, A., Adan, A., and Forero, D.A. (2014). Study of a Functional Polymorphism in the PER3 Gene and Diurnal Preference in a Colombian Sample. Open Neurol. J. 8, 7–10.

Pereira, D.S., Tufik, S., Louzada, F.M., Benedito-Silva, A.A., Lopez, A.R., Lemos, N.A., Korczak, A.L., D'Almeida, V., and Pedrazzoli, M. (2005). Association of the length polymorphism in the human Per3 gene with the delayed sleep-phase syndrome: Does latitude have an influence upon it? Sleep 28, 29–32.

Pfeiffenberger, C., and Allada, R. (2012). Cul3 and the BTB Adaptor Insomniac Are Key Regulators of Sleep Homeostasis and a Dopamine Arousal Pathway in Drosophila. PLoS Genet. 8, e1003003.

Phillips, M.A., Long, A.D., Greenspan, Z.S., Greer, L.F., Burke, M.K., Villeponteau, B., Matsagas, K.C., Rizza, C.L., Mueller, L.D., and Rose, M.R. (2016). Genome-wide analysis of long-term evolutionary domestication in Drosophila melanogaster. Sci. Rep. 6, 39281.

Phillips, M.A., Rutledge, G.A., Kezos, J.N., Greenspan, Z.S., Talbott, A., Matty, S., Arain, H., Mueller, L.D., Rose, M.R., and Shahrestani, P. (2018). Effects of evolutionary history on genome wide and phenotypic convergence in Drosophila populations. BMC Genomics 2018 191 19, 1–17.

Pittendrigh, C.S. (1954). On temperature independence in the clock system controlling emergence time in Drosophila. Proc. Natl. Acad. Sci. U. S. A. 40, 1018–1029.

Pittendrigh, C.S. (1966). The circadian oscillation in Drosophila pseudoobscura pupae: a model for the photoperiodic clock. Zeitschrift Fur Pflanzenphysiologie 54, 275–307.

Pittendrigh, C.S. (1967). Circadian systems. I. The driving oscillation and its assay in Drosophila pseudoobscura. Proc. Natl. Acad. Sci. U. S. A. 58, 1762–1767.

Pittendrigh, C.S. (1993). Temporal Organization: Reflections of a Darwinian Clock-Watcher. Annu. Rev. Physiol. 55, 17–54.

Pittendrigh, C., and Bruce, V. (1959). Photoperiodism and Related Phenomena in Plants and Animals. Dly. Rhythm. as Coupled Oscil. Syst. Their Relat. to Thermo- Photoperiod. 475–505.

Pittendrigh, C.S., and Daan, S. (1976). A functional analysis of circadian pacemakers in nocturnal rodents - IV. Entrainment: Pacemaker as clock. J. Comp. Physiol. A 106, 291–331.

Pittendrigh, C.S., and Minis, D.H. (1972). Circadian systems: longevity as a function of circadian resonance in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 69, 1537-1539.

Pittendrigh, C.S., and Takamura, T. (1987). Temperature dependence and evolutionary adjustment of critical night length in insect photoperiodism. Proc. Natl. Acad. Sci. 84, 7169-7173.

Pittendrigh, C.S., Bruce, V.V., and Kaus, P. (1958). On the significance of transients in daily rhythms. Proc. Natl. Acad. Sci. U. S. A. 44, 965–973.

Pogue-Geile, K.L., Lyons-Weiler, J., and Whitcomb, D.C. (2006). Molecular overlap of fly circadian rhythms and human pancreatic cancer. Cancer Lett. 243, 55–57.

Prabhakaran, P.M., and Sheeba, V. (2012). Sympatric Drosophilid species melanogaster and ananassae differ in temporal patterns of activity. J. Biol. Rhythms 27, 365–376.

Prabhakaran, P.M., De, J., and Sheeba, V. (2013). Natural conditions override differences in emergence rhythm among closely related drosophilids. PLoS One. 8, e83048.

Qin, B., Humberg, T.H., Kim, A., Kim, H.S., Short, J., Diao, F., White, B.H., Sprecher, S.G., and Yuan, Q. (2019). Muscarinic acetylcholine receptor signaling generates OFF selectivity in a simple visual circuit. Nat. Commun. 10, 1–16.

Qiu, J., and Hardin, P.E. (1996). Developmental state and the circadian clock interact to influence the timing of eclosion in Drosophila melanogaster. J. Biol. Rhythms 11, 75–86.

R Core Team (2020). R: A Language and Environment for Statistical Computing.

Randler, C., Faßl, C., and Kalb, N. (2017). From Lark to Owl: developmental changes in morningness-eveningness from new-borns to early adulthood. Sci. Rep. 7, 45874.

Ravikumar, B., Berger, Z., Vacher, C., O'Kane, C.J., and Rubinsztein, D.C. (2006). Rapamycin pre-treatment protects against apoptosis. Hum. Mol. Genet. 15, 1209–1216.

Ravikumar, B., Imarisio, S., Sarkar, S., O'Kane, C.J., and Rubinsztein, D.C. (2008). Rab5 modulates aggregation and toxicity of mutant huntingtin through macroautophagy in cell and fly models of Huntington disease. J. Cell Sci. 121, 1649–1660.

Redlin, U., and Mrosovsky, N. (1999a). Masking of locomotor activity in hamsters. J. Comp. Physiol. A. 184, 429–437.

Redlin, U., and Mrosovsky, N. (1999b). Masking by light in hamsters with SCN lesions. J. Comp. Physiol. A Sensory, Neural, Behav. Physiol. 184, 439–448.

Reed, L.K., Lee, K., Zhang, Z., Rashid, L., Poe, A., Hsieh, B., Deighton, N., Glassbrook, N., Bodmer, R., and Gibson, G. (2014). Systems genomics of metabolic phenotypes in wild-type Drosophila melanogaster. Genetics 197, 781-793.

Refinetti, R., Wassmer, T., Basu, P., Cherukalady, R., Pandey, V.K., Singaravel, M., Giannetto, C., and Piccione, G. (2016). Variability of behavioral chronotypes of 16 mammalian species under controlled conditions. Physiol. Behav. 161, 53–59.

Refinetti, R., Earle, G., and Kenagy, G.J. (2019). Exploring determinants of behavioral chronotype in a diurnal-rodent model of human physiology. Physiol. Behav. 199, 146–153.

Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., and Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 44.W1, W83-W89.

Remolina, S., Chang, P., and Leips, J. (2012). Genomic Basis of Aging and Life History Evolution in Drosophila Melanogaster. Evolution (N. Y). 66, 3390-3403.

Rensing, L. (1989). Is “masking” an appropriate term? Chronobiol. Int. 6, 297–300.

Richard, B., Pamela, C., and Lyman, J.L. (1997). The Earth and Its Peoples.

Richter, C.P. (1922). A behavioristic study of the activity of the rat. Comp. Psychol. Monogr.

Rieger, D., Stanewsky, R., and Helfrich-Förster, C. (2003). Cryptochrome, compound eyes, Hofbauer-Buchner eyelets, and ocelli play different roles in the entrainment and masking pathway of the locomotor activity rhythm in the fruit fly Drosophila melanogaster. J. Biol. Rhythms. 18, 377-391.

Robilliard, D.L., Archer, S.N., Arendt, J., Lockley, S.W., Hack, L.M., English, J., Leger, D., Smits, M.G., Williams, A., Skene, D.J., et al. (2002). The 3111 Clock gene polymorphism is not associated with sleep and circadian rhythmicity in phenotypically characterized human subjects. J. Sleep Res. 11, 305–312.

Roenneberg, T. (2012). Internal Time (Harvard University Press).

Roenneberg, T., and Merrow, M. (2016). The circadian clock and human health. Curr. Biol. 26, R432–R443.

Roenneberg, T., Daan, S., and Merrow, M. (2003). The art of entrainment. J. Biol. Rhythms 18, 183–194.

Roenneberg, T., Kuehnle, T., Pramstaller, P.P., Ricken, J., Havel, M., Guth, A., and Merrow, M. (2004). A marker for the end of adolescence. Curr. Biol. 14, R1038–R1039.

Roenneberg, T., Kumar, C.J., and Merrow, M. (2007). The human circadian clock entrains to sun time. Curr. Biol. 17, 1–3.

Roenneberg, T., Kuehnle, T., Juda, M., Kantermann, T., Allebrandt, K., Gordijn, M., and Merrow, M. (2007b). Epidemiology of the human circadian clock. Sleep Med. Rev. 11, 429–438.

Roenneberg, T., Allebrandt, K. V., Merrow, M., and Vetter, C. (2012). Social jetlag and obesity. Curr. Biol. 22, 939–943.

Roenneberg, T., Kantermann, T., Juda, M., Vetter, C., and Allebrandt, K. V. (2013). Light and the human circadian clock. Handb. Exp. Pharmacol. 217, 311–331.

Roenneberg, T., Pilz, L.K., Zerbini, G., and Winnebeck, E.C. (2019). Chronotype and social jetlag: A (self-) critical review. Biology (Basel). 8, 54.

Rose, M., Rutledge, G., Phung, K., Phillips, M., Greer, L., and Mueller, L. (2014). An Evolutionary and Genomic Approach to Challenges and Opportunities for Eliminating Aging. Curr. Aging Sci. 7, 54–59.

Rudolph, T., Yonezawa, M., Lein, S., Heidrich, K., Kubicek, S., Schäfer, C., Phalke, S., Walther, M., Schmidt, A., Jenuwein, T., et al. (2007). Heterochromatin Formation in Drosophila Is Initiated through Active Removal of H3K4 Methylation by the LSD1 Homolog SU(VAR)3-3. Mol. Cell 26, 103–115.

Ruf, F., Mitesser, O., Mungwa, S.T., Horn, M., Rieger, D., Hovestadt, T., and Wegener, C. (2019). Natural Zeitgebers cannot compensate for the loss of a functional circadian clock in timing of a vital behaviour in Drosophila. BioRxiv 0–3.

Ryder, E., Blows, F., Ashburner, M., Bautista-Llacer, R., Coulson, D., Drummond, J., Webster, J., Gubb, D., Gunton, N., Johnson, G., et al. (2004). The DrosDel collection: A set of P-element insertions for generating custom chromosomal aberrations in Drosophila melanogaster. Genetics 167, 797–813.

Ryder, E., Ashburner, M., Bautista-Llacer, R., Drummond, J., Webster, J., Johnson, G., Morley, T., Yuk, S.C., Blows, F., Coulson, D., et al. (2007). The DrosDel deletion collection: A Drosophila genomewide chromosomal deficiency resource. Genetics 177, 615–629.

Samis, H. V (1978). Aging and Biological Rhythms (Advances in Experimental Medicine and Biology) 108, 1-4.

Sarkar, S., Perlstein, E.O., Imarisio, S., Pineau, S., Cordenier, A., Maglathlin, R.L., Webster, J.A., Lewis, T.A., O'Kane, C.J., Schreiber, S.L., et al. (2007). Small molecules enhance autophagy and reduce toxicity in Huntington's disease models. Nat. Chem. Biol. 3, 331–338.

Sarkar, S., Krishna, G., Imarisio, S., Saiki, S., O'Kane, C.J., and Rubinsztein, D.C. (2008). A rational mechanism for combination treatment of Huntington's disease using lithium and rapamycin. Hum. Mol. Genet. 17, 170–178.

Saunders, D.S., Gillanders, S.W., and Lewis, R.D. (1994). Light-pulse phase response curves for the locomotor activity rhythm in Period mutants of Drosophila melanogaster. J. Insect Physiol. 40, 957–968.

Von Schantz, M., Taporoski, T.P., Horimoto, A.R.V.R., Duarte, N.E., Vallada, H., Krieger, J.E., Pedrazzoli, M., Negrão, A.B., and Pereira, A.C. (2015). Distribution and heritability of diurnal preference (chronotype) in a rural Brazilian family-based cohort, the Baependi study. Sci. Rep. 5, 1-6.

Schlichting, M., and Helfrich-Förster, C. (2015). Photic entrainment in drosophila assessed by locomotor activity recordings. In Methods in Enzymology, (Academic Press Inc.), pp. 105–123.

Schmal, C., Herzog, E.D., and Herzel, H. (2018). Measuring Relative Coupling Strength in Circadian Systems. J. Biol. Rhythms 33, 84–98.

Schwartz, M.D., and Smale, L. (2005). Individual differences in rhythms of behavioral sleep and its neural substrates in nile grass rats. J. Biol. Rhythms 20, 526–537.

Seebens, H., Blackburn, T.M., Dyer, E.E., Genovesi, P., Hulme, P.E., Jeschke, J.M., Pagad, S., Pyšek, P., Winter, M., Arianoutsou, M., et al. (2017). No saturation in the accumulation of alien species worldwide. Nat. Commun. 8, 1–9.

Sehgal, A., Price, J.L., Man, B., and Young, M.W. (1994). Loss of circadian behavioral rhythms and per RNA oscillations in the Drosophila mutant timeless. Science (80-.). 263, 1603–1606.

Selcho, M., Millán, C., Palacios-Muñoz, A., Ruf, F., Ubillo, L., Chen, J., Bergmann, G., Ito, C., Silva, V., Wegener, C., et al. (2017). Central and peripheral clocks are coupled by a neuropeptide pathway in Drosophila. Nat. Commun. 8, 1-13.

Senthilan, P.R., Piepenbrock, D., Ovezmyradov, G., Nadrowski, B., Bechstedt, S., Pauls, S., Winkler, M., Möbius, W., Howard, J., and Göpfert, M.C. (2012). Drosophila auditory organ genes and genetic hearing defects. Cell 150, 1042–1054.

Seugnet, L., Suzuki, Y., Merlin, G., Gottschalk, L., Duntley, S.P., and Shaw, P.J. (2011). Notch signaling modulates sleep homeostasis and learning after sleep deprivation in drosophila. Curr. Biol. 21, 835–840.

Shafer, O.T., and Keene, A.C. (2021). The Regulation of Drosophila Sleep. Curr. Biol. 31, R38–R49.

Shakhmantsir, I., Nayak, S., Grant, G.R., and Sehgal, A. (2018). Spliceosome factors target timeless (tim) mRNA to control clock protein accumulation and circadian behavior in Drosophila. Elife 7, 1–27.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.

Shaw, P.J., Cirelli, C., Greenspan, R.J., and Tononi, G. (2000). Correlates of Sleep and Waking in Drosophila melanogaster. Science (80-.). 287, 1834–1837.

Sheeba, V., Nihal, M., Mathew, S.J., Swamy, N.M., Chandrashekaran, M.K., Joshi, A., and Sharma, V.K. (2001). Does the difference in the timing of eclosion of the fruit fly Drosophila melanogaster reflect differences in the circadian organization? Chronobiol. Int. 18, 601-612.

Sheppard, A.D., Hirsch, H.V.B., and Possidente, B. (2015). Novel masking effects of light are revealed in Drosophila by skeleton photoperiods. Biol. Rhythm Res. 46, 275–285.

Sinha, C., Sinha, V.D.S., Zinken, J., and Sampaio, W. (2011). When time is not space: The social and linguistic construction of time intervals and temporal event relations in an Amazonian culture. Lang. Cogn. 3, 137–169.

Skopik, S.D., and Pittendrigh, C.S. (1967). Circadian systems, II. The oscillation in the individual Drosophila pupa; its independence of developmental stage. Proc. Natl. Acad. Sci. U. S. A. 58, 1862–1869.

Spitschan, M., Aguirre, G.K., Brainard, D.H., and Sweeney, A.M. (2016). Variation of outdoor illumination as a function of solar elevation and light pollution. Sci. Rep. 6, 1–14.

St Johnston, D. (2002). The art and design of genetic screens: Drosophila melanogaster. Nat. Rev. Genet. 3, 176–188.

Stuber, E.F., Dingemanse, N.J., Kempenaers, B., and Mueller, J.C. (2015). Sources of intraspecific variation in sleep behaviour of wild great tits. Anim. Behav. 106, 201–221.

Sulzman, F.M., Ellman, D., Fuller, C.A., Moore-Ede, M.C., and Wassmer, G. (1984). Neurospora circadian rhythms in space: A reexamination of the endogenous-exogenous question. Science 225, 232–234.

Swade, R.H. (1969). Circadian rhythms in fluctuating light cycles: Toward a new model of entrainment. J. Theor. Biol. 24, 227–239.

Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 49, D605–D612.

Taillard, J., Philip, P., Coste, O., Sagaspe, P., and Bioulac, B. (2003). The circadian and homeostatic modulation of sleep pressure during wakefulness differs between morning and evening chronotypes. J. Sleep Res. 12, 275–282.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Takahashi, K.H., Teramura, K., Muraoka, S., Okada, Y., and Miyatake, T. (2013). Genetic correlation between the pre-adult developmental period and locomotor activity rhythm in Drosophila melanogaster. Heredity (Edinb). 110, 312–320.

Tanaka, K., and Watari, Y. (2009). Is early morning adult eclosion in insects an adaptation to the increased moisture at dawn? Biol. Rhythm Res. 40, 293–298.

Tang, H.W., Wang, Y.B., Wang, S.L., Wu, M.H., Lin, S.Y., and Chen, G.C. (2011). Atg1-mediated myosin II activation regulates autophagosome formation during starvation-induced autophagy. EMBO J. 30, 636–651.

Thakurdas, P., Sharma, S., Vanlalhriatpuia, K., Sinam, B., Chib, M., Shivagaje, A., and Joshi, D. (2009). Light at night alters the parameters of the eclosion rhythm in a tropical fruit fly, Drosophila jambulina. Chronobiol. Int. 26, 1575–1586.

Thibault, S.T., Singer, M.A., Miyazaki, W.Y., Milash, B., Dompe, N.A., Singh, C.M., Buchholz, R., Demsky, M., Fawcett, R., Francis-Lang, H.L., et al. (2004). A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. Nat. Genet. 36, 283–287.

Thummel, C.S. (2001). Molecular mechanisms of developmental timing in C. elegans and Drosophila. Dev. Cell 1, 453–465.

Tobler, R., Franssen, S.U., Kofler, R., Orozco-terWengel, P., Nolte, V., Hermisson, J., and Schlötterer, C. (2014). Massive habitat-specific genomic response in D. melanogaster populations during experimental evolution in hot and cold environments. Mol. Biol. Evol. 31, 364–375.

Toda, H., Williams, J.A., Gulledge, M., and Sehgal, A. (2019). A sleep-inducing gene, nemuri, links sleep and immune function in Drosophila. Science (80-.). 363, 509–515.

Troein, C., Locke, J.C.W., Turner, M.S., and Millar, A.J. (2009). Weather and seasons together demand complex biological clocks. Curr. Biol. 19, 1961–1964.

Turcotte, D.L., Cisne, J.L., and Nordmann, J.C. (1977). On the evolution of the lunar orbit. Icarus 30, 254–266.

Turner, T.L., and Miller, P.M. (2012). Investigating natural variation in drosophila courtship song by the evolve and resequence approach. Genetics 191, 633-642.

Turner, T.L., Stewart, A.D., Fields, A.T., Rice, W.R., and Tarone, A.M. (2011). Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in Drosophila melanogaster. PLoS Genet. 7, e1001336.

Vaze, K.M., and Sharma, V.K. (2013). On the adaptive significance of circadian clocks for their owners. Chronobiol. Int. 30, 413–433.

Vaze, K., Nikhil, K.L., Abhilash, L., and Sharma, V.K. (2012a). Early-and late-emerging Drosophila melanogaster fruit flies differ in their sensitivity to light during morning and evening. Chronobiol. Int. 29, 674–682.

Vaze, K.M., Kannan, N.N., Abhilash, L., and Sharma, V.K. (2012b). Chronotype differences in Drosophila are enhanced by semi-natural conditions. Naturwissenschaften 99, 967–971.

Vaze, K.M., Nikhil, K.L., and Sharma, V.K. (2014). Circadian rhythms: 4. Why do living organisms have them? Resonance 19, 175-189.

Venken, K.J.T., Simpson, J.H., and Bellen, H.J. (2011). Genetic manipulation of genes and cells in the nervous system of the fruit fly. Neuron 72, 202–230.

Viola, A.U., Archer, S.N., James, L.M.M., Groeger, J.A., Lo, J.C.Y., Skene, D.J., von Schantz, M., and Dijk, D.J. (2007). PER3 Polymorphism Predicts Sleep Structure and Waking Performance. Curr. Biol. 17, 613–618.

Vivanco, P., Rol, M.Á., and Madrid, J.A. (2009). Two steady-entrainment phases and graded masking effects by light generate different circadian chronotypes in Octodon degus. Chronobiol. Int. 26, 219–241.

Vivanco, P., Rol, M.A., and Madrid, J.A. (2010). Pacemaker phase control versus masking by light: Setting the circadian chronotype in dual octodon degus. Chronobiol. Int. 27, 1365–1379.

Vlachos, C., Burny, C., Pelizzola, M., Borges, R., Futschik, A., Kofler, R., and Schlötterer, C. (2019). Benchmarking software tools for detecting and quantifying selection in evolve and resequencing studies. Genome Biol. 20, 1–11.

Volk, S., Dyroff, J., Georgi, K., and Pflug, B. (1994). Subjective sleepiness and physiological sleep tendency in healthy young morning and evening subjects. J. Sleep Res. 3, 138–143.

Vosshall, L.B., and Young, M.W. (1995). Circadian rhythms in drosophila can be driven by period expression in a restricted group of central brain cells. Neuron 15, 345–360.

Wagh, D.A., Rasse, T.M., Asan, E., Hofbauer, A., Schwenkert, I., Dürrbeck, H., Buchner, S., Dabauvalle, M.C., Schmidt, M., Qin, G., et al. (2006). Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in Drosophila. Neuron 49, 833–844.

Walch, O.J., Cochran, A., and Forger, D.B. (2016). A global quantification of “normal” sleep schedules using smartphone data. Sci. Adv. 2, e1501705.

Wang, M., Zhao, Y., and Zhang, B. (2022). SuperExactTest: Exact Test and Visualization of Multi-Set Intersections.

Watari, Y. (2002). Comparison of the circadian eclosion rhythm between nondiapause and diapause pupae in the onion fly, Delia antiqua. J. Insect Physiol. 48, 83–89.

Wheeler, D.A., Hamblen-Coyle, M.J., Dushay, M.S., and Hall, J.C. (1993). Behavior in light-dark cycles of Drosophila mutants that Are arrhythmic, blind, or both. J. Biol. Rhythms 8, 67–94.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).

Wickham, H., Chang, W., Henry, L., Pedersen, T.L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., and Dunnington, D. (2020). ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics.

Winfree, A.T. (1987). The timing of biological clocks (Macmillan).

Wittmann, M., Dinich, J., Merrow, M., and Roenneberg, T. (2006). Social jetlag: Misalignment of biological and social time. In Chronobiology International, (Chronobiol Int), pp. 497–509.

Wood, J.G., Schwer, B., Wickremesinghe, P.C., Hartnett, D.A., Burhenn, L., Garcia, M., Li, M., Verdin, E., and Helfand, S.L. (2018). Sirt4 is a mitochondrial regulator of metabolism and lifespan in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. 115, 1564–1569.

Wu, G., Anafi, R.C., Hughes, M.E., Kornacker, K., and Hogenesch, J.B. (2016). MetaCycle: An integrated R package to evaluate periodicity in large scale data. Bioinformatics 32, 3351–3353.

Yadav, P., Choudhury, D., Sadanandappa, M.K., and Sharma, V.K. (2015). Extent of mismatch between the period of circadian clocks and light/dark cycles determines time-to-emergence in fruit flies. Insect Sci. 22, 569–577.

Yang, Z., Emerson, M., Su, H.S., and Sehgal, A. (1998). Response of the timeless protein to light correlates with behavioral entrainment and suggests a nonvisual pathway for circadian photoreception. Neuron 21, 215–223.

Yukilevich, R., Turner, T.L., Aoki, F., Nuzhdin, S. V., and True, J.R. (2010). Patterns and processes of genome-wide divergence between North American and African Drosophila melanogaster. Genetics 186, 219–239.

Zavada, A., Gordijn, M.C.M., Beersma, D.G.M., Daan, S., and Roenneberg, T. (2005). Comparison of the Munich Chronotype Questionnaire with the Horne-Östberg's morningness-eveningness score. Chronobiol. Int. 22, 267–278.

Zhang, W., Yan, Z., Jan, L.Y., and Jan, Y.N. (2013). Sound response mediated by the TRP channels NOMPC, NANCHUNG, and INACTIVE in chordotonal organs of Drosophila larvae. Proc. Natl. Acad. Sci. U. S. A. 110, 13612–13617.

Zheng, X., and Sehgal, A. (2010). AKT and TOR Signaling Set the Pace of the Circadian Pacemaker. Curr. Biol. 20, 1203–1208.

Zheng, X., Yang, Z., Yue, Z., Alvarez, J.D., and Sehgal, A. (2007). FOXO and insulin signaling regulate sensitivity of the circadian clock to oxidative stress. Proc. Natl. Acad. Sci. U. S. A. 104, 15899–15904.

Zhou, D., Udpa, N., Gersten, M., Visk, D.W., Bashir, A., Xue, J., Frazer, K.A., Posakony, J.W., Subramaniam, S., Bafna, V., et al. (2011). Experimental selection of hypoxia-tolerant Drosophila melanogaster. Proc. Natl. Acad. Sci. 108, 2349-2354.

Zimmerman, W.F., Pittendrigh, C.S., and Pavlidis, T. (1968). Temperature compensation of the circadian oscillation in Drosophila pseudoobscura and its entrainment by temperature cycles. J. Insect Physiol. 14, 669–684.

Zitnan, D., and Adams, M.E. (2005). Neuroendocrine Regulation of Insect Ecdysis. In Comprehensive Molecular Insect Science, (Elsevier), pp. 1–60.

List of publications

***Asterisk: First author**

1. ***Arijit Ghosh**, and Vasu Sheeba. “VANESSA – Shiny apps for accelerated time-series analysis and visualization of Drosophila circadian rhythm and sleep data”. Journal of biological rhythms, accepted, in press.
2. ***Arijit Ghosh**, Pragya Sharma, Shephali Dansana, and Vasu Sheeba. “Evidence for coevolution of masking and circadian phase in Drosophila melanogaster”. Journal of biological rhythms 36, no. 3 (2021): 254-270.
3. Lakshman Abhilash, **Arijit Ghosh**, and Vasu Sheeba. “Selection for timing of eclosion results in co-evolution of temperature responsiveness in Drosophila melanogaster”. Journal of biological rhythms 34, no. 6 (2019): 596-609.

Google scholar: https://scholar.google.co.in/citations?user=Nuaw_FoAAAAJ&hl=en

ORCID id: https://orcid.org/0000-0002-7910-3170