Investigating Circadian Outputs of the *Caenorhabditis elegans* Transcriptome

Jack Munns

Doctor of Philosophy

University of York

Biology

July 2019

Abstract

Circadian rhythms can be observed across phyla and are interlinked with many aspects of health, disease and survival. A partially-conserved genetic basis for the generation of circadian rhythms has been well-characterised in common laboratory models: mouse (*Mus musculus*), fly (*Drosophila melanogaster*) and zebrafish (*Danio rerio*). In the nematode *Caenorhabditis elegans* however, the molecular basis of circadian rhythms remains poorly understood. This is despite reports of circadian rhythms in *C. elegans* behaviour, physiology and gene expression. The work described here focused on detecting and characterising gene expression rhythms in *C. elegans* to advance understanding of the molecular clock.

First, the use of the promoters of two putative rhythmic genes, *nlp-36* and *sur-5*, were explored as luciferase reporters of the circadian clock. These genes did not provide an effective readout of a molecular clock using entrainment conditions of light, temperature or in combination. Expression of *nlp-36*, but not *sur-5*, was found to oscillate under entrainment (environmental cycles) and free-running (constant) conditions. However, this required exposure to both light and temperature cycles in phase, and revealed high variation between biological replicates and a small fold change.

To identify new candidate genes and improve understanding of the extent of genetic circadian rhythms in *C. elegans*, this dual light and temperature protocol was applied over a two-day RNA sequencing time series. This experiment revealed 263 genes, mostly previously unidentified, to significantly oscillate over the full time series. Further analysis of this time series also revealed non-circadian patterns including diurnally driven genes and novel 16-hour oscillations in gene expression.

Collectively these results suggest *C. elegans* to be a circadian clock-regulated organism on a smaller scale than other species, detail a new approach to effective entrainment and highlight challenges that must be overcome to understand the *C. elegans* molecular clock.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I'd firstly like to thank my family (some of whom might have had a vague idea of what I was doing here) for their support: my parents, Lorraine and Shane Munns, my brothers, Alex and Oliver, my grandparents, Patricia and Brian Kaye, my uncles Rafael Munns and Mark Kaye and my cousin Darcy Kaye.

In addition, come what may, writing this thesis has been a large undertaking, and I can't envisage producing anything else quite so substantial any time soon. For that reason, it seems fitting to dedicate it to the memories of my grandparents Madge and Alan Munns and my cousin Nikki Kaye.

I'd like to thank my supervisors, Dr Sangeeta Chawla and Professor Seth Davis for their support, guidance, patience and friendliness. Having two such generous and accommodating supervisors was invaluable to my PhD work, and made my academic experience a hugely enjoyable one.

I'd also like to thank the Department of Biology for funding my PhD and my Thesis Advisory Panel members, Dr Sean Sweeney and Professor Calvin Dytham, for their helpful advice.

My PhD work also benefited from the kind scientific assistance of a number of others. For their help with this project, I'd like to thank Dr Katherine Newling, Dr Sally James, Dr John Davey and Dr Lesley Gilbert (University of York, Technology Facility) and Professor Ian Hope (University of Leeds).

My time at York has been an absolute joy, thanks to a great many wonderful people. Thanks to Kayla McCarthy, Laura Covill (ugh, bridge), my Golden Ball quiz brothers, Alex Haworth and Dr Nick Johnson, Dr Lotte van Beek (my thesis buddy), Emma Stewart, Dr Iulia Gherman (boardgames!), Tess Leslie, Nathaniel Holman (Gradshare!), Dr Iain Hartnell, Annie Smith, Grace Cowen, Dan Jackson, Ed Sonia, Phil Brailey-Jones, Caroline Pearson (Gradshare!), James Robson, Aritha Dornau (+ Nagini, Houdini and Tahini). The many friendly faces from D1: Laura Fort (the best), Dr Chris Ugbode, Alison Fellgett, Egle Beigaite, Valentina Galassi Deforie, Dr Chris Elliot, Katy Hyde, Khalid Jambi, Katherine Hanlon, Dr Adam Middleton, Amy Cording and Dr Ryan West (I suppose). The friendly faces up on L2: Dr Rachael Oakenfull, Dr Manuela Iovinella, Mandi Davis, James Ronald, Sarah Lock, Paula Avello Fernández, Dr Jess Hargreaves and Jessie Downing.

Special thanks to my London pals, the nicest people in the world: Zammy Fairhurst, Shaan Subramaniam, Magda Meier, Charlotte Flanagan, Lucy Nevard, Emily Heathward and Marc Beloussov.

A quick shout out to the Brixworth diaspora: Richard, Joe, Tegan, Paul, Dave, Karina, Amy. (they'll never read this).

And last but not least, thanks to the immortal Ronnie James Dio:

"Long Live Rock 'n' Roll!"

Declaration

I declare that this thesis is a presentation of original work and I am the sole author, except where stated in the text. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Some of the data contained in Chapter 3 have been accepted for publication in the following article, on which I am a listed author:

Hargreaves, J. K., Knight, M. I., Pitchford, J. W., Oakenfull, R. J., Chawla, S., Munns, J. & Davis, S. J. Wavelet spectral testing: Application to nonstationary circadian rhythms. *Ann. Appl. Stat.* (2019).

Chapter 1. Introduction

Circadian rhythms describe daily cycles in physiology and behaviour that are endogenously controlled, but highly responsive to environmental change. They are the product of circadian clocks, subcellular timing mechanisms that oscillate over approximately 24-hour periods. Circadian rhythms reflect a common evolutionary drive to adapt to a diurnal world, and can be observed in vastly diverse organisms, ranging from animals, plants and fungi to photosynthetic prokaryotes (Takahashi, 2017; Hardin, 2005; Harmer, 2009; Hurley et al., 2014; Liu et al., 1995). In many organisms, circadian clocks are integral to essential biological processes and may be a necessity for health and survival. In humans for example, the timing of sleep, feeding and insulin sensitivity are all partially governed by the clock (Scheer et al., 2009). Conversely, chronic disruption and aberrant clock functioning have been associated with a wide range of diseases, including sleeping disorders, obesity and diabetes, psychiatric conditions, neurodegenerative disease and cancers (Zee et al., 2013; Jagannath et al., 2017; Sahar and Sassone-Corsi, 2009). A strong understanding of the molecular basis of circadian rhythms has been uncovered in most widely-used animal models of health and disease, with mice, fruit flies (*Drosophila melanogaster*) and zebrafish (*Danio rerio*) all being established molecular clock models (Yu and Hardin, 2006; Vatine et al, 2011). A notable exception exists however, in the nematode *Caenorhabditis elegans*, for which understanding of the circadian clock machinery is poor. *C. elegans*, as a visually blind nematode, raises a fascinating question as to the evolutionary importance of a clock, in that it may occupy an ecological niche in which diurnal timing is no longer essential, and may therefore lack robust ancestral circadian rhythms. Conversely, as an important model of developmental biology, ageing and disease (Corsi et al., 2015; Denzel et al., 2019; Alexander et al., 2014), an understanding of circadian timing in *C. elegans* could illuminate the roles a circadian clock plays throughout the life of an organism. This chapter provides a background to circadian rhythms in established models, explores

the potential evolutionary and ecological basis for a *C. elegans* clock and discusses existing literature documenting circadian rhythms in these nematodes.

1.1 Background to circadian rhythms

Circadian clocks are characterised by a number of shared features, common to the biologically diverse species in which they have been modelled. The similarities that exist between metazoan models have enabled the construction of a clear picture of circadian biology in these species, defining a number of hallmarks of circadian rhythmicity, along with widely applicable experimental approaches for the detection of rhythms.

1.1.1 The conserved molecular basis of circadian rhythms: models of transcriptional oscillators in model species

The most established mechanism by which circadian rhythms are produced (and on which this thesis is based) is through transcription-translation negative feedback loops (TTFLs), which generate approximately 24-hour rhythms in gene expression. In animals, TTFLs have been best characterised in *Drosophila* and mammals (principally mice), in which many of the molecular components of the core feedback loop show considerable genetic and mechanistic conservation (Figure 1; Yu and Hardin, 2006). In both of these models, illustrated in Figure 1.1, heterodimeric transcription factors regulate the expression of their own inhibitors; a primary negative feedback loop is formed as these transcription factors bind common promoter sequences (containing E-box elements), enabling the transcription and translation of genes encoding the negative arm of this loop. The translated proteins themselves form heterodimers and translocate to the nucleus, where they promote displacement of transcription factors from E-boxes. The eventual degradation of these negative repressors allows a new cycle of transcription to begin approximately 24 hours later, resulting in transcription patterns with circadian frequency in E-box-promoted genes.

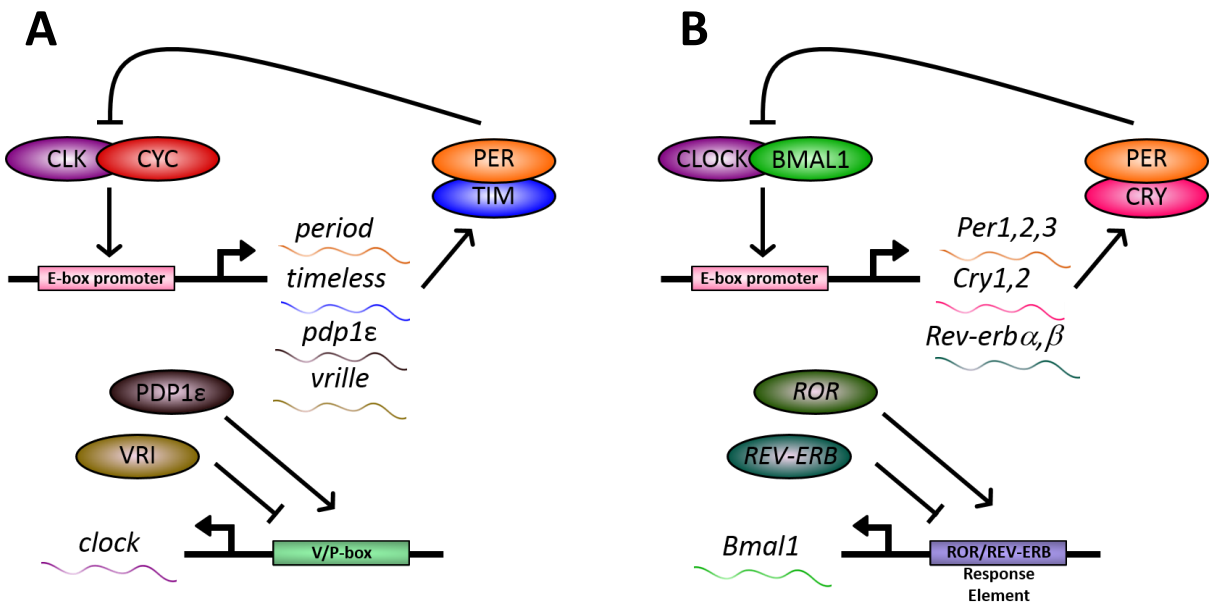


Figure 1.1: Model of primary and secondary transcriptional feedback loops.
A) *Drosophila* **B)** mammals**.** Primary feedback loops consist of semi-homologous heterodimers (CLK/CYC, encoded by *clock* and *cycle*) in *Drosophila* and CLOCK/BMAL1 in mammals, which bind to E-box sequences. This activates transcription and translation of proteins that in turn form heterodimers (PER/TIM and PER/CRY) and repress their own transcription by inhibiting CLK/CYC and CLOCK/BMAL1. These loops generate approximately 24-hour endogenous cycles in gene expression under constant conditions. Primary loops are augmented by E-box-governed stabilising loops in which competing elements facilitate or repress the positive element of the primary loop; PDP-1ε and VRILLE competitively regulate *clock* transcription in *Drosophila*, while RORs and REV-ERBs regulate *Bmal1* transcription in mammals (TTFLs are comprehensively reviewed in Yu and Hardin, 2006; Hardin 2005; Takahashi 2017). The TTFL is also highly conserved in zebrafish (*Danio rerio*), in which the components of the primary and secondary loops are largely homologous to the mammalian system, with notable additional paralogues of the *clock*, *bmal*, *per* and *cry* genes (reviewed in Vatine et al., 2011).

TTFLs are characteristically robust in that they compensate for external temperature to consistently maintain an approximately 24-hour period (Kurusawa and Iwasa, 2005), enabling organisms to endogenously anticipate diurnal changes. However, they are also highly sensitive to changes in environmental signals (zeitgebers), which act to shift the phase of the clock, enabling organisms to synchronise to seasonal variation in day length. Light is the best-characterised zeitgeber in all models of circadian rhythms, and in *Drosophila* and mice acts to upregulate transcription of the

genes encoding the negative repressor element. In *Drosophila* this occurs through the activation of a light-receptive cryptochrome (homologous to the mammalian *Cry* genes, but differing in protein function), which degrades TIM proteins and allows transcription of *period* and *timeless* to recommence (Yu and Hardin, 2006). In mice, transcription of the *Per* gene is actively upregulated as light triggers a signalling cascade that results in activation of the *Per*-regulating transcription factor CREB (Jagannath et al., 2017).

The approximately 24-hour rhythms of the TTFL persist cell-autonomously in most tissues in these models (Ito and Tomioka, 2016; Yoo et al., 2004; Whitmore et al., 2000). In mammals, individual cellular clocks are coordinated by a master pacemaker, the suprachiasmatic nucleus (SCN) in the hypothalamus. This enables light signals detected in the eye, predominantly by melanopsin-containing intrinsically photosensitive retinal ganglion cells (ipRGCs), to reach the majority of cells and tissues that are not light responsive (Peirson et al., 2009). Conversely, *Drosophila*, while also exhibiting central coordination of the clock (via clusters of lateral and dorsal neurons in the brain), can entrain to its diurnal environment cell-autonomously through light-responsive peripheral tissues (Ito and Tomioka, 2016). Zebrafish, while sharing conserved core clock genes with mammals, are also capable of extraocular photoreception in directly-light entrainable peripheral tissues (Whitmore et al., 2000; Peirson et al., 2009), and the extent to which *D. rerio* relies on a centrally-governed pacemaker is unclear (Ben-Moshe Livne et al., 2016).

The robust endogenously-generated rhythms in gene expression generated by TTFLs can be experimentally observed in constant (free-running) conditions at the level of tissues in mice (Yoo et al., 2004) and at the whole-organism level in *Drosophila* and developing zebrafish (Figure 1.2; Stanewsky et al., 1997; Weger et al., 2013). An additional point of consideration for circadian study however, is that these entrained tissue or organism-wide rhythms in gene expression do not persist indefinitely. Instead, expression rhythms show a characteristic damping effect, generally after

several days in the absence of environmental stimuli (illustrated in Figure 1.2 using the *Drosophila period* gene as an example). Rather than a loss of rhythmicity, this damping is suggested to be a consequence of cells and tissues falling out of phase in the absence of a zeitgeber; *in vitro* experiments in mammalian fibroblasts have suggested a lack of coupling between individual cells, which still maintain individual rhythms independently but with divergent phases (Welsh et al., 2004). Accordingly, experiments investigating circadian rhythms (including most of those performed as part of this thesis, and those referenced in this chapter) typically follow a similar protocol in most organisms, beginning with entrainment, in which an organism is exposed to several environmental cycles in order to synchronise and set the phase of the TTFLs, before release into constant conditions, in which the endogenous oscillator is allowed to free-run and rhythmicity can be measured.

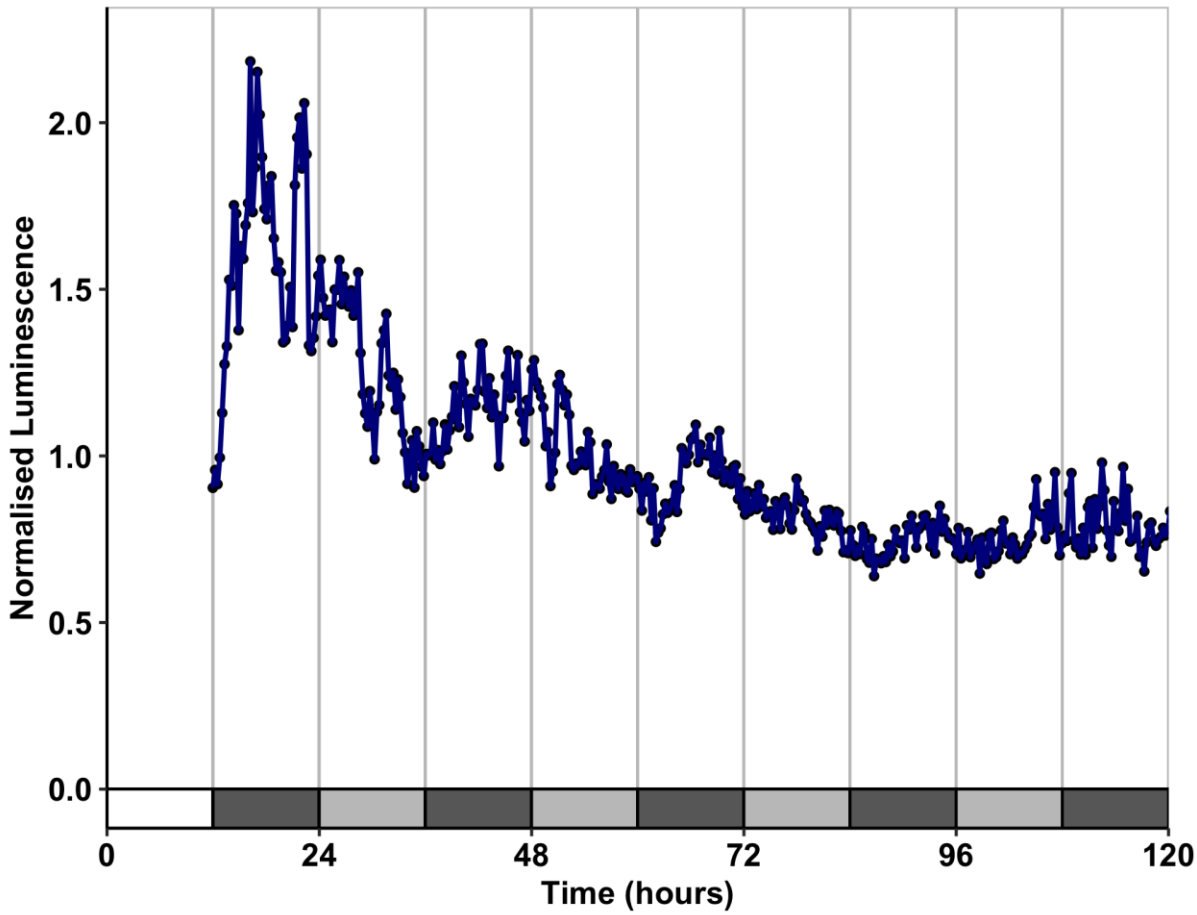


Figure 1.2: Damping of circadian rhythms of the *period* gene in *Drosophila melanogaster* in constant darkness.
Data show average normalised luminescence over time, generated from transgenic *D. melanogaster* (males, n = 6) that express luciferase downstream of part of the *period* gene promoter sequence (Stanewsky et al., 1997). Recordings took place in constant darkness following 3 days entrainment to 12:12 hour light:dark cycles at constant temperature. Light grey bars indicate the subjective day, when the light phase would be expected in cycling conditions. Further methodological details on luminescence recording is given in Chapter 3 methods, section 3.4.

1.2 Evolutionary and ecological potential for circadian rhythms in *C. elegans*

Despite the seemingly ubiquitous presence of circadian clocks and their conservation in animals, the evolutionary conservation, mechanisms and significance of circadian timing in *C. elegans* remain open questions in chronobiology. *C. elegans* is short-lived (with a variable laboratory lifespan of 12-18 days at 20°C; Kenyon, 1997), and a relatively anatomically simple animal (adult hermaphrodites comprise only 959 cells,

including 302 neurons). As to whether it is capable of, or likely to exhibit circadian rhythms could depend on a wide range of factors relating to its evolutionary history and life cycle. Discussed here are the evolutionary conservation of TTFL components and environmental sensing in *C. elegans*, its ecology and the environmental niche it occupies, and parallels between circadian rhythms and the genetics of *C. elegans* development, along with the implications these aspects of its biology might have for a circadian clock.

1.2.1 Homologues of TTFL genes and potential zeitgebers in *C. elegans*

From an evolutionary perspective, there is clear potential for a circadian oscillator to persist in *C. elegans*. By molecular phylogenetics, *C. elegans* has been placed alongside *D. melanogaster* in the Ecdysozoa clade within the protostomes (Borner et al., 2014), while mammals are distally positioned in the deuterostomes (Bourlat et al., 2006). Given the high degree of genetic and functional conservation between the circadian systems of *Drosophila* and mammals, a transcriptionally-regulated circadian system in *C. elegans* would be expected to retain some shared elements of the *Drosophila* and mammalian clocks. In particular, *C. elegans* should share a common ancestor with *D. melanogaster* and mammals with homologues of the CLK and PER proteins (Figure 1.1). This is reflected in the *C. elegans* genome: *C. elegans* expresses homologues of the majority of the TTFL elements that are conserved between *Drosophila* and mammals, along with homologues of genes unique to each (summarised in Table 1.1).

Table 1.1: *C. elegans* homologues of core clock-related genes in *Drosophila* and mice. (adapted from Romanowski et al., 2014)

***Caenorhabditis elegans* Gene**	***Drosophila melanogaster* clock gene**	***Mus musculus* clock gene**
lin-42	*period*	*Per1, Per2, Per3*
tim-1	*timeless*	-
aha-1	*cycle* and *clock*	*Bmal1 and Clock*
ces-2	*pdp1ε*	-
atf-2	*vrille*	-
kin-20	*doubletime*	*Casein kinase 1δ, Casein kinase 1ε*
nhr-23	-	*Rora*, *Rorb*
nhr-85	-	*Rev-erbα, Rev-erbβ*

Notably absent in the genome of *C. elegans* however, are genes with significant homology to mammalian *Cry1, Cry2* or *Drosophila cryptochrome* (Romanowski et al., 2014). These cryptochromes have different, but essential roles in mammals and *Drosophila*, acting as part of the core repressor element of the mammalian clock, and serving as a light input to the clock in *Drosophila* (Hardin, 2005). This provides a potentially vital evolutionary difference between *C. elegans* (a visually blind animal) and established circadian clock models, in which light sensing and phototransduction are key to synchronising and phase-adjusting the clock (Jagannath et al., 2017; Harper et al., 2017). The ipRGC and melanopsin-dependent mechanism of light detection used by mammals to set the clock is distinct from visual photoreception (Peirson et al., 2009), and therefore the lack of a visual system in *C. elegans* should not preclude it from possessing a light-oriented clock. However, the degree to which *C. elegans* is responsive to light is open to question. *C. elegans* has been demonstrated to exhibit negative phototaxis in response to light of varying wavelengths (Burr, 1985; Ward et al., 2008), but only possesses one known photoreceptor: LITE-1, a seven-transmembrane taste-receptor homologue (Gong et al., 2016). LITE-1 is highly sensitive to UV light (with absorbance peaks at wavelengths at 280 and 320 nm), shorter than the absorbance peak of mammalian melanopsin (480 nm; Peirson et al., 2009) and the peak sensitivity of *Drosophila*

cryptochrome (suggested to be approximately 390 nm; VanVickle-Chavez and Van Gelder, 2007). The potential for a photically-entrained clock in *C. elegans* is therefore quite unclear.

However, while light is a pervasive environmental signal and the most well-understood zeitgeber in the established animal clock models, numerous other diurnal inputs have been shown to synchronise circadian rhythms. Non-photic zeitgebers that have been demonstrated to entrain animal clocks include temperature, feeding (Stokkan et al., 2001; Cavallari et al., 2011), social interaction (Fuchikawa et al., 2016) and mechanosensation (Simoni et al., 2014). Of these, temperature is the most well-studied; temperature is well-established in synchronising transcriptional oscillations in plant circadian systems (Salome and McClung, 2005), and the three major animal models discussed in section 1.1, mice, *Drosophila* and zebrafish, have all been shown to entrain to thermal cycles to varying extents. In endotherms, temperature may only play a limited role; behavioural rhythms in mice have been shown to be only weakly entrainable to temperature cycles (Refinetti, 2010), while *in vitro*, gene oscillations have been shown to be phase-adjustable to temperature, but predominantly only in peripheral, non-SCN tissues (Buhr et al., 2010). In *Drosophila* and zebrafish however (the peripherally entrainable, poikilothermic models), rhythms in both TTFL gene expression and behaviour have been reported to entrain to temperature cycles in invariant light conditions (Glaser and Stanewsky, 2005, 2007; Currie et al, 2009; Yoshii et al., 2005; Lahiri et al., 2014; Lahiri et al., 2005; López-Olmeda et al., 2006; López-Olmeda and Sánchez-Vázquez, 2009). An increased ability to entrain to temperature in poikilothermic animals, or invertebrates, could certainly suggest *C. elegans* might be capable of using temperature as a zeitgeber.

As to whether *C. elegans* might use temperature preferentially or in place of light to entrain, it would be the first species documented to do so. However, work in *Drosophila* has suggested that temperature and light entrainment mechanisms,

while likely integrated, represent partially independent systems. Under constant light, which causes arrhythmicity in flies, temperature cycles can restore locomotor and TTFL protein rhythms (Yoshii et al., 2005). Conversely, flies in which temperature synchronisation is abolished (*nocte* strain mutants) have been shown to still entrain *period* gene rhythms to light:dark cycles (Glaser and Stanewsky, 2005). Further, in the absence of a fully-functioning CRY protein, the most vital clock element for which *C. elegans* expresses no homologous genes, *Drosophila* can still entrain to temperature (Glaser and Stanewsky, 2005). Typically, light entrainment appears to take precedence over temperature, as when both zeitgebers are applied in conflicting phases, *Drosophila* (and zebrafish) locomotor activity has been shown to phase-adjust to light:dark cycles away from temperature (Yoshii et al., 2010; Harper et al., 2016; 2017; López-Olmeda and Sánchez-Vázquez, 2009). However, flies lacking functioning CRY have been shown to remain entrained to a temperature cycle (Yoshii et al., 2010). The presence of a distinct input system to *Drosophila* clocks that can persist when light sensing is perturbed, suggests temperature is a potent environmental cue, and if it lacks complex photoreception or a cryptochrome homologue, one that *C. elegans* could still use to entrain a clock. Discussed in section 1.3, both light and temperature have been explored as zeitgebers in *C. elegans* circadian rhythm experiments.

1.2.2 Ecology of *C. elegans*

As to whether *C. elegans* might use temperature, light or other zeitgebers to entrain its clock, a confounding issue is that despite being one of the most widely-used laboratory model organisms, relatively little is known about its ecology. As such, the diurnal signals that nematodes would likely be exposed to in nature are not immediately clear. Populations of *C. elegans* (including the laboratory N2 strain on which most research is based), have been isolated in compost and soil (Sterken et al., 2015; Nicholas et al., 1959). If *C. elegans* is primarily a soil-dwelling bacterivore, exposure to changes in light may be rare. Nematodes in soil would likely be more

exposed to temperature oscillations, which largely correspond to air temperatures throughout the day, albeit with amplitude and phase differences at different depths and with seasonal variation (Kristensen, 1959; Parton and Logan, 1981; Sagalovich et al., 2002). However, *C. elegans* preferred habitat may not be a subterranean one. In conditions of overcrowding, starvation or high temperatures, developing *C. elegans* larvae can enter a long-lived metabolically inactive dauer state (Androwski et al., 2017; Golden and Riddle, 1984; life-cycle illustrated in Figure 1.3). Researchers searching for wild populations of *C. elegans* in compost heaps have reported *C. elegans* to be primarily found in the dauer state (Barriere and Félix, 2005), while largely proliferating populations have been isolated on fruit, plant stems and in association with small invertebrates (Félix and Duveau, 2012). This might suggest that the more favourable environment for *C. elegans* is a highly diurnal one. Furthermore, *C. elegans* are likely to commensally utilise invertebrates for migration and dispersal (Petersen et al., 2015), and interactions with other species may drive the nematode into day-night cycling environments. Overall, while the natural habitat of *C. elegans* remains a question that requires further study, it should not be assumed that *C. elegans* does not encounter diurnal variation in its environment and with it the selective pressure from which circadian clocks have evolved.

1.2.3 Developmental roles of clock gene homologues in *C. elegans*

In investigating *C. elegans* in the context of conserved rhythms in other species, a further evolutionary consideration that could have bearing on a potential clock exists in its developmental biology. Although the presence of a *C. elegans* circadian oscillator has not been established, it does exhibit well-defined timing events in its developmental program (illustrated in Figure 1.3), in which each of several larval stages have temperature-dependent defined durations (Byerly et al., 1976). Developmental biology and circadian timing are also biologically intertwined in other species; in *Drosophila* for example, the first *period* gene mutants with altered circadian rhythm lengths were discovered through pupal eclosion phenotypes

(Konopka and Benzer, 1971), and also show an altered developmental duration (Kyriacou et al., 1990). In *C. elegans*, several of the homologues of conserved core clock genes in other species (given in Table 1.1) have been shown to have roles in governing progression through the developmental system, and a number of mechanistic parallels between the processes are apparent. The genetics of developmental timing in *C. elegans* could thus provide potential insights into the extent to which the molecular circadian clock is likely to be conserved in *C. elegans*.

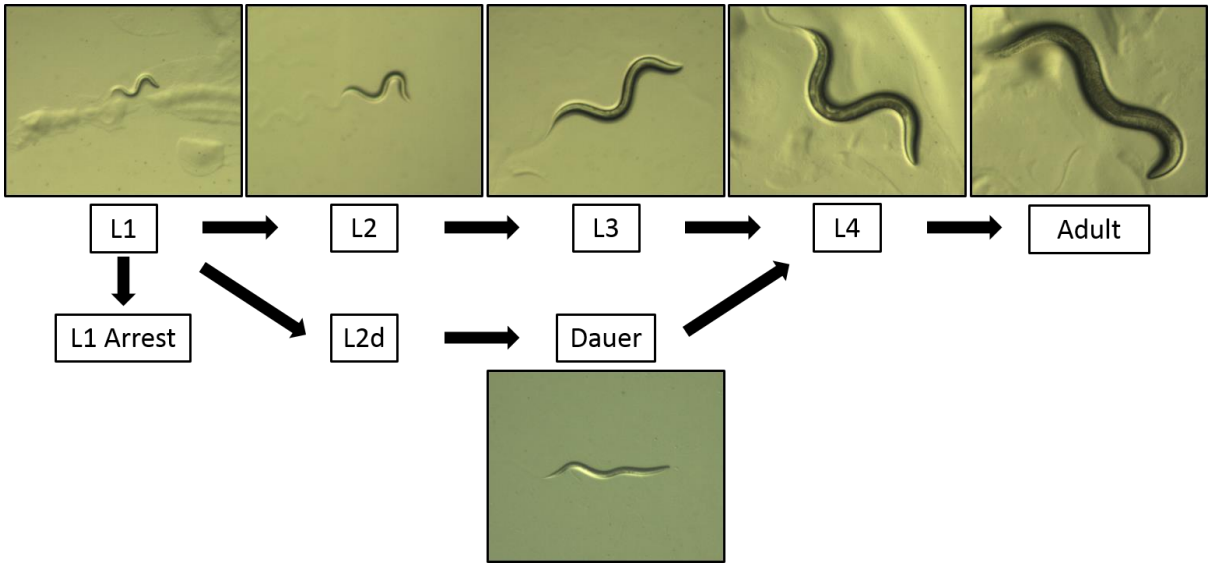


Figure 1.3: Developmental stages of *Caenorhabditis elegans*.
Development comprises four larval stages prior to adulthood: L1, L2, L3 and L4, each of which has a defined duration at a given temperature (Byerly et al., 1976). Progression is separated by moults, coupled with a period of lethargus (Lažetić and Fay 2017). In unfavourable conditions, *C. elegans* can enter two longer-lived, stress resistant states. Firstly, if starved, *C. elegans* can reversibly arrest growth at the L1 stage (Baugh 2013). Secondly, if overcrowded, starved or exposed to high temperatures, larvae can enter a non-feeding, morphologically distinct and metabolically inactive alternative L3 stage, known as a dauer. Dauer entry occurs following a decision at a pre-dauer L2d stage, after which a nematode can arrest or undergo normal developmental progression (Androwski et al., 2017; Golden and Riddle, 1984).

Most notably, numerous studies have found developmental regulatory roles for *lin-42*, the *C. elegans* homologue of the *period*/*Per* genes (which encode the conserved portion of the heterodimeric repressor element in the *Drosophila* and mammalian TTFLs; Figure 1.1). *lin-42*, like the *period*/*Per* genes, shows oscillatory expression patterns, but relative to moulting cycles throughout development rather than circadian time (Jeon et al., 1999). Additionally, LIN-42 protein has been suggested to bind extensively to regulatory regions of protein-coding and non-coding genes and, again analogous to PER, to act as a negative regulator of transcription (Perales et al., 2014). Appropriate *lin-42* expression is required for normal development; nematodes expressing truncated *lin-42* have been shown to develop more slowly, with increased growth arrest, erratic lethargus and reduced survival to adulthood (Edelman et al., 2016). Dauer entry is also increased in *lin-42* loss-of-function mutants, while increasing *lin-42* expression has been shown to prevent dauer formation (Tennessen et al., 2010). These phenotypes are likely the complications of mistiming in the developmental program; a loss of functioning LIN-42 protein causes accelerated development in certain tissues, leading to premature cell differentiation (Banerjee et al., 2005), morphogenic events (Tennessen et al., 2006; Berardi et al., 2018) and consequent tissue abnormalities in adults. The basis of these observations has been suggested to be a result of the role of *lin*-42 as a transcriptional regulator, specifically of microRNAs that themselves govern developmental progression; with loss of *lin-42* function, the heterochronic miRNAs let-7, lin-4, miR-48, miR-84 and miR-241 are all overexpressed (Perales et al., 2014; McCulloch and Rougvie, 2014). *lin-42* therefore appears to resemble its mammalian and *Drosophila* homologues as an oscillatory transcriptional regulator, vital to a timing process. However, its considerably different function raises the question of any role *lin-42* might play in regulating circadian rhythms.

Alongside *lin-42*, two other homologues of core clock elements have been shown to have similar effects on developmental timing: the *C. elegans* homologue of a highly conserved kinase (*doubletime*/*Casein kinase 1ε*) that is involved in regulating the 24-

hour period of TTFLs in *Drosophila* and mice (Isojima et al., 2009; Top et al., 2018), and *tim-1*, the homologue of the *Drosophila* TTFL repressor element *timeless* (Figure 1.1). Independent RNA silencing of *kin-20*, *tim-1* and *lin-42* has been shown to trigger some of the same premature differentiation events while also ameliorating the heterochronic phenotypes and lethality seen in let-7 loss-of-function mutants (Banerjee et al., 2005). Like *lin-42*, *kin-20* loss-of-function has also been found to slow development, and also lead to reduced brood sizes (Rhodehouse et al., 2018). While more work would be needed to fully characterise these genes, the similar and profound effects found by perturbing three vital clock gene homologues could potentially indicate an evolutionary functional shift in which the clock network has been adapted to provide accurate timing to heterochronic events and ensure that development proceeds without error.

In addition to homologous genes, other mechanistic parallels have been drawn between circadian clocks and *C. elegans* development. The TTFL model of circadian rhythms (Figure 1.1) suggests that a proportion of the transcriptome should oscillate in circadian time (the basis of Chapter 4. Results 2 in this thesis). One RNA-seq experiment across larval development found that a fifth of expressed genes oscillated along with the larval stages, which are largely consistently timed at 25°C (Hendriks et al., 2014). The lethargus phases that precede moulting have also been likened to sleep, a cyclical circadian output (Borbély et al., 2016), in that lethargus is a reversible state of behavioural quiescence characterised by reduced responsiveness to stimulation (Raizen et al., 2008). The pigment dispersing factor (PDF) hormone, which is necessary for circadian rhythms in *Drosophila* behavioural quiescence (Renn et al., 1999), has also been shown to mediate quiescence in *C. elegans* (Choi et al., 2013). Further, a novel gene, *nlp-22* (for which no homologues exist outside the *Caenorhabditis* genus) oscillates relative to lethargus and if overexpressed, triggers behavioural quiescence and reduced feeding and movement (Nelson et al., 2013). *nlp-22* has itself been suggested to be regulated by *lin-42*, highlighting a mechanistic analogy between lethargus and the larval stages (a

possible larval clock) and the circadian clock that is potentially underpinned by genetic conservation. The numerous parallels and genetic links observed between *C. elegans*-specific development and circadian timing in other organisms again raise the possibility of *C. elegans* exhibiting a modified circadian system in order to optimally govern its development.

In considering whether the parallels between circadian clocks and development in *C. elegans* are indicative of an organism that has diverged from its diurnally rhythmic metazoan relatives, or simply reflect two convergent aspects of its biology, an important point of note is that the comparisons drawn between circadian and developmental timing are also associated with numerous substantial genetic and mechanistic differences. *lin-42* and *period*/*Per* genes are negative transcriptional regulators but have different action; *period*/*Per1*,*2* and the TTFL act to generate oscillations, while *lin-42* appears to suppress oscillations in the microRNAs it regulates (McCulloch and Rougvie, 2014; Perales et al., 2014). *kin-20* and *tim-1* have been shown to have similar roles as *lin-42*, but there is no evidence of the genes forming a developmental network akin to the circadian system. Instead, *kin-20* and *lin-42* are suggested to affect let-7 largely independently (Rhodehouse et al., 2018), while *tim-1* has been shown to have an entirely different function in regulating chromosome cohesion (Pasqualone et al., 2003). Finally, while many genes oscillate in time with lethargus, the different larval stages generally proceed for different lengths of time, only having similar durations at 25°C (Hendriks et al., 2014; Byerly et al., 1976) and the system thus lacks temperature compensation, a key feature of robust circadian timing (see section 1.1). As such, these developmental observations could indicate a potentially superficial resemblance to a circadian system. While clock gene homologues have clear functions in development and mechanistic analogies can be drawn between the two systems (and are frequently placed in that context by authors), the evolutionary significance of these parallels is unclear in terms of the presence of circadian timing in *C. elegans*.

1.3 Evidence for circadian rhythms in *C. elegans*

The uncertainties discussed in section 1.2 surrounding the capability of *C. elegans* for photoreception, the extent of its exposure to diurnal cycles in nature and whether aspects of an ancestral circadian system have been modified to time development, could point to an organism without a functioning circadian clock. However, the lack of a known molecular oscillator in *C. elegans* could also arise from difficulties in modelling rhythms. Discussed below, numerous examples of circadian rhythms have been reported in *C. elegans* behaviour, physiology and gene expression. Experiments investigating these rhythms collectively serve to provide evidence of a *C. elegans* circadian clock, but also reveal a challenging organism for measuring circadian outputs.

1.3.1 Circadian rhythms in behaviour and physiology

A major difficulty in studying circadian rhythms in *C. elegans* is that it lacks an obvious detectable rhythmic phenotype. Fundamental outputs of the clock in the established animal models are cycles of activity and rest, and circadian rhythms can be easily studied in *Drosophila* using simple infrared activity monitors to record robustly cycling locomotor activity (Fogg et al., 2014), and similarly using running wheels in mice and other rodents (Siepka and Takahashi, 2005). These methods enable straightforward detection of animals that show abnormal or arrhythmic patterns of behaviour, and in doing so have enabled mutations in genes important for proper clock function to be uncovered. Such approaches have led to the discovery and characterisation of core clock genes and important regulatory elements (Ralph and Menaker, 1988; Vitaterna et al., 1994; Lowrey et al., 2000; Ko et al., 2002; Fogg et al., 2014). Detecting rhythms in *C. elegans* is less simplistic, but numerous phenotypic approaches have been developed that have reported the presence of circadian rhythms in the nematode, some of which could potentially enable the dissection of the molecular clock.

The majority of these phenotypic methods have attempted to follow the approaches that have proven valuable in other models in focusing on locomotor activity. *C. elegans* locomotor assays have indicated the presence of a circadian clock that can be entrained by both light and temperature, while also highlighting difficulties in measuring rhythms in nematodes. As regards to light entrainment, multiple publications have reported evidence from infrared beam-crossing assays, similar to those used in *Drosophila*, suggesting wild-type *C. elegans* in liquid culture exhibit significant circadian rhythms in locomotor activity under cycles of light and dark and in subsequent constant darkness (Simonetta and Golombek 2007; Simonetta et al., 2009; Herrero et al., 2015). Across these experiments however (which all utilised variations of the same protocol), only around half or fewer nematodes were reported to demonstrate significant locomotor rhythms; one study identified significant rhythmicity in 52.1% of individuals under entrainment conditions and 37.5% in free-running (Simonetta and Golombek, 2007), while another in looking at small populations (5-7 nematodes) found close to one-third to be significantly rhythmic in both entrainment and free-running conditions (Herrero et al., 2015). These proportions are indicative of a less behaviourally rhythmic organism than other models, but perhaps reflect an inevitable challenge of measuring *C. elegans* locomotion in that unlike mice and *Drosophila*, *C. elegans* in liquid (or solid) media do not show continuous periods of rest. The published data indeed show movement and activity spikes across the 24-hour cycle, requiring the authors to use several data filtering approaches to detect rhythms (Simonetta and Golombek 2007). Another published example of circadian locomotor rhythms in light-entrained nematodes tracked movement in liquid media over time and reported significant circadian fluctuations in average movement speeds (Saigusa et al., 2002). Unlike the infrared studies, this approach did not report arrhythmic individuals despite using very small populations of 2-3 nematodes. However, despite being published nearly two decades ago, no further studies have utilised this method to interrogate the clock, and the approach has yet to be documented in large samples of nematodes. A further point

of note concerning these reports of light-entrained rhythms collectively, is that while all using 12:12-hour light:dark cycles to entrain the same N2 laboratory strain, studies show a lack of consistency with regards to the phase of the circadian oscillations. Different infrared assay studies found nematodes to have activity peaks during the night-day transition or between early and the middle of the night, while the approach measuring movement speed found activity peaks in the middle of the day (Simonetta et al., 2009; Herrero et al., 2015; Saigusa et al., 2002). These phases were also highly divergent within the infrared studies with a broad range of activity peaks, which is generally in contrast to results in *Drosophila* and mice (Harper et al., 2016; Vitaterna et al., 1994). Taken together, these data provide evidence that *C. elegans* can entrain endogenously-driven behaviour to light:dark cycles, but also suggest a difficult to assay organism that is either highly variable in its diurnal patterns of activity, or in which light is not sufficient to generate robust patterns of activity.

In considering temperature as a zeitgeber, both one aforementioned infrared beam crossing study and one video-tracking approach (calculating the velocity of nematodes on solid media) have reported temperature-entrained locomotor activity that persists in free-running conditions (Simonetta et al., 2009; Winbush et al., 2015). Comparable to the light entrainment studies, the former approach resulted in significant rhythms in 57% of nematodes, while the latter study reported 24-hour rhythms in small populations (5-20) as a proof of principle. In contrast to light, phase appears to be quite consistent in response to temperature, with nematodes being (predictably as an ectothermic organism) more active in the warm phases. However, daytime peaks were also reported in free-running conditions in both approaches, and the infrared study reported far less variation in free-running peaks of activity following entrainment by temperature than by light (Simonetta et al., 2009). This may suggest temperature to be more effective in entraining behaviour in nematodes than light, contrasting with observations in other models (López-Olmeda and Sánchez-Vázquez, 2009; Refinetti, 2010; discussed in section 1.2.1). Further, tracking

movement velocity found nematodes to demonstrate anticipation, a key hallmark of endogenously-driven rhythms, with nematode velocity increasing prior to the application of increased temperature. Velocity measurements did however also suggest another circadian hallmark in rapidly damping rhythms, only persisting for one day in free-running conditions. While quick damping is seen in gene expression rhythms in other model organisms (illustrated by the *Drosophila* example given in Figure 1.2), locomotor activity rhythms can persist for over a week in constant darkness without damping out (Fogg et al., 2014; Renn et al., 1999). As such, it would appear that while the hallmarks of a circadian clock are present in nematodes and that temperature entrainment is no less effective than light in generating circadian rhythms, behavioural activity cycles are relatively non-robust in *C. elegans*.

These behavioural approaches have generally been presented as *in vivo* assays that could be used to study the molecular clock. Despite this, only very limited insights into the genetic basis of *C. elegans* rhythms have been uncovered in these studies, and only a small number of genes have been suggested to be necessary for proper circadian cycling. The *period*/*Per* homologue *lin-42* is the only core TTFL gene (see Figure 1.1; Table 1.1) reported to have an effect on circadian cycling, with two partial deletion mutants having been found to display a significantly lengthened average period in locomotor activity rhythms, relative to wild-type animals in response to light entrainment (Simonetta et al., 2009). However, the reported period increase was just 1.5 hours, a relatively small effect in comparison to substantial phenotypic effects shown to occur in cases of *Drosophila period* loss of function (Konopka and Benzer, 1971). Given that *lin*-42 loss of function leads to considerable developmental abnormalities (Banerjee et al., 2005; Tennessen et al., 2006; Berardi et al., 2018; discussed in section 1.2.3), differences in periodicity could also potentially be attributed to a non-circadian function of the gene. A more drastic effect on rhythms was reported in looking at the two nematode homologues of *Drosophila* PDF and its receptor (*pdf-1*, *pdf-2* and *pdfr-1*), which play an important role in circadian synchronisation in flies (Renn et al., 1999; Goda et al., 2019). Light-entrained

locomotor activity of *C. elegans pdf-1* deletion mutants was completely abolished, while genetic rescue experiments could return rhythmicity to locomotor cycles. Nematodes with deletions in *pdf-2* and *pdfr-1* were also reported to show weaker rhythms although, somewhat contradictorily, *pdf-1*;*pdf-2* double mutants were reported to show no entrainment issues, implying an absence of these gene products does not itself abolish rhythmicity (Herrero et al., 2015). These results could imply a conserved role of PDF between *Drosophila* and *C. elegans* in governing the molecular clock, albeit a non-essential one given the lack of effects in double mutants. Finally, temperature-entrained nematode velocity was also found to be dampened in *tax-2* mutants (Winbush et al., 2015), a gene involved in phototaxis, thermotaxis and other sensory responses (Ward et al., 2008; Hedgecock and Russell, 1975; Coburn and Bargman, 1996). These effects were subtle however, and may have reflected lower overall activity than dampened rhythms. Overall, the evidence from these locomotor approaches has yet to provide clarity as to the molecular basis of circadian rhythms, with only subtle or unclear effects arising from the perturbation of the investigated genes. Additionally, loss of function of the homologue of the central conserved TTFL repressor element in *Drosophila* and mammals, *lin-42*, had a relatively non-profound impact on rhythms suggests that it has an involvement in circadian timing, but not a definitive role, and raises the question as to whether the core TTFL could be conserved in *C. elegans*.

Along with activity patterns, evidence of circadian rhythmicity in *C. elegans* has also been reported in other phenotypic responses. Endogenous rhythms have been reported in nematode olfaction, with one publication reporting that during and subsequent to temperature cycles, populations of 100 nematodes showed sinusoidal, circadian patterns in their aversion to a chemorepellent (1-octanol). Nematodes were found to show reduced avoidance during the night and the subjective night (the equivalent time period in constant conditions, see Glossary; Olmedo et al., 2012). Another study found *C. elegans* to show temperature-compensated circadian oscillations in stress resistance; over the course of

entrainment to light:dark cycles and in constant darkness, the proportion of nematodes surviving hyperosmotic stress steadily increased during the day and decreased during the night (Kippert et al., 2002). This result is particularly unique in that it was reported in starved, growth arrested L1 larvae (see Figure 1.3), in contrast to the adults used in most other studies discussed here, and in that cycles persisted for five days in constant conditions, in contrast with the rapid damping seen in the aforementioned nematode velocity experiment (Winbush et al., 2015). In comparison to the locomotor assays, these two approaches are somewhat intensive in terms of the number of nematodes required, and given that they are not designed to work on an individual nematode level (as well as the latter measuring survival or death), may be ineffective for the purposes of dissecting the molecular clock. They do however, highlight the metabolic importance of the clock and beyond activity and rest, suggesting wider diurnal adaptation in nematodes. Chemosensation and stress resistance may require expensive energetic investment, which nematodes may not be able to maintain throughout the day, or perhaps trade-off for metabolic efficiency.

1.3.2 Circadian rhythms in transcription

As illustrated in Figure 1.1, the primary outputs of TTFLs are 24-hour rhythms in transcription. A number of studies in *C. elegans* have successfully reported circadian rhythmicity by examining gene expression directly or through fluorescence and luminescence reporter approaches. Circadian expression rhythms were first reported in a genome-wide microarray study, in which approximately 2% of the *C. elegans* genome was suggested to endogenously oscillate with approximately 24-hour rhythms, following entrainment to either warm:cold or light:dark cycles (van der Linden et al., 2010). In addition, data from this work and other studies have shown select genes to oscillate in quantitative reverse transcription PCR (RT-qPCR) experiments following entrainment to temperature, light or a combination of both (van der Linden et al., 2010; Olmedo et al., 2012; Goya et al., 2016). Uncovering

these reported circadian oscillations in gene expression also led to the first circadian gene expression reporters in *C. elegans*. Firstly, nematodes expressing Green fluorescent protein (GFP) under the promoter of a temperature-entrained gene, *nlp-36*, have been reported to show cycling fluorescence, demonstrating the potential of a non-behavioural method for *in vivo* rhythm detection in small populations of 15-20 nematodes (van der Linden et al., 2010). More recent work showed rhythms in firefly (*Photinus pyralis*) luciferase, expressed under the promoter of the *sur-5* gene, to also report rhythms in nematodes in response both light and temperature (Goya et al., 2016). This approach used larger populations of nematodes (n = 100), but also could report rhythmicity in individuals, and could therefore have considerable potential in high throughput screens to identify circadian mutants. Interestingly, these nematodes were found to only be entrained by light and temperature in reverse orientation to diurnal conditions, with a light phase paired with a cryophase and a dark phase paired with a thermophase (hence referred to as antiphasic entrainment). Given that light would be expected to co-occur with warmth in nature, this offers a somewhat unexpected answer to the question discussed in section 1.2.1 of how *C. elegans* could entrain a clock. There is also some contradiction over temperature entrainment in the literature; studies looking at temperature entrainment alone under similar but varying temporal protocols and analytical approaches have found differing results in the expression patterns of the same genes (Olmedo et al., 2012; van der Linden et al., 2010). These reports of circadian rhythmicity in gene expression and *in vivo* detection systems strongly suggest the presence of a transcriptional clock in nematodes, but also highlight the outstanding questions relating to how *C. elegans* entrains. These ideas form the basis of Chapter 3. Results 1 in this thesis.

Considering the core TTFL itself, research into circadian expression patterns has resulted in no positive findings. Despite the evidence presented in locomotor assays that *lin-42* loss-of-function mutants have an elongated period (Simonetta et al., 2009), *lin-42* mRNA has not been found to show rhythmic oscillations in adult

nematodes in two studies that employed temperature and light entrainment (Olmedo et al., 2012; van der Linden et al., 2010). Other clock gene homologues in *C. elegans* have largely remained unexplored in published literature, but the aforementioned genome-wide array study did not detect rhythmicity in *tim-1*, *aha-1* or *atf-2*, *C. elegans* homologues of core clock genes (Table 1.1) that cycle in mammals, *Drosophila* or both (van der Linden et al., 2010; Pizarro et al., 2013; Yu and Hardin, 2006). The fact that a proportion of the transcriptome has been found to be rhythmically expressed, but the clock homologues do not appear to oscillate in adult nematodes, raises the central question of the extent of functional conservation of the circadian clock and how transcriptional rhythms might be generated in *C. elegans*.

1.3.3 Non-transcriptional molecular rhythms in *C. elegans*

A final note on the molecular basis of circadian rhythms is that they are not exclusively the product of transcriptional clocks. Peroxiredoxins (PRXs) are anti-oxidant enzymes that reduce potentially damaging oxidising intracellular hydrogen peroxide and enter an inactive state before being recycled. This reduction-oxidation recycling has been shown exhibit an approximately 24-hour cycle that is remarkably conserved, being present in *Drosophila*, *Arabidopsis*, *Synechococcus* (photosynthetic prokaryotes)*, Halobacterium salinarum* (an archaeon) and in mammalian tissues (Edgar et al., 2012), notably including anucleate red blood cells (O'Neill and Reddy, 2011). This conservation has been found to extend to nematodes; temperature-entrained *C. elegans* have been reported to show significant, approximately 24-hour rhythms in the level of over- and hyper-oxidised PRX protein (Olmedo et al., 2012). This ultra-conserved rhythm might indicate a pre-TTFL evolutionary origin of circadian rhythms across domains of life which still persists in *C. elegans*, and potentially one that could offer an explanation as to a lack of apparent conservation of clock genes.

1.4 Summary and outline of experiments

In summary, circadian clocks are seemingly ubiquitous among species for which diurnal cycles are a fundamental aspect of life and survival, and show remarkable evolutionary conservation within the metazoa. This is exemplified by TTFLs, highly genetically and mechanistically similar molecular oscillators that persist in the distantly-related *Drosophila*, mammalian and zebrafish models of circadian biology (Figure 1.1). Widely-applicable common experimental approaches of entrainment, followed by free-running measurements enable relatively straightforward study of endogenous clocks in these models. Despite the genetic conservation of the TTFL and common experimental approaches however, the molecular basis of a circadian clock in the widely-used laboratory model *C. elegans* has thus far eluded chronobiology research.

The biology of *C. elegans* differs in many respects to the established circadian models, raising questions as to the evolutionary and ecological potential for a circadian clock to persist in these nematodes. A prevailing factor is that light is the primary zeitgeber in circadian models (Harper et al., 2017; López-Olmeda and Sánchez-Vázquez, 2009; Refinetti, 2010), but *C. elegans* lacks homologues of the cryptochromes and complex photoreceptors that enable other organisms to entrain to it (Romanowski et al., 2014). It is also possible that *C. elegans* occupies an ecological niche where light and other environmental cycles are uncommon and diurnal timing is not required (Sterken et al., 2015), reducing the selective importance of maintaining a clock. Further, proposed roles of *C. elegans* core clock homologues in developmental timing and parallels between the two timing systems could suggest that over the course of *C. elegans* evolution, elements of the circadian system have been adapted to regulate development, shifting away from functions that enable environmental synchronisation (discussed in section 1.2.3). However, none of these factors necessarily preclude *C. elegans* from possessing a circadian clock; *C. elegans* could primarily utilise non-photic zeitgebers to entrain, with

temperature in particular having been shown to effectively synchronise a number of metazoan clocks (Glaser and Stanewsky, 2007; Lahiri et al., 2014). The natural habitat in which *C. elegans* proliferates has also been suggested to be one in which it may frequently be exposed to the same diurnal cycles that led to the evolution of rhythms in other models (Félix and Duveau, 2012). Finally, that aspects of *C. elegans* development are mechanistically comparable, potentially clock-based and require clock gene homologues, may not indicate a reduced capability for circadian timing. Overall, the ecological and evolutionary considerations discussed here do not provide a strong indication as to the presence or absence of a *C. elegans* circadian clock, but rather highlight an array of potential differences that could exist between molecular oscillators in nematodes and other metazoan models.

While the molecular detail of a *C. elegans* clock remains largely unknown, questions as to the presence of one have been answered in a wide variety of literature reporting rhythms in nematodes. Endogenous circadian rhythms have been documented in *C. elegans* locomotor behaviour, in its olfactory responses, its resistance to hyperosmotic stress and in its gene expression through microarray, RT-qPCR and reporter approaches (discussed in section 1.3). Work has identified rhythms using protocols of light or temperature cycles, suggesting both as zeitgebers *C. elegans* can use to entrain, and has also reported examples of temperature compensation and anticipation (Winbush et al., 2015; Kippert et al., 2002; Goya et al., 2016), key hallmarks of circadian rhythms. However, both phenotypic and gene expression experiments have highlighted difficulties in measuring rhythms, in some cases only being able to report rhythms in populations or in proportions of nematodes, finding variation and inconsistencies in phase and reporting fast damping oscillations. Work in *C. elegans* therefore suggests the presence of a clock, but potentially a non-robust one. In terms of the genes involved, these approaches have only yielded partial insights into the molecular basis of circadian timing; loss-of-function mutations in several genes have been shown to perturb rhythms in behaviours and gene expression to varying extents (Simonetta et al., 2009; Herrero

et al., 2015; Winbush et al., 2015; van der Linden et al., 2010; Goya et al., 2016), but no core clock genes have been identified.

The work contained in this thesis focuses on gene expression, the primary output of a transcriptional clock in established models, in trying to improve understanding of the central problem in *C. elegans* chronobiology: the existence of a molecular clock and the degree to which genetic rhythmicity is preserved in this species. Experiments performed as part of this work aimed to address a number of the wide-ranging questions that arise from the literature discussed in this chapter. These include how *C. elegans* entrains, the extent and robustness of its rhythms and the identities of circadian-oscillating genes. Chapter 3. Results 1 details primary experiments exploring circadian entrainment and luciferase reporting of two putative rhythmic genes; *sur-5* and *nlp-36*. The purpose of this work was to build upon previous literature in developing accurate reporting to enable molecular characterisation of the clock. However, these experiments provided limited success in identifying robust circadian rhythms in expression of the two genes and in developing reporter systems. The work that follows in Chapter 4. Results 2 therefore aimed to better characterise circadian gene expression in *C. elegans*, doing so at the level of the whole transcriptome. Results 2 describes the first circadian RNA-sequencing time series in *C. elegans*, with the objectives of identifying and characterising oscillating genes and defining the extent of circadian gene regulation in *C. elegans*. Described in this work are previously unidentified circadian and non-circadian oscillations, an absence of conserved rhythms and wider insights into circadian regulation in nematodes. As a whole, this thesis serves to advance understanding of genetic oscillations and circadian rhythms in *C. elegans*, build a clearer picture of the utility and challenges of the model, and provide direction for the further characterisation of the *C. elegans* clock.

Chapter 2. General Methods and Tables

2.1 Tables: *C. elegans* strains and primers

Table 2.1: Nematode strains used throughout thesis.

Strain name	Genotype	Source	Notes	References
N2	Wild type, laboratory strain	Caenorhabditis Genetics Center (CGC; see notes on strains)	Laboratory strain widely used as basis for most research in *C. elegans.* Isolated Bristol, UK, 1951.	Sterken et al., 2015; Nicholas et al., 1959
PE254	*feIs4 [sur-5p::luciferase::GFP + rol-6(su1006)]* V	CGC	Contains a stable integrated transgene which expresses a luciferase-GFP in-frame fusion protein downstream of the *sur-5* promoter sequence. N2 background.	Lagido et al., 2008
PE255	*feIs5 [sur-5p::luciferase::GFP + rol-6(su1006)]* X	CGC	As above.	Lagido et al., 2008
AB1	Wild type, wild isolate	CGC	Isolated in Adelaide, Australia, 1983.	Caenorhabditis Genetics Center (CGC) (also see http://www.wormbook.org/wli/wbg8.2p52/)
MY9	Wild type, wild isolate	CGC	Isolated in July 2002, Münster, Germany. Before use here, frozen within 5 generations.	CGC
PE254 x AB1	*feIs4 [sur-5p::luciferase::GFP + rol-6(su1006)]* V	Progeny of PE254 and AB1 strains	Generated by crossing male PE254 nematodes into AB1 background.	-
PE254 x MY9	*feIs4 [sur-5p::luciferase::GFP + rol-6(su1006)]* V	Progeny of PE254 and MY9 strains	Generated by crossing male PE254 nematodes into MY9 background.	-
NLIH13 (informal name)	*Ex [nlp-36p::luciferase::GFP + rol-6(su1006)]*	Generated with Ian Hope, University of Leeds	Contains an extrachromosomal array which expresses a luciferase-GFP in-frame fusion protein of the *nlp-36* promoter sequence. N2 background.	-

NLIH2 (informal name)	*Ex [nlp-36p::luciferase::GFP + rol-6(su1006)]*	Generated with Ian Hope, University of Leeds	As above, generated from a different microinjection event.	-
MEG strain (No name assigned by authors)	*Is [sur-5p::luciferase::GFP + rol-6(su1006)]*	Diego Golombek, National University of Quilmes	Contains a stable integrated transgene (spontaneously integrated) which expresses a luciferase-GFP in-frame fusion protein downstream of the *sur-5* promoter sequence. N2 background.	Goya et al., 2016

Notes on strains

N2, PE254, PE255, AB1 and MY9 strains provided by the Caenorhabditis Genetics Center (CGC; https://cgc.umn.edu), which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). PE254 and PE255 express the same *Psur-5::luc::gfp* transgene integrated into the genome at different locations (Lagido et al., 2008). The MEG strain expresses a different *Psur-5::luc::gfp* transgene, spontaneously integrated into the genome. This strain was kindly provided by Diego Golombek (National University of Quilmes). *Pnlp-36::luc+::gfp* strains (NLIH13 and NLIH2) were generated for this project by myself and Professor Ian Hope (University of Leeds). Each strain is derived from single nematode microinjection events of the same *Pnlp-36::luc+::gfp* transgene.

Table 2.2: Primers used for PCR and RT-qPCR.

Usage	Gene	Assigned Name	Forward Sequence	Reverse Sequence
Transgene construction	*nlp-36*	nlp-36	AGTAGGTCTAGAGAAAGGCTGGCATAAG	TATCTGCAGGGGAAACATGAGAATATC
RT-qPCR (reference)	*pmp-3*	Q2 pmp-3	GTTCCCGTGTTCATCACTCAT	ACACCGTCGAGAAGCTGTAGA
RT-qPCR (reference)	*cdc-42*	Q2 cdc-42	CTGCTGGACAGGAAGATTACG	CTCGGACATTCTCGAATGAAG
RT-qPCR	*sur-5*	Q2 sur-5	CACCCCAAGGTTTTGTTCAC	TGAAGGTGTCGGATACAACG
RT-qPCR	*nlp-36*	Q2 nlp-36	GACGATGACGTCACTGCTCT	AGTTTGTGGAATTTGGCGGG

Notes on primers

Primers for established *C. elegans* reference genes taken from published literature (Hoogewijs et al., 2008). *sur-5* primer sequence also taken from prior work for accuracy of comparison (Goya et al., 2016). All primers tested for single product amplification.

2.2 Generic methods

2.2.1 Media composition and preparation

Nematode growth media (NGM) plates

NGM plates (Stiernagle, 2006) were made with 17 g agar, 3 g NaCl, 2.5 g bacto-peptone per L deionised H_2O, which was then autoclaved before aseptic addition of 1 mL 1 M $CaCl_2$, 1 mL 1 M $MgSO_4$, 1 mL 5 mg/mL cholesterol solution (dissolved in 95% ethanol) and 25 mL 1 M KPO_4 buffer (108.3 g KH_2PO_4 and 35.6 g K_2HPO_4 in 1 L H_2O, adjusted to pH 6). Constituent solutions all autoclaved prior to use with the exception of cholesterol. Plates were poured to approximately two thirds capacity of 60 mm or 90 mm diameter petri plates.

LB

LB broth (Bertani, 1951) contained 10 g tryptone, 5 g yeast extract and 10 g NaCl added to 1 L deionised H_2O, subsequently autoclaved. To make LB Plates, 15 g agar was also added and the solution autoclaved then poured aseptically.

S buffer

S Buffer contained 5.85 g NaCl added to a mixture of 129 mL 0.05 M K_2HPO_4 and 871 mL 0.05 M KH_2PO_4, subsequently autoclaved.

NGM plate preparation

Plates for general maintenance were seeded by adding 30-100 µL overnight OP50 strain *Escherichia coli* cultured in LB broth. Where noted, some experimental plates were seeded with 10X concentrated overnight culture by centrifugation and resuspension in LB. Plates were then incubated at 37°C overnight. Plates containing 25 µM 5-Fluorodeoxyuridine (FUDR) utilised in some experiments were twice seeded by adding 10X culture, leaving at room temperature overnight and UV-killing *E. coli*

the following day (using a Stratagene UV Stratalinker 1800 set to 9999 Joules) and repeating one time. Once seeded, all plates were stored at 4°C until needed.

***Escherichia coli* preparation**

E. coli OP50 strain (Stiernagle, 2006; obtained from Caenorhabditis Genetics Center) were grown in 50-150 mL LB (sterilised by autoclaving) cultures, grown at 37°C overnight. LB was either directly inoculated from glycerol stocks directly or by growing glycerol stock on LB plates and picking a colony to inoculate LB. Glycerol stocks were prepared by mixing 0.2-0.3 volumes glycerol (sterilised by autoclave) with 1 mL overnight cultured *E.coli* in LB and freezing at -80°C.

2.2.2 Maintenance and manipulation of nematodes

***C. elegans* husbandry**

Unless otherwise stated, *C. elegans* were maintained at room temperature on 60 mm or 90 mm diameter petri plates containing standard NGM, seeded with OP50 strain *E. coli*.

Transferring to new plates

For general maintenance of strains, nematodes were transferred by chunking; a small NGM square was cut and placed on a fresh seeded plate.

To transfer nematodes to new plates in experiments, plates were washed by adding 2-3 mL of sterile S buffer or H_2O, centrifuging at 13200 RPM for 30 seconds to 1 minute, aspirating to approximately 200 µL, resuspending nematodes, pipetting onto new plates and spreading by gentling rotating the plate.

Freezing Nematode Strains

For long-term storage, all strains were frozen soon after arrival or generation. Plates with large populations of recently starved L1/L2 stage nematodes were washed with S buffer and into 1.5 or 2 mL microcentrifuge tubes and kept on ice for 15 minutes.

Subsequently 1 volume Soft Agar Freezing Solution (0.58 g NaCl, 0.68 g KH_2PO_4, 30 g glycerol, 0.56 mL 1 M NaOH, 0.4 g agar per 100 mL deionised H_2O, autoclaved and kept molten at 50°C until use) was added and thoroughly mixed by pipetting. Samples were then frozen at -80°C.

Egg preparation by bleaching

NGM plates housing gravid nematodes were vigorously washed with 3.5 mL sterile H_2O into 50 mL conical centrifuge tube. Solution was made up to 14 mL with sterile H_2O and 2 mL 5 M NaOH and 4 mL thin bleach (sodium hypochlorite) were subsequently added. Solution was vortexed for a few seconds and allowed to sit for 2 minutes. This was repeated every 2 minutes for 12 minutes, or longer if a large number of nematodes still appeared intact. Tubes were centrifuged for 1 minute at 20°C and solution was aspirated by pipetting to approximately 5 mL and made up to approximately 35 mL with sterile H_2O. This process was then repeated once. Solution was then aspirated to 100 µL x the number of plates required and briefly vortexed. Solution was pipetted onto plates and spread by gentling rotating the plate.

2.2.3 Imaging

Images in Chapter 1, Figure 1.3 were taken using an AxioCam ERc 5s (ZEISS), mounted on a Stemi 508 stereo microscope (ZEISS), and processed using ZEN 2 (blue edition) software (ZEISS).

Chapter 3. Results 1: Exploring Luminescence Reporters as Readouts of the *Caenorhabditis elegans* circadian clock

3.1 Introduction

As discussed in Chapter 1. Introduction, numerous examples of circadian rhythms have been reported in *C. elegans* behaviour, physiology and gene expression, with most having also been presented as potential reporter assays (section 1.3). Despite the availability of these tools, no assay has led to the discovery of a core *C. elegans* clock gene, and no forward genetic screen to identify circadian mutants exists in 20 years of published literature. In addition, studies that have examined circadian gene expression in nematodes have reported rhythms under different entrainment protocols of light, temperature and in combination (van der Linden et al., 2010; Olmedo et al., 2012; Goya et al., 2016), highlighting the important question of the zeitgebers that *C. elegans* could use to optimally entrain its clock (discussed in Introduction, section 1.2). The work described in this chapter aimed to both build on prior work in developing gene expression reporters, and to investigate and optimise entrainment in *C. elegans*.

Effective reporters of circadian genes could be hugely useful to *C. elegans* chronobiology; an accurate *in vivo* reporter with the sensitivity required to accurately delineate circadian and non-circadian or perturbed expression patterns could lead to the genetic characterisation of the *C. elegans* molecular clock. To investigate the potential for accurate *in vivo* reporting of circadian rhythms in *C. elegans*, the work detailed here explored the use of luciferase as a readout of the nematode clock. Luciferase reporters have been used with success in real-time measurements of circadian gene expression in *Arabidopsis* (Southern and Millar, 2005), *Drosophila* (Stanewsky et al., 1997), zebrafish (Vallone et al., 2004), the fungus *Neurospora* (Hurley et al., 2014) and in mammalian cell culture (Yoo et al., 2004). The

work in these models all similarly generated transgenic organisms that express luciferase under the promoter of a core TTFL oscillating clock gene (see Chapter 1, Figure 1.1). Using nematodes in which luciferase expression is driven by the promoter sequences of genes that had previously been reported to show rhythmic expression, the experiments described here examine how effectively the luciferase reporter approach that has proven to be successful across kingdoms could be adapted for use in *C. elegans*.

Prior to the start of this work, no evidence for circadian luciferase reporters in *C. elegans* had been published. However, real-time reporting of circadian rhythms in gene expression was described in *C. elegans* by expressing green fluorescent protein (GFP) under the promoter of the gene *nlp-36* (van der Linden et al., 2010). While not necessarily a core clock gene like the reporters used in the models discussed above, *nlp-36* had been reported to show circadian expression patterns in response to temperature cycles and in subsequent free-running conditions (van der Linden et al., 2010). Luciferase offers a considerable advantage over GFP for *in vivo* circadian study in that it requires no excitation, while GFP fluorescence requires blue light (488 nm peak absorbance in the case of the published nematodes; Cormack et al., 1996). Blue light excitation could interfere with entrainment protocols (Gong et al., 2016) and in the measurement of the endogenous clock in free-running conditions (constant dark), while also making automation and regular *in vivo* recordings challenging. With no excitatory requirement, a higher dynamic range and a catalytic half-life of approximately two hours (Ignowski and Shchaffer, 2004), luciferase could enable higher fidelity, automated circadian reporting of *nlp-36* expression. To build upon prior work and further develop *in vivo* reporting of *nlp-36* expression, I generated transgenic nematodes expressing firefly (*Photinus pyralis*) luciferase under the control of the *nlp-36* promoter (generating the strains NLIH13 and NLIH2).

In addition to these nematodes, circadian luciferase reporting was more recently documented in *C. elegans* using transgenic animals expressing luciferase under the

promoter of the *sur-5* gene (Goya et al., 2016). The authors reported circadian rhythms in luminescence and gene expression in response to entrainment by light and temperature in antiphasic combination (pairing a light phase with a cryophase, and a dark phase with a thermophase; see Glossary). Experiments performed here also utilised this published strain (referred to as the MEG strain here, as it was not named in Goya et al., 2016), in addition to other previously available transgenic lines that express luciferase under the control of the *sur-5* promoter (the PE254 and PE255 strains; Lagido et al., 2008).

In this chapter I describe experiments investigating the luminescence responses of transgenic nematodes to light and temperature cycles and whether combinations of these signals can effectively drive endogenous rhythms. These experiments, using nematodes expressing luciferase under either *nlp-36* or *sur-5* promoter sequences, examine the potential utility of these respective strains for studying the *C. elegans* molecular clock. In further investigating *nlp-36* and *sur-5* as circadian genes and to better characterise and optimise entrainment in *C. elegans*, I also detail expression of these genes, directly measured under different combination protocols of light and temperature through RT-qPCR experiments.

3.2 Results and discussion

In all luciferase reporter experiments, luminescence was measured *in vivo* over time, in nematodes expressing either *Pnlp-36::luc+::gfp* or one of two *Psur-5::luc+::gfp* transgenes (strains listed in Chapter 2, Table 2.1; these nematodes also express GFP as a transgenic marker, but this should not affect the catalytic half-life of luciferase). Luminescence was measured in liquid media in 96-well plates, containing one nematode per well unless noted. Luminescence was read individually from each well using a TopCount NXT scintillation counter (Perkin-Elmer), a well-established approach for circadian measurement in *Arabidopsis* and *Drosophila* (Hargreaves et al., 2019; Glaser and Stanewsky, 2005) and previously used for developmental timing assays in *C. elegans* (Olmedo et al., 2015).

3.2.1 Nematodes expressing luciferase under the promoters of *nlp-36* and *sur-5* show consistent luminescence responses to temperature, but not light cycles

The expression of *nlp-36* has previously been reported to be rhythmic in response to temperature and not to light cycles (van der Linden et al., 2010), while *sur-5* rhythms have been reported under combined light and temperature cycles in antiphase (Goya et al., 2016). Initial experiments using strains expressing *Pnlp-36::luc+::gfp* and *Psur-5::luc+::gfp* transgenes assessed their responsiveness to these environmental signals independently. Luminescence outputs of two *C. elegans* strains, NLIH13 and PE254 (which express *Pnlp-36::luc+::gfp* and *Psur-5::luc+::gfp* transgenes respectively; all strains detailed in Chapter 2, Table 2.1), over the course of 12:12-hour, 20°C:15°C temperature cycles are illustrated in Figure 3.1. Luminescence of PE254-strain nematodes in response to 12:12-hour light:dark cycles are illustrated in Figure 3.2.

An important point of note regarding these experiments is that luminescence recordings from self-fertilising adult hermaphrodites could be prone to interference from developing larvae (of which each adult can generate up to 300; Corsi et al., 2015). The experiments described here used two approaches to resolve these issues: firstly, 5-Fluorodeoxyuridine (FUDR), which inhibits DNA synthesis, was used to prevent hermaphrodite reproduction (Figure 3.1A, 3.2A), and secondly, male nematodes were also used in some experiments (Figure 3.1B-D, 3.2B). *C. elegans* is an androdioecious species in which self-fertilising hermaphrodites are the vastly more commonly occurring sex (> 99.8%; Chasnov and Chow, 2002), and the majority of *C. elegans* research is carried out using them. Work performed in hermaphrodites is therefore more widely-applicable to the species as a whole and for comparisons to prior circadian studies, which also used FUDR (van der Linden et al., 2010; Goya et al., 2016). However, given its potent biological effects, also using males in the absence of FUDR serves to investigate any potential circadian perturbation resulting from FUDR treatment. To generate *Pnlp-36::luc+::gfp* males in sufficient number for analysis, NLIH13 strain hermaphrodites were first crossed to wild-type N2 males, as

NLIH13 males exhibit a roller phenotype (*rol-6*; Higgins and Hirsh, 1970; Mello et al., 1991) that reduces male mating success. The F1 male and hermaphrodite progeny were then used for experiments. This was not necessary for PE254 strain (*Psur-5::luc+::gfp*) nematodes in which male mating was unimpeded.

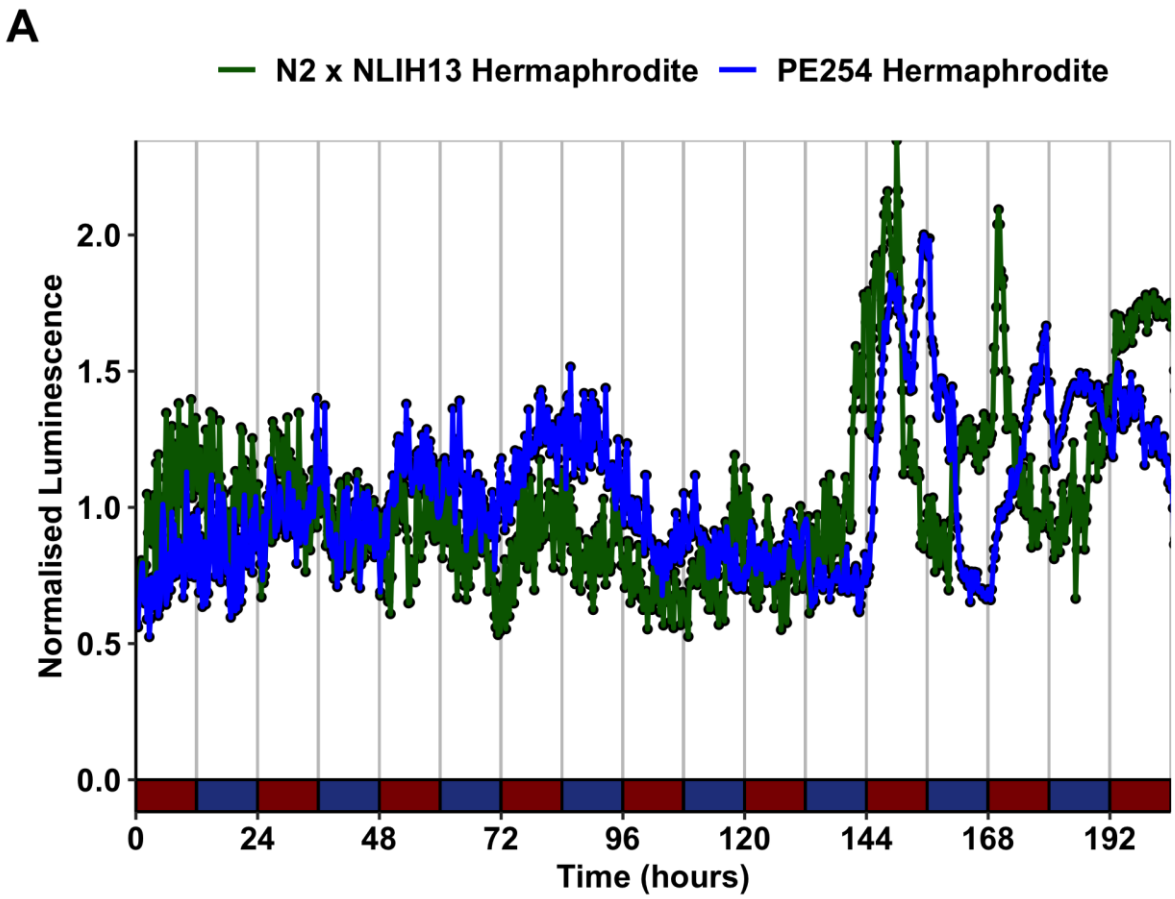
Shown in Figure 3.1A and 3.1B, *Pnlp-36::luc+::gfp* and *Psur-5::luc+::gfp*-expressing hermaphrodites and males showed a clear fluctuation in luminescence in response to temperature cycles, mostly following the daily patterns in temperature changes. In both strains and sexes, oscillations proceed for approximately six days before becoming more erratic. From these plots, no obvious differences are apparent between strains.

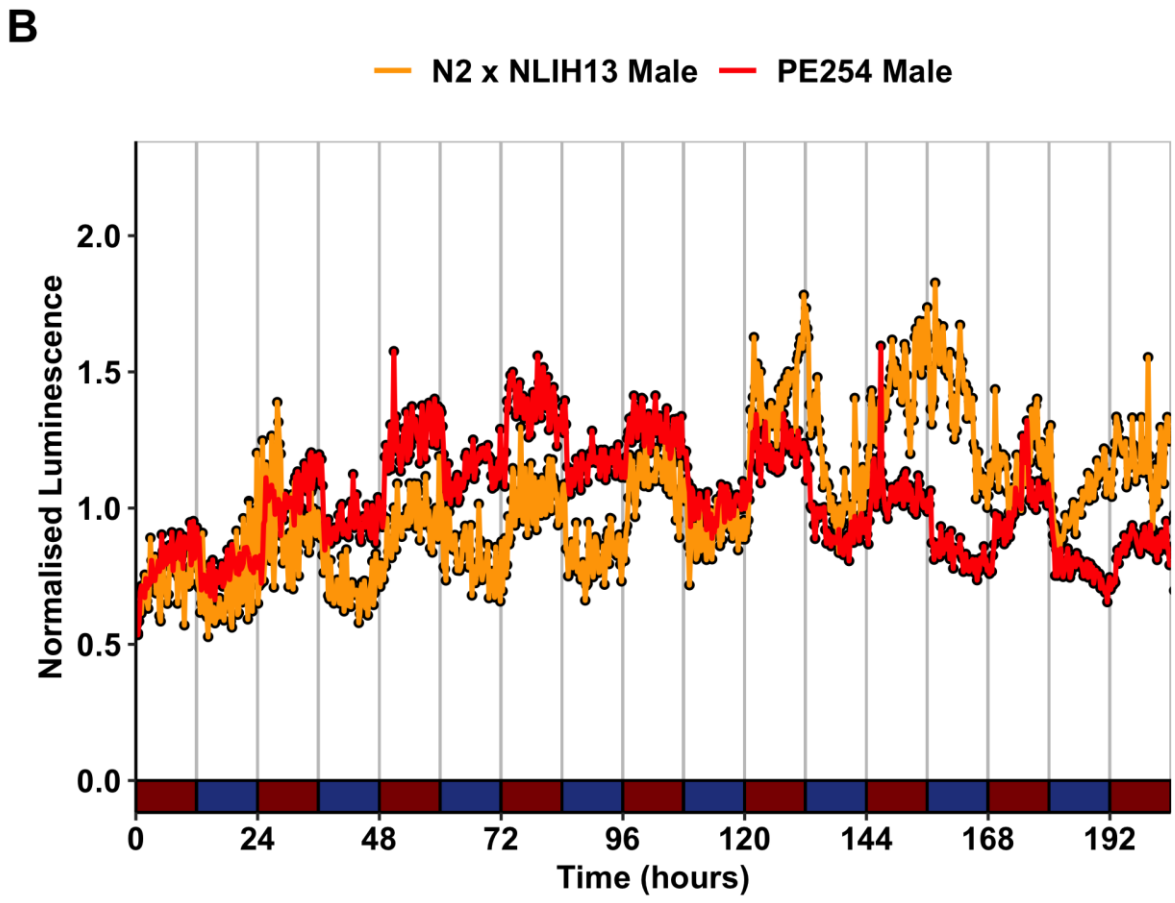
These oscillations could reflect rhythmic gene expression, but could also be a result of other metabolic or behavioural activity of the nematodes, or could perhaps be an artefact of the TopCount apparatus itself. One possibility is that increased luminescence during warm phases reflects an increased catalytic activity of luciferase at higher temperatures. Mammalian experiments have found circadian rhythmicity to dwarf the effects of temperature changes on the luciferase enzyme (Feeney et al., 2016), but this may not be the case in *C. elegans*, an ectotherm. That luminescence oscillations noticeably become more erratic after approximately six days (144 hours) does highlight a probable biological basis for the oscillations. This erraticism does not reflect death, as deceased nematodes were excluded from analyses, but could reflect ageing or an exhausted food (or ATP, required for luciferase activity) supply. While the biological bases of the oscillations observed in Figure 3.1A and 3.1B are not apparent from these experiments, these results do suggest the first six days of adulthood are likely an appropriate time to perform circadian experiments, after which luminescence recordings will be less reliable.

Also notable in Figure 3.1A and 3.1B, is that within the prevailing trend in normalised luminescence following the 12:12-hour cycles, the sampling rate in these experiments (approximately every 15 minutes) reveals higher frequency fluctuations

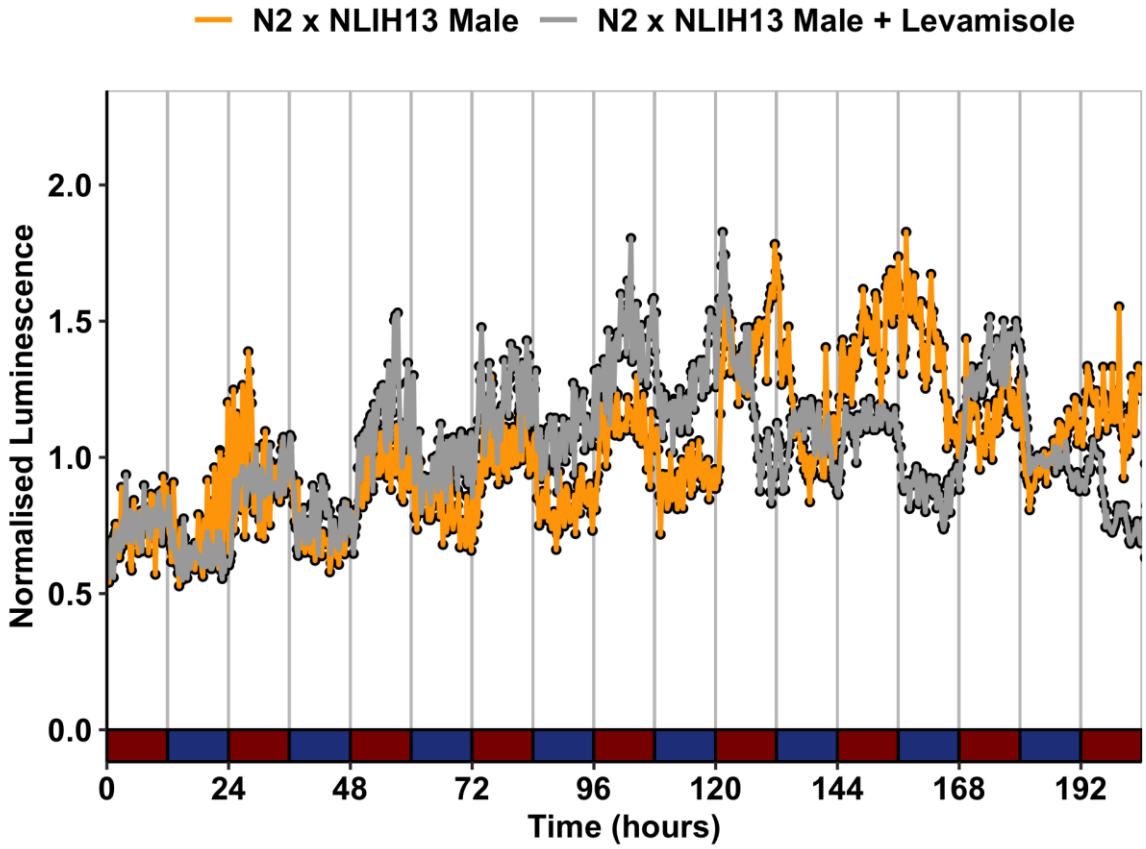
between timepoints. To investigate whether these patterns reflect nematode movement, possibly towards and away from detectors, *C. elegans* were also treated with the cholinergic agonist levamisole (100 µM), which causes nematodes to be immobilised (Figure 3.1C and 3.1D). Levamisole-treated nematodes do not show substantial differences to untreated nematodes by visual inspection, as regards to both 24-hour oscillations and high frequency variation between timepoints, suggesting a limited role of nematode movement in the observed data trends. Rather than exhibiting reduced variation between timepoints, levamisole-treated PE254 nematodes appear more erratic in their luminescence profiles (Figure 3.1D), potentially reflecting increased survival difficulty when movement is prevented.

Collectively, the experiments illustrated in Figure 3.1 suggest a biological responsiveness of *Pnlp-36::luc+::gfp* and *Psur-5::luc+::gfp*-expressing nematodes to temperature. These responses appear to be similar in both genotypes and can be measured consistently for the first six days of adulthood. Levamisole-treatment suggests that the observed trends in these data do not strongly correlate to movement in liquid culture.


A
N2 x NLIH13 Hermaphrodite
PE254 Hermaphrodite
Normalised Luminescence
2.0
1.5
1.0
0.5
0.0
0
24
48
72
96
120
144
168
192
Time (hours)


B
N2 x NLIH13 Male
PE254 Male
Normalised Luminescence
2.0
1.5
1.0
0.5
0.0
0
24
48
72
96
120
144
168
192
Time (hours)

C


N2 x NLIH13 Male
N2 x NLIH13 Male + Levamisole
Normalised Luminescence
2.0
1.5
1.0
0.5
0.0
0
24
48
72
96
120
144
168
192
Time (hours)

D

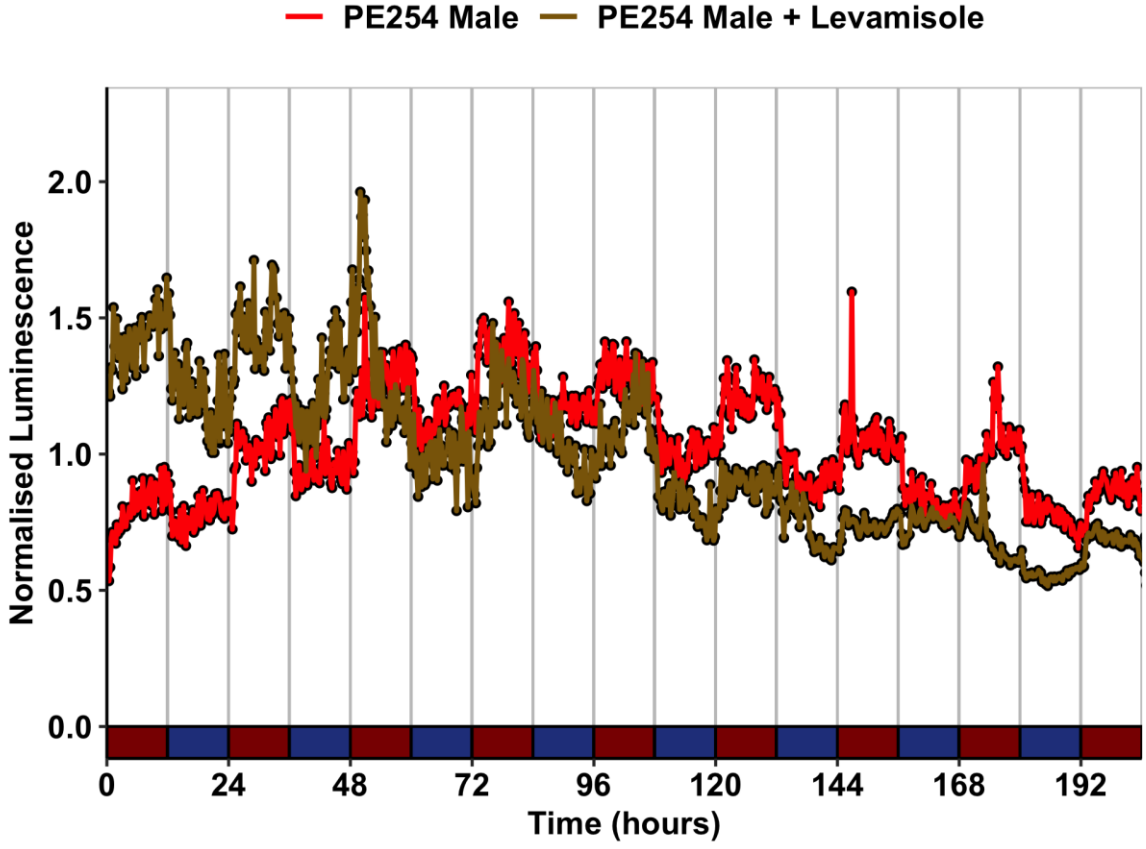

PE254 Male
PE254 Male + Levamisole
Normalised Luminescence
2.0
1.5
1.0
0.5
0.0
0
24
48
72
96
120
144
168
192
Time (hours)

Figure 3.1: Luminescence measurements of reporter nematodes expressing luciferase under the promoters of *nlp-36* (NLIH13) and *sur-5* (PE254) are strongly responsive to changes in temperature for approximately six days of adulthood.
A) Hermaphrodites of the N2 x NLIH13 F1 progeny (n = 12) and PE254 strain (n = 12), containing 50 µM FUDR. **B)** Males of the N2 x NLIH13 F1 progeny (n = 12) and PE254 strain (n = 28), without FUDR. **C)** Comparison of N2 x NLIH13 F1 males in the presence (n = 12) and absence (n = 12) of 100 µM levamisole. **D)** Comparison of males of the PE254 strain in the presence (n = 12) and absence (n = 28) of 100 µM levamisole. Coloured bars indicate 12:12-hour temperature cycles (red = 20°C, blue = 15°C). Constant darkness was maintained throughout. All plots taken from a single 96-well plate experiment in which each well contained an individual late-L4/young adult stage nematode. Nematodes that died (based on luminescence readings indistinguishable from background signal) were removed from analysis.

To investigate the potential requirement of light for gene expression oscillations driven by the *sur-5* promoter, suggested by prior work (Goya et al., 2016), *Psur-5::luc+::gfp*-expressing PE254-strain nematodes were also measured under 12:12-hour light:dark cycles. Nematodes were exposed to red and blue LEDs outside the TopCount apparatus and recorded at a lower sampling frequency (approximately every 40 minutes), as sampling required being measured in darkness. Unlike temperature, light appears to have no clear or consistent impact on the entrainment of nematodes (Figure 3.2). FUDR-untreated males do appear to show some oscillations in luminescence, with an approximately 12-hour cycle (Figure 3.2C), but this pattern lacks repetition before or after, offering little evidence to suggest it is a response to the light:dark cycles. Additionally, some temperature fluctuation took place over the course of this experiment (approximately 3.5°C), which could have interfered with or contributed to these results.

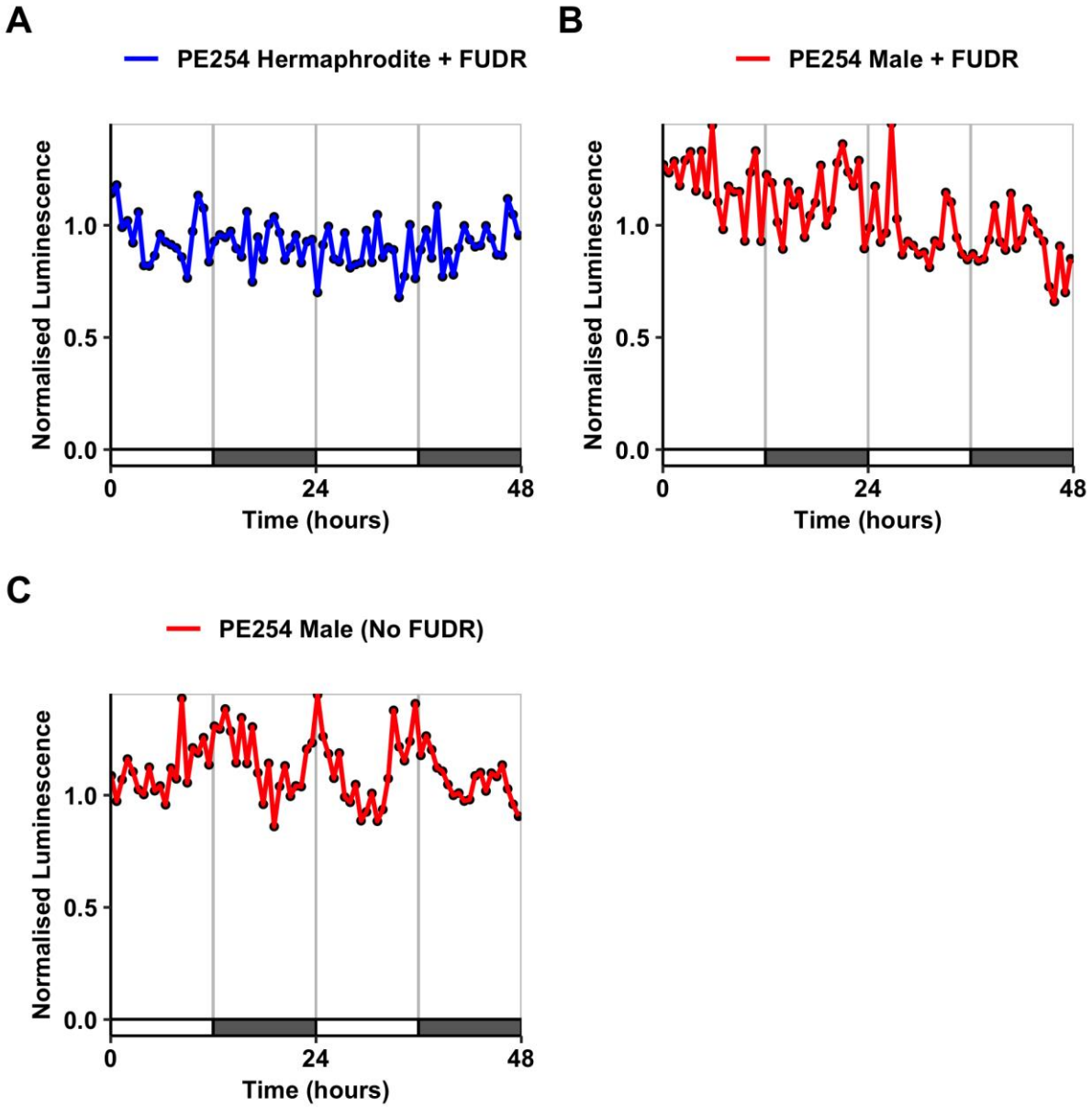


Figure 3.2: Nematodes expressing luciferase under the *sur-5* promoter (PE254) show no clear responsiveness to light entrainment.

A) PE254 strain hermaphrodites, treated with 50 µM FUDR (n = 12). **B)** PE254 strain males, treated with 50 µM FUDR (n = 10). **C)** PE254 strain males, without added FUDR (n = 25). White and grey bars indicate 12:12-hour light cycles (white = light, grey = dark). Temperature was 23.5°C ± 1.75°C throughout. All plots taken from a single 96-well plate experiment in which each well contained an individual late-L4/young adult stage nematode. Nematodes that died (based on luminescence readings indistinguishable from background signal) were removed from analysis.

3.2.2 Luminescence outputs do not provide strong evidence of endogenous oscillations in *nlp-36* and *sur-5* promoter driven luciferase following temperature or antiphasic light and temperature entrainment

Given the increased responsiveness of nematodes to temperature (as well as the capability to measure both entrainment in a more controlled environment), further temperature-entrainment experiments were performed to investigate whether *Pnlp-36::luc+::gfp* and *Psur-5::luc+::gfp*-expressing nematodes also exhibit free-running oscillations. In these experiments adult nematodes were exposed to two 12:12-hour 20°C:15°C cycles before being released into constant conditions (15°C), and luminescence was measured throughout. The luminescence responses to these conditions of two *nlp-36*-reporter strains, NLIH13 and NLIH2 (which both express the same *Pnlp-36::luc+::gfp*-transgene), and males of the *Psur-5::luc+::gfp*-expressing PE254 strain and are illustrated in Figure 3.3.

The primary finding of these experiments, observable in all Figure 3.3 plots, is a lack of overt evidence of robust circadian rhythms in luminescence patterns. Taken together with data shown in Figure 3.1, this suggests that while luminescence can be driven in these nematodes by cycles of temperature, this entrainment protocol is not sufficient to entrain an endogenous clock such to produce continued oscillations in strains expressing either transgene.

From this experiment it is not clear whether the lack of apparent rhythms reflects a lack of circadian gene expression or other biological factors. Notable from Figure 3.3 however is a consistent pattern in luminescence over time that differs between the two promoter constructs; the average luminescence patterns of PE254 males with and without levamisole (Figure 3.3C) display a highly consistent pattern that differs to those of the *Pnlp-36::luc+::gfp* hermaphrodites (Figure 3.3A and 3.3B). Specifically, the latter nematodes show more variation between timepoints, somewhat masking the effect of temperature during entrainment, and show a general upward trend in luminescence in free-running rather than a downward one. This could suggest that

while circadian rhythms are not immediately apparent, that these nematodes are nevertheless accurate reporters of *nlp-36* and *sur-5* expression.

Another apparent trend in Figure 3.3C is a steady increase in PE254-strain luminescence from the start of the experiment. Given that this point marks the first exposure of nematodes to luciferin, this may reflect an initial catalytic depletion of luciferase that has been transcribed and translated prior to the experiment. Through this trend however, is the clear effect of temperature, similar to luminescence responses in Figure 3.1, in which luminescence reporters show fluctuations reflecting the daily cycle.

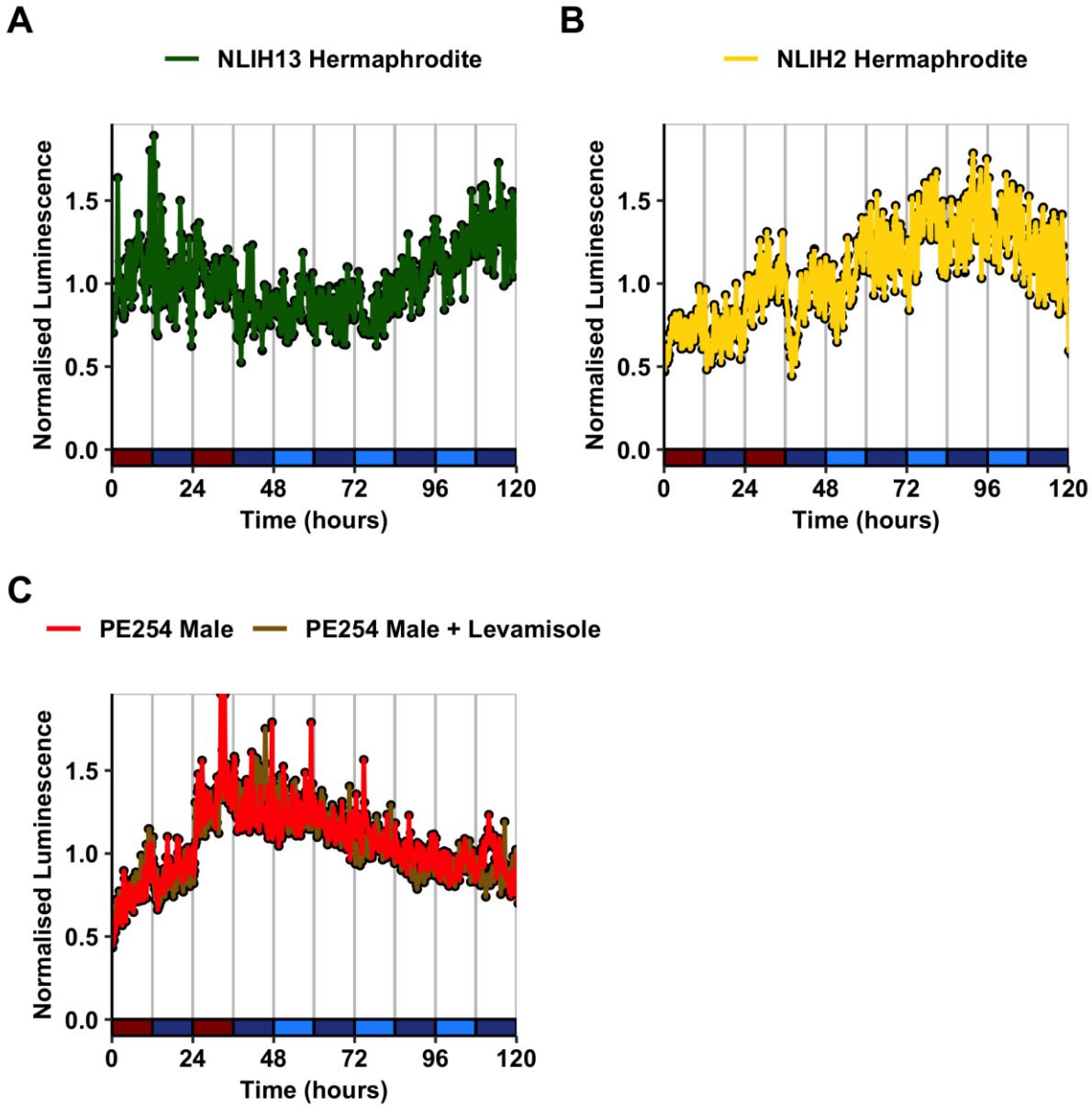


Figure 3.3: *nlp-36* (NLIH13 and NLIH2) and *sur-5* (PE254) luminescence reporters show distinct, but non-circadian patterns in constant conditions following two days of temperature entrainment.

Nematodes were exposed to two 12:12-hour temperature cycles (indicated by coloured bars; red = 20°C, dark blue = 15°C), then released into constant temperature (15°C, light blue bars indicate subjective day). Constant darkness was maintained throughout**. A)** NLIH13 hermaphrodites (n = 12). **B)** NLIH2 hermaphrodites (n = 11). **C)** PE254 strain males in the presence (n = 10) and absence (n = 9) of 100 µM levamisole. All wells containing hermaphrodites were supplemented with 50 µM FUDR. All plots taken from a single 96-well plate experiment in which each well contained an individual late-L4/young adult stage nematode. Nematodes that were not visibly alive or that reproduced were removed from analysis (based on luminescence signal and visual inspection at the conclusion of the experiment).

As discussed in section 3.1, *sur-5* gene expression and *sur-5* promoter-driven luciferase have been reported to show circadian rhythms in response to light and temperature together as dual zeitgebers, but in antiphasic conformation (Goya et al., 2016). To further investigate this approach, TopCount experiments were performed here also using an antiphasic entrainment protocol (12:12-hour cycles of light at 15°C and dark at 20°C) to entrain both the same nematode strain used in the publication, referred to here as the MEG strain (Figure 3.4), as well as the previously used *sur-5* and *nlp-36* reporter strains, PE254, NLIH13 and NLIH2 (Figure 3.5; strains fully detailed in Chapter 2, Table 2.1). In these experiments, entrainment within the TopCount was not possible, so organisms were pre-entrained in a growth chamber and measured only subsequent to entrainment in free-running conditions of constant dark and constant warmth.

Firstly considering the previously published MEG-strain nematodes, the luminescence data presented a major issue in a weakness of signal. Figure 3.4A shows a normalised average of 11 MEG-strain nematodes. It is possible that a circadian oscillating pattern may be visible in these nematodes, with average luminescence increasing from the start of the subjective light/cold phase before decreasing again over a 24-hour period. However, in these luminescence assays, nematodes with average luminescence < 40 counts per second (CPS) were excluded as potentially reflecting background signal (in accordance with the standard TopCount approach in plant biology). Only 17.5% of nematodes that did not reproduce or die during the experiment met this threshold (in contrast to all other strains referenced in this chapter, in which individual nematodes could easily be differentiated from background noise). Further, as shown in Figure 3.4B, individuals with > 40 average CPS often drop below this threshold, sometimes hitting zero, meaning patterns in the average data reflect undetectable luminescence at times. As such, the accurate reporting of potential sinusoidal rhythmic patterns using individuals from this nematode strain may not be achievable using the TopCount system employed here.

While the prior work describing the MEG-strain nematodes did report rhythms in individuals as a proof of principle, it largely used populations of nematodes (n = 100), and also reported inconsistencies, with over half of entrained populations not showing rhythms in free-running (Goya et al., 2016). Results in these experiments may be consistent with this finding in some respects. Shown in Figure 3.4C-F, using populations of 50 nematodes per well in a 96-well plate did ameliorate the issue of low detectability, and comparable to the results in Figure 3.3, largely revealed consistent strain-specific patterns of luminescence. These nematode populations, to varying extents, show an increase in luminescence towards the start of the first day of recording, before decreasing and damping out. This could potentially indicate endogenously-driven rhythmicity. However, the amplitude of this pattern is much greater in Figure 3.4D than 3.4C, E and F, and when these four populations are averaged (shown in Figure 3.4G), while the increased activity in the first day of recording is preserved, the overall oscillation is largely masked. These singular occurrence, faint cycles that are more detectable in some populations than others, could be recapitulating the findings of the original work in detecting rhythms in a proportion of nematodes (Goya et al., 2016). The implication of this might be that these results reflect the extent of clock regulation of *sur-5* and ultimately point to a non-robust, rapidly damping endogenous clock in *C. elegans*. However, the evidence of rhythms provided here is very limited; a second free-running oscillation would make a strong case for *sur-5* being a rhythmic gene (which was reported in the prior work), but rhythms here only persist for one day before completely damping out. Therefore, these results do not entirely contradict previous data in identifying rhythms in MEG-strain nematodes, but the rapidly damping observed luminescence patterns do not fully recapitulate the prior work or provide strong evidence of clock regulation, and could represent another biological process.

One possible confounding factor in these experiments, which could have contributed to the spiking in luminescence patterns in Figure 3.4C-F, is that hermaphrodite reproduction may have occurred and influenced expression patterns. While FUDR

does prevent reproduction in most nematodes, a minority of individuals in these experiments can reproduce, possibly due to having fertilised gametes prior to FUDR application. Wells containing reproducing nematodes are typically excluded from analysis in the experiments throughout this chapter by scoring fluorescence by eye (detailed in Chapter 3 methods, section 3.4), or if a rapid increase in luminescence is observed. However, the dimness of both fluorescence and luminescence in the MEG-strain nematodes made accurate detection improbable.

Overall, antiphasic entrainment of the MEG strain could be indicative of circadian rhythms, but the variability and rapid damping of luminescence patterns, along with the technical detection issues in this strain, both in the requirement of large populations and the accurate scoring of reproduction, make it difficult to draw firm conclusions from these data. Looking forward to further study of circadian rhythms, these issues would make it difficult to measure perturbed rhythms or score mutants and perform forward genetics approaches, suggesting a limited utility of this strain for understanding the *C. elegans* clock, at least using the TopCount method applied here.

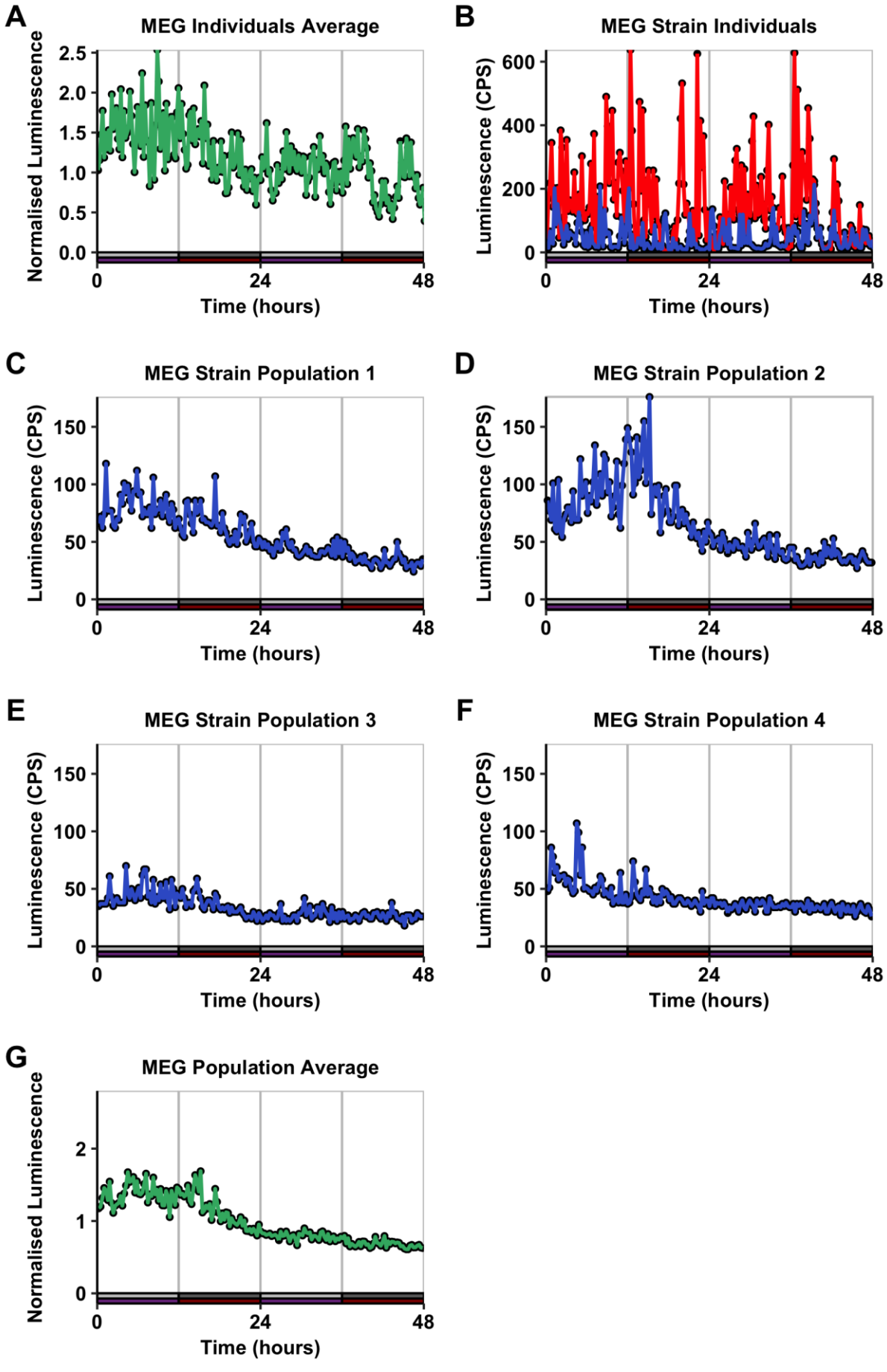

A
MEG Individuals Average
Normalised Luminescence
2.5
2.0
1.5
1.0
0.5
0.0
0
24
48
Time (hours)
B
MEG Strain Individuals
Luminescence (CPS)
600
400
200
0
0
24
48
Time (hours)
C
MEG Strain Population 1
Luminescence (CPS)
150
100
50
0
0
24
48
Time (hours)
D
MEG Strain Population 2
Luminescence (CPS)
150
100
50
0
0
24
48
Time (hours)
E
MEG Strain Population 3
Luminescence (CPS)
150
100
50
0
0
24
48
Time (hours)
F
MEG Strain Population 4
Luminescence (CPS)
150
100
50
0
0
24
48
Time (hours)
G
MEG Population Average
Normalised Luminescence
2
1
0
0
24
48
Time (hours)

Figure 3.4: Luminescence measurements of MEG-strain (*Psur-5::luc::gfp*) nematodes following antiphasic light and temperature entrainment.
Data from two separate experiments showing luminescence over two days in constant darkness at 20°C, subsequent to entrainment by three 12:12-hour cycles of light at 15°C and dark at 20°C prior to recording. **A)** Average normalised luminescence from wells containing individual nematodes (n = 11). **B)** Raw luminescence of two representative individuals from **A**, highlighting frequent troughs of undetectable luminescence. **C-F)** Raw luminescence data from individual wells containing populations of 50 individuals, measured in counts per second (CPS). **G)** Average normalised luminescence of C-F. Light grey and purple bars indicate the subjective light and cryophase, while dark grey and red indicate the subjective dark and thermophase. Plots taken from two separate single 96-well plate experiments. Recordings follow 12 hours at 20°C in constant darkness in the TopCount. Wells all supplemented with 50 µM FUDR.

Identical experiments using antiphasic entrainment were also performed in the previously used *Psur-5::luc+::gfp* and *Pnlp-36::luc+::gfp*-expressing nematodes (Figure 3.5) to compare with the MEG-strain. In contrast to the MEG-strain luminescence profiles in Figure 3.4, the *Psur-5::luc+::gfp*-expressing PE254-strain males show nearly flat expression profile, seemingly not showing any response to the antiphasic entrainment protocol.

As to why these differences are observed between strains, possible explanations could relate to the fact that the MEG and PE254 strains express different, independently generated *Psur-5::luc+::gfp* transgenes, inserted into the *C. elegans* genome at different chromosomal locations. As regards to chromosomal location, the MEG and PE254 strains both express the transgenes integrated into the genome by random insertion (Goya et al., 2016; Lagido et al., 2008), and as such could be under the influence of different proximal regulatory genetic elements. Regarding the different transgenes themselves, luciferase could similarly be affected by genetic regulatory elements that relate directly to the normal *sur-5* expression; the primary difference between the transgenes is that the MEG strain was generated by using 1052 base pairs (bp) upstream of the *sur-5* start codon to drive luciferase expression, while the PE254 strain expresses luciferase with 3700 bp of upstream sequence. The 2648 bp of additional sequence in the PE254 strain could contain important

regulatory elements that define accurate *sur-5* expression. Whether the inter-strain differences relate to the promoter sequence or the site of insertion, the implication of differing luminescence profiles in Figure 3.4 and Figure 3.5 would be that the luminescence profile of at least one of these strains does not reflect true, wild-type *sur-5* expression patterns.

The three strains included in Figure 3.5 also unambiguously show no evidence of circadian rhythmicity. The two *Pnlp-36::luc+::gfp*-expressing strains, NLIH13 and NLIH2 do however show a consistent pattern of changing expression. Both strains (generated from independent microinjection events of the same transgene) have luminescence profiles that gradually increase over time before peaking at 72 hours and starting to decline. Both strains also show increasing fluctuations between timepoints, becoming much more variable throughout the time series. Given that Figure 3.5 represents very small numbers of nematodes (n = 5 and 7) and only includes those did not die or reproduce, these results do suggest accuracy and consistency in the TopCount approach in recording luminescence over time. The patterns of expression observed in Figure 3.5 do also resemble those in Figures 3.3A and 3.3B, and as such may not be a result of the specific entrainment conditions, but rather could reflect *nlp-36* expression over time in ageing nematodes. The NLIH13 and NLIH2 strains therefore, if not of use for circadian study, could potentially offer insights into *nlp-36* expression over time.

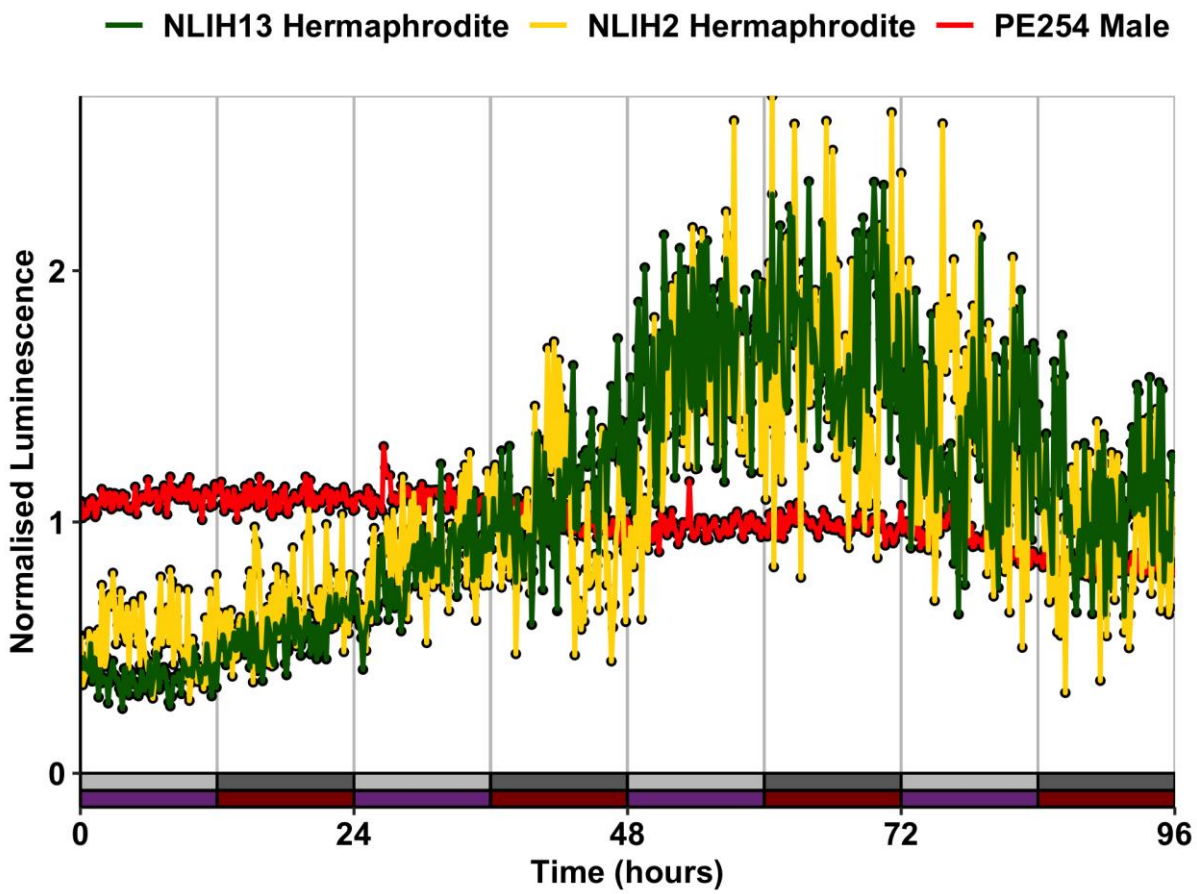


Figure 3.5: Luminescence measurements of NLIH13 (*Pnlp-36::luc::gfp*), NLIH2 (*Pnlp-36::luc::gfp*) and PE254 (*Psur-5::luc::gfp*) strains following antiphasic entrainment. NLIH13 (n = 5), NLIH2 (n = 7) and PE254 (n = 49) nematodes were entrained by three 12:12-hour cycles of light at 15°C and dark at 20°C in a growth chamber before being recorded in a TopCount in constant darkness at 20°C. Light grey and purple bars indicate the subjective light and cryophase, while dark grey and red indicate the subjective dark and thermophase. Plots taken from single 96-well plate experiment in which each well contained an individual late-L4/young adult stage nematode. Recordings follow 12 hours at 20°C in constant darkness in the TopCount. Wells containing hermaphrodites were supplemented with 50 µM FUDR. Nematodes that were not visibly alive or that reproduced were removed from analysis (based on luminescence signal and visual inspection at the conclusion of the experiment).

3.2.3 Expressing *Psur-5::luc+::gfp* in different genetic backgrounds has limited effects on temperature-entrained luminescence patterns

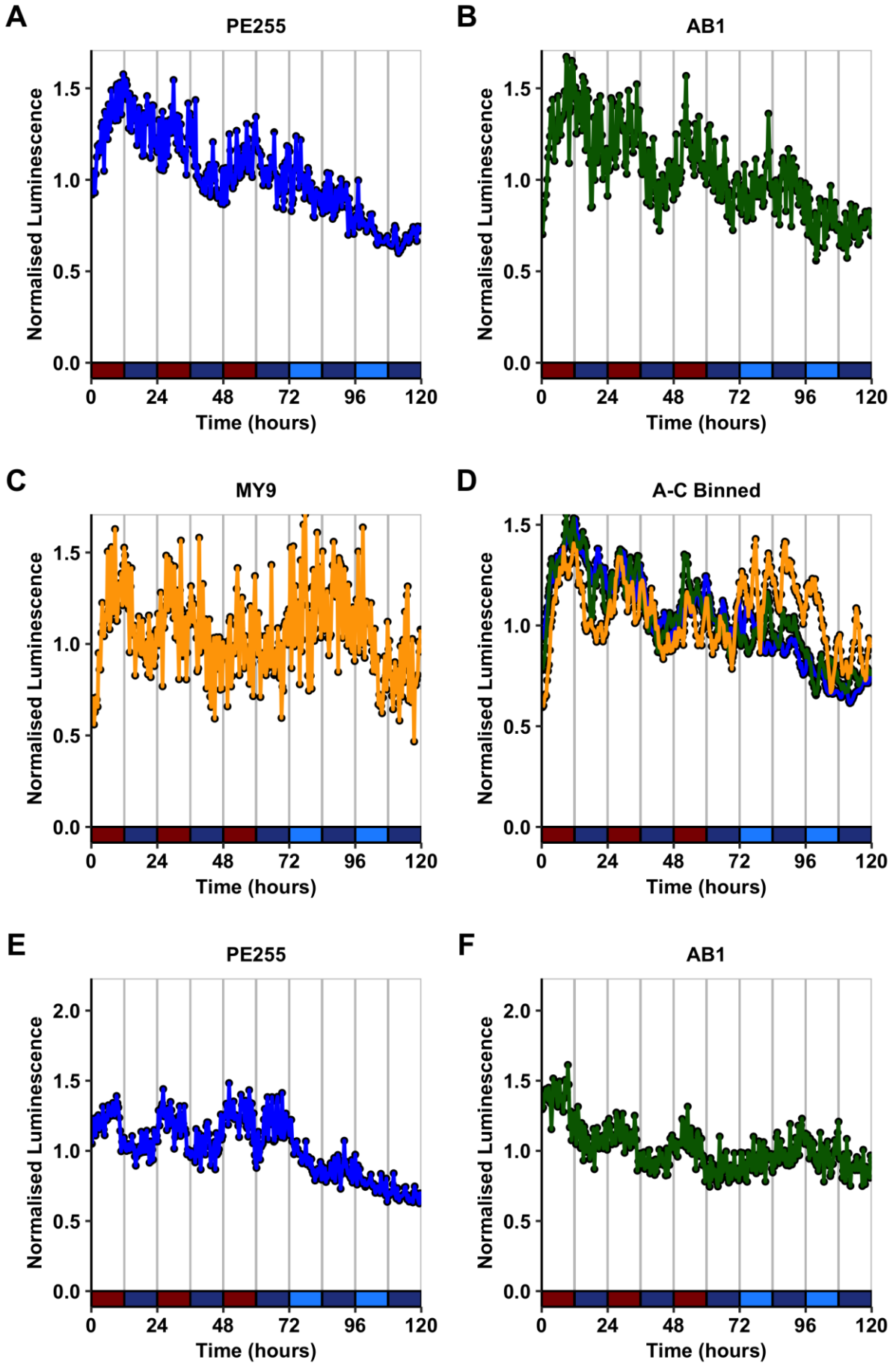
The vast majority of *C. elegans* research is carried out in the N2 strain background, which since its isolation in 1951 has been maintained on NGM plates (See General

Methods, section 2.2.1) in laboratories or in freezers (Sterken et al., 2015; Nicholas et al., 1959). One further possibility investigated through luminescence reporter approaches here was whether laboratory adaptation might have diminished or affected a circadian clock in the N2 strain. To answer this question, the *Psur-5::luc+::gfp* transgene expressed in PE255-strain nematodes (a different strain that expresses the same integrated transgene as PE254, both of which were generated from an N2 laboratory strain-background; Lagido et al., 2008, see Table 2.1, strains used) was introduced into two different strain backgrounds. PE255 males were crossed with wild isolate hermaphrodites collected from two different geographical locations: Germany (MY9 strain) and Australia (AB1 strain), with the former having been frozen within five generations of collection, prior to use in these experiments (Table 2.1). Transgenic male progeny were then crossed with wild-isolate hermaphrodites for at least eight generations, selecting for transgenic progeny each time, to express the *Psur-5::luc+::gfp* transgene in the wild-isolate genetic background.

Experiments using these nematodes measured responses to temperature cycles (as these measurements can be performed during entrainment and free-running phases and temperature has the most pronounced effect on luminescence; Figure 3.1). Figure 3.6 shows results from two replicate temperature entrainment experiments, showing normalised average luminescence of nematodes from the original N2 background and AB1 and MY9 wild-isolate backgrounds.

The primary result of these experiments is that wild-isolate strains did not reveal overt evidence of circadian rhythmicity in free-running. This suggests that the lack of endogenous rhythms in *Psur-5::luc+::gfp*-expressing nematodes does not relate to the laboratory N2 background having suppressed rhythms relative to those found in nature. However, these experiments could suggest some impact of strain background on luminescence patterns in that across two biological replicates (Figure 3.6A-D and Figure 3.6E-H) the three backgrounds show consistent within-strain

expression patterns and somewhat divergent between-strain expression; during temperature cycles all three strains appear to follow the diurnal cycles, but become more distinct once released into constant conditions (exemplified in Figure 3.6D and 3.6H). The German MY9 strain also shows higher amplitude fluctuations than the AB1 and N2 backgrounds between timepoints throughout the time series, along with a broader increase once released into free-running conditions. It should therefore be concluded from these experiments, that while the N2 laboratory background has not diminished any obvious circadian expression that might be seen using this *Psur-5::luc+::gfp* transgene, that strain background could possibly affect future luminescence reporter experiments, perhaps using other potentially circadian-regulated transgenes.


A
PE255
B
AB1
C
MY9
D
A-C Binned
E
PE255
F
AB1
Normalised Luminescence
Time (hours)
0.0
0.5
1.0
1.5
2.0
0
24
48
72
96
120

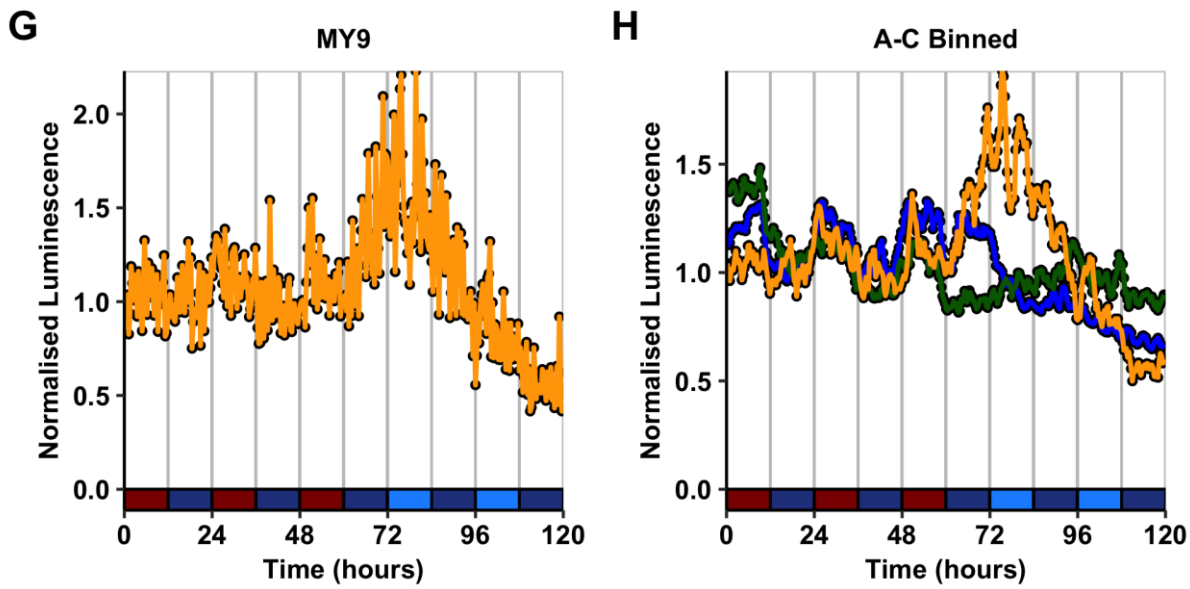


Figure 3.6: *sur-5* luminescence reporters expressed in different strains show differences in luminescence under temperature entrainment and free-running conditions.
Nematodes were exposed to three 12:12-hour temperature cycles (indicated by coloured bars; red = 20°C, dark blue = 15°C), then released into constant conditions (15°C, light blue bars indicate subjective day). Constant darkness was maintained throughout**. A)** Hermaphrodites of the PE255 strain (N2 background, n = 17). **B)** Hermaphrodites expressing the PE255 transgene in the AB1 background (n = 29). **C)** Hermaphrodites expressing the PE255 transgene in the MY9 background (n = 26). **D)** Comparison of A-C, with binned sequential five-point averages for clarity of comparison and noise reduction. **E-H)** Biological replicates of A-D (PE255: n = 20, AB1: n = 24, MY9: n = 22). All plots in A-D and E-H taken from respective single 96-well plate experiments in which each well contained an individual late-L4/young adult stage nematode. All wells supplemented with 50 µM FUDR. Nematodes that were not visibly alive or that reproduced were removed from analysis (based on luminescence signal and visual inspection at the conclusion of the experiment).

3.2.4 Luminescence reporting in nematodes using pharmacological inhibitors of circadian rhythms

Along with strain background, any circadian basis to the luminescence profiles of nematodes was further explored by treatment with known pharmacological inhibitors of circadian rhythms. Pharmacological perturbation of luminescence profiles was investigated using the two *Psur-5::luc+::gfp* transgene-expressing strains, PE254 and PE255, treated with two pharmacological agents: PF-670462 and SB203580. PF-670462 inhibits *Casein kinase 1ε* and *Casein kinase 1δ* in mammals

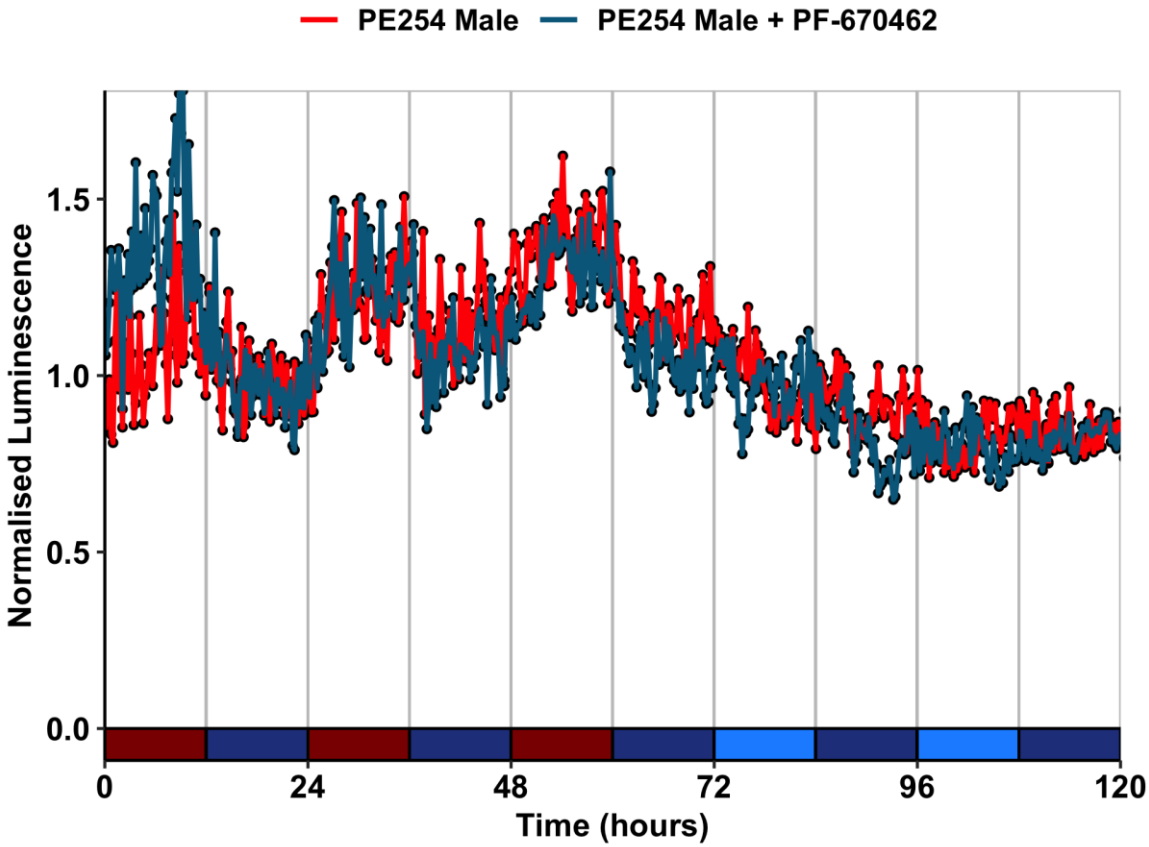
(notable as a core TTFL accessory kinase, homologous to *Drosophila doubletime* and *C. elegans kin-20*, see Chapter 1, Table 1.1) and has been shown to lengthen circadian periods in locomotor and transcriptional rhythms (Meng et al., 2010). In *C. elegans*, treatment with PF-670462 has notably been shown to lengthen the circadian period of *sur-5* promoter-driven luminescence rhythms (Goya et al., 2016). SB203580 is an inhibitor of p38 MAP kinases (p38Ks), which are expressed in clock neurons in *Drosophila*. Flies deficient for p38K have been shown to have aberrant free-running locomotor rhythms (Dusik et al., 2014). SB203580 has also been shown to lengthen circadian rhythms in melatonin release and expression of *Bmal1* (a core mammalian clock gene; Chapter 1, Figure 1.1) in mammalian cells (Hayashi et al., 2003; Kon et al., 2015).

Luminescence experiments exploring the effects of PF-670462 and SB203580 on *C. elegans* under temperature entrainment and free-running conditions are shown in Figure 3.7. Collectively, these experiments again reveal a remarkable within-experiment similarity when comparing nematodes of the same genotype and sex; those in the presence and absence of inhibitors show very similar luminescence profiles despite being an average of often relatively few individual nematodes ($n \geq 5$), providing the best example of the intra-strain consistency of these luciferase reporters of gene expression discussed throughout this chapter. The obvious similarities between nematodes treated and untreated with pharmacological inhibitors of circadian rhythms would appear to imply a lack of involvement of a circadian clock in the observed entrainment patterns presented throughout this chapter (although insufficient dosage and the possibility that these inhibitors do not affect a *C. elegans* clock would provide equally logical alternative explanations).

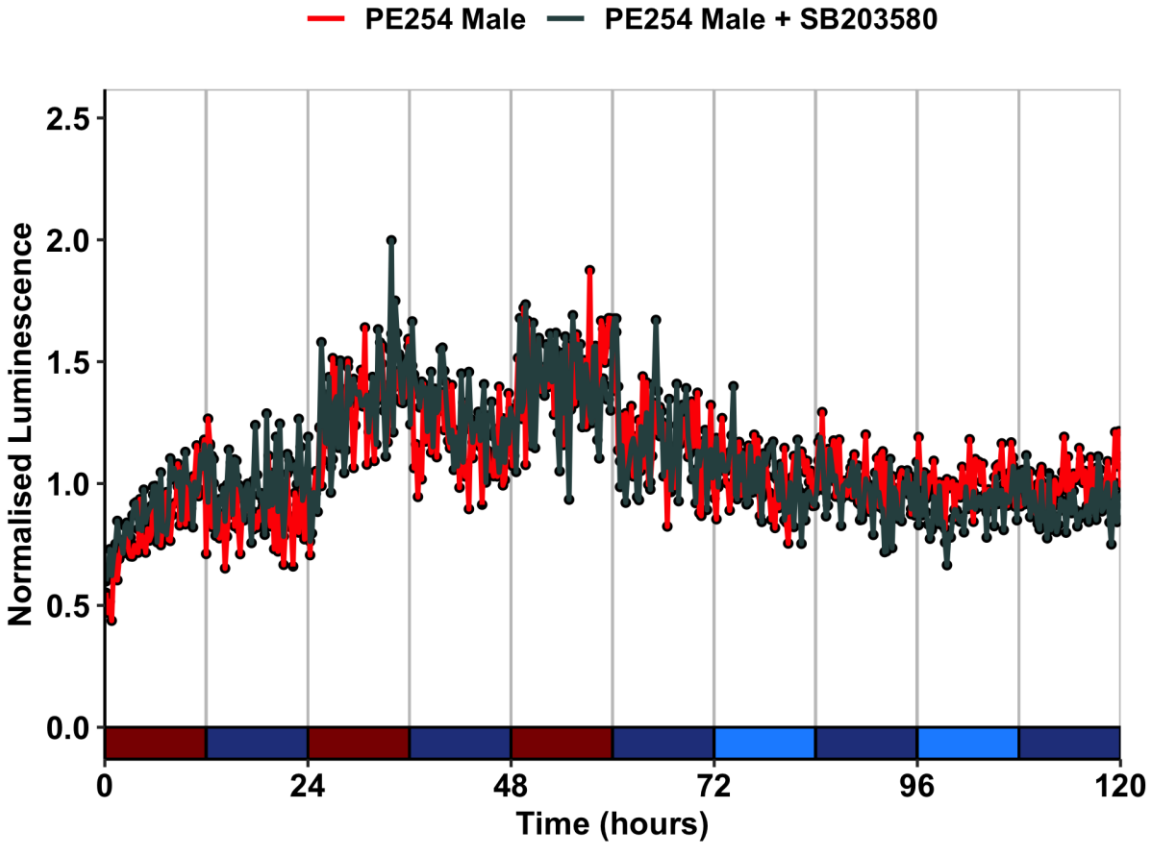
Despite the visually stark similarities observed in Figure 3.7 however, subtle effects of the pharmacological inhibitors on rhythmic observations may be evident. Wavelet spectral analysis of the data in Figure 3.7F, performed by Jessica Hargreaves, suggested that while the temporal profiles of the treated and untreated males likely

do not differ, the intensity of the profiles do. These data have since been accepted for publication (Hargreaves et al., 2019). As such, while the presence or influence of a circadian clock is not clear from these studies, reported inhibitors of the clock in *C. elegans* and other models (Goya et al., 2016; Meng et al., 2010; Dusik et al., 2014; Kon et al., 2015) do result in a statistically detectable effect on the nematode oscillations presented here, which certainly leaves the potential for a nematode clock to play a role in the observed luminescence patterns.

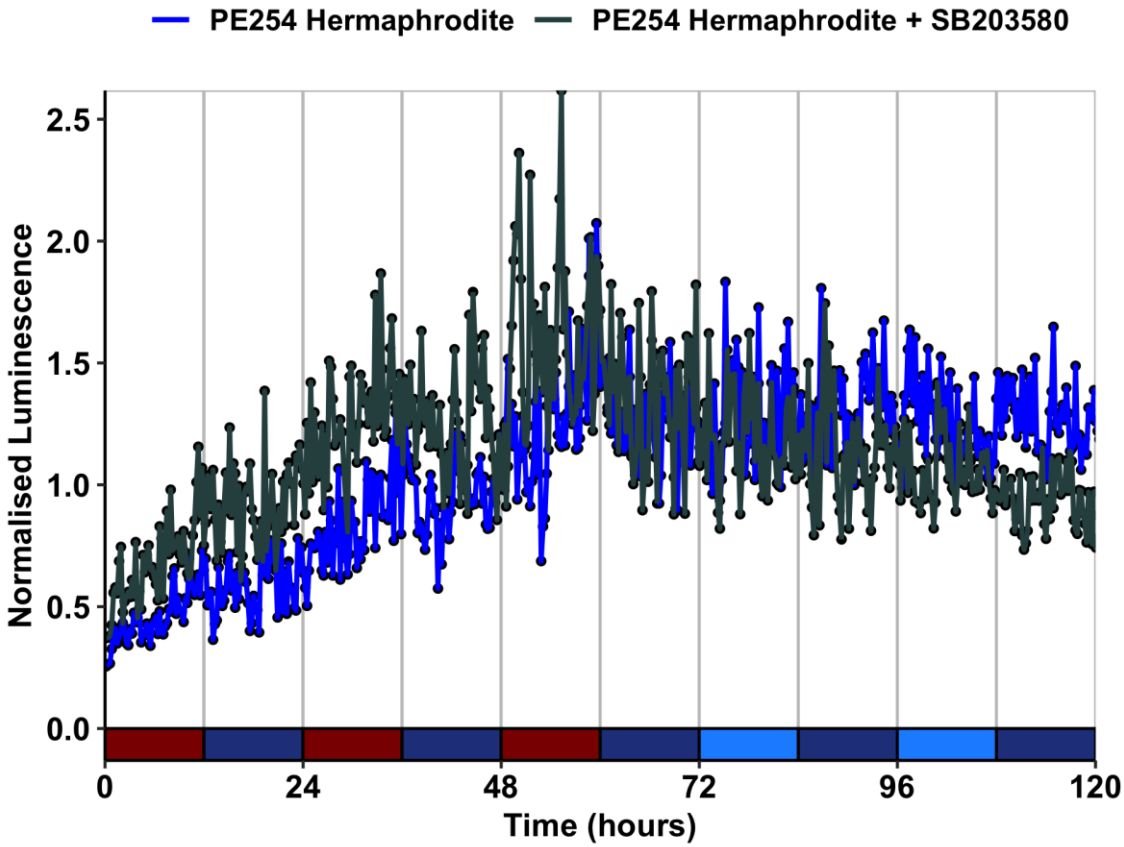
A


PE254 Male
PE254 Male + PF-670462
Normalised Luminescence
0.0
0.5
1.0
1.5
0
24
48
72
96
120
Time (hours)

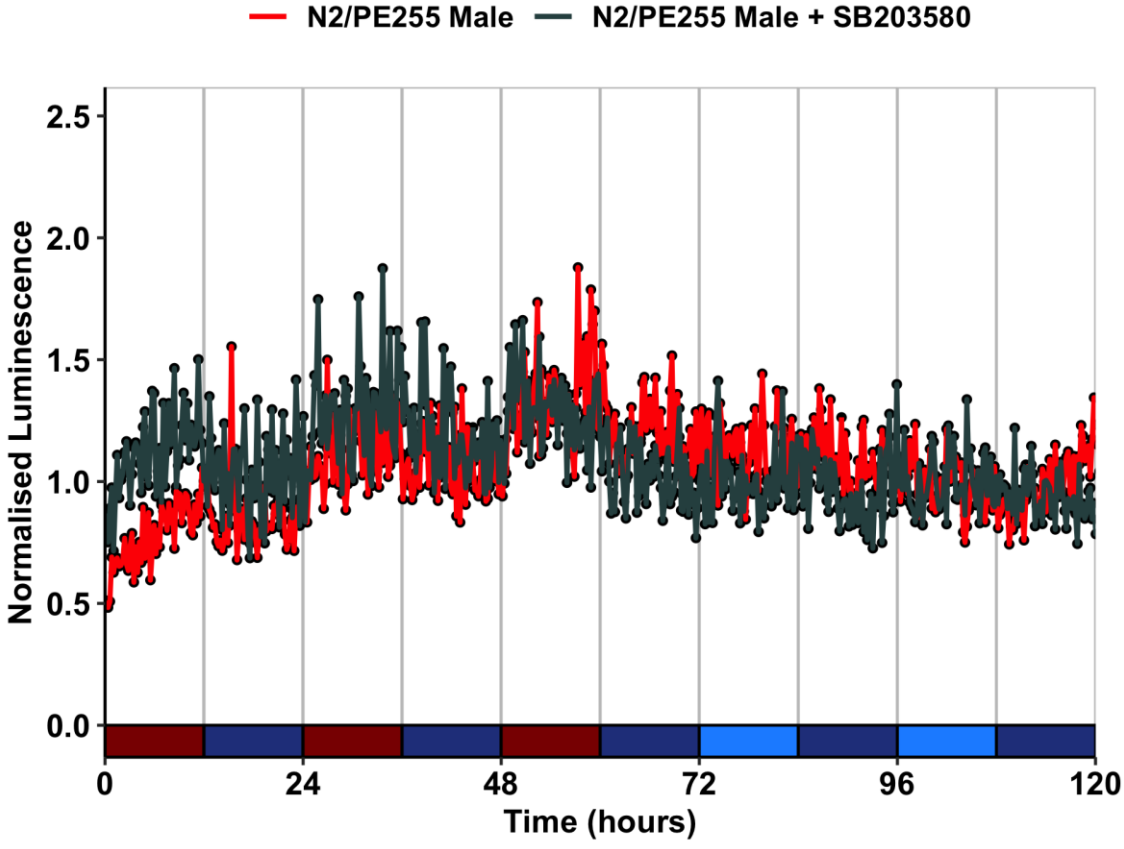
Bi


PE254 Male
PE254 Male + SB203580
Normalised Luminescence
0.0
0.5
1.0
1.5
2.0
2.5
0
24
48
72
96
120
Time (hours)

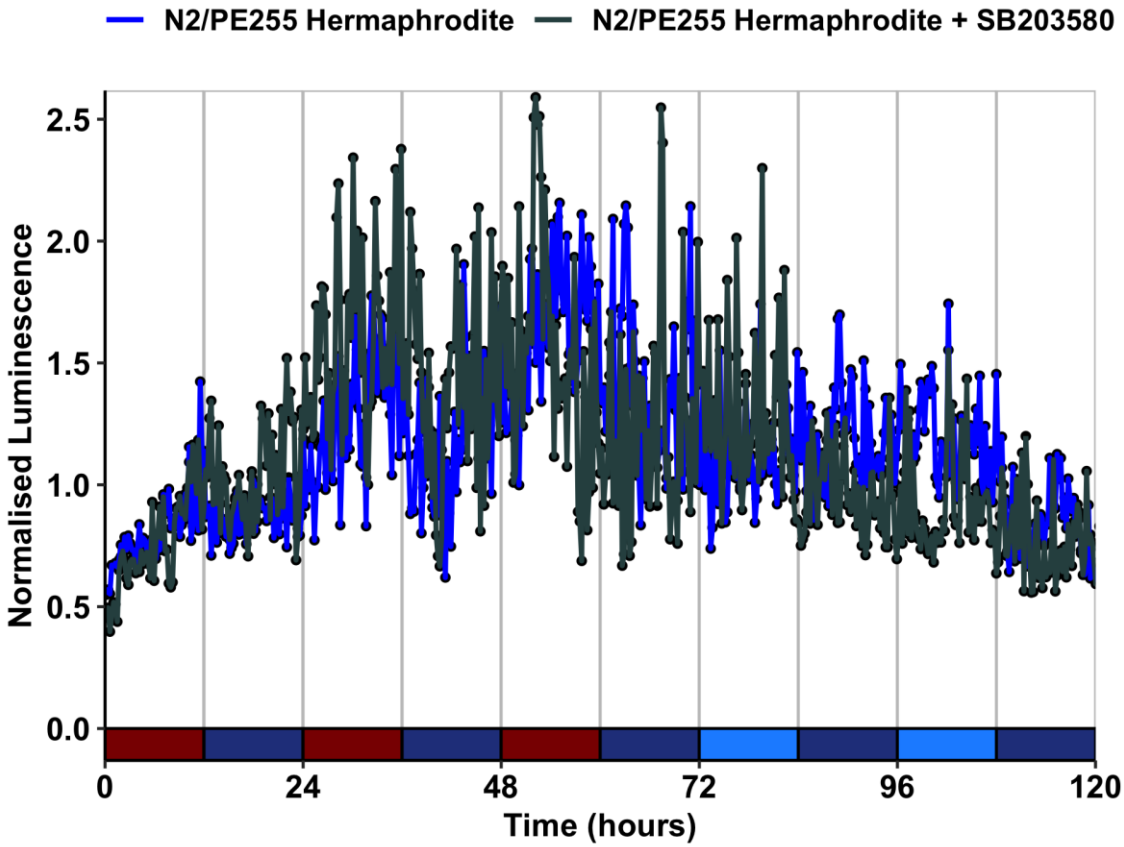
Bii


PE254 Hermaphrodite
PE254 Hermaphrodite + SB203580
Normalised Luminescence
2.5
2.0
1.5
1.0
0.5
0.0
0
24
48
72
96
120
Time (hours)

Biii


N2/PE255 Male
N2/PE255 Male + SB203580
Normalised Luminescence
2.5
2.0
1.5
1.0
0.5
0.0
0
24
48
72
96
120
Time (hours)

Biv


N2/PE255 Hermaphrodite
N2/PE255 Hermaphrodite + SB203580
Normalised Luminescence
2.5
2.0
1.5
1.0
0.5
0.0
0
24
48
72
96
120
Time (hours)

C

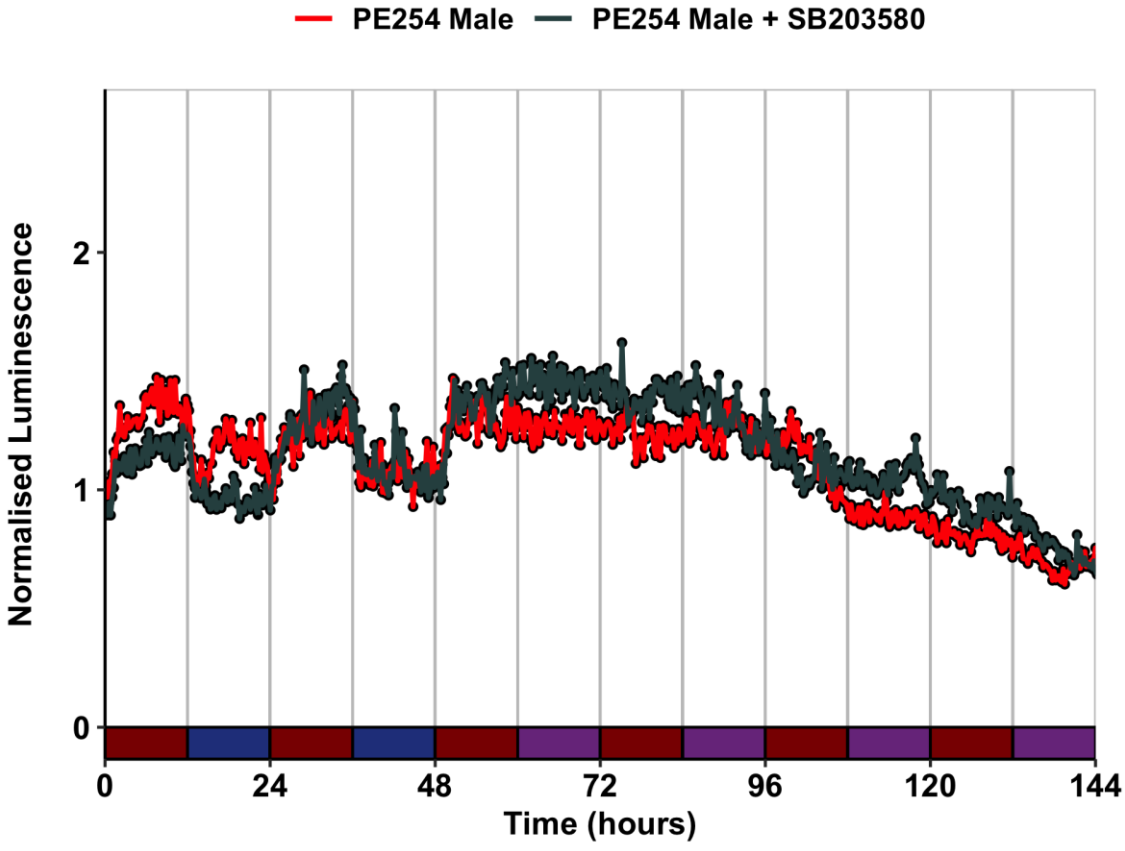

PE254 Male
PE254 Male + SB203580
Normalised Luminescence
2
1
0
0
24
48
72
96
120
144
Time (hours)

Figure 3.7: Pharmacological inhibitors further highlight consistent patterns of expression in *sur-5* luminescence reporter nematodes under temperature entrainment and free-running conditions.
Nematodes were exposed to two or three 12:12-hour temperature cycles (indicated by coloured bars; red = 20°C, dark blue = 15°C), then released into constant conditions (either 15°C, in which light blue bars indicate subjective day, or 20°C where purple bars indicate subjective night; see Glossary). Constant darkness was maintained throughout all experiments**. A)** PE254 males in the presence (n = 22) and absence (n = 21) of PF-670462. **B)** Nematodes in the presence and absence of SB203580, allowed to free-run at 15°C. **Bi)** PE254 males (treated: n = 12, untreated: n = 12). **Bii)** PE254 hermaphrodites (treated: n = 6, untreated: n = 5). **Biii)** N2 x PE255 male progeny (treated: n = 5, untreated: n = 6. **Biv)** N2 x PE255 hermaphrodite progeny (treated: n = 6, untreated: n = 8). **C)** PE254 males in the presence (n = 46) and absence (n = 41) of SB203580, allowed to free-run at 20°C. Figures **A**, **B** and **C** represent three independent single 96-well plate experiments. All hermaphrodite wells were supplemented with 50 µM FUDR. N2 x PE255 crosses were performed due to the *rol-6* marker (Mello et al., 1991) expressed by PE255 nematodes reducing male mating success. Nematodes that were not visibly alive or that reproduced were generally removed from analysis.

3.2.5 Chi-square periodogram analysis of luciferase assays

To attempt to objectively quantify potential rhythmicity in the luciferase experiments described above, the chi-square periodogram was used to analyse luminescence patterns. This was performed using ActogramJ software (Schmid et al., 2011), which provides estimates of periodicity of a given luminescence profile based on frequency peaks as well as the associated periodogram power value (Qp). Significance is defined as a peak with a Qp value greater than a chi-square distribution significance threshold (p = 0.05 here). Such periodogram approaches are generally applied to activity data, which can be erratic and subject to noise (Sokolove and Bushell, 1978; Brown et al., 2019), which could also the case for the observed luminescence patterns presented throughout this chapter. To identify potential circadian rhythms, the chi-square periodogram was applied to detect the strongest peak frequency within the range of 20 to 28 hours. These results are summarised in Table 3.1, which gives the period and power (Qp) of the tallest periodogram peak and statistical threshold values.

To test the effectiveness of the method, the chi-square periodogram was initially applied to temperature entrainment data over a five-day period (hour 0 to hour 120 on the samples shown in Figure 3.1). This five-day window was selected to reflect the approximate duration over which clear oscillations over the temperature cycles were apparent before starting to become erratic. As shown in Table 3.1, approximate circadian periods were identified in the majority of samples over the five-day timeframe, but were only significant in two of the six conditions illustrated in Figure 3.1. To provide a basis for direct comparison with later experiments, in which entrainment took place over shorter timeframes, the same data were also used for analysis of two-day periods of temperature cycles (hour 24 to hour 72 in Figure 3.1). In these cases, while 24-hour periodicity was detected, none of the treatment conditions were suggested to have a significant circadian rhythm. Collectively, these results may therefore suggest that the chi-square periodogram approach is more effective over longer time series. It should be noted however, that over the two-day reduced timespan, ActogramJ identified different period lengths in the same samples compared to the five-day analysis window, and Qp values increase, rather than decrease in four of the six strain conditions that presented in Figure 3.1. In these cases, stronger rhythms, albeit still nonsignificant, have been identified over the shorter timeframe. Taken together, these examples suggest the chi-square periodogram approach is capable, but perhaps somewhat ineffective in identifying the observed luminescence patterns as circadian rhythms, particularly over shorter time series. However, while not significantly rhythmic, the two-day Qp values presented in Table 3.1 do provide reference values to which free-running data and other experiments using alternate entrainment approaches can be compared.

In two-day light-entrained nematodes (experiments illustrated in Figure 3.2), the chi-square periodogram also did not identify significant circadian rhythms. In these samples, where 24-hour patterns were not overtly visible, peaks reflecting periods between 20 and 28 hours were still identified, but Qp values were substantially lower than temperature-entrained data of Figure 3.1 (Table 3.1). Given that in one case

(PE254 males without FUDR), Figure 3.1C shows a possible sinusoidal waveform with a frequency less than 20 hours, these data were also analysed for peaks in expression between one and 28 hours. An 11.4-hour peak was detected in the case of PE254 males without FUDR, but its Qp value is also non-significant. Given that the periodogram analysis of light:dark entrainment generally identified 20-28 hour rhythms but did not identify statistical significance greater Qp values or than the temperature-entrainment data described above, this analysis approach does not suggest light entrainment to be superior to using temperature cycles in entraining *C. elegans*.

To assess rhythmicity in free-running experiments, the chi-square periodogram was applied independently to the final 48 hours of entrainment-phase data and the initial 48 hours of free-running data in each genotype or treatment condition. With respect to temperature-entrained nematodes expressing either *Pnlp-36::luc+::gfp* or *Psur-5::luc+::gfp* transgenes (Figure 3.3), once again no significant rhythms were identified, but approximate 24-hour peaks were again detected in both entrainment and free-running, with Qp values generally increasing in the free-running phase. This could potentially suggest that while not visually overt like the responses to warm:cold cycles, underlying, non-significant 24-hour rhythms may be present in the free-running data.

As regards to antiphasic entrainment to both temperature and light (Figure 3.4A, C-G), a significant rhythm was identified in one of four MEG-strain populations, with rhythms approaching significance in the other three (Figure 3.4C-F) over a 48-hour free-running phase. Non-significant rhythms with approximate circadian frequency were also identified in *Pnlp-36::luc+::gfp* and other *Psur-5::luc+::gfp* nematodes (Figure 3.5). These results could provide some evidence of a weak clock in *C. elegans,* entrainable by temperature and light signals in antiphase and only detectable in select populations. This finding would be consistent with the prior MEG-strain

literature, in which free-running rhythms in luminescence could be detected, but in fewer than 50% of populations (Goya et al., 2016).

Temperature-entrained nematodes expressing a *Psur-5::luc+::gfp* transgene in different genetic backgrounds (Figure 3.6) or in response to pharmacological inhibitors of circadian rhythms (Figure 3.7) once again suggest many examples of non-significant rhythms with approximate 24-hour periods, alongside select examples of 24-hour peaks identified as significant. Of the wild-isolate strains, nematodes expressing *Psur-5::luc+::gfp* in an AB1 genetic background (Australian origin) were identified as rhythmic in the entrainment phase in one (Figure 3.6F), but not the other (Figure 3.6B), of two biological replicates. In the pharmacological treatment experiments, in one instance, PE254-strain male control nematodes showed a rhythm in warm:cold cycles, but not when treated with the inhibitor SB203580 (Figure 3.7C), possibly suggesting the inhibitor perturbs normal rhythms. Notably however, males treated with SB203580 were found to have Qp values approaching significance. Additionally, no significance was identified in PE254 males in 48-hour warm:cold cycles in the absence of an inhibitor in four other experiments (Figures 3.1D, 3.3C, 3.7A and 3.7Bi). Finally, perhaps the most interesting finding from this analysis is that in one 96-well experiment, *Psur-5::luc+::gfp*-expressing PE254-strain nematodes were identified as significantly rhythmic in both entrainment and free-running when the PF-670462 inhibitor was present (Figure 3.7A), but not in its absence. The presence of significant rhythmicity in entrainment and free-running could potentially suggest pharmacological treatment resulted in a stronger rhythm, although it is unclear what the mechanism of action would be. Collectively these analyses do suggest the potential for rhythms under and following temperature cycles in *Psur-5::luc+::gfp*-expressing nematodes, although as with the antiphasic experiments, cases of non-significance were far more numerous than cases of significant 24-hour periodicity.

Table 3.1: Summary of luciferase assay chi-square periodogram analysis results.

Figure	Sub-figure	Genotype/Treatment	Conditions	Analysis Timeframe (Hours)	Period Parameters (Hours)	Period (strongest peak; Hours)	Qp	Significance Threshold (p = 0.05)	Significant (Yes/No)
3.1	**A**	PE254 Hermaphrodite/FUDR	WC	0-120	20-28	26.87	95.1987	167.0372	No
	A	N2 x NLIH13 Hermaphrodite/FUDR	WC	0-120	20-28	24.05	140.6176	151.8484	No
	B, C	N2 x NLIH13 Male	WC	0-120	20-28	22.97	146.9801	145.9723	Yes
	B, D	PE254 Male	WC	0-120	20-28	22.75	103.8763	144.7945	No
	C	N2 x NLIH13 Male/Levamisole	WC	0-120	20-28	20.15	80.672	130.5914	No
	D	PE254 Male/Levamisole	WC	0-120	20-28	23.40	181.957	148.3251	Yes
3.1	**A**	PE254 Hermaphrodite/FUDR	WC	24-72	20-28	19.93	99.8347	129.4015	No
	A	N2 x NLIH13 Hermaphrodite/FUDR	WC	24-72	20-28	23.62	130.0052	149.5004	No
	B, C	N2 x NLIH13 Male	WC	24-72	20-28	23.40	147.4017	148.3251	No
	B, D	PE254 Male	WC	24-72	20-28	26.00	120.0809	162.3766	No
	C	N2x NLIH13 Male/Levamisole	WC	24-72	20-28	20.15	86.4049	130.5914	No
	D	PE254 Male/Levamisole	WC	24-72	20-28	21.23	129.5783	136.5259	No
3.2	**A**	PE254 Hermaphrodite/FUDR	LD	0-48	20-28	22.17	38.5494	58.9237	No
			LD	0-48	1-28	1.90	5.6078	10.6321	No
	B	PE254 Male/FUDR	LD	0-48	20-28	20.90	34.4773	56.2905	No
			LD	0-48	1-28	1.27	0.071	7.8674	No
	C	PE254 Male	LD	0-48	20-28	22.17	50.561	58.9237	No
			LD	0-48	1-28	11.40	30.8943	35.7072	No
3.3	**A**	NLIH13 Hermaphrodite/FUDR	WC	0-48	20-28	24.05	143.8314	151.8484	No
			Free-run	48-96	20-28	22.75	107.1649	144.7945	No
	B	NLIH2 Hermaphrodite/FUDR	WC	0-48	20-28	19.93	71.1565	129.4015	No
			Free-run	48-96	20-28	21.02	99.7434	135.3409	No
	C	PE254 Male	WC	0-48	20-28	24.05	106.0894	151.8484	No

			Free-run	48-96	20-28	24.48	129.7003	154.1932	No
	C	PE254 Male/Levamisole	WC	0-48	20-28	24.05	111.326	151.8484	No
			Free-run	48-96	20-28	24.27	124.3421	153.0212	No
3.4	**A**	MEG Individuals Average/ FUDR	Free-run	0-48	20-28	23.40	132.825	148.3251	No
	C	MEG Strain Population 1/ FUDR	Free-run	0-48	20-28	24.00	107.7931	127.0185	No
	D	MEG Strain Population 2/ FUDR	Free-run	0-48	20-28	24.00	119.9783	127.0185	No
	E	MEG Strain Population 3/ FUDR	Free-run	0-48	20-28	24.27	126.807	128.2105	No
	F	MEG Strain Population 4/ FUDR	Free-run	0-48	20-28	24.00	140.9394	127.0185	Yes
	G	MEG Population Average/FUDR	Free-run	0-48	20-28	24.00	109.8702	127.0185	No
3.5	**-**	PE254 Male	Free-run	0-48	20-28	20.27	114.365	199.3885	No
			Free-run	0-96	20-28	22.93	128.3776	222.2532	No
	-	NLIH13 Hermaphrodite/FUDR	Free-run	0-48	20-28	24.00	131.1227	231.3518	No
	-	NLIH2 Hermaphrodite/FUDR	Free-run	0-48	20-28	20.13	117.6761	198.2404	No
3.6	**A**	PE255 Hermaphrodite/FUDR	WC	24-72	20-28	25.33	102.291	110.2071	No
			Free-run	72-120	20-28	24.00	84.2219	105.3564	No
	B	AB1 Hermaphrodite/FUDR	WC	24-72	20-28	24.00	103.872	105.3564	No
			Free-run	72-120	20-28	24.00	79.4123	105.3564	No
	C	MY9 Hermaphrodite/FUDR	WC	24-72	20-28	24.00	85.7887	105.3564	No
			Free-run	72-120	20-28	24.33	73.5495	106.5713	No
	E	PE255 Hermaphrodite/FUDR	WC	24-72	20-28	23.67	91.7313	104.1401	No
			Free-run	72-120	20-28	24.00	84.0727	105.3564	No
	F	AB1 Hermaphrodite/FUDR	WC	24-72	20-28	23.33	114.5085	102.9221	Yes
			Free-run	72-120	20-28	21.33	61.4298	95.5799	No
	G	MY9 Hermaphrodite/FUDR	WC	24-72	20-28	20.33	39.3435	91.8847	No
			Free-run	72-120	20-28	24.00	85.1059	105.3564	No

3.7	**A**	PE254 Male	WC	24-72	20-28	21.88	125.7792	140.075	No
			Free-run (Constant Cold)	72-120	20-28	27.08	160.641	168.2007	No
	A	PE254 Male/PF-670462	WC	24-72	20-28	23.83	176.5391	150.6748	Yes
			Free-run (Constant Cold)	72-120	20-28	22.10	155.983	141.2562	Yes
	Bi	PE254 Male	WC	24-72	20-28	23.83	146.6731	150.6748	No
			Free-run (Constant Cold)	72-120	20-28	20.15	115.9306	130.5914	No
	Bi	PE254 Male/SB203580	WC	24-72	20-28	19.93	122.3982	129.4015	No
			Free-run (Constant Cold)	72-120	20-28	25.13	151.5172	157.7048	No
	Bii	PE254 Hermaphrodite/FUDR	WC	24-72	20-28	19.93	75.0893	129.4015	No
			Free-run (Constant Cold)	72-120	20-28	23.40	139.4509	148.3251	No
	Bii	PE254 Hermaphrodite/SB203580, FUDR	WC	24-72	20-28	19.93	106.498	129.4015	No
			Free-run (Constant Cold)	72-120	20-28	22.32	132.6952	142.4365	No
	Biii	N2 x PE255 Male	WC	24-72	20-28	19.93	94.5639	129.4015	No
			Free-run (Constant Cold)	72-120	20-28	22.53	106.1188	143.616	No
	Biii	N2 x PE255 Male/SB203580	WC	24-72	20-28	23.18	130.8763	147.1491	No
			Free-run (Constant Cold)	72-120	20-28	19.93	123.7487	129.4015	No
	Biv	N2 x PE255 Hermaphrodite/FUDR	WC	24-72	20-28	22.97	112.4227	145.9723	No
			Free-run (Constant Cold)	72-120	20-28	22.10	86.5708	141.2562	No

	Biv	N2 x PE255 Hermaphrodite/SB203580, FUDR	WC	24-72	20-28	21.45	129.6262	137.7098	No
			Free-run (Constant Cold)	72-120	20-28	22.53	135.6307	143.616	No
	C	PE254 Male	WC	24-72	20-28	25.13	168.7643	157.7048	Yes
			Free-run (Constant Warm)	72-120	20-28	21.88	69.2895	140.075	No
	C	PE254 Male/SB203580	WC	24-72	20-28	24.05	143.1254	151.8484	No
			Free-run (Constant Warm)	72-120	20-28	22.32	110.354	142.4365	No

Overall, the chi-square periodogram approach suggests the presence of weak circadian rhythms in the *C. elegans* luciferase-reporter experiments described in this chapter, but the approach may not represent an effective method for analysing these data. As shown in Table 3.1, the ActogramJ software frequently identifies periods of approximately 24 hours in the entrainment and free-running data. These periods are largely non-significant, but include some examples of significance and other peaks that are close to a significance threshold. However, within these results are both cases where the approach does not find significance when oscillations are visibly apparent, and cases where luminescence patterns are identified as significant where hallmarks of circadian rhythms are lacking.

Regarding cases of non-significance, luminescence patterns in Figure 3.1 show visible, approximately 24-hour oscillations, but the chi-square periodogram approach does not find significant rhythms when applied to two-day data, and only two genotype conditions are found to be significant over five-day sampling. One potential issue with this analysis approach is therefore that a longer sampling period may be necessary to effectively detect rhythms; the chi-square periodogram method has generally been shown to be effective using considerably longer sampling durations than two days (Sokolove and Bushell, 1978; Brown et al., 2019). However, an alternate explanation for a lack of significant rhythms being detected in the Figure 3.1 data could be that the luminescence oscillations do not show conventional circadian expression patterns; the observed recordings more closely resemble a binary square wave than the sinusoidal patterns typified by core clock genes in other species (Pizarro et al., 2013; Li et al., 2017; Leming et al., 2014). The analysed waveforms may be indicative of a direct and fast response to temperature rather than a 24-hour trend, and the chi-square method may be ineffective in detecting this. If these patterns are a direct response and do not relate to circadian gene expression, a failure to detect significance may be an issue only in the sense that true rhythms during entrainment are likely to be masked by these patterns.

A potentially greater concern from this analysis is in cases where significant rhythmicity was detected in luminesce data, but rhythms were not visibly apparent, particularly in considering free-running data (for example in Figures 3.4F, 3.7A). While this could mean the approach is detecting rhythms the human eye cannot, circadian hallmarks, such as a repeating pattern, continuation of the entrained expression pattern in constant conditions and sinusoidal oscillations (discussed in Chapter 1, section 1.1.1) are largely absent in these data. These cases could therefore indicate subtle rhythms, or could be false positives, indicative of a statistical approach that cannot effectively detect rhythmicity in these data, possibly due to noise. With no basis for comparison, these significant results make it difficult to discern the appropriateness of the periodogram analysis and to draw conclusive evidence of rhythmicity from the luminescence data.

3.2.6 RT-qPCR experiments reveal circadian oscillations in *nlp-36* in response to light and temperature entrainment in phase

Given the lack of overt circadian rhythms found through luminescence reporting using two previously published reporter genes, and further considering the question of how to entrain *C. elegans*, I investigated expression of *nlp-36* and *sur-5* directly by RT-qPCR under different entrainment conditions of light and temperature (Figure 3.8). To identify significant rhythms, data were analysed using the meta2d function within the R package MetaCycle (Wu et al., 2016; R Core Team, 2018), which performs two circadian rhythm detection algorithms, JTK_Cycle and Lomb-Scargle (Hughes et al., 2010; Glynn et al., 2006), and generates integrated p-values.

Interestingly, under the protocols in which *nlp-36* and *sur-5* were found to be rhythmic (entrainment by temperature, antiphasic light and temperature respectively; van der Linden et al., 2010; Goya et al., 2016; see Glossary), no significant evidence of rhythmicity was detected (Figure 3.8B and 3.8D; Table 3.2). This is despite measuring expression of *sur-5* using the same qPCR primers as previous work. However, when light and temperature entrainment were used

together in phase (under 12:12-hour light/warm:dark/cold conditions, with free-running in constant dark/cold; 15°C), *nlp-36* revealed an expression pattern in phase with the prior work (van der Linden et al., 2010), with a decline in expression during the day, and an increase during the night, which then repeats itself in free-running conditions (Figure 3.8A), exhibiting a fold change of 1.49. Three biological replicates were performed to confirm this pattern, which revealed a highly consistent expression pattern during the light and 20°C thermophase (indicated by low SEM in Figure 3.8A), but with much increased variation in darkness at 15°C. JTK_Cycle identified a highly significant (q = 0.00718, Benjamini-Hochberg corrected for multiple testing) 24-hour period in this dual-entrained *nlp-36* expression profile (Table 3.2), although Lomb-Scargle did not, resulting in a meta2d q-value significant at the 5%, but not 1% threshold. Evidence for these rhythms under in-phase light and temperature cycles was also examined in *sur-5*, but in this case no significant rhythm or clear driving effect of the environmental cycles was apparent (Figure 3.8C).

As to why *nlp-36* expression becomes more variable between biological replicates over the time series in Figure 3.8A, such a result might be expected in that in circadian biology, diurnal signals are a much more powerful driver of rhythmicity at the organism level than endogenous clocks. This is evidenced by the susceptibility of organism and tissue-level rhythms to damping over time as cellular and tissue clocks rapidly become desynchronised (illustrated in Chapter 1, Figure 1.2; Welsh et al., 2004). However, given that the increased variation in *nlp-36* expression between replicates also occurs in the first night phase, the increased standard error could also be indicative of less consistent gene expression in conditions of darkness at 15°C. It could be the case that nematodes are more metabolically variable in colder conditions in particular. From an experimental design perspective however, these environmental conditions should not be a particularly unrealistic or uncommon representation of those that nematodes would encounter in nature, based on locations in which they have been isolated (discussed in Chapter 1, section 1.2.2).

Overall, these RT-qPCR data provide examples of both agreement and contradiction with published literature. Expression patterns of *nlp-36 in* Figure 3.8A are in agreement with published literature in oscillating across entrainment and free-running days (van der Linden et al., 2010). However, these results were only apparent using light and temperature to entrain, as opposed to temperature alone as in the prior work, which was not sufficient to drive rhythms here (Figure 3.8B). This work also found no evidence of rhythmicity in *sur-5*, in contradiction to prior work (Goya et al., 2016). As to why these differences were observed, a major difference between RT-qPCR experiments here and previously published *nlp-36* and *sur-5* expression data (van der Linden et al., 2010; Goya et al., 2016), is that for these experiments FUDR was not used, meaning adults were allowed to develop and reproduce normally, rather than halting development prior to embryo development. FUDR halts development just prior to adulthood through inhibiting DNA synthesis and cell division, which could cause unknown effects on gene expression in nematodes still undergoing a developmental program. The approach taken here enabled the avoidance of a potent pharmacological agent, but did necessitate manual collection of individual nematodes, meaning a relatively small number were used (n = 50). Given that prior work has frequently reported rhythms in large populations of nematodes, and sometimes shown rhythms only in a proportion of animals (Simonetta and Golombek, 2007; Herrero et al., 2015; Goya et al., 2016; discussed in Chapter 1, section 1.3), the small number of nematodes used here could possibly explain the observed differences in rhythmicity.

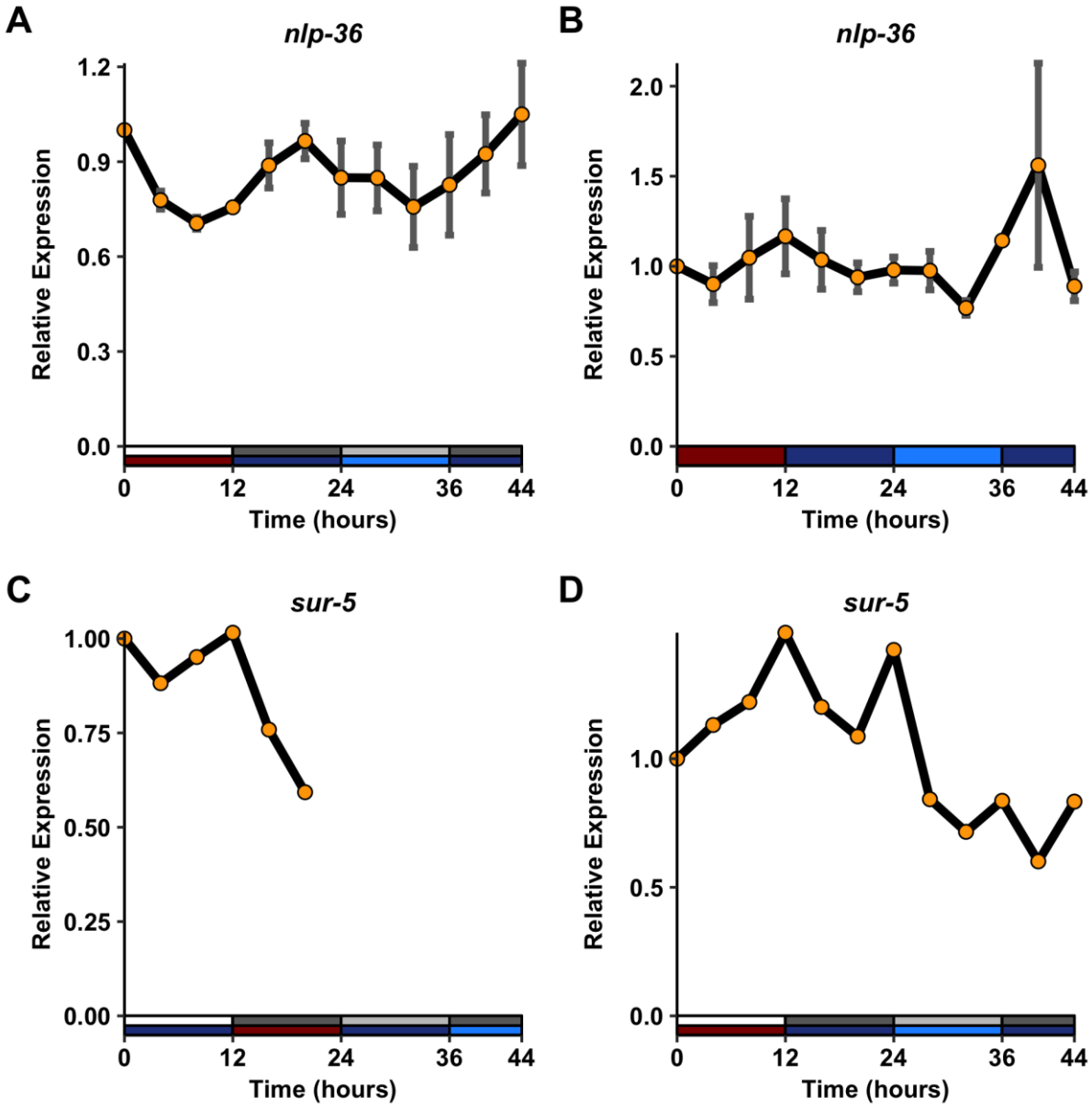


Figure 3.8: RT-qPCR experiments suggest circadian oscillations in *nlp-36* but not *sur-5* expression.

Experiments using different entrainment protocols of photic and 20°C:15°C thermal cycles. **A)** *nlp-36* expression under light and temperature entrainment in phase, average of three biological replicates. **B)** *nlp-36* expression under temperature entrainment only, in constant darkness, average of two biological replicates. **C)** *sur-5* expression under light and temperature entrainment in antiphase, entrainment phase only. **D)** *sur-5* expression under light and temperature entrainment in phase. In all experiments, free-running was under constant dark/cold. Coloured bars indicate photic and thermal conditions (white = light, dark grey = dark, light grey = dark, subjective light phase, red = warm, dark blue = cold, light blue = cold, subjective warm phase). Each timepoint sample represents an individual NGM plate from which adult hermaphrodites (n = 50) were manually picked. To calculate relative expression, Ct values at each timepoint were normalised to the reference gene *pmp*-3, and

then to time 0, and subsequently averaged in the cases of multiple biological replicates. Where present, error bars indicate S.E.M. All experiments performed in the PE255 strain.

Table 3.2: Significance values and circadian periods of *nlp-36* and *sur-5* expression over RT-qPCR time series using different entrainment conditions.

Gene + Entrainment Conditions	meta2d q-value	JTK_Cycle q-value	Lomb-Scargle q-value	meta2d Period (Hours)	JTK_Cycle Period (Hours)	Lomb-Scargle Period (Hours)
nlp-36 **in phase**	0.013	0.00718	0.841	23.6	24	23.2
nlp-36 **temperature**	NA	NA	> 1	NA	NA	24.6
sur-5 **in phase**	1	1	1	28	28	28
sur-5 **antiphase**	0.328	0.851	0.23	26	24	28

3.3 Summary

The work described here aimed to build upon prior work that suggested the genes *nlp-36* and *sur-5* both show circadian expression patterns and can be effectively used as *in vivo* reporters of the circadian clock in *C. elegans*. Luciferase reporters of these genes, using transgenic strains generated here and from prior published work did not however reveal overt evidence of circadian rhythms. These strains all responded to temperature cycles with 24-hour luminescence patterns (Figure 3.1), but these observations may not reflect gene expression, but rather other metabolic responses. Subsequent to entrainment by temperature and combined temperature and light entrainment (Figure 3.3, 3.4, 3.5), endogenous rhythms were not generally apparent in these strains. Reasons as to the lack of overt rhythms in these experiments could be that rhythms are subtle or masked by stochastic factors, that their detection is impeded by methodological issues, or simply that the transgenes and strains explored here do not show circadian regulation of luciferase expression.

The idea of subtle or masked rhythms was potentially indicated in (Figure 3.4C-G), in which nematodes showed increased activity in constant conditions for approximately 24 hours, but with variation in amplitude and a reduced peak when averaged. The idea of subtle or masked rhythms was also further explored in experiments expressing the transgenes in wild-isolate backgrounds and by treating nematodes putative pharmacological inhibitors of rhythms. These experiments respectively served to investigate whether laboratory adaptation of the N2 strain had diminished circadian rhythms in expression (Figure 3.6), and to identify changes in expression patterns that might indicate that ambiguous or masked circadian signals are present. Neither approach revealed compelling, unambiguous evidence of circadian rhythms, but wavelet spectral analysis of the pharmacological data did suggest subtle changes in luminescence intensity as a result of pharmacological inhibitors (Figure 3.7C; Hargreaves et al., 2019). The observations of potential rhythms following antiphasic entrainment and the subtle effects of pharmacological inhibition could indicate that if rhythms are present in *C. elegans*, their detection will be non-trivial relative to other models (in which luciferase rhythms can be convincingly discerned; illustrated in Chapter 1, Figure 1.2). This notion could suggest that gene expression reporter approaches (using luciferase or other reporters) will not have the same utility in dissecting the molecular basis of *C. elegans* circadian rhythms as in other models, due to increased difficulty in the delineation of clear circadian rhythms from perturbed expression.

Based on experimental results here, methodological issues are perhaps not the most likely explanation for the absence of detected circadian rhythms. Throughout this chapter repeated observations were made of consistent changing expression patterns in transgene-specific luminescence profiles. Different strains showed different profiles, but patterns were remarkably consistent within nematodes of the same strain (and between strains expressing the same transgene; Figure 3.5). These observations were particularly notable where patterns persisted even when very few nematodes were assayed (Figures 3.5 and 3.7). These results suggest transgene

expression drives observed luminescence, and each transgene generates specific luminescence patterns over the course of the time series. From this perspective, there would be potential for an effective circadian reporter transgene in *C. elegans*, based on luciferase, provided it is driven by an appropriate promoter.

To answer the question of whether *nlp-36* and *sur-5* are rhythmically expressed and should be used in circadian reporter constructs, as well as to fulfil the other objective of this chapter and ascertain the environmental conditions to which *C. elegans* optimally entrains, a series of RT-qPCR experiments were performed under temperature and combined light and temperature cycles. These experiments were able to recapitulate previous data in finding significant rhythms in *nlp-36* expression, but did not identify rhythmic expression in *sur-5* (Figure 3.8). Further, rather than clarify temperature or antiphasic dual entrainment as most effective for *C. elegans* entrainment (the two approaches suggested in prior work; van der Linden et al., 2010; Goya et al., 2016), entrainment here was only found to be effective when using light and temperature in phase, conditions which had not been used in previous work.

Overall, the data presented in this chapter were ultimately unsuccessful in developing robust circadian reporters of gene expression, instead largely not being able to replicate prior work. This chapter highlighted a range of issues for such approaches, which could stem from choice of rhythmic gene promoter, but more broadly could arise from an overall lack of robust rhythms in *C. elegans* compared to other model organisms. As such, this work raises the question of whether luciferase reporter approaches could ever be effective in the study of the *C. elegans* molecular clock. The only strong indication of rhythmicity in this chapter was found with respect to the *nlp-36* gene, and only in response to light and temperature cycles, not previously used for circadian study of *C. elegans*. However, a combination of small fold change and high variability between replicates could generate uncertainty as to whether rhythms are present (and thus the utility of using the *nlp-36* gene in

reporter approaches). The work that follows in Chapter 4, Results 2 applies light and temperature entrainment in phase to the *C. elegans* transcriptome as a whole, both to isolate further candidate genes of interest for future study and to better characterise the extent to which *C. elegans* as an organism shows rhythmic gene expression.

3.4 Chapter 3 methods

Generation of transgenic nematodes

To generate the NLIH13 and NLIH2 strains, a 3000 bp region upstream of the first exon of *nlp-36* was amplified from genomic library cosmid B0464 (obtained from Wellcome Trust Sanger Institute, Cambridge) using PCR (primer sequences given in Chapter 2, Table 2.2). **PCR reaction contents:** 100 ng DNA, 0.5 µM forward primer, reverse primer, 0.2 mM dNTPs, 1 µL *Pfu* DNA Polymerase (Stratagene), 5 µL 10X *Pfu* reaction buffer, H_2O to 50 µL. **PCR conditions:** 94°C (45 seconds), 35 cycles of {94°C (45 seconds), 55°C (45 seconds), 72°C (210 seconds)}, 72°C 10 minutes.

This fragment was inserted in place of the *sur-5* promoter region into a modified pSLGCV vector (pSLGCV was generated by Anne Glover & Jonathan Pettitt; Addgene plasmid # 49862; http://n2t.net/addgene:49862; RRID:Addgene_49862; Lagido et al., 2008), generating the *Pnlp-36::luc+::gfp* transgene. Promoter sequences were inserted into pSLGCV by restriction digest subcloning. This vector was used to transform N2 strain nematodes by microinjection with assistance from Professor Ian Hope, University of Leeds. Plasmids (at concentrations of 200-900 mg/mL) were co-injected with pRF4 plasmid (0.12 mg/mL), containing the *rol-6* marker gene (Mello et al., 1991). Progeny expressing the *rol-6* phenotype were selected over subsequent generations and allowed to reproduce.

Introduction of *Psur-5::luc+::gfp* into wild isolate strains

PE254 strain males expressing *Psur-5::luc+::gfp* were generated by maintaining plates with abundant larvae at 30°C for 4-6 hours. Approximately 10 male PE254 strain nematodes were placed on an NGM plate with approximately five AB1 or MY9 strain (Table 2.1) hermaphrodites and allowed to reproduce. Some male progeny of these crosses, selected for fluorescence, were then mated again to AB1 and MY9 hermaphrodites. This process was repeated for at least eight generations. Subsequently, individual hermaphrodites were self-fertilised, progeny were scored for fluorescence, and homozygous parental individuals were maintained as separate strains.

Luciferase detection experiments

Nematodes were grown on standard NGM plates. Individual nematodes were selected and placed into a 96-well plate containing standard S buffer (Chapter 2, section 2.2.1) with 100 µM D-luciferin and 1 g/L (wet weight) cultured *E. coli* OP50. All males were sampled as adults, while hermaphrodites were generally sampled at the late L4/young adult stage. All hermaphrodite experiments also contained 50 µM FUDR (to prevent reproduction). Where applicable, p38k and CK1 inhibitors, SB203580 (Sigma-Aldrich) and PF-670462 (Sigma-Aldrich) were used at concentrations of 10 µM. Luminescence was measured in counts per second (CPS), once every 5-15 minutes in most experiments, using a Packard TopCount NXT scintillation counter. Raw data were exported into Toptemp II (A. Hall; unpublished data). All experiments were maintained in the TopCount at 15°C for approximately 12 hours prior to the onset of recording, except those illustrated in Figure 3.4 and 3.5, which were maintained at 20°C. Generally, nematodes that had reproduced or were deceased at the conclusion of experiments, identified either from CPS values indistinguishable from empty wells or from visual scoring for fluorescence and movement, were excluded from experimental analysis and results.

Drosophila luminescence in Figure 1.2 (Chapter 1) was measured using the same recording and analysis methods. Male flies were placed in individual wells in 96-well plate, each containing 150 µL solid media consisting of 1% Agar, 5% sucrose, 100 mM luciferin, and covered by a domed 20 µL microcentrifuge tube lid to restrict movement, pierced with a fine needle to enable air flow. Flies were entrained to three 12:12-hour light:dark cycles at 19°C in a growth chamber prior to the start of TopCount recording. Recording took place in constant darkness at 19°C.

Reverse transcription quantitative PCR experiments: sample collection and processing

Note: entrainment conditions and sample collection and processing methods differed for and RT-qPCR and RNA-seq (Chapter 4: Results 2) experiments. RNA-seq sampling and processing methods are given in (Chapter 4 methods, section 4.4).

Experimental nematode eggs were isolated either by the bleaching method (see Chapter 2, section 2.2.2) or by allowing 8-10 adults to lay eggs on NGM plates seeded with *E. coli* for 4-6 hours. Plates were then kept at 15°C in constant darkness until entrainment began, except those under antiphasic entrainment (Glossary) which were kept at 20°C in constant darkness. Experiments were sometimes performed in two growth chambers 12 hours out of phase to simplify sample collection. Samples were collected every four hours over the final entrainment and first free-running days from dawn (time 0).

RNA sampling entailed manually picking 50 nematodes into 100 or 250 µL TRIzol (Ambion) and immediately freezing at -80°C. To extract RNA, samples in TRIzol were defrosted then subjected to three freeze/thaw cycles in which they were snap frozen in liquid nitrogen, defrosted at 37°C and vortexed for 30 seconds. Subsequently, samples were left to sit for 5 minutes at room temperature, 0.2 vols chloroform was added and samples were centrifuged for 15 minutes at 13200 RPM at 4°C. Most of the aqueous phase was transferred to a new tube and 0.1 vols 3M sodium acetate

pH 5.2, 2.5 vols ethanol and sometimes 1 µL GlycoBlue (Invitrogen) were added. Samples were left at -20°C for 1 hour and centrifuged at 4°C at 13200 RPM for 1 hour. Most supernatant was aspirated and 200 µL 70% ethanol was added and centrifuged for 15 mins at 4°C. ethanol was removed and pellet was allowed to air dry for approx. 10 mins, before resuspension in 14-20 µL DEPC-treated H_2O. RNA samples were then stored at -80°C until needed.

cDNA synthesis and qPCR

Some samples were DNase-treated using RQ1 RNase-Free DNase (Promega) following the manufacturers' protocol. RNA was reverse transcribed using SuperScript III or SuperScript IV Reverse Transcriptase (Invitrogen), following the manufacturers' protocols.

Quantitative PCR experiments were performed on StepOnePlus Real-Time PCR System using MicroAmp Fast Optical 96-Well Reaction Plates (Applied Biosystems). cDNA was diluted to 1/10 starting concentration using DEPC-treated H_2O. Each well contained 5 µL diluted cDNA, 2 µL primers (0.05 µM final concentration), 2X Fast SYBR Green Master Mix (Applied Biosystems) and 3 µL DEPC-treated H_2O. Each cDNA reaction was performed in triplicate. Where the Ct standard deviation > 0.5 for each replicate, the outlier sample was excluded from the average. Relative expression was calculated by normalising to the reference gene and then to the first timepoint using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Analysis of RT-qPCR data

Rhythmic genes were detected using the meta2d function in MetaCycle (Wu et al., 2016) in within R (R Core Team, 2018). The four gene/entrainment condition combinations were analysed independently and Benjamini-Hochberg-corrected q-values (Benjamini and Hochberg, 1995) calculated from meta2d p-values. Analysis utilised JTK_Cycle and Lomb-Scargle only. Fold-change was calculated by dividing the

maximum relative expression by the minimum, using average relative expression in cases where multiple replicates were used.

Chapter 4. Results 2: Insights into circadian and non-circadian oscillations in *C. elegans* gene expression from RNA sequencing

4.1 Introduction

An expectation of the transcriptional circadian clock found in model species (the TTFL, illustrated in Chapter 1, Figure 1.1), is that a proportion of the genome should show oscillatory patterns of expression with a frequency approximating 24 hours. At a basic level, this proportion should include genes directly regulated by the TTFL (genes driven by promoter regions containing E-box elements in the case of the *Drosophila* and mammalian models). However, through downstream transcription factors and serial gene interactions, the rhythmic transcriptome could also represent a complex regulatory network in which many genes exhibit 24-hour expression cycles (Bozek et al., 2009; Lehmann et al., 2015; Meireles-Filho et al., 2013). Consequently, the degree to which organisms' genomes are clock-regulated, along with the identities and functions of rhythmic genes, form central questions in chronobiology that help to define the wider importance of circadian rhythmicity.

The advent of transcriptome-wide approaches to measuring gene expression (initially microarrays, and more recently RNA sequencing) has enabled chronobiologists to begin to answer these questions, and also yielded wider insights into the complexities of diurnal regulation. In *Drosophila*, numerous microarray time series have been published reporting circadian rhythmicity in the fly genome (Appendix Table 6.2; Meta-analysed in Keegan et al., 2007), and transcriptomic approaches have also been used to characterise rhythmic microRNAs (Yang et al., 2008), and to compare messenger RNA and nascent RNA (Rodriguez et al., 2013). In mice, studies have particularly highlighted differences in rhythmic genes across brain regions and peripheral tissues (Hughes et al., 2007; Hughes et al., 2009; Kasukawa et al., 2011, Brown et al., 2017), and have demonstrated substantial rhythmic mRNA expression;

one time series analysis of microarray and RNA-seq data reported 43% of transcripts to show endogenous circadian rhythms in at least one organ (Zhang et al., 2014).

Transcriptome-wide approaches have also been used to identify rhythmic genes in a plethora of other organisms that, like *C. elegans*, are not conventionally used to model circadian rhythms, but for which a better understanding of circadian regulation could offer insights to the fields of research in which they are used. Among these are species used to study interspecific dynamics and disease, ranging from the mosquito vectors of human malaria and yellow fever (Rund et al., 2011; Rund et al., 2013; Ptitsyn et al., 2011; Leming et al., 2014) to behaviour-modulating fungal parasites of ants (De Bekker et al., 2017). In honey bees, a model of eusociality, microarray time-series analysis has suggested intraspecific differences in temporally rhythmic genes between sibling workers with different roles in the colony (Rodriguez-zas et al., 2012). A less well-studied species in which microarray time series analysis has identified circadian rhythms is *Ciona intestinalis*, an animal that shares a commonality with *C. elegans* in that it expresses homologues of core TTFL genes that have themselves not been found to be rhythmic (Matsumae et al., 2015; van der Linden et al, 2010; Olmedo et al., 2012; see Chapter 1, section 1.2.1). Transcriptomic time series approaches have thus proven to be malleable tools in identifying clock-regulated genes, both in circadian model organisms and less-established species. These experiments have enabled a deeper understanding of complex genomic regulation by circadian clocks and the biological functions they serve.

Despite the broad utility of transcriptomic approaches in characterising circadian clocks, only one transcriptome-wide gene expression study in *C. elegans* exists in published literature (van der Linden et al., 2010). Through separate microarray time series in which nematodes were entrained by light or temperature, the authors reported 294 and 88 genes to oscillate in the respective conditions over combined entrainment and free-running phases. In total, these oscillating genes account for

around 2% of expressed genes in *C. elegans*, a generally lower proportion than those reported in other species (Appendix Tables 6.1, 6.2, 6.3), and considerably smaller than the aforementioned 43% of genes reported across tissues in mice (Zhang et al., 2014). These results likely suggest a substantially reduced role of a transcriptional clock in regulating *C. elegans* gene expression. However, transcriptomics studies have often painted an incomplete picture of circadian gene expression in other organisms; microarray data from different, contemporary experiments within the same species have previously uncovered different gene lists, including when analysed by the same methods (Keegan et al., 2007; revisited in Chapter 5. Discussion, section 5.3.1). As such, understanding of the true extent of diurnal regulation in *C. elegans* will require further transcriptome-wide analyses.

In this chapter I present the first circadian RNA-seq time series in *C. elegans*. While no microarray or RNA-seq experiment is likely to identify every rhythmic transcript in any model of circadian rhythms (without a financially improbable number of time-points, replicates and sequencing depth; Schurch et al., 2016; Li et al., 2015), RNA-seq is likely to be a more effective approach to rhythm detection. In particular, RNA-seq offers improved signal-to-noise ratio relative to microarray chips and in comparative studies has been found to perform better in identifying lower-expressed transcripts, while producing generally correlated results (Zhao et al., 2014; Chen et al., 2017). In this experiment, six biological replicates were sequenced per timepoint in order to capture as many true positive results as possible (Schurch et al., 2016; six replicates being the maximum achievable number and in excess of any prior metazoan circadian time series). To entrain the clock, attempting to maximise diurnal variation in gene expression, this experiment utilised dual zeitgebers of light and temperature in phase (in accordance with the RT-qPCR conditions that generated significant rhythmicity in *nlp-36* expression in Chapter 3, Results 1; Figure 3.8A). Further, experimental nematodes were maintained in these conditions as adults, throughout development, and parental generation nematodes were also entrained, meaning nematodes were exposed to environmental cycles from embryogenesis. In

describing this experiment, I present novel insights into the rhythmicity of the *C. elegans* transcriptome, including previously unreported genes with circadian expression patterns and novel non-circadian oscillations.

4.2 Results and discussion

RNA-seq gene expression data was generated using *C. elegans* RNA, harvested at 4-hour intervals over a 12-point time series spread across an entrainment day (a 12:12-hour light/20°C:dark/15°C cycle) and a free-running day (constant dark/15°C; full time series conditions illustrated in Appendices, Figure 6.1). Each sample comprised approximately 100-200 age-synchronised hermaphrodite adult nematodes from one NGM plate.

Sequencing of cDNA was performed using the Illumina HiSeq 3000 system (Illumina, Inc.), which enables a high sequencing depth (approximately 300 million reads per sequencing lane), low cost relative to this depth and a low error rate (Reuter et al., 2015). Fragments of length 150 bp were sequenced, which should be sufficient to detect differentially expressed genes, utilising the full *C. elegans* reference transcriptome (Chhangawala et al., 2015; reference transcriptome WBcel235 obtained from GenBank, Benson et al., 2013).

Paired-end sequencing of the samples, spread across two sequencing lanes on a flow cell, resulted in 12.4 to 32.8 million reads per each of 71 samples, equating to an approximate exome coverage of 68.7x to 181.7x (based on exon regions representing an estimated 27% of the 100 megabase *C. elegans* genome; The *C. elegans* Sequencing Consortium, 1998). This range of 12.4 to 32.8 million reads for each sample would be expected to be sufficient to detect a majority of circadian rhythms, based on estimates from data in mice and *Drosophila* (Li et al., 2015). All total RNA samples passed quality control measures for integrity (clear 60S and 40S peaks in every sample, Agilent Bioanalyzer RIN score ≥ 9 for 75% of samples; Schroeder et al., 2006; see section 4.4, methods) prior to mRNA isolation, cDNA synthesis and

sequencing. Sequencing was successful in all but one of the 72 samples (one biological replicate at the 44-hour, final timepoint), for which very few reads were generated. As a result, five biological replicates were used for all downstream analyses at this timepoint and six replicates for all others. Quality Control by Multi-QC did not highlight issues with any sequenced sample (average Phred score base quality > 30 for all samples, normally distributed GC content).

Sequence mapping and quantification of transcript abundance was performed using Salmon (Patro et al., 2017), which uses an accurate, computationally efficient 'quasi-mapping' procedure (in which reads are assigned to a transcriptome index) to determine read counts, while performing an in-built correction for GC-content bias, a documented issue found to occur using the Illumina platform (Chen et al., 2013).

4.2.1 Analysis by MetaCycle reveals transcripts with 24-hour periodicity under entrainment and free-running conditions

To identify genes with circadian expression patterns, time series data were analysed using the meta2d function within MetaCycle, as in Chapter 3, Results 1 (Wu et al., 2016; R Core Team, 2018). Core circadian genes in most species typically exhibit sinusoidal expression patterns with an approximate 24-hour frequency (Pizarro et al., 2013; Li et al., 2017; Leming et al., 2014). The meta2d function performs two algorithms developed to detect these circadian patterns, JTK_Cycle and Lomb-Scargle, and generates integrated p-values for each gene, which are then corrected for sample size using the Benjamini-Hochberg (BH) method (Benjamini and Hochberg, 1995). Genes were first filtered to include only those for which 50% of all samples showed an expression value greater than zero (measured in Transcripts Per Kilobase Million; TPM, see section 4.4, methods), providing an initial list of 16,176 genes. Genes were then pre-screened by ANOVA to identify those that showed significant differences in TPM between timepoints. MetaCycle was used to identify genes within this list with an approximately 24-hour period.

Over the full 44-hour time series, 2526 genes were suggested to show significant differences in expression (ANOVA $p < 0.05$), of which 263 were reported to show significant circadian rhythmicity, accounting for multiple testing (Benjamini-Hochberg adjusted $q < 0.05$). However, when the entrainment and free-running days were analysed independently (dividing the 12-point time series into two 6-point datasets), the rhythmic gene list dramatically increased in the diurnal entrainment set (1159 genes). Conversely, considering the free-running day alone resulted in much smaller gene sets passing the ANOVA pre-screen and MetaCycle tests (406 and 26 respectively; full results are summarised in Table 4.1).

Table 4.1: Number of genes reported to show significant expression changes over time and circadian rhythmicity in 12-timepoint dataset and 6-timepoint subsets.

Dataset	No. differentially expressed (ANOVA $p < 0.05$)	No. significantly rhythmic ($q < 0.05$)
Two-day Time Series	2526	263
Entrainment Day	2377	1159
Free-running Day	406	26

Zeitgebers (time givers) are by definition, stronger drivers of rhythms than an endogenous clock, exemplified by a characteristic damping of rhythms seen under free-running conditions (illustrated in Figure 1.2; Chapter 1: Introduction. Discussed in section 1.1). More robust periodicity would therefore be expected under environmental cycles than constant conditions, as is seen in these data. Curiously however, only four genes (*fat-6*, *eif-3.E*, *gstk-1* and *pbo-1*) were identified as significantly rhythmic by MetaCycle through the independent analysis of the entrainment day, free-running day and the full time series, and only a further four (*sws-1*, *ZK185.5*, *K09H9.7* and *grd-5*) were shared between significant gene lists from the two individual days (illustrated in Figure 4.1). This suggests that despite significant periodicity in expression, there is a lack of consistency across the time series, in which free-running gene expression patterns largely do not replicate those that occur under environmental cycles.

Figure 4.1: Overlapping and non-overlapping rhythmic genes between the 12-timepoint time series and 6-timepoint subsets.
Numbers indicate genes with a meta2d q-value < 0.05.

Most of the genes identified as rhythmic over the full time series were also identified as rhythmic when considering the entrainment day alone (234 of 263), while in contrast, only seven of the 263 were identified as rhythmic during the free-running day (Figure 4.1). The stark contrast in overlapping genes between the subsets and the full time series likely implies that the cycling environmental conditions of the entrainment day provided the primary driver of rhythmicity over the full 44-hour time series. Ultimately, the results of the separate analyses are not entirely comparable; the free-running subset contains a smaller starting pool of genes generated by the ANOVA pre-screening approach (Table 4.1), possibly due to an increased variance in expression values without the environmental stimuli (as observed in the RT-qPCR data presented in Figure 3.8A; Chapter 3: Results 1). However, that a majority of the 44-hour oscillating genes could be identified during the entrainment day alone certainly suggests a key role of the environmental stimuli.

A possible explanation for the majority of the rhythms detected in these analyses could therefore be that a large number of genes show expression changes as a direct response to changes in light and temperature, rather than synchronised circadian rhythms. Conversely however, the abundance of genes suggested to oscillate under entrainment conditions could still be a result of a circadian clock; the observed circadian periodicity could be indicative of an organism that is not only responsive to light or temperature, but one that does so with an adaptive rhythm. It could be the case that *C. elegans* has a circadian clock but cannot effectively maintain rhythms endogenously under constant conditions. Given that rapid damping in gene expression rhythms is observed in other animal models of circadian clocks during free-running conditions (see Chapter 1, Figure 1.2), and that these invariant conditions are largely atypical of the natural environment in which evolution has provided organisms with a circadian clock, the idea of a non-robust, endogenously weak or imprecise clock in *C. elegans* could be a conceivable evolutionary outcome.

Despite the clear effect of environmental inputs on the number of observed rhythmic genes, a key result of the MetaCycle analyses is that circadian expression patterns were detected in a small number of genes that were not found to significantly oscillate in entrainment conditions. As shown in Figure 4.1, 29 significant genes were detected over the full time series that were not found to be significantly rhythmic in analysing the entrainment day data alone. A further 15 genes were found to significantly oscillate only within the constant conditions of the free-running day. These significantly-oscillating genes suggest the presence of an endogenous driver in the generation of periodic expression patterns. In addition, the MetaCycle approach used here integrates two community standard approaches to rhythm detection (JTK_Cycle and Lomb-Scargle; Zhang et al., 2014; Refinetti, 2010) and JTK_Cycle in particular has been used in a number of transcriptomic studies in mice and *Drosophila* (Appendices, Tables 6.1 and 6.2). As such, these results represent evidence of circadian rhythms in *C. elegans* detected with a method comparable to reports of circadian oscillations in established models. These data are therefore also

consistent with other literature that suggests the presence of a circadian clock in *C. elegans* that regulates transcription (Goya et al., 2016; Olmedo et al., 2012), and does so on a reduced scale than the mammalian and *Drosophila* models (van der Linden et al., 2010; Zhang et al., 2014; Hughes et al., 2012; Appendices, Tables 6.1, 6.2 and 6.3).

Considering the identities and characteristics of the detected rhythmic genes, a primary observation is that the most significantly rhythmic genes over the full time series generally display a periodicity greater than 24 hours. q-values and period estimates of the 12 most statistically significant genes according to MetaCycle are given in Table 4.2 (those with the smallest Benjamini-Hochberg adjusted meta2d q-values over the full two-day time series; full list given in Appendices, Table 6.4). Most of these genes have periods greater than 24 hours according to estimates generated by MetaCycle and both the JTK_Cycle and Lomb-Scargle approaches individually. MetaCycle parameters were set to identify gene expression patterns with frequencies of 20-28 hours, and most genes according to Table 4.2 are close to this upper limit. This could indicate a slow-running clock in *C. elegans*, perhaps poorly temperature-compensated against the 15°C night and free-running phases.

Table 4.2: Summary of meta2d, JTK_Cycle and Lomb-Scargle q-values (Benjamini-Hochberg adjusted p-values), period estimates and fold change in expression of the most significantly rhythmic genes over 44 hours as identified by meta2d.

Gene	meta2d q-value	JTK_Cycle q-value	Lomb-Scargle q-value	meta2d Period (Hours)	JTK_Cycle Period (Hours)	Lomb-Scargle Period (Hours)	TPM Fold Change
fat-6	0	6.33E-11	0.00114	27.27	28	26.54	4.36
acl-12	1.06E-10	2.95E-07	0.0131	24.06	24	24.11	2.57
R07E5.4	3.52E-08	3.90E-05	0.0135	28	28	28	2.86
maoc-1	9.39E-08	3.90E-05	0.0376	23.76	24	23.51	1.59
col-93	1.72E-07	3.36E-05	0.117	27.13	28	26.26	4.08
lea-1	2.36E-06	0.000242	0.117	26.97	28	25.95	1.74
ruvb-1	5.02E-06	0.000242	0.146	28	28	28	1.29
alh-9	5.26E-06	0.000145	0.237	27.63	28	27.27	1.47
Y82E9BL.18	6.79E-06	0.000471	0.125	28	28	28	3.05
F46F2.3	6.79E-06	0.000145	0.237	28	28	28	2.64
cnc-4	8.21E-06	0.000411	0.184	27.63	28	27.27	7.45
eif-3.E	8.21E-06	0.000471	0.146	28	28	28	1.23

Also apparent within Table 4.2, are differences in significance between the JTK_Cycle and Lomb-Scargle tests that are collectively used to generate the meta2d q-values. Of the 12 most significant MetaCycle genes, all are significant according their JTK_Cycle q-values, but only four are significant according to Lomb-Scargle ($q < 0.05$). This is not a result of the Lomb-Scargle algorithm identifying different genes to JTK_Cycle, as the four most significant genes identified by meta2d are also the four most significant according to Lomb-Scargle, and four of the six most significant according to JTK_Cycle, suggesting the approaches converge on the same significant genes. Rather, the reduced significance in Lomb-Scargle q-values suggest it to be a

more stringent method than JTK_Cycle under the sampling protocol used here. This highlights a substantial impact of algorithm or statistical test choice on the number of transcripts identified as rhythmic. This finding is consistent with a number of other transcriptomic studies; Lomb-Scargle and JTK_Cycle are two of a considerable number of methods and algorithms developed for detecting circadian rhythms (Hughes et al., 2017), and several studies using more than one approach have generally reported differences in the number of genes identified (Hughes et al., 2007; 2009; 2012; Ptitsyn et al., 2011, see Appendices, Tables 6.1, 6.2 and 6.3). As well as stringency, a factor in the small number of rhythmic genes identified by Lomb-Scargle is the effect of sample size; it should be noted that the Lomb-Scargle q-values represent Benjamini–Hochberg adjusted p-values accounting for a large sample of 2256 genes, and individually 155 genes had p-values < 0.05 according to Lomb-Scargle analysis. These observations emphasise that caution should be applied in drawing conclusions from individual approaches and individual transcriptomic studies concerning the overall extent of circadian regulation.

Considering the expression patterns of these significant genes, a number of hallmarks of circadian rhythmicity are apparent. The expression patterns of the six most significantly rhythmic genes according to meta2d q-value are shown in Figure 4.2. In particular, *acl-12* and *maoc-1* (Figure 4.2B and 4.2D) show a dampened repetition of the trend in the entrainment day, with similar phases (a key characteristic of expression rhythms; discussed in Chapter 1, section 1.1). The *acl-12* gene also shows a sufficiently high fold-change in expression (2.57; Table 4.2), such that the observed trend is unlikely to be masked by variation between replicates. From a more general perspective, all of the gene expression profiles shown in Figure 4.2 are consistent with the principle of anticipation; the evolutionary idea that the advantage of endogenous, self-maintaining rhythms is that they allow an organism to prepare itself for an oncoming environmental change and to distribute energetic resources efficiently (Vaze and Sharma, 2013; McClung, 2006). Most genes in Figure 4.2 could be said to fulfil this minimum requirement in that in constant conditions,

largely around the 24-hour timepoint, the trends in gene expression change direction and expression begins to mimic the trends seen in environmental cycles. This incomplete repetition of the environmentally-driven expression pattern in the free-running phase would be in keeping with the idea discussed above of a rapidly damping oscillator in *C. elegans*.

These observations are subjective however, and the case could also be made that the genes do not all appear to be strongly rhythmic. Statistically, the most compelling evidence for circadian rhythmicity is in *fat-6* expression (Figure 4.2A), which meta2d identified as rhythmic over the full time series as well as on the individual and free-running days, and was found to be highly significant by both of the JTK_Cycle and Lomb-Scargle algorithms individually (Table 4.2). Subjectively however, it does not appear to show an overt repeating periodicity of approximately 24 hours. This highlights difficulties in the circadian analysis of sequencing data, in that genes recognised by statistical algorithms may differ from those that appear most rhythmic by observation. As discussed above, and shown in Table 4.2, most of the genes identified here showed long running rhythms with periods greater than 24 hours, and so the most statistically significant results may be genes demonstrating highly significant single cycles rather than the repeating oscillations that would appear most rhythmic by the characteristic of repetition.

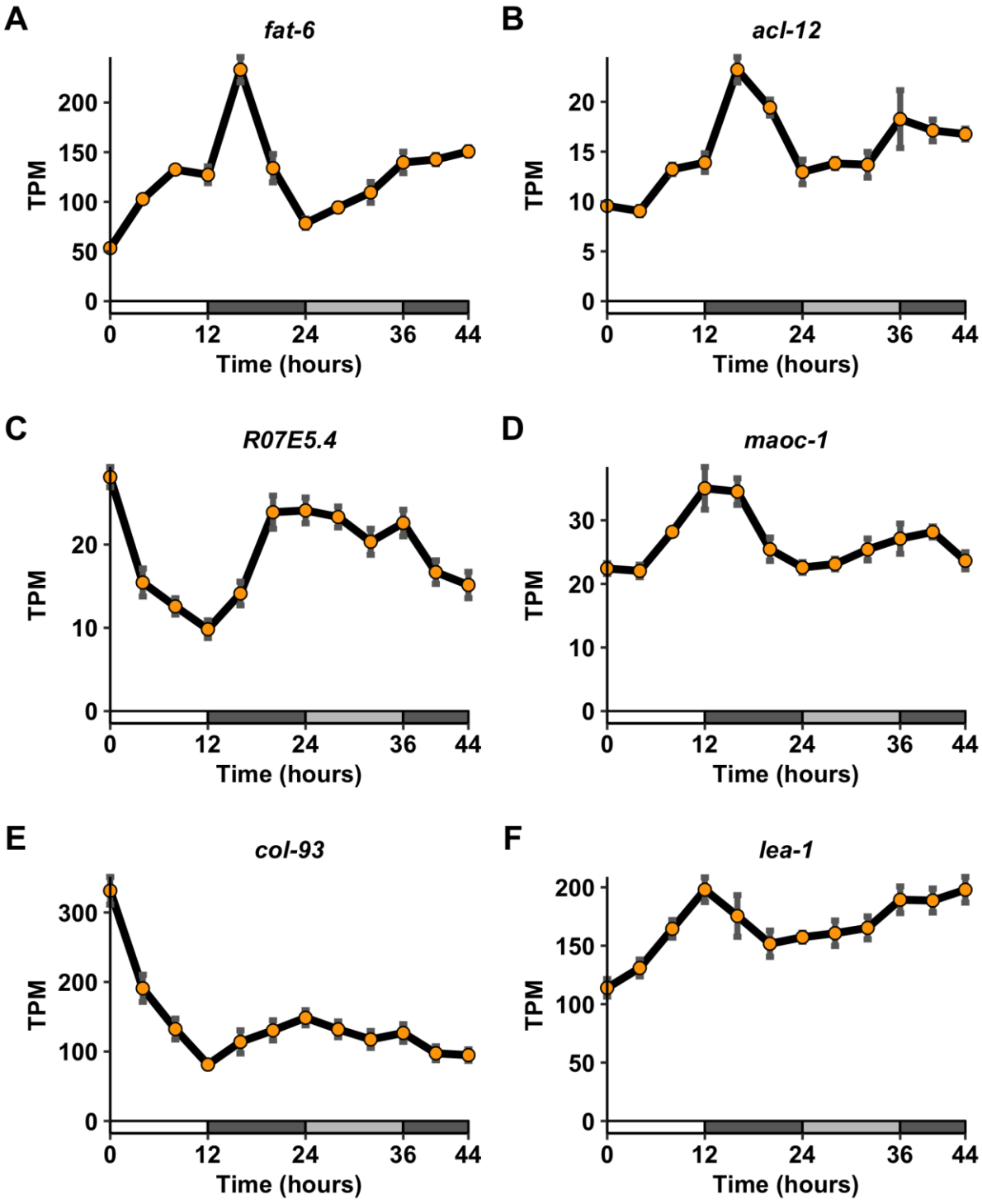


Figure 4.2: Genes identified by MetaCycle as significantly rhythmic over 44-hour time series.

Each point represents the mean Transcripts Per Kilobase Million (TPM) of all replicates at a given timepoint. Error bars show the SEM. Diurnal conditions are indicated by bars at the base of each graph (white = light/20°C, dark grey = dark/15°C, light grey indicates subjective day; see Glossary; dark/15°C). q-values and periodicity listed in Table 4.2.

Also notable from the majority of the genes in Figure 4.2 (A, B, D and F) is a trend previously highlighted with respect to *nlp-36* expression in RT-qPCR experiments (Figure 3.8A in Chapter 3: Results 1): relatively little variation is apparent in the light/warm phase of the experiments, followed by an obvious SEM increase in the dark/cold phases. In measuring trends in gene expression, particularly in this experiment, analytical approaches must detect rhythms in expression profiles that are subject to a myriad of stochastic factors and potential noise; measured expression trends could be influenced by non-circadian gene regulation, feeding and ageing, interactions with other nematodes, and many other variables that are not easily controlled for in *C. elegans*, as well as in other species. As previously discussed, even oscillating core clock genes in established circadian models characteristically dampen in constant conditions (shown here in Chapter 1, Figure 1.2). This provides a challenge for mathematical approaches, which must detect these rhythms and provide effective high-throughput methods for separating rhythmic from non-rhythmic expression in large datasets, while limiting false positive and negative results. In doing so they leave results, with particular reference to those presented in Table 4.2 and Figure 4.2, somewhat open to interpretation, and consequently make it difficult to establish the presence, or absence, of circadian rhythms in *C. elegans*.

4.2.2 Functional analysis of genes identified as rhythmic by MetaCycle

In understanding the biological importance of a potential *C. elegans* circadian clock, a key question is that of the biological functions served by rhythmically expressed genes. Firstly, to gain insights into the primary outputs of the *C. elegans* clock, the Gene Ontology (GO) categories (The Gene Ontology Consortium, 2000; The Gene Ontology Consortium, 2019) of the six most rhythmic genes according to MetaCycle were examined (those genes highlighted in Figure 4.2, Table 4.2). These GO categories highlight a number of diverse biological processes, cellular components and molecular functions potentially served by the *C. elegans* circadian clock (summarised in Table 4.3, GO terms obtained from WormBase WS272, available at

wormbase.org; Lee et al., 2017). These GO terms include involvement in the immune response to Gram-negative bacteria (*maoc-1*), acting as a structural component of the nematode cuticle (*col-93*) and responsiveness to heat, desiccation and hyperosmotic stress (*lea-1*). The hyperosmotic stress response is particularly notable, given that it is a process that has previously been suggested to show endogenous circadian rhythms (Kippert et al., 2002; discussed in Chapter 1, section 1.3.1). The most striking observation from Table 4.3 however, are a large number of ontology terms relating to lipids and particularly, fatty acids. All five genes for which GO terms have been assigned (*fat-6*, *acl-12*, *maoc-1*, *col-93* and *lea-1*) are associated with at least one enzymatic function related to lipid or fatty acid metabolism and a cellular component term relating to phospholipid cellular membranes.

Table 4.3: Gene Ontology categories of the most significant genes identified by MetaCycle.

Gene	GO category	GO Terms
fat-6	Biological Process	GO:0006629 lipid metabolic process, GO:0006631 fatty acid metabolic process, GO:0006633 fatty acid biosynthetic process, GO:0006636 unsaturated fatty acid biosynthetic process, GO:0007275 multicellular organism development, GO:0042759 long-chain fatty acid biosynthetic process, GO:0045087 innate immune response, GO:0055114 oxidation-reduction process.
	Cellular Component	GO:0005789 endoplasmic reticulum membrane, GO:0016020 membrane, GO:0016021 integral component of membrane, GO:0030176 integral component of endoplasmic reticulum membrane.
	Molecular Function	GO:0004768 stearoyl-CoA 9-desaturase activity, GO:0005506 iron ion binding, GO:0016491 oxidoreductase activity, GO:0016717 oxidoreductase activity, acting on paired donors, with oxidation of a pair of donors resulting in the reduction of molecular oxygen to two molecules of water.
acl-12	Biological Process	GO:0006629 lipid metabolic process, GO:0006644 phospholipid metabolic process, GO:0008654 phospholipid biosynthetic process, GO:0016024 CDP-diacylglycerol biosynthetic process, GO:0036498 IRE1-mediated unfolded protein response, GO:0036499 PERK-mediated unfolded protein response.
	Cellular Component	GO:0016020 membrane, GO:0016021 integral component of membrane.
	Molecular Function	GO:0003841 1-acylglycerol-3-phosphate O-acyltransferase activity, GO:0016740 transferase activity, GO:0016746 transferase activity, transferring acyl groups.
R07E5.4	Biological Process	None.
	Cellular Component	None.
	Molecular Function	None.

maoc-1	Biological Process	GO:0006635 fatty acid beta-oxidation, GO:0045087 innate immune response, GO:0050829 defence response to Gram-negative bacterium.
	Cellular Component	GO:0005778 peroxisomal membrane.
	Molecular Function	GO:0016508 long-chain-enoyl-CoA hydratase activity, GO:0042802 identical protein binding.
col-93	Biological Process	None.
	Cellular Component	GO:0004768 stearoyl-CoA 9-desaturase activity, GO:0005506 iron ion binding, GO:0016491 oxidoreductase activity, GO:0016717 oxidoreductase activity, acting on paired donors, with oxidation of a pair of donors resulting in the reduction of molecular oxygen to two molecules of water.
	Molecular Function	GO:0042302 structural constituent of cuticle.
lea-1	Biological Process	GO:0006869 lipid transport, GO:0006972 hyperosmotic response, GO:0009269 response to desiccation, GO:0009408 response to heat, GO:0042157 lipoprotein metabolic process.
	Cellular Component	GO:0005576 extracellular region.
	Molecular Function	GO:0008289 lipid binding.

While GO terms are in many cases inferred and based on homology, three of these five genes have been directly studied in *C. elegans* literature, with two shown to have key functions in fatty acid metabolism. Of the two genes that showed the clearest repeating expression patterns, *acl-12* and *maoc-1* (Figure 4.2B,D; discussed in section 4.2.1), the former has not been directly studied in published literature, while *maoc-1* has a defined role in a fatty acid catabolism pathway, with loss-of-function mutants having been shown to have increased levels of branched-chain fatty acids (Zhang et al., 2010; Wang et al., 2013). Secondly, the most significantly rhythmic gene according to MetaCycle, *fat-6* (Table 4.2), is one of three delta-9 desaturase genes, which are compensatory, but collectively vital for survival (Brock et al., 2006). The potential importance of the *fat-6* gene to health is further evidenced by studies that have found normal *fat-6* expression to be key to pharmacological inhibition of fat accumulation (Peng et al., 2016), and from longevity experiments; constitutive expression of the *xbp-1s* transcription factor has been shown to increase lifespan through increased lipase activity, and *fat-6* expression is upregulated with, and required for, full extension of lifespan (Imanikia et al., 2019). In addition to these genes, another of the five genes with assigned GO terms, *lea-1*, has not been directly studied in *C. elegans* with regards to functions in lipid metabolism, but has been in terms of its involvement in responses to desiccation, having been shown to be transcriptionally upregulated in dehydration conditions, and when silenced, to result in reduced survival to osmotic or heat stress (Gal et al., 2004). This gene could therefore play a role in the previously reported circadian rhythm in survival to hyperosmotic stress conditions (Kippert et al., 2002). These direct investigations into *maoc-1*, *fat-6* and *lea-1* function, serve to highlight specific processes, important for health and survival, that are likely to be impacted by a *C. elegans* circadian clock, based on the MetaCycle data presented here.

Interestingly, a number of these genes have homologues in other species that have been studied in published literature, most notably including mammalian model systems and humans, with conserved or similar roles in lipid metabolism (based on

WormBase WS272 BLASTP matches; Lee et al., 2017). The *C. elegans acl-12* and *maoc-1* genes are orthologues of both mouse and human *Lpgat1*/*LPGAT1* (lysophosphatidylglycerol acyltransferase 1) and *Hsd17b4*/*HSD17B4* (hydroxysteroid 17-beta dehydrogenase 4) respectively, with both human orthologues having roles in fatty acid metabolism (Yang et al., 2004; Möller et al., 1999). *C. elegans fat-6* is also a conserved gene, with orthologous delta-9 desaturase genes in mammals including human *SCD* and *SCD5* and *Scd3* in mice, all of which function in fatty acid metabolism (Ntambi and Miyazaki, 2004; Sinner et al., 2012). Additionally, the mouse orthologue of *lea-1*, *Plin4* (*PLIN4* in humans), has also been studied in the context of lipid metabolism, having a role in lipid droplet formation (Čopič et al., 2018).

Given that several of the most significantly rhythmic *C. elegans* genes have roles in lipid metabolism and have mammalian homologues with established similar or conserved roles, an immediate question is whether lipid metabolism might be a process governed by the circadian clock that is conserved between *C. elegans* and other species. To investigate mammalian homologues in previous transcriptomic data, the closest mouse (*Mus musculus*) homologues of the six most significantly rhythmic *C. elegans* genes were entered into CircaDB, a database containing 13 transcriptomic gene expression datasets from circadian time series analysis of various mouse tissues (Pizarro et al., 2013; available at: http://circadb.hogeneschlab.org). Strikingly, five of the six genes were found to be rhythmic with 20 to 28-hour periods based on JTK_Cycle results from previous time series (Table 4.4). As shown in Table 4.4, some of these genes were also found to be significantly rhythmic according to Lomb-Scargle analysis, although comparably to the *C. elegans* results described in this chapter, many fall below 5% significance when p-values are corrected for multiple testing. CircaDB also provides analysis of transcriptomic data in 13 human tissues (albeit using a different statistical approach and artificially ordered samples, as time series sampling is not possible in humans; Ruben et al., 2018). In this case, CircaDB suggested significant oscillations occur in three of the human homologues of the six queried genes: *SCD5* (homologous to *C.*

elegans fat-6), *HSD17B4* (*maoc-1*) and *COL24A1* (homologous to *col-93*, the only queried gene without a reported rhythmic mouse homologue). Further, while less comprehensively studied, the *Drosophila* homologues of two of the six genes in question have also been found to show circadian expression patterns in primary literature. The *Drosophila* genes *desat1* and *CG13185*, homologous to *C. elegans fat-6* and *lea-1* respectively, have all been found to show circadian oscillations in published literature (Krupp et al., 2013; Xu et al., 2011). The fact that all six of the most significantly rhythmic genes identified in *C. elegans* in this work have mammalian homologues that have been suggested to show circadian rhythms, and two have *Drosophila* homologues, suggests that key circadian clock-governed processes in established models, particularly as relates to lipid metabolism, may be indeed be conserved in *C. elegans*.

Table 4.4: Homologous genes found to show significant circadian rhythms in *C. elegans* and *M. musculus*.

***C. elegans* Gene**	***M. musculus* Homologue**	***M. musculus* Tissue**	**JTK_ Cycle p-value**	**JTK_ Cycle q-value**	**JTK_ Cycle period (hours)**	**Lomb-Scargle p-value**	**Lomb-Scargle q-value**
fat-6	*Scd3*	Pituitary	1.35e-07	8.14e-05	25.0	0.00224	0.379798
		Lung	0.000681626	0.0160184	28.0	0.0186	0.468372
		Pituitary	0.000103491	0.0178833	23.0	0.0109	0.833222
acl-12	*Lpgat1*	Liver	1.72e-09	1.68e-07	24.0	7.23e-06	0.000770121
		Liver	9.05964e-05	0.00229107	24.0	0.019	0.297868
R07E5.4	*Ifi30*	Liver	0.00302354	0.0315263	26.0	0.0474	0.466857
maoc-1	*Hsd17b4*	Lung	0.00043513	0.0112848	26.0	0.0278	0.509777
lea-1	*Plin4*	Aorta	4.39243e-06	0.00109558	24.0	0.0068	0.77544
		Adrenal gland	0.000584053	0.0416164	22.0	0.017	1.0

To investigate whether lipid metabolism and the other functional roles highlighted in discussed above are representative of the full list of rhythmic genes identified by MetaCycle, and to further examine the principal biological functions that may be clock-governed in *C. elegans*, functional annotation was performed using DAVID (Database for Annotation, Visualization and Integrated Discovery; Huang et al., 2009a; Huang et al., 2009b). DAVID combines GO terms (The Gene Ontology Consortium, 2000; The Gene Ontology Consortium, 2019), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kanehisa et al., 2016) and other functional descriptors in producing annotation clusters. These clusters are given an enrichment score, a ranking based upon the average significance of the annotation terms within. Here, DAVID was used to identify processes and pathways enriched in the two-day

significantly rhythmic MetaCycle dataset relative to the transcriptome as a whole (all 16,176 genes found to be expressed in at least 50% of sequenced samples). Figure 4.3 provides a simplified illustration of the results of this analysis, highlighting just the significant GO terms (which are generally representative of clusters as a whole) in clusters for which ≥ 50% of all annotation terms had an FDR q-value < 0.1 (refer to section 4.4, methods).

The two annotation clusters that met these criteria suggest two quite different functional categories are significantly overrepresented in the rhythmic gene list. The first cluster contains genes relating to ribosome biogenesis and nucleolar processes, possibly suggesting that global translation is diurnally regulated in *C. elegans*. The second cluster contains broad groups of genes relating to biosynthesis and the enzymatic processing of drugs and carbohydrates, potentially suggesting important diurnal regulation of the processing of ingested material. The genes within these annotation groups could be directly circadian-regulated, or perhaps show indirect diurnal regulation through rhythmic necessity (*C. elegans* potentially being relatively metabolically active and dormant at different times throughout the day, leading to an upregulation of necessary genes).

These results also suggest that the predominant lipid-metabolism related roles discussed with regards to the most significant genes are not reflective of the rhythmic dataset as a whole, with the only similarities relating to the structure of membranes and general processes of metabolism. Given that DAVID only identified two enrichment clusters with FDR < 0.1 for ≥ 50% of annotation terms, this analysis could indicate that the 263 genes identified as rhythmic are highly variable in their functions, performing an array of roles rather than relating to key clock-regulated processes. With respect to these results, it should be noted that in other organisms, a substantial proportion of the genome exhibits circadian rhythms, across different tissues with different roles (Zhang et al., 2014). As such, examining whole-organism overrepresentation of genes may also miss some key tissue-specific roles of circadian

clocks.

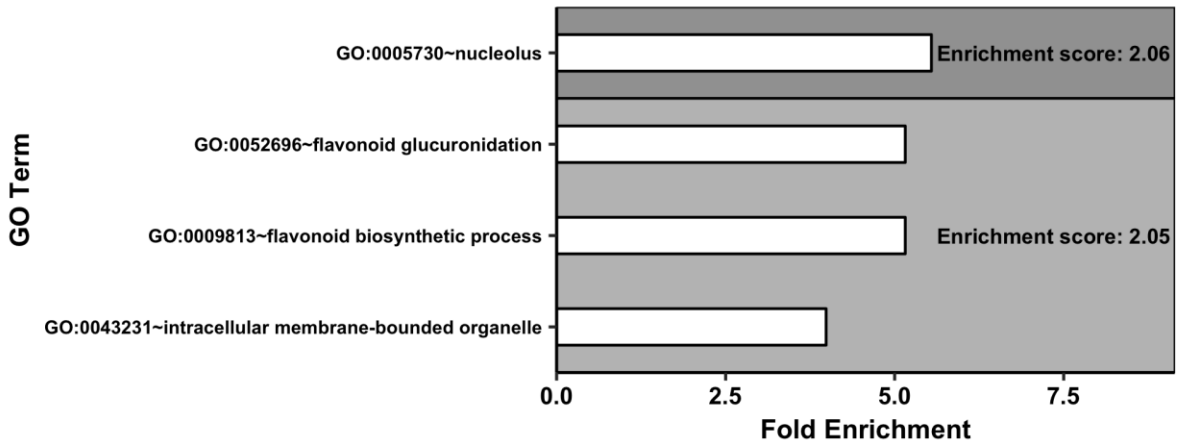


Figure 4.3: Significant GO terms enriched in MetaCycle data.
DAVID analysis clusters depicted are all clusters for which ≥ 50% of all annotation terms had an FDR < 0.1. GO terms displayed are those with an FDR < 0.1. Fold enrichment indicates the ratio of the proportion of genes within the meta2d dataset that relate to each GO term relative to the proportion within all expressed genes. Enrichment scores rank clusters based upon the average significance of the annotation terms within each annotation cluster.

4.2.3 Comparison of MetaCycle data to previous transcriptomic analysis

As previously discussed in section 4.1, while this work represents the first RNA-seq circadian time series in *C. elegans*, one previous microarray study was performed entraining nematodes to warm:cold (WC) cycles and light:dark (LD) cycles in separate experiments (van der Linden et al., 2010). Transcriptomic studies within the same organisms are notable for producing lists of circadian genes with limited overlapping genes (discussed later in Chapter 5. Discussion, section 5.3.1), including in the *C. elegans* microarrays, for which only two genes were found to show rhythms in constant conditions following both light and temperature entrainment protocols. However, studies in the established *Drosophila* and mammalian models, despite largely divergent gene lists, do typically retrieve core genes involved in the oscillating TTFL (those illustrated in Chapter 1, Figure 1.1; Keegan et al., 2007; Zhang et al., 2014, Appendices, Tables 6.1 and 6.2). As such, strong candidates for oscillating genes directly involved in circadian timing in *C. elegans* might be expected to co-

occur in both the MetaCycle data described here and the previous microarray time series.

A small, proportion of genes identified by MetaCycle were also found to oscillate in the previous study. Shown in Table 4.5, 27 genes of the 263 that comprise the full two-day MetaCycle dataset (a proportion of 10.3%) also occur in one of six gene lists produced by the previous microarrays. These previous gene lists are based on analysis of appended entrainment data, free-running data and both in combination using either light or temperature, and amount to approximately 3267 transcripts (including some co-occurring genes). The majority of genes that are overlapping between studies appear in the entrainment-driven rather than free-running gene sets of the prior work; in particular 22 of the 27 overlapping genes were found to oscillate in WC cycles in the microarray study, but no genes uncovered in this study overlap with the WC + CC gene set. This is perhaps expected, given that in the prior work, the entrained microarray gene sets were much larger than those identified in constant conditions, as was overwhelmingly the case in the RNA-seq MetaCycle data presented here (Figure 4.1, discussed in section 4.2.1). This does raise the question as to whether these genes are clock-regulated or directly driven in their expression by light or temperature. However, any tightly clock-regulated genes would also be expected to oscillate under diurnal cycles, and given that the external environment sets the phase of circadian rhythms, would be expected to oscillate with higher significance in these conditions.

Within the set of 27 overlapping genes are a number of findings of interest. Firstly, five of the 12 most significant genes (Table 4.2) identified by meta2d analysis over the two-day time series appear on this list of 27 genes (highlighted in green in Table 4.5). Secondly, one of the four genes identified within the two-day time series and the MetaCycle analysis of the individual days (highlighted in purple, Figure 4.1). The convergence of two studies with different entrainment conditions and different analysis approaches on these genes suggests that they would be strong candidates

for further study of the circadian clock, potentially in the development of circadian reporters. That two *C. elegans* studies would show limited convergence but still detect some of the most significant genes is also consistent with transcriptomic studies in other models in identifying core clock genes (Keegan et al., 2007). These six genes could therefore be regarded as candidates for direct involvement in the regulation of the core circadian clock and targets for further study.

Table 4.5: All genes identified by meta2d analysis of the two-day time series that were also previously identified in a microarray time series.
Previous data taken from van der Linden et al., 2010.

Gene	meta2d q-value	Prior Study Cycling Category	Prior Study F_{24} p-value
ttr-30	0.007241773	LD cycles + WC cycles	0.0002 (LD) 0.0042 (WC)
clec-1	0.000295488	LD cycles	0.0147
abce-1	0.000176173	WC cycles	0
acl-12	1.06E-10	WC cycles	0.0002
acs-7	0.000709098	WC cycles	0.0065
anmt-3	0.012255521	WC cycles	0
aqp-4	0.000873584	WC cycles	0.0001
cnc-4	8.21E-06	WC cycles	0
col-142	0.032574113	WC cycles	0.0022
cyp-33C9	0.00514216	WC cycles	0
dct-11	0.001722106	WC cycles	0.0013
dhs-28	0.001301404	WC cycles	0
dod-20	0.0043219	WC cycles	0
fbxa-163	0.001277371	WC cycles	0.0005
fkh-6	0.047393313	WC cycles	0.0039
maoc-1	9.39E-08	WC cycles	0
mct-6	0.000202921	WC cycles	0
pbo-1	0.002898404	WC cycles	0
pgp-1	0.00956616	WC cycles	0.0147
pgp-9	0.010567933	WC cycles	0
ppw-2	0.003495263	WC cycles	0.0016
ruvb-1	5.02E-06	WC cycles	0.0002
ugt-63	0.01907707	WC cycles	0
col-93	1.72E-07	Combined LD and DD cycles + DD alone	0.0001 (LDDD) 0.0004 (DD)
col-94	0.000201034	Combined LD and DD cycles	0.0008
hpo-6	0.002898404	Combined LD and DD cycles	0.0005
acp-6	0.000495758	CC alone	0.0164

Three of these six genes highlighted in Table 4.5, *acl-12*, *maoc-1* and *col-93* were previously discussed in section 4.2.2, with the former two having been found to have rhythmic homologues in mammals. Of the remaining three, *pbo-1* and *ruvb-1* have homologues in other phyla, while *cnc-4* is nematode specific (WormBase WS272, Lee

et al., 2017). *Ruvbl1*, the mammalian homologue of *ruvb-1*, was also found to be rhythmic according to CircaDB (in mouse liver, JTK_Cycle p-value: 2.26744e-05, period: 26 hours; Pizarro et al., 2013; http://circadb.hogeneschlab.org).

In considering candidates for further study, *acl-12*, *maoc-1* and *ruvb-1*, the three *C. elegans* genes that were identified in two transcriptomic experiments and have significantly oscillating mammalian homologues, should all be immediate candidates for further investigation. However, one *C. elegans* gene in particular, *acl-12*, and its *M. musculus* homologue *Lpgat1*, are notable in showing similar expression patterns with similar phases in this time series and the two mouse time series shown in Table 4.4 (with a bathyphase around four hours into the day, and an acrophase around four hours into the night; nematode data shown in Figure 4.2, mouse data available at: http://circadb.hogeneschlab.org). The highly statistically significant rhythmicity in *acl-12*/*Lpgat1* expression, as identified in both *C. elegans* transcriptomic studies (Table 4.5) and in two mammalian experiments (Table 4.4), coupled with this similarity in expression patterns across different organisms, and the fact that the gene exhibits one of the clearest repeating circadian profiles of any gene in this work, with a high, 2.57-fold change (Figure 4.2D; Table 4.2), altogether suggest *acl-12* is likely the strongest candidate for further study identified in this RNA-seq experiment.

4.2.4 MetaCycle analysis indicates an absence of rhythms in TTFL gene homologues

The high degree of conservation of core TTFL genes between *Drosophila*, mice and other models has been central to understanding circadian biology in the metazoa, and as discussed above, transcriptomic approaches do often converge on those relating to the core clock (Keegan et al., 2007; Zhang et al., 2014). Prior work in *C. elegans* has not found TTFL homologues to oscillate however (van der Linden et al., 2010; Olmedo et al., 2012). Here, most known conserved clock homologues and associated genes were found to be expressed (meeting the criterion of reads being detected in > 50% of RNA-seq samples) but were also found not to show significant changes in expression (ANOVA p-values > 0.05). This included most of the major *C.*

elegans homologues of TTFL-associated genes detailed in Chapter 1, Table 1.1 (*lin-42*, *tim-1*, *aha-1*, *ces-2*, *kin-20*, *nhr-23*, *nhr-85*), along with *pdf-1* (*nlp-74*) and *pdf-2* (*nlp-37*), homologues of *Drosophila* PDF previously suggested to be necessary for *C. elegans* circadian locomotor rhythms (Herrero et al., 2015; discussed in Chapter 1, section 1.2.1). Two exceptions to this were *atf-2*, a homologue of the *Drosophila* secondary-loop gene *vrille*, and *pdfr-1*, the *C. elegans* PDF receptor, which were found to show significant expression changes over time in the RNA-seq data (ANOVA $p < 0.05$). However, neither of these genes showed any indication of rhythmicity by JTK_Cycle, Lomb-Scargle or meta2d integrated q-values (meta2d q-value = 1).

This lack of significant oscillations in TTFL homologues, while in concordance with prior work (van der Linden et al., 2010; Olmedo et al., 2012), is a substantial result because it makes a strong implication that circadian rhythmicity in *C. elegans* does not have a genetically conserved basis. As discussed in Chapter 1, section 1.1.1, the TTFL, while varying in some respects between models, shares key conserved elements that the last common ancestor of *C. elegans,* mice and *Drosophila* would be expected to express (based on molecular phylogenetics; Borner et al., 2014). The expected basis for rhythms in *C. elegans* would therefore be an at least partially conserved one. The *Drosophila*/mammalian *period*/*Per* genes are of particular importance because their rhythmic expression is vital to the TTFL model in *Drosophila,* mice and other models (in conjunction with binding partners; section 1.1.1). The fact that *lin-42*, the *C. elegans period*/*Per* homologue does not oscillate according to this and prior work, would seem to suggest the transcriptional circadian outputs observed here have a molecular basis outside that of a conserved TTFL.

4.2.5 Weighted gene correlation network analysis reveals diurnally responsive gene expression and new oscillating transcripts

A fundamental further consideration in quantifying the extent of rhythmic gene expression in the *C. elegans* transcriptome is that sinusoid-detection approaches, like MetaCycle, may ultimately have limited scope in identifying rhythmic genes. Many

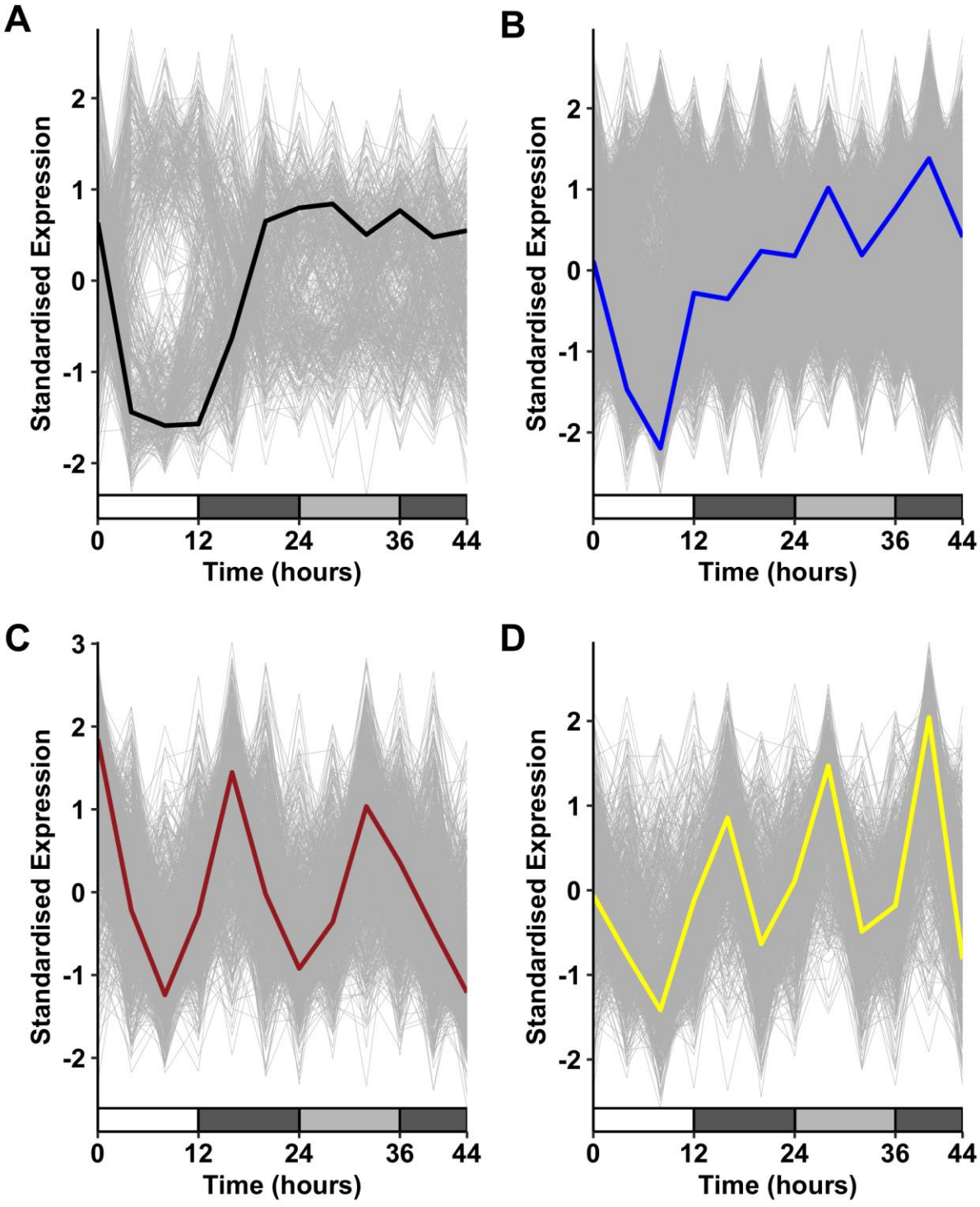
approaches have been developed that principally identify sinusoidal trends in data (Hughes et al., 2017), and MetaCycle was used in the above analyses based on its integration of two different approaches to maximise the accuracy of rhythm detection. However, the fundamental requirement of a circadian expression pattern is a repetitive oscillation rather than a sinusoidal one. This is the principal difference between the statistical approaches employed by MetaCycle and subjective observations of the data discussed in section 4.2.1 with regards to Figure 4.2. Sinusoid detection would be expected to detect a core clock, with transcriptomic studies in *Drosophila*, mice and other organisms generally successfully detecting the oscillatory TTFL genes (Chapter 1, Figure 1.1) through JTK_Cycle (and other similar approaches, see Appendices, Tables 6.1, 6.2, 6.3). However, such approaches may miss downstream genes that are non-sinusoidal, yet still diurnally rhythmic. Examples of these expression patterns identified in other studies include genes with twin-peaking patterns over the daily cycle (Pembroke et al., 2015) and pulsatile spikes in expression with a 24-hour frequency (Rund et al., 2013). Identifying different approaches for detecting circadian rhythms may therefore be essential to understanding rhythmic gene expression in *C. elegans*.

To move beyond sinusoid detection and further investigate diurnal patterns of expression in the RNA-seq dataset, I utilised weighted gene co-expression network analysis (WGCNA; Langfelder and Horvath, 2008). WGCNA is a multifaceted approach to mining big, multidimensional datasets based on pairwise correlations. Here, it was used as a data reduction method to identify prevailing gene expression trends over time, without making prior assumptions as to the expression patterns. To do this, WGCNA was used to generate an adjacency matrix of all genes based on their pairwise co-expression (absolute Pearson correlation coefficient) across the 71 samples (excluding outliers). The data were then separated into modules (clusters) using a hierarchical clustering approach, generating groups of genes based on the similarity of their expression profiles.

WGCNA analysis of the full two-day dataset (16,176 genes, those for which 50% of samples had > 0 TPM) was used to generate 13 modules, each named with an assigned colour. Each module is represented by an eigengene, with standardised expression values generated from the values of the first principle component for each sample within each module, averaged by timepoint. Eigengenes therefore indicate the dominant expression trends of the genes contained within modules. The number of genes assigned to each module is given in Table 4.5, while the expression patterns of the eigengenes and the genes within each model are shown in Figure 4.4. Many of the eigengenes suggest expression patterns that could offer insights into temporal variation in *C. elegans* gene expression, and four were selected for further analysis on the basis of showing trends of particular interest to a circadian experiment. Two of these modules appear to exhibit diurnal regulation (Black and Blue, Figure 4.4A and 4.4B), and two appear to show ultradian rhythmic oscillations (Brown and Yellow, Figure 4.4C and 4.4D).

Table 4.6: List of WGCNA modules and numbers of genes therein.

Module	**No. genes**
Black	511
Blue	2499
Brown	1244
Green	615
Greenyellow	137
Magenta	313
Pink	368
Purple	222
Red	561
Salmon	63
Tan	137
Turquoise	5808
Yellow	803
Grey (no module)	3435



(Part E Overleaf)

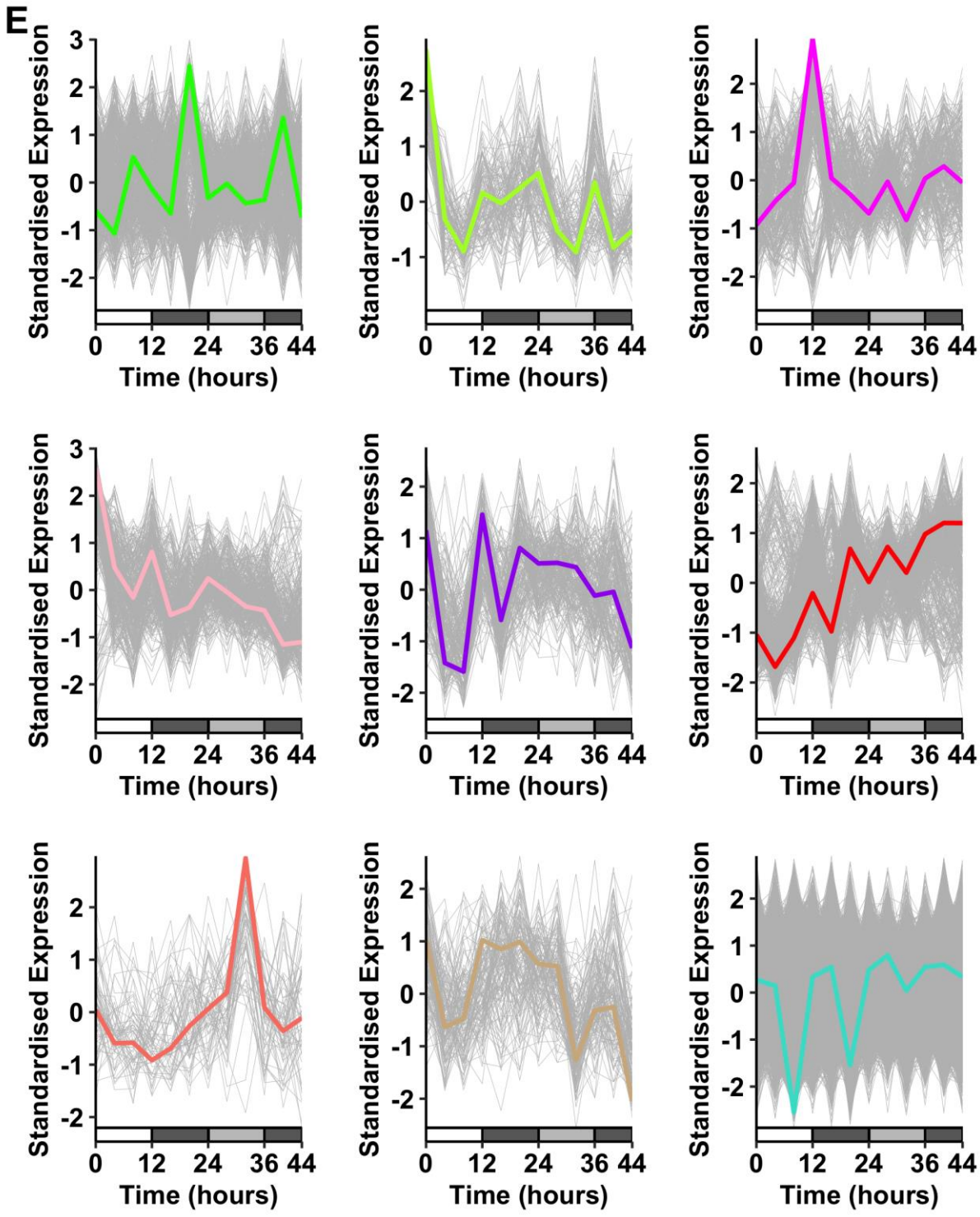


Figure 4.4: Standardised expression of genes by WGCNA module over time.
Grey lines indicate standardised average expression of genes within each module, coloured lines indicate the module colour and represent the module eigengene, averaged by timepoint.

The Black and Blue modules appear to indicate sets of genes that are diurnally expressed, but only under an environmental cycle. The eigengenes show a decline through the day (light and warm conditions), and an increase at the onset of night (dark and cold), after which expression either remains steady (Black, Figure 4.4A), or shows comparatively reduced expression changes (Blue, Figure 4.4B). These modules appear to highlight sets of genes that are downregulated or inhibited by light, temperature or both. Additionally, in both modules the decline and restoration of expression levels occurs close to a 24-hour period, potentially indicating diurnally-regulated gene expression. This pattern is somewhat comparable to the MetaCycle results in which a large number of genes were suggested to show oscillations across the entrainment phase but not in the free-running phase (Table 4.1). These modules could support the prior discussed idea of a non-robust circadian clock; that the genes within these modules show circadian regulation, controlled by a weak endogenous oscillator that is unable to sustain rhythms in constant conditions.

In contrast, the Brown and Yellow modules (Figure 4.4C and 4.4D) are represented by eigengenes that oscillate throughout the entire time series, potentially providing evidence of true oscillatory gene expression in the adult nematode. Notably, trends in the expression of the eigengenes spanning multiple time-points are apparent, rather than simple fluctuations in expression between single timepoints. In the Brown module, these form a 16-hour repeated pattern over the first 36 hours, while in the Yellow module, periodicity fluctuates between 16 and 12 hours. These modules could both indicate the presence of novel ultradian rhythms in nematodes or alternatively, these results could suggest a circadian clock that is short-running in the dark and cold conditions used here. This latter notion would provide a contrast with the most significant genes identified by MetaCycle, in which predominantly long-running periods were reported (Figure 4.2)

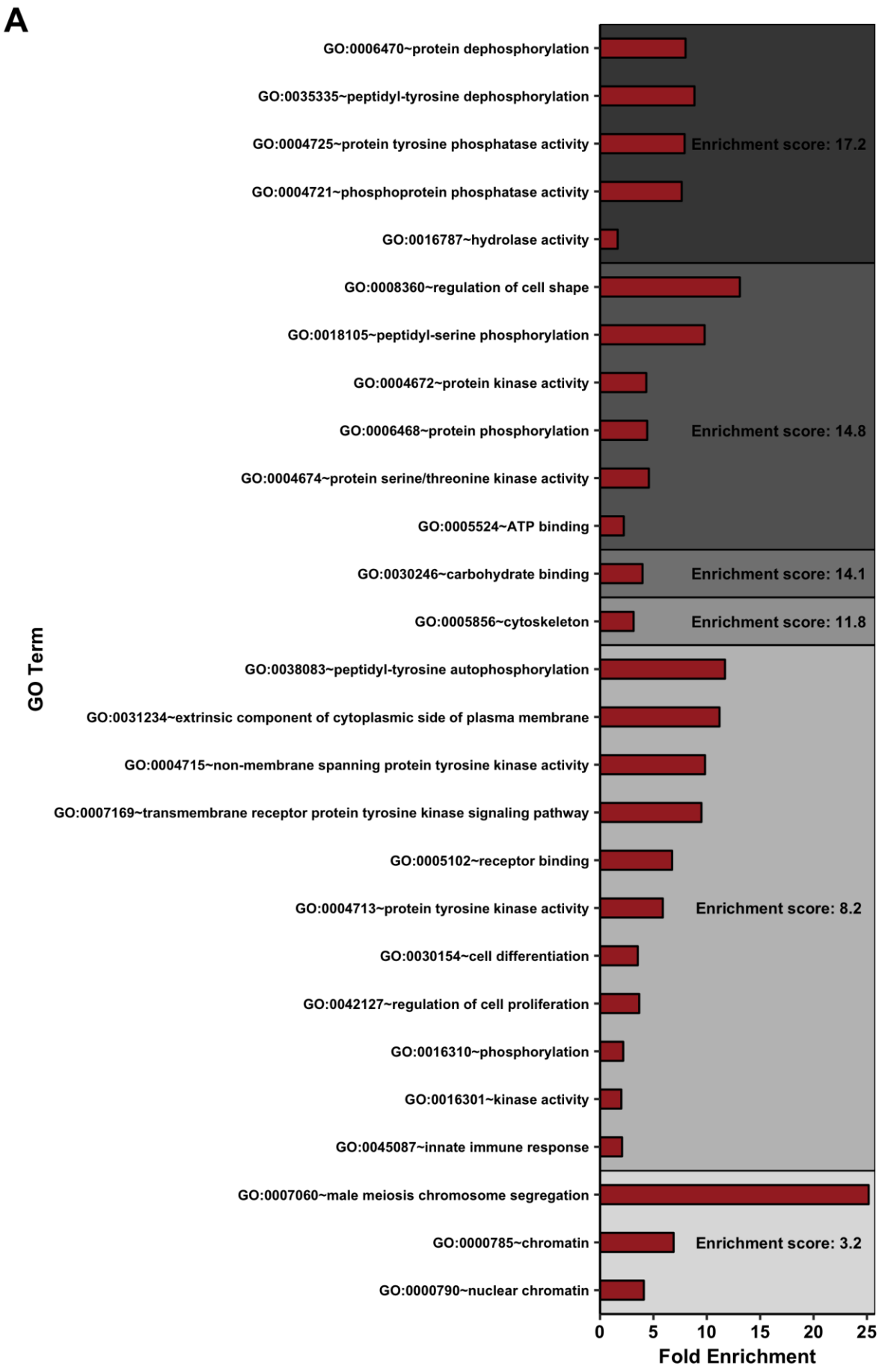
4.2.6 Functional analysis of genes in Brown and Yellow modules

To further interrogate the ultradian oscillations seen in the Brown and Yellow modules, DAVID analysis (described In detail in section 4.2.2 above) was performed to search for enriched gene functions in these subsets relative to all expressed genes. Figure 4.5 shows enrichment clusters for the Brown and Yellow modules for which ≥ 50% of all annotation terms had an FDR q-value < 0.1, highlighting just the GO terms (as described in section 4.2.2).

Regarding the Yellow module, this approach resulted in annotation clusters containing genes relating to cell membrane components and cell communication. As such, genes relating to these broad functions are overrepresented in the Yellow module and these functions may be regulated in part by the ultradian oscillatory expression patterns indicated by the module eigengene (Figure 4.4D). While these results are statistically significant and do warrant further study, some caution should be applied in interpreting (and not over-interpreting) the biological significance of these functional enrichment clusters. In the case of this module, DAVID analysis may offer only limited insights into the functions of the oscillations in eigengene expression, as the numbers actual genes involved are relatively small. For example, the most enriched GO category shown in Figure 4.5B, cell communication, represents an 8.37-fold enrichment, but represents only five genes (Appendices, Table 6.7). The cell communication cluster does contain other gene categories (only GO terms are depicted in Figure 4.5 for graphical simplicity; others shown in Appendices, Table 6.7), but these contain similarly low numbers of genes. Additionally, the low overall enrichment score for the cluster is indicative of the fact that not all annotation terms within the enrichment cluster are statistically significant. The enrichment score in the cluster relating to cell membranes is much larger (6.15), indicating more consistent significance between terms, but the GO terms themselves are only enriched by a factor of < 1.5, meaning the genes involved are not hugely overrepresented in this module, relative to the genome-wide average. Overall, DAVID analysis of the Yellow

module therefore highlights cell membrane components and cell communication as potential areas in which oscillating gene expression is important, and provides lists of genes for further investigation, but likely has not explained a wider function of the oscillations.

In contrast to the Yellow module, DAVID analysis of the Brown module returned a raft of significant functional categories within annotation clusters with much higher enrichment scores. These enrichment clusters also appear to converge on similar functions; the two clusters with maximum enrichment scores (Figure 4.5B) contain terms relating to phosphorylation and dephosphorylation of proteins (with a third cluster containing terms relating to tyrosine kinase signalling also being highly enriched). These terms are of particular interest to a chronobiological study in that in circadian model species, phosphorylation events via a number of accessory proteins are thought to maintain the 24-hour periodicity TTFL (Isojima et al., 2009; Top et al., 2018). Unlike the Yellow module, significant FDR values are apparent across the functional terms in each group (Appendices, Tables 6.6 and 6.7) and almost all functional terms in the top three clusters contain > 20 genes, with some being considerably higher. The genes contained in these modules and their roles in phosphorylation and dephosphorylation are therefore substantially enriched in this module and provide a target area for further study.


A
GO:0006470~protein dephosphorylation
GO:0035335~peptidyl-tyrosine dephosphorylation
GO:0004725~protein tyrosine phosphatase activity
GO:0004721~phosphoprotein phosphatase activity
GO:0016787~hydrolase activity
Enrichment score: 17.2
GO:0008360~regulation of cell shape
GO:0018105~peptidyl-serine phosphorylation
GO:0004672~protein kinase activity
GO:0006468~protein phosphorylation
GO:0004674~protein serine/threonine kinase activity
GO:0005524~ATP binding
Enrichment score: 14.8
GO:0030246~carbohydrate binding
Enrichment score: 14.1
GO:0005856~cytoskeleton
Enrichment score: 11.8
GO:0038083~peptidyl-tyrosine autophosphorylation
GO:0031234~extrinsic component of cytoplasmic side of plasma membrane
GO:0004715~non-membrane spanning protein tyrosine kinase activity
GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway
GO:0005102~receptor binding
GO:0004713~protein tyrosine kinase activity
GO:0030154~cell differentiation
GO:0042127~regulation of cell proliferation
GO:0016310~phosphorylation
GO:0016301~kinase activity
GO:0045087~innate immune response
Enrichment score: 8.2
GO:0007060~male meiosis chromosome segregation
GO:0000785~chromatin
GO:0000790~nuclear chromatin
Enrichment score: 3.2
GO Term
0
5
10
15
20
25
Fold Enrichment

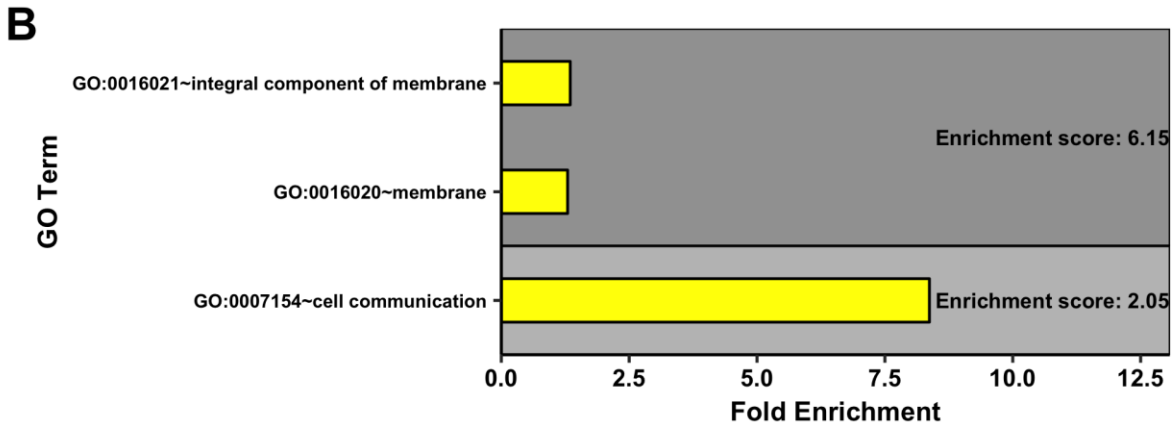


Figure 4.5: Significant GO terms enriched in WGCNA modules.
A) Brown and **B)** Yellow DAVID analysis clusters depicted are all clusters for which ≥ 50% of all annotation terms had an FDR < 0.1. GO terms displayed are those with an FDR < 0.1. Fold enrichment indicates the ratio of the proportion of genes within the respective module gene lists that relate to each GO term relative to the proportion within all expressed genes. Enrichment scores are a rank clusters based upon the average significance of the annotation terms within each annotation cluster.

4.2.7 Representativeness of module eigengenes

In assessing the biological significance of the expression patterns and functions indicated by WGCNA modules, a vital question not yet addressed is whether the eigengenes reflect the actual expression patterns of the genes within the respective modules. Given that eigengenes represent the average first principal component value at each timepoint, they should be representative of a proportion of the variation within the data, but further examination is required to determine the extent to which the characteristics of the eigengene are reflective of the true expression of any individual genes. It is apparent from the eigengene plots (Figure 4.4) that the genes within each module (the grey lines) do show substantial variation in standardised expression and do not all conform to the trends of the eigengene. For example, evident from the Black module (Figure 4.4A) are many genes with high or increasing expression during the entrainment day (the first 12 hours), in absolute contrast to the reduced expression exhibited by the eigengene. However, while a number of genes do deviate from the eigengene at various timepoints, the

eigengenes do appear to represent the broad overall trend in gene expression (best exemplified by modules that contain the fewest genes, two of the bottom panels in Figure 4.4E for example). Each module therefore likely contains a number of genes with expression patterns that are well-summarised by the characteristics of the eigengene, and inevitably some genes with expression patterns that do not resemble the key eigengene changes over time.

To further investigate the representativeness of eigengenes, and to examine the identities of genes within the Brown and Yellow modules, genes that most closely reflected the eigengene expression were identified. This was achieved firstly by plotting the expression of the 12 genes whose sum standardised expression values deviated least from the eigengene for each module. For the Brown module, this identified 12 genes with standardised expression patterns that very closely matched the eigengene values (illustrated in Figure 4.6A). Examining relative expression in TPM revealed some highly expressed genes (five of 12 showed 40 < maximum TPM < 50), while the remainder generally showed greater than average expression and few showed low expression (10 of 12 had > 5 max. TPM, > 1 minimum. TPM). Additionally, many of these genes showed > 2 fold-change in expression (Figure 4.6B). As such these genes are likely to be affected by stochastic factors, and be effective, detectable candidates for further investigation. Interestingly, the six genes with the highest expression values all came from the msp (major sperm protein) gene class (shown in green in Figure 4.6B), suggesting that these genes and may be driving the eigengene trend. However other genes in this set of 12 (*gipc-2*, *T23G11.1* and *gsp-3*; shown in blue in Figure 4.6B) also showed considerable expression and fold-change values (> 4 average TPM, > 1.5 fold-change), highlighting that these oscillations may be representative of wider expression patterns and genetic functions.

Contrastingly, many of the 12 genes that least deviated from the Yellow eigengene showed low expression and did not closely follow the eigengene pattern (data not

shown). To isolate genes that oscillated with high overall expression, the genes with lowest deviations from the Yellow eigengene were plotted with the added criteria of average TPM > 5 and fold-change > 2. However, the nine genes in the Yellow module that met these criteria deviated far more in their standardised expression from the eigengene pattern (shown in Figure 4.6C) than the genes that most closely followed the Brown eigengene (Figure 4.6A). Additionally, the relative expression patterns of these genes do not show consistent oscillations (Figure 4.6D); rather than closely mimicking the eigengene at every timepoint, as seen with respect to the Brown module (Figure 4.6C), these genes correspond only to parts of the eigengene time series and show flat expression in places.

The approach taken here in using minimised differences from WGCNA eigengenes represents a novel method for identifying candidate genes. The approach also provided utility as a validation method for eigengenes of interest, suggesting the Brown module to reflect true, 16-hour oscillatory expression throughout the full time series, and the Yellow module to mostly represent a composite of expression patterns that closely mimic the eigengene in some parts of the time series but not others. Therefore, based on the two examples presented here, individual gene expression patterns appear to be faithfully represented by the module eigengene in some instances. The genes of most interest in analysing eigengene expression are those that correspond to the eigengene most closely, while still oscillating with high amplitude. These genes can be simply isolated by measuring least differences from the eigengene and if necessary, applying expression level and fold-change filters to the data. As such, this method may provide an effective pre-screening approach for modules, prior to analysis of more of the hundreds of genes that can be contained in each.

The most stark result from the WGCNA approach was the identification of high amplitude expression oscillations with a 16-hour frequency, particularly in genes of the msp class. This was the extent of the work performed here and would have

continued with the further characterisation of the oscillations. Future experiments would have included whether the entrainment conditions were necessary to stimulate the oscillations, whether periodicity is affected at different temperatures and whether oscillations are also reflected in protein levels. The potential significance of these oscillations is further considered in the Chapter 5. Discussion (section 5.4).

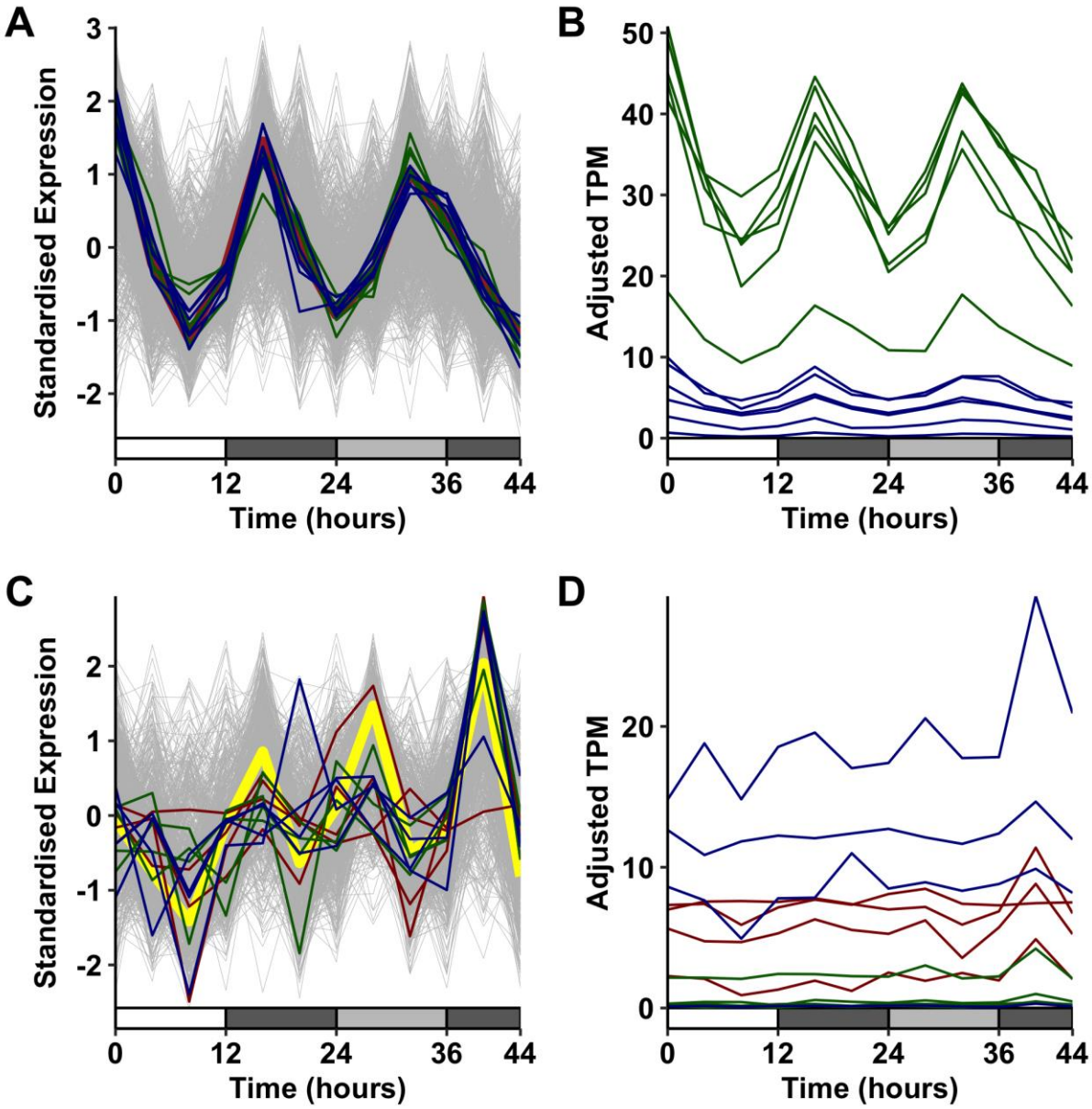


Figure 4.6: Using module eigengenes to identify candidate genes for further analysis. **A)** standardised expression of brown eigengene and 12 genes with the smallest total deviations from the eigengene. Genes of the msp class are highlighted in green, other genes in blue. **B)** TPM values, adjusted for batch effects of genes highlighted in A. **C)** Yellow eigengene and nine genes with average expression > 5 TPM and fold-change > 2. **D)** Batch effect-corrected TPM values of genes highlighted in A. In all panels grey lines indicate standardised average expression of remaining genes within each module.

4.3 Summary

The work described in this chapter utilised RNA-seq over circadian time to generate insights into the oscillatory transcriptome of adult *C. elegans*. Transcriptome-wide

gene expression was measured over the course of two days, consisting of an entrainment day (using a previously unreported approach of dual cycles of light and temperature, in phase; Glossary) and a subsequent free-running day (under conditions of constant dark/cold). Two distinct analytical approaches were employed to identify oscillations: MetaCycle (incorporating JTK_Cycle and Lomb-Scargle algorithms), a conventional, sinusoid detection approach developed for identifying circadian genes in transcriptomic data, and WGCNA, a hierarchical clustering method used here to identify prevailing expression trends.

MetaCycle analysis revealed a considerable number of genes that oscillate in environmental cycles and over the full time series, but relatively few in constant conditions (Table 4.1; Figure 4.1). Interestingly, relatively few of these genes overlapped between the full time series data and the free-running day alone. This suggested a powerful role of the environment over the full time series, and a much more limited role of endogenous drivers of rhythms. However, highly significant endogenous genes were identified that could not be described as rhythmic over the entrainment day alone, suggesting the presence of endogenously rhythmic genes, albeit on a much smaller scale than similar experiments in conventional circadian models (Appendices, Tables 6.1, 6.2). As such, based on the statistical evidence presented, this work suggests the presence of an endogenous circadian clock in *C. elegans* that regulates a relatively small number of genes.

The MetaCycle analysis revealed a number of potential candidate genes for further study. Most of the genes identified as oscillating by MetaCycle were novel, having not been identified in the only prior *C. elegans* transcriptome-wide time series (van der Linden et al., 2010). However, a small proportion of genes were identified by both studies (Table 4.5) and were also found to have oscillating homologues in mammals (Table 4.4). One gene in particular, *acl-12*, was highlighted as a candidate for further study in that it met these criteria, as well as being one of the most

significant genes identified by MetaCycle and showing a repeating expression pattern in entrainment and free-running conditions.

The most significant genes identified by MetaCycle also generally showed longer periods than 24 hours and did not all show clear, repeating oscillatory expression. MetaCycle analysis may therefore have not optimally reported genes with circadian expression profiles. WGCNA was subsequently used as a non-specific, trend-based approach to identify broad oscillations in gene expression. WGCNA highlighted many trends of potential interest (Figure 4.4), including genes that appeared to cycle in entrainment but not free-running (the Black and Blue modules; Figure 4.4A and 4.4B) and potentially novel oscillations with ultradian periods (Brown and Yellow modules; Figure 4.4C and 4.4D).

Further analysis of the Brown module in particular highlighted genes enriched for particular functional roles, as well as genes with high amplitude, 16-hour cycling expression. Functional analysis using DAVID suggested enrichment of genes with roles in phosphorylation to be highly significantly overrepresented in the Brown module gene set (Figure 4.5A), highlighting the potential for the processes of phosphorylation and dephosphorylation to be driven by oscillatory gene expression. Analysis of genes that least deviate from the eigengene indicated the module eigengene to be a true representation of oscillatory novel 16-hour oscillations expression patterns (Figure 4.6A, 4.6B). The accuracy with which the Brown module eigengene represented some of the genes within highlights the potential utility of WGCNA both as a candidate gene discovery approach (here particularly highlighting the msp genes for further study), and as an effective method to identify oscillations in time series data.

4.4 Chapter 4 methods

RNA-seq experimental conditions and sample collection

Nematodes were maintained on 90 mm diameter NGM plates throughout and kept under entrainment conditions until the final collection day. Entrainment conditions consisted of 12:12-hour cycles of light (10 µmol m^{-2} s^{-1}) at 20°C and darkness at 15°C. Entrainment and subsequent free-running conditions were all performed in a single growth chamber, in which warming took approximately 85 mins ± 10 mins and cooling took 55 minutes ± 5 mins. Parental generation nematodes were initially placed onto plates seeded with 10X concentrated *E.coli* OP50 (Stiernagle, 2006) by chunking from previous plates (see Chapter 2, section 2.2.2). Following two days of entrainment, eggs were prepared by bleaching (section 2.2.2) and added to 4 unseeded plates for 2 days to cause L1 larval arrest for age synchronisation. Nematodes were transferred to 40 seeded plates to develop for two days and transferred again to 40 plates containing 25 µM 5-Fluorodeoxyuridine (FUDR) to prevent reproduction. All manipulations were performed around the start of the light/20°C warm phase (dawn). Nematodes were entrained for a further 3 days and then constant darkness at 15°C for 1 day. Full entrainment scheme is illustrated in Appendices, Figure 6.1).

RNA Collections took place at dawn (time 0) and every subsequent 4 hours for 2 days (12 timepoints in total). Each biological sample comprised one plate with approximately 100-200 animals. Nematodes were harvested by washing in 2 mL S buffer (see section 2.2.1), centrifugation for 1 min, aspiration to approximately 100 µL, resuspension in 250 µL TRIzol Reagent (Ambion), mixed by pipetting, and immediately frozen at -70°C. Three samples were collected simultaneously at each timepoint and two independent time series utilising the same conditions were performed, resulting in six replicates in total across two batches.

RNA Processing and sequencing

Samples for RNA-seq were processed by batch in randomised order in sets of 12. To extract RNA, samples in TRIzol were defrosted then subjected to 3 freeze/thaw cycles in which they were snap frozen in liquid nitrogen, defrosted at 37°C and vortexed for 30 seconds. Subsequently, samples were left to sit for 5 minutes at room temperature, 50 µL chloroform was added and samples were centrifuged for 15 minutes at 13200 RPM at 4°C. Most of the aqueous phase was transferred to a new tube and then processed following the protocol of the QIAGEN RNeasy Micro kit, including on-column RNase-Free DNase treatment.

Further processing was performed by Dr Sally James and Dr Lesley Gilbert in the Bioscience Technology Facility at the University of York. Sample quality was checked using the Agilent 2100 Bioanalyzer on RNA Nano Chips, with a majority having a RIN value of 10. 100 ng total RNA for each sample for library preparation using the NEBNext RNA Ultra II RNA directional library prep kit. Unique 8 bp dual indices were added to each sample. Sample quality was again checked using 2100 Bioanalyzer. Sequencing was then performed across two lanes on an Illumina HiSeq 3000 machine.

Sequence Processing

Initial sequence processing was principally performed by Dr Katherine Newling in the Bioscience Technology Facility at the University of York. Sequence quality was checked by MultiQC software (Ewels et al., 2016). Sample depth ranged from 12.4 to 32.8 million reads per sample, except for one sample at the 44-hour timepoint, which contained very few reads and was excluded from further analysis, resulting in 5 replicates at this time-point.

Adapter sequences were trimmed using Cutadapt (Martin, 2011), and samples reanalysed using MultiQC. Salmon (Patro et al., 2017) was used for quasi-mapping of sequencing reads to the *C. elegans* genome (assembly WBcel235, obtained from

GenBank, Benson et al., 2013), and for quantification of gene expression. TPM values were obtained using Sleuth (Pimentel et al., 2017), within R (R Core Team, 2018) following preparation of files using wasabi (COMBINE-lab, 2018).

MetaCycle and WGCNA Analysis

Data was initially filtered to include only genes with > 0 TPM in ≥ 36 of all 71 samples (using the goodSamplesGenes function in the WGCNA R package; Langfelder and Horvath, 2008). One-way ANOVA was performed in R to pre-screen the data for genes that showed significant changes in expression ($p < 0.05$). This was performed independently over the whole filtered dataset and dividing the dataset into two sets of six timepoints representing the entrainment and free-running days. Rhythmic genes were detected using the meta2d function in MetaCycle (Wu et al., 2016) on all three ANOVA gene lists, setting period limits to 20 and 28 hours. Expression fold change for each gene was calculated by dividing the maximum expression (TPM) averaged at each timepoint by the minimum.

Weighted gene co-expression network analysis (WGCNA; Langfelder and Horvath, 2008) was performed utilising the complete filtered dataset (> 0 TPM in > 50% of samples), corrected for batch effects using ComBat within the sva R package (Leek et al., 2018). WGCNA settings generally did not deviate from those used in online tutorial section I (Langfelder and Horvath, 2016). The soft-threshold power was set manually at 6 such to minimise the numbers of genes assigned to no module (grey) and the largest module (turquoise).

All MetaCycle and WGCNA analyses were performed in R (R Core Team, 2018). Gene expression figures were plotted using the ggplot2 R-package (Wickham, 2016).

Functional Enrichment

Enrichment analysis was performed using Database for Annotation, Visualization and Integrated Discovery (DAVID) version 6.8 (Huang et al., 2009a; Huang et al., 2009b).

FDR values are multiplied by 100 by the DAVID program. Significance values reported here are DAVID FDR values divided by 100.

Chapter 5. Discussion

5.1 Introduction and summary of experiments

Circadian rhythms are a seemingly inexorable feature in the lives of organisms that have evolved under exposure to diurnal cycles. The fundamental evolutionary importance of circadian clocks is highlighted by the independent evolution of complex, mechanistically similar transcriptional oscillators across kingdoms and domains (Takahashi, 2017; Hardin, 2005; Harmer, 2009; Hurley et al., 2014; Liu et al., 1995), as well as the remarkable within-kingdom conservation of these oscillators between distantly-related organisms (Chapter 1, Figure 1.1). The contextual background to this thesis was that this intra-kingdom conservation has enabled the extensive characterisation of conserved molecular oscillators in most core metazoan laboratory models (*Drosophila,* mice and other mammals and zebrafish; Yu and Hardin, 2006; Vatine et al., 2011), but not *C. elegans*. This is despite numerous reports describing circadian rhythms in behaviour, physiology and gene expression in these nematodes in published literature (discussed in Chapter 1. Introduction, section 1.3). The absence of a known molecular clock or TTFL in nematodes gives rise to a plethora of questions concerning whether *C. elegans* has or needs a circadian clock, the environmental signals it entrains to, the extent to which its biology is clock-regulated and the robustness of rhythms, and the identities of genes that regulate and are regulated by the clock.

The experiments performed as part of this thesis aimed to contribute answers to these questions by focusing on gene expression, the primary output of the transcriptional clock. In Chapter 3. Results 1, experiments were performed addressing the use of light, temperature and combined signals for entrainment, focusing on two specific genes and their potential as reporters of circadian rhythms in nematodes. Luminescence reporting of transgenic *C. elegans* strains expressing luciferase driven by the promoter sequences of *sur-5* or *nlp-36* did not however,

provide evidence of circadian rhythmicity. Direct measurements of expression of the two genes by RT-qPCR only showed significant rhythms in *nlp-36* and only in response to combined light and temperature entrainment, in phase (Figure 3.8A). These *nlp-36* mRNA oscillations also lacked robustness, being highly variable in dark/cold conditions between replicate samples and only showed a small, 1.49-fold change. As such, the potential for *nlp-36* luminescence reporter nematodes to consistently report the activity of the clock, and consequently the utility of *nlp-36* as a circadian read-out, may be limited. Given that the luciferase reporters did not demonstrate rhythmicity, and the expression patterns of neither gene provided a firm basis on which to build towards effective reporters, the work in Chapter 4. Results 2 took a wider approach in investigating circadian expression across the *C. elegans* transcriptome. This entailed a two-day RNA-seq time series, which suggested the presence of highly significant rhythms in gene expression, but also suggested that they occur in a very small proportion of genes. The identified genes were largely novel, but also included some previously identified genes and some homologues of rhythmic genes in other models, providing a number of candidates for further study. Additionally, analysis of the RNA-seq data by WGCNA revealed a number of different diurnal trends in gene expression data and enabled the identification of novel oscillating genes with non-circadian periods. This chapter comprises appraisals of the experiments performed in both results chapters, the wider implications of results, and gives direction for further research.

5.2 Developing reporters of circadian genes: potential issues and future strategies

Experiments performed in Chapter 3. Results 1 described the responses of nematodes expressing a luciferase reporter under the promoter sequences of *nlp-36* or *sur-5* to different entrainment protocols of temperature and light. With a view towards forward genetics applications and uncovering clock-regulating genes, these experiments served to investigate the potential for *in vivo* reporting of endogenously-driven gene expression on an individual nematode basis. The principal

findings of these experiments were a clear response of luminescence activity to temperature cycles (Figure 3.1), and consistent strain-specific patterns in luminescence in the absence of external cues (Figure 3.5, 3.7). However, no effective reporters were identified and none of these experiments revealed compelling evidence of circadian rhythms, with most suggesting an absence of any rhythmicity.

An important question posed by these results is whether luciferase reporters could, with modifications, provide an effective strategy for understanding the molecular clock in *C. elegans* (as has been suggested in previous literature; Goya et al., 2016). The lack of detected rhythmicity in the examined nematode strains could be attributed to range of genetic factors, the most vital of which being whether the promoter sequence used is clock-regulated (discussed below). However, the luminescence profiles observed could also be a result of more complex issues relating to the transgenes. The *C. elegans* strains used in Results 1 differed as to whether transgenes were genomically-integrated or extrachromosomal, their insertion site, their anatomical and tissue-specific expression patterns and in their signal strength, all of which could have contributed to the observed luminescence patterns. As to whether the issues found in these experiments could be resolved with further work are therefore likely to be both gene and strain specific.

The proposed basis of potential rhythmicity in *sur-5*-driven luciferase reporters was based on prior work that identified rhythmic luciferase oscillations under antiphasic entrainment in the MEG strain (Goya et al., 2016). These results were not replicated here in similar antiphasic and other entrainment conditions in a number of *sur-5* reporter strains (MEG, PE254 and PE255; Figure 3.4, 3.5). A substantial experimental issue with the MEG strain was a low and variable signal strength; luminescence could not be detected consistently by the TopCount apparatus (Figure 3.4), and the associated GFP marker was only detectable by eye in some nematodes. The previous experimenters likely had the same issue, having manually picked the visually brightest nematodes for study (Goya et al., 2016). Perhaps more confounding

however, was that transgene expression was more visible (often exclusively) within developing embryos than somatic tissues. Experiments were therefore likely to be recording embryonic expression of luciferase rather than that of the entrained adults. Experiments using this strain are therefore prone to the raft of non-circadian oscillation patterns and gene expression changes that occur in larval development (Hendriks et al., 2014). Further, given that these hermaphrodite experiments necessitate the use of FUDR to prevent reproduction, embryonic development is halted, which could itself result in abnormal gene expression in the germline. One possible solution to the interference of embryonic expression would be to use males for study, as was done with other strains here, but the MEG nematodes also express a roller phenotype (*rol-6*; Higgins and Hirsh, 1970; Mello et al., 1991), which can impede male mating and make maintaining male populations difficult. These MEG-strain issues collectively highlight the challenges of accurate expression reporting in individual transgenic nematodes before gene expression itself can be interrogated.

Experiments using the PE254 and PE255 strains (expressing an alternative *Psur-5::luc+::gfp* transgene) did not suffer the same issues as the MEG-strain, being easily detectable in individual nematodes and with consistent expression in the intestine, without predominant expression in oocytes. High mating success also enabled easy generation of males in the PE254 strain. The primary challenge in the PE254 and PE255 reporters was therefore that rhythmicity could not be detected, despite comprehensive assessment of the transgene (in hermaphrodites, males and different wild-isolate backgrounds). This could simply be because *sur-5* is not rhythmically expressed; despite being reported to be rhythmic in prior literature (Goya et al., 2016), no evidence was found in this work for rhythmic *sur-5* expression in RT-qPCR experiments (Figure 3.8C and 3.8D) or RNA-seq (data not shown) under any of the environmental cycles. Reasons as to why promoters of the same gene could result in the differential expression seen in MEG, PE254 and PE255 strains could be a result of genomic regulatory elements; the MEG strain transgene was constructed with a much shorter genomic sequence to serve as a promoter than the PE254/PE255

transgene (1052 and 3700 bp upstream of the *sur-5* start codon respectively). As a result, the MEG strain transgene may not encapsulate the full *sur-5* promoter or may lack key *sur-5* regulatory elements that are present in the PE254/PE255 transgene. Additionally, all three strains express the transgenes in different genomic locations so each could be influenced by different positional effects from proximal regulatory elements. These differences could explain the variation in signal strength between strains and the differing luminescence profiles. Further work using the *Psur-5::luc+::gfp* transgenes could explore their expression in different genomic locations to test for positional effects and fully characterise their expression.

Unlike *sur-5*, *nlp-36* gene expression did show significant circadian expression in RT-qPCR experiments, with an expression pattern in phase with previous data (van der Linden et al., 2010). Luminescence experiments did not reveal evidence of rhythms, however. The simplest explanation for this disparity could be that the previously unreported entrainment conditions of dual light and temperature in phase were required to generate RNA rhythms here (Figure 3.8A), as opposed to temperature alone in prior work (van der Linden et al., 2010). These conditions were not tested in *nlp-36* luciferase reporter strains, and it could be that *C. elegans* entrains poorly and requires both light and temperature to do so effectively. However, an alternative, transgene-specific explanation could also account for absence of luminescence rhythms in that the *Pnlp-36::luc+::gfp*-expressing NLIH13 and NLIH2 strains (unlike the *sur-5* reporter strains) express the transgene as an extrachromosomal array, not integrated into the genome. Extrachromosomal arrays result in mosaicism; they are only semi-stably transmitted through meiosis and mitosis and only a proportion of progeny carry the transgene (Stinchcomb et al., 1985; Yuen et al., 2011). This mosaicism could result in inconsistent expression of the transgene between animals and between individual tissues. Notably however, average expression patterns in nematodes were largely consistent across the two strains in these experiments (Figure 3.5).

To circumvent some of the issues presented in the above strains, future circadian reporters could utilise targeted insertion of single-copy transgenes. The MosSCI method for insertion and potentially CRISPR/Cas9 gene editing could be used for targeted single-copy transgene integration (Frøkjær-Jensen et al., 2008; Philip et al., 2019). These approaches could possibly ameliorate some difficulties in luminescence detection and help to identify the issues that prevent the strains used here from accurately reporting circadian rhythms. All nematode strains used in these experiments were generated from a similar microinjection approach in which transgene DNA assembles into multicopy concatemers. If integrated into the genome, as was the case in all the *sur-5* strains, these transgene sequence repeats can be silenced (Leyva-Díaz et al., 2017). Silencing may account for the variable and low detectability of the MEG strain, and single copy insertion could prevent this. Additionally, targeting the transgene could resolve whether the role of positional effects account for differences between the *sur-5* strains. Finally, stable inheritance of an integrated transgene could resolve any mosaicism-related difficulties present in the *nlp-36* strains.

The optimal transgene reporter of circadian rhythms would be one in which luminescence could be measured on an individual level, and in which individuals of interest could be used to generate progeny. Such a strain would enable the detection of mutants with abnormal circadian rhythms, and their propagation for genetic analysis. While individual-level detection was achieved in a number of strains explored in Chapter 3. Results 1, the ideal of progeny generation is much less straightforward to achieve in *C. elegans* than in other species, with the primary difficulty being that experiments must contend with the ability of hermaphrodites to self-fertilise. FUDR, a common solution to this problem, prevents DNA replication and cell division, allowing individual nematode luminescence to be recorded. However, FUDR use, in halting development and preventing reproduction, also means that any nematodes of interest cannot be used to generate progeny for further study. As such, using FUDR raises an issue for potential mutagenesis-based

forward genetics approaches; any large-scale screen for circadian mutants would require every mutagenised nematode to be genetically replicated and testing to be performed on progeny, not all of which would carry the mutation of interest as a result of meiotic crossing over. The ideal approach therefore (as discussed above with respect to the MEG-strain) would be to use males for study. Males were used in many experiments presented in Results 1, but not in any prior circadian studies. Males are the rarer sex in nature and in a research context (< 0.2% in the laboratory N2 strain; Chasnov and Chow, 2002), and are less representative of *C. elegans* as a research model. However, given their use eliminates the need for FUDR, males may represent the ideal approach, and a possible necessity for effective dissection of the molecular circadian clock.

Finally, to push forward the use luciferase reporters in circadian study, the major consideration should be the gene promoter used. A central aim of the RNA-seq experiment detailed in Chapter 4. Results 2 was to identify new genes that could have utility in circadian research. A large number of potential candidates were generated (Figure 4.1), the strongest of which would likely be *acl-12*. The *acl-12* gene was found to be highly significantly rhythmic (having the second lowest meta2d q-value; Table 4.2), with a 2.57-fold change in expression. This gene also showed the most overt repeating oscillation of the genes examined in Results 2 (Figure 4.2) and showed consistent expression between replicates in entrainment (as indicated by small S.E.M values). *acl-12* was also identified as rhythmic within the only prior *C. elegans* circadian time series (van der Linden et al., 2010), and its mammalian homologue, *Lpgat1*, has also been found to be significantly rhythmic (Table 4.4). The guidance provided by this work for further experiments in developing luciferase reporters would therefore be to utilise the *acl-12* promoter sequence. Initial work should first be to fully validate and optimise the entrainment of *acl-12* oscillations, before generation of transgenic nematodes, ideally transformed by single-copy transgene insertion into the *C. elegans* genome.

5.3 MetaCycle analysis of RNA-seq data suggested *C. elegans* is a transcriptionally rhythmic organism

When analysed by MetaCycle (meta2d function; Wu et al., 2016), 263 statistically significant rhythmic genes were identified over the full two-day RNA-seq time series, representing around 1.6% of expressed genes detected in this study. This is a smaller, but comparable proportion to previous *C. elegans* microarray data in which 380 genes were suggested to cycle over entrainment and free-running phases (using either light or temperature entrainment protocols; van der Linden et al., 2010). The data presented here therefore reinforce the notion that *C. elegans* is a circadian organism based on the metric of rhythmic transcription.

5.3.1 MetaCycle analysis suggested fewer genes are rhythmic in *C. elegans* than other animal models

An important question in placing the transcriptome-wide results in context, and in defining the degree to which *C. elegans* is a clock-regulated animal, is how the number of rhythmic transcripts identified here compares with other organisms with well-characterised transcriptional oscillators. Given that the extent of transcriptomic rhythms are a central question in chronobiology, a diverse base for comparison exists in that a considerable number of microarray and RNA-seq circadian time series have been performed in mammals, *Drosophila* and other circadian organisms (the key results of which are summarised in Appendices, Tables 6.1, 6.2 and 6.3). These studies in the same organisms vary greatly in their experimental and analytical approaches, and in the number of genes suggested to be clock-controlled. However, almost all, to widely varying extents, identify more rhythmically-expressed genes in other organisms than seen here in *C. elegans*.

As regards to mammals, the contrast with *C. elegans* is often stark; data from 16 experiments on various mouse tissues isolated in free-running conditions, subsequent to entrainment, report between 334 and > 5000 genes to be rhythmic.

These mammalian experiments are based on sets of expressed genes generally comparable in size to the 16,176 genes found to be expressed in *C. elegans* in Results 2 (and for the most part, but not always, report at a 5% significance threshold; Table 6.1). Most of these studies identify considerably more oscillating genes than the 263 identified here by MetaCycle over the full two-day time series, and dwarf the 26 genes identified as significantly rhythmic over the free-running day alone (Table 6.1; Results 2, Figure 4.1). A number of prior studies also report rhythms in cultured murine cells, generally finding lower numbers of rhythmic genes (ranging from < 100 to > 1000), but that still generally exceed the number found in these *C. elegans* data by a considerable margin (Table 6.1). The most straightforward and likely explanation for these differences is that mouse and rat genomes are much more widely regulated by a circadian clock than *C. elegans*. However, it is also of note that considerable differences are seen in the identities of rhythmic genes between tissues in mammalian studies, including examples in which the same experimental and analytical approaches were used; in mice, one time series reported 43% of genes to be significantly rhythmic across 12 tissues (8504 of 19788 genes, the highest estimate of any work), but no individual tissue showed rhythmic expression of more than 3500 genes (Zhang et al., 2014). The fact that different cells and tissues appear to express different genes under circadian clock-control could provide a second explanation for the small number of rhythms detected in nematodes; a lack of rhythmicity seen across the organism as a whole could potentially be a result of expression differences between tissues masking a rhythmic signal emanating from a specific population of cells.

As regards to transcriptomic time series in *Drosophila* and other insects (including mosquitoes and honey bees; Appendices, Table 6.2 and 6.3), the breadth of *C. elegans* rhythms identified here is somewhat more comparable in scale. While the numbers of rhythmic genes identified by transcriptomic approaches in these species are still generally in excess of the 263 *C. elegans* genes identified by meta2d analysis, insect time series typically report fewer rhythmic genes than mammalian studies.

Many insect time series have also reported gene sets from both entrainment and free-running phases, to which the *C. elegans* MetaCycle data can be more closely compared. The greatest difference between *C. elegans* and these insect models is again the 26 genes that could be identified as significantly rhythmic from the *C. elegans* free-running data alone; transcriptome-wide studies in insects report numbers of rhythmic transcripts between 100 and approximately 1000 in insect heads or bodies (Leming et al., 2014; Hughes et al., 2012a; Table 6.1 and Table 6.2). Conversely though, from the *C. elegans* entrainment day data, MetaCycle analysis, identified 1159 genes as significantly oscillating (Results 2, Table 4.1) when corrected for multiple tests. This is a number consistent with and greater than several *Drosophila* and mosquito studies, in which the majority of work reports between 200 and 1300 oscillating genes (Appendices, Tables 6.2 and 6.3; excluding some high and low outliers). One possible implication of the greater reduction in rhythmic gene number in free-running conditions in *C. elegans* relative to insect models could be the idea frequently referred to throughout this thesis of a non-robust clock; *C. elegans* could have a clock that responds effectively to environmental changes, resulting in similar rhythmic transcription to the insect models referred to here, but fails to maintain rhythms in endogenous conditions, resulting in a vastly reduced number of oscillating genes relative to other models.

Overall, the proportion of oscillating genes observed in Chapter 4. Results 2 is strikingly low in *C. elegans*, particularly in comparing the 26 genes identified by MetaCycle in free-running conditions to the highest estimates of rhythmicity in mammals (Zhang et al., 2014). This observation could lead to the hypothesis that *C. elegans* lacks a functioning circadian clock. However, transcriptome-wide rhythmic gene lists in model organisms with fully established clocks are variable in range and do not all substantially exceed *C. elegans* estimates when considering entrainment phase and over the full 44-hour time series performed here. While necessary caution should be applied in comparing transcriptomic studies, given their divergent

methodologies and results, collective comparisons of these studies certainly suggest that clock regulation of the *C. elegans* transcriptome should not be discounted.

Comparing the identities of the 263 rhythmic genes identified in the RNA-seq analysis here to the 380 identified in the prior *C. elegans* microarray work (van der Linden et al., 2010), a notable observation is a distinct lack of overlap between the two studies. Only 27 of the 263 genes identified over the full time series performed here were also identified in the prior data. Most of these genes were also only previously shown to oscillate in environmental cycles and not in constant conditions (Results 2, Table 4.1; discussed in section 4.2.1). This lack of consistency arising from two distinct experimental approaches in the same organism could potentially imply non-robust or weak circadian rhythms. In comparing to prior work in other models again however, a lack of overlapping genes is not an unexpected finding; the plethora of circadian transcriptomics studies in published literature reveal a considerable precedent for limited agreement between rhythmic gene lists in individual studies. Some of these distinct gene lists almost certainly result from tissue differences as discussed above, but another source of inter-study intraspecific variation in rhythmic genes is that different analytical methods produce considerably different results. This has been reported in prior work in mammals (Hughes et al., 2009; Table 6.1) and *Drosophila* (Hughes et al., 2012a; Table 6.2), and was observed here in the RNA-seq data in that Lomb-Scargle was far more stringent than JTK_Cycle in defining significantly rhythmic genes (Table 4.2; discussed in section 4.2.1). This aspect of circadian transcriptomic studies is perhaps best exemplified by a *Drosophila* meta-analysis of five different microarray time series, which reported a maximum gene co-occurrence of 27.8% between two studies, and when reanalysing the data through a different approach, reported 214 rhythmic genes, only 81 of which had been previously identified (Keegan et al., 2007). As such, analytical approaches appear to play a key role in determining not only the number of genes, but their identities too. It is important to note that some of these *Drosophila* microarray studies are early transcriptome-wide approaches, published soon after the initial genome sequence (Adams et al., 2000),

but numbers of rhythmic genes have not necessarily shown increases in more recent work (Appendices, Table 6.2). Therefore, when placed into the context of prior work, the lack of overlap in gene identities in comparing *C. elegans* results could be a probable outcome from two different approaches, and should not be interpreted as an experimental issue or a specific characteristic of *C. elegans* circadian biology.

A perhaps more confounding issue than the differences in rhythmic gene identities observed between studies, is the occurrence of within-study differences in gene identities, specifically those occurring between entrainment and free-running data. The MetaCycle data in Chapter 4. Results 2 revealed only 8 of 26 genes identified in free-running to also oscillate over the entrainment day (Figure 4.1). Similar results have generally also been found in other organisms when entrainment and free-running phases are analysed independently. In transcriptomics literature, three microarray and RNA-seq time series in *Drosophila* and one microarray study in mosquitoes report separate entrainment and free-running gene lists, with only a small proportion of genes overlapping (5-28%; Hughes et al., 2012a; Lin et al., 2002; Ueda et al., 2002; Leming et al., 2014). The prior *C. elegans* microarray data also found a lack of overlaps between gene lists generated from light or temperature cycles and subsequent free-running conditions (nine and two genes or 7.76% and 0.97% respectively; van der Linden et al., 2010). These observations in prior work demonstrate that the data described in Results 2 are not unusual in finding very limited overlap between entrained and free-running gene lists, but do highlight what is perhaps a fundamental weakness of circadian analysis methods. Diurnal signals (zeitgebers) should be a stronger driver of rhythmicity than the endogenous clock, the presumed driver of rhythms in the free-running phase. Consequently, as observed, smaller numbers of genes would be expected to oscillate in constant conditions. However, free-running rhythms should be a continuation of entrainment rhythms, and the genetic identities of oscillating transcripts should be expected to be the same. That within-study overlaps between diurnal conditions are so small in these experiments strongly suggests that the conventional analytical approaches

(generally sinusoid detecting algorithms; Hughes et al., 2017) are not effectively retrieving rhythmic genes. These approaches could be producing either false negative or false positive results; non-overlapping rhythmic gene lists could reflect true rhythmic transcripts not being properly identified due to excessively stringent statistical thresholds in large datasets, or conversely, could result from non-cycling genes being mischaracterised as circadian due to ineffective modelling of rhythms. Very few rhythmic genes were found in free-running conditions in this experiment, which could suggest weak endogenous control of circadian timing, but the fact these genes were largely not also found to cycle in the entrainment phase may also imply an inadequate analytical approach.

5.3.2 Methodological considerations for analysing transcriptomic rhythmicity in *C. elegans*

As is evident from the huge variation in the numbers and identities of rhythmic genes reported in studies described in section 5.3.1, aspects of experimental design can have a profound impact on the results of circadian transcriptomic studies. One major consideration in particular is whether to measure rhythms in the entrainment or free-running phase of a circadian experiment, so as to maximise the detection of genuine endogenously-driven circadian rhythms. The approach used here was to measure entrainment in the final entrainment day and the first free-running day, and then to analyse the full time series and each day independently. This approach revealed a large number of rhythmic genes in diurnal cycles (1159), a very small estimate of rhythmicity in constant conditions (26), and an intermediate number (263) over the full time series (Figure 4.1). These results effectively represent three, highly varying estimates of the extent of rhythmicity in *C. elegans*.

Of the three estimates, the free-running rhythms could be viewed as the truest estimate of nematode rhythmicity due to being produced endogenously in invariant environmental conditions. However, in this study, free-running genes were identified from six timepoints, and MetaCycle algorithms were therefore being applied over a

20-hour experimental timespan to detect approximately 24-hour (between 20-28 hours) rhythms. This could potentially have contributed to the low number of significant rhythms identified. To capture the true extent of endogenous clock regulation, and exclude the influence of environmental signals, rhythms could be recorded over two free-running days, as done in some mammalian studies (Zhang et al., 2014; Hughes et al., 2012b; Appendices, Table 6.1). These endogenous rhythms are subject to fast damping however (described in Introduction chapter; section 1.1.1), and clock-regulated genes could consequently be missed by analyses due to rhythmic expression quickly flattening out. This issue could be exacerbated in *C. elegans* in that behavioural and luminescence reporter literature indicates *C. elegans* rhythms may dampen faster and be less individually detectable than those in other species (Simonetta and Golombek, 2007; Herrero et al., 2015; Goya et al., 2016; Fogg et al., 2014). The only previous *C. elegans* transcriptomic study (van der Linden et al., 2010) avoided the issue of damping in reporting multiple-day free-running rhythms by appending single free-running days from different samples. However, this may compromise statistical independence, and could lead to false positives being identified (Hughes et al., 2017). As such, reporting free-running rhythms alone may not be an effective or reliable option in *C. elegans*.

It should also be considered that approaches that ignore environmental cycles when reporting circadian rhythms may not be optimal for investigating rhythmic transcriptomes. While environmental signals can trigger expression changes in many genes independent of a circadian clock, they are still a vital component of circadian rhythms; the environment sets the phase of the clock, and organisms have not evolved to maintain rhythms for prolonged periods without environmental inputs. Additionally, free-running rhythms should be a continuation of environmentally-driven oscillations and as such, not sampling during entrainment may effectively miss a core component of the expression rhythm. To maximise the number of true rhythmic genes identified, circadian experiments should therefore perhaps not seek to avoid sampling across environmental cycles and instead, aim to detect continuous

patterns across entrainment and free-running, as was done here. Such an approach may be particularly necessary in an organism like *C. elegans*, where the damping of oscillations may be fast. Working across entrainment and free-running may be necessary to capture the key aspect of rhythms: repetition. To better detect rhythms under these cycles, improved accuracy could be delivered by alternative detection methods centred towards detecting repeating patterns in data (which was one of the reasons WGCNA was subsequently used on this data; discussed below in section 5.4).

Regarding the design of the RNA-seq experiment described in Results 2, potential improvements could have been made by increased sampling, both by increasing frequency during the time series and by lengthening the sampling period. Insufficient sampling frequency could result in false negatives; the 4-hour resolution used here is standard in most transcriptomic studies (Appendices, Tables 6.1, 6.2 and 6.3), but has been suggested to be underpowered (Hughes et al., 2017), and JTK_Cycle has been suggested to miss 20% of true positive rhythms at this resolution (Hughes et al., 2010). The issue of detecting few rhythms over six timepoints in free-running could have also been mitigated by a longer overall sampling period. The most significant rhythms identified by MetaCycle oscillated with periods that generally exceeded 24 hours (Results 2, Table 4.2). The addition of one or two extra timepoints could have resulted in higher numbers of rhythmic genes (true positives) and improved the accuracy of the analysis approaches. The limiting factor determining sample number in this and other experiments is inevitably one of cost, and the maximum possible number of samples were sequenced here. To improve this experiment however, some compromise could have been made with the number of replicates at each timepoint. A maximal number of replicates (six) were used here, informed by literature suggesting a substantial likelihood of false negatives in a large proportion of low fold change genes if fewer were used (Schurch et al., 2016). Having more replicates also reduces the impact of stochastic factors, and may be particularly important if *C. elegans* rhythms dampen more quickly than other models (Hughes et al., 2017). Given that false negative results are a consequence of both insufficient

sampling and insufficient replicate number, this experiment may (or may not) have more accurately detected *C. elegans* rhythms with additional timepoints at the expense of having one fewer replicate per timepoint.

A final point of note on the MetaCycle approach employed here is that if *C. elegans* indeed does have a rapidly damping or poorly endogenously maintained clock, the JTK_Cycle and Lomb-Scargle algorithms may be poorly suited to detecting its rhythms. These sinusoid-detecting approaches, in particular JTK_Cycle, represent community standards for circadian rhythm analysis (see Appendices, Table 6.1, 6.2, 6.3; Pizzaro et al., 2013). However, in common circadian model organisms, repeating rhythms can be observed over multiple cycles before damping out (Hughes et al., 2012a; Hughes et al., 2012b), and the JTK_Cycle algorithm is widely used because of its capability to detect these rhythms. In the *C. elegans* RNA-seq experiment described in Results 2, collections took place over an entrainment and free-running day to maximise rhythm detection, but examples of resultant rhythms (indicated by Results 2, Figure 4.2B, D and F) appear to show a waveform that substantially dampens between the first and second cycle. If such rhythms are present throughout the transcriptome, perhaps including more extreme examples of damping, the approaches employed by MetaCycle may be ineffective in detecting them. Given that so few significantly rhythmic genes (26) were identified during the free-running phase alone in the RNA-seq time series (Figure 4.1), sampling during both entrainment and free-running phases may have been a necessary approach to identify potential circadian rhythms. However, JTK_Cycle and Lomb-Scargle as used here may be suboptimal for reporting them in the rapidly damping data. These concerns informed the decision to use WGCNA to identify expression trends that may not conform to regular sinusoids, but further work could consider potentially modelling the *C. elegans* data around a damping sinusoidal waveform, which may better represent expected circadian trends in the data.

5.4 WGCNA revealed novel ultradian oscillations in *C. elegans*

The gene list inconsistencies discussed in section 5.3.1 suggest a potential underperformance of various circadian analysis approaches in generating full or accurate circadian expression profiles. However, because conventional circadian approaches like MetaCycle function to detect sinusoidal trends in expression data (Wu et al., 2016; Hughes et al., 2010; Glynn et al., 2006), such approaches, even if optimised, could also miss a wealth of circadian expression patterns in data that do not conform to sinusoidal patterns of expression. Previously described examples of these patterns include twin-peaking genes (Pembroke et al., 2015) and pulsatile spikes (Rund et al., 2013). Further analysis was performed on all expressed genes in the RNA-seq dataset using WGCNA (Langfelder and Horvath; 2008), a hierarchical clustering approach, in order to identify any prevailing trends in gene expression patterns.

The expression trends of interest identified by WGCNA (Chapter 4, Results 2, Figure 4.4) are perhaps the most novel data presented in this thesis. The Brown WGCNA module revealed 16-hour oscillations in eigengene expression, closely matched by the expression patterns in TPM of a number of genes. The genes identified included some with high expression levels (TPM values) and substantial fold change (Figure 4.6). As such, WGCNA was successfully used as a candidate gene discovery approach, identifying real gene expression patterns that potentially demonstrate novel ultradian rhythms with a frequency previously unidentified in *C. elegans*.

The genes most closely matching the trend of the eigengene and demonstrating these high amplitude oscillatory expression patterns included a number of genes in the msp (major sperm protein) class. Loci encoding these msp genes are spread across chromosomes, but the proteins share a commonality in a highly conserved domain; a single exon that encodes variants of the major sperm protein. Dual functions have been identified in msps; the proteins act both as components of filaments that enable sperm motility, and as secreted hormones involved in

stimulating oocyte maturation (Han et al., 2010). Genes of the msp class are exclusively and highly expressed in spermatocytes, where they have been estimated to represent 17% of total protein (Klass and Hirsh; 1981). That the strongest signal identified by WGCNA concerned genes only expressed in one cell type could be of relevance to the idea that tissue-specific oscillations could be masked by the absence of oscillations in other tissues (Discussed in section 5.3.1). Further work investigating circadian or other oscillations in *C. elegans* might therefore consider tissue-specific approaches where possible.

As to the functional purpose of these undefined 16-hour expression oscillations, the msp-class genes are also nematode specific, and so no directly comparable rhythm would be expected to exist in models from other phyla. However, one of the two known msp functions, oocyte maturation, effectively represents successful progression through the cell cycle (Han et al., 2010). The eukaryotic cell cycle is a checkpoint-based process with variable durations (Koreth and van den Heuvel, 2005), that is tightly-governed by cyclins, proteins that cycle in expression. The cell cycle has been likened to a clock in that it is circular and can be temporally regular if it proceeds through checkpoints without issue (Tyson and Novak, 2008). The process is also regulated by kinases, which interestingly, were shown by functional analysis to be substantially enriched in the Brown WGCNA module, along with other genes involved in phosphorylation and dephosphorylation of proteins (discussed in Results 2, section 4.2.6). Secreted MSP signalling to oocytes forms part of a bidirectional communication process required to progress through the cell cycle (Han et al., 2010). Oocytes have been shown to arrest in meiotic prophase in the absence of sperm (McCarter et al., 1999), a function that serves to stop too many unfertilised oocytes being produced at energetic expense. One hypothesis for the basis of these novel 16-hour oscillations could therefore be a timing mechanism for the efficient production of fertilised embryos, with MSP proteins possibly playing a cyclin-like role in the progression of oocyte maturation.

Considering developmental biology more generally, developmental processes are often characterised by stringent timekeeping. This is perhaps best exemplified by the clock and wavefront model in vertebrate development used to explain somitogenesis and segmentation. In this model, repeating patterns of somites (embryonic segments that eventually differentiate into skeletal structures) are formed with a frequency of 30 minutes to 2 hours by oscillatory transcriptional feedback loops that coordinate developmental pathways (Uriu, 2016). While this process is chordate specific and the oscillatory duration is substantially shorter than the 16-hour msp oscillation periods identified in Results 2 (Figure 4.6), there is clear evolutionary precedent for the importance of timing in developmental processes.

The purpose of this work was to identify circadian rhythms, and while not circadian in duration, the msp expression oscillations reported here were observed during and following a circadian entrainment protocol. Numerous mechanistic and genetic parallels do exist between circadian timing and *C. elegans* development, discussed in Chapter 1, Introduction (section 1.2.3). The msp gene oscillations uncovered here could offer a novel further example of oscillations in developmental timing and, if these observed rhythms were entrained to the environment, a further parallel between development and the circadian clock.

The extent of the work performed here was the identification of oscillations in gene expression. An immediate aim of further work should therefore be to understand the function of these oscillations. An initial aim of experiments could be to establish whether proteins also oscillate and how the MSP receptor VAB-1 (Cheng et al., 2008) is affected. It is also particularly important to consider here that nematodes were treated with FUDR in this experiment, with the effect being that gene expression and protein production could take place, but DNA synthesis and consequently, developmental progression, could not. As a result, any feedback to msp expression that occurs as part of normal development and involves DNA synthesis would not be observed in these experiments. Given the particular role of msps in oocyte

maturation (Han et al., 2010), a process of DNA replication, a vital question to be answered by further study is whether these oscillations persist in the absence of FUDR.

A further immediate question, in considering the principle aims of this work in understanding circadian biology, is the relationship these msp genes have to the entrainment conditions in which they were sampled and whether these oscillations are clock-controlled or independent. The observed oscillations do not appear to strongly oscillate in phase with the entrainment cycles, but could be driven by the onset of the light and warm phase. One potential hypothesis could be that these oscillations are a result of nematodes exhibiting an evolutionarily depreciated fast-running circadian clock. On the other hand, the consistency of the oscillations could suggest they are resistant to changes in light and temperature. First experiments to answer these questions should examine expression of the msp genes over time in constant conditions, investigating whether the duration of oscillations alters or if rhythms become less synchronised under different temperature conditions. Experiments should also aim to characterise the responsiveness of oscillations to diurnal cycles, under different conditions of varying light and temperature.

5.5 Existing challenges and future directions in investigating circadian rhythms in *C. elegans*

The primary goal of the work here and most other *C. elegans* chronobiology research is to understand the molecular basis of circadian rhythms. Both the reporter approaches employed in Results 1, and the transcriptome-wide screen for rhythmic genes in Results 2 were performed with a view to ultimately uncovering core clock genes. Considering the reasons for which the molecular clock has proven challenging for circadian study, the experiments performed here, along with previous literature, serve to highlight some of the key difficulties of the *C. elegans* model. Substantial issues in *C. elegans* chronobiology that may act as barriers to uncovering the molecular clock include a potential non-robustness of *C. elegans* rhythms, a lack of

knowledge of the conditions required for entrainment, the likelihood that elements of the *C. elegans* clock are not conserved with other models and aspects of its reproductive biology that make *C. elegans* refractory to circadian study.

5.5.1 *C. elegans* may lack robust circadian rhythms

The idea of a non-robust clock was a concept frequently referred to throughout this thesis. It could be the case that *C. elegans* demonstrates circadian rhythms, but challenges in studying the molecular clock arise because they are weakly maintained and rapidly dampen, perhaps as a result of an ineffective endogenous timer. This notion could be supported by prior work in which significant rhythms in expression or behaviour were reported, but could only be observed in a proportion or populations of nematodes (Simonetta and Golombek, 2007; Herrero et al., 2015; Goya et al., 2016). Examples of expression data from experiments in Results 1 and Results 2 could also be consistent with this idea; RT-qPCR and RNA-seq both revealed evidence of circadian rhythmicity, but also found high variation between individual biological replicate samples. A reduced capacity for maintaining rhythms could reflect *C. elegans* diverging from a rhythmic ancestor and retaining only a residual clock, perhaps due to adapting to an environmentally arrhythmic habitat or evolutionarily favouring developmental timing roles for circadian genes (concepts discussed in Introduction chapter, section 1.2).

To further explore the idea of robustness in *C. elegans*, two key considerations would be sample size and temperature. RT-qPCR experiments described in Results 1 used 50-nematode samples, while RNA-seq in Results 2 used 100-200 nematodes. A greater number than used here could be required to generate consistent outputs and optimise circadian RNA experiments in *C. elegans*. Conversely, triplicate sampling in RT-qPCR generated a significant rhythm in *nlp-36* expression and the average of six replicates in RNA-seq did not. It could certainly also be the case therefore that no effective sample size would generate robust *nlp-36* expression rhythms in *C. elegans*. As regards to temperature, a source of the variation between

biological replicates could be that *C. elegans* regulates poorly at colder temperatures (15°C here), and further work might consider a higher temperature for the cryophase to decipher whether the root of a lack of rhythms stems from a non-robust circadian oscillator, or simply poor circadian regulation at lower temperatures. Circadian biology in *C. elegans* would perhaps benefit from performing comparative experiments under differing temperatures and in different sized populations.

If it is the case that rhythms are present, but simply poorly maintained in nematodes, the idea explored in Results 1 of using genetic reporters to measure rhythms in individual nematodes may ultimately have limited utility; the rationale for using this approach was to be able to identify mutants and explore the clock through mutagenesis and forward genetics (similar to approaches that led to the discovery of clock genes in other organisms; Benzer and Konopka, 1971). However, non-robust rhythms that require large populations to detect rhythms would make characterising individual circadian mutants difficult (discussed above in section 5.2). Circadian reporters could still be used for reverse genetics however, if strong candidate genes were isolated, and if large populations of genetically identical nematodes were used.

5.5.2 Entrainment in *C. elegans* is poorly understood, and *C. elegans* may lack conservation with other circadian models

The absence of robust observed rhythms here and in literature could also be a result of other difficulties in interrogating the nematode clock, one of which being that basic knowledge of *C. elegans* chronobiology is still lacking. The crux of the issue for molecular understanding could be that despite a body of literature that documents circadian rhythms, work has yet to gain a strong foothold on either robustly oscillating genes or the environmental signals to which nematodes can entrain. Organism-specific issues that beleaguer circadian study in *C. elegans* include core clock homologues not appearing to be rhythmically expressed (Results 2, section 4.2.4; van der Linden et al., 2010; Olmedo et al., 2012), as well as *C. elegans*' distinct and relatively primitive system for photoreception (Gong et al., 2016) and the

question of whether it could service a circadian oscillator akin to those of other models (discussed in Introduction chapter, section 1.2.1). These factors create a dual issue for experimental design; understanding entrainment of expression cycles requires a robust circadian reporter, and establishing a robust circadian reporter requires ecologically relevant entrainment conditions. This problem is well-represented by the work in Results 1, in that experiments aimed to optimise entrainment based on previously reported circadian genes, but under a range of entrainment conditions, few circadian rhythms could be identified.

The issues surrounding entrainment in particular were a focal point of this thesis. The question of how different zeitgebers affect the transcriptional circadian clock arises, in part, from a sparse literature regarding gene expression rhythms, in which different studies offer contrasting approaches for successful entrainment. Of the three prior studies that report circadian expression rhythms, one entrained nematodes using 12:12-hour cycles of light and temperature independently (van der Linden et al., 2010), another using 16:8-hour temperature cycles (Olmedo et al., 2012) and the third through 12:12-hour antiphasic cycles of both zeitgebers (Goya et al., 2016). These approaches are not inherently contradictory; nematodes could possibly entrain to both zeitgebers independently or in combination. However, the 16:8-hour temperature entrainment approach raised replication issues, only identifying expression rhythms in one of nine re-examined putative temperature entrained genes (Olmedo et al., 2012; van der Linden et al., 2010). Additionally, light and temperature signals out-of-phase have been shown to represent a conflicting signal in *Drosophila* and zebrafish (Yoshii et al., 2010; Harper et al., 2016; 2017; López-Olmeda and Sánchez-Vázquez, 2009). As such, the evidence for the two signals being effective *C. elegans* zeitgebers in antiphase, raises the prospect of fundamental differences between *C. elegans* and established biological models. For these reasons, experiments were performed here under a range of entrainment conditions (Chapter 3, Results 1) with a view to both optimising approaches and better characterising environmental sensing with respect to *C. elegans* rhythms.

The work in this thesis may serve to further confound the issue of circadian entrainment in being the first study to report light and temperature cycles, in phase (see Glossary), to be an effective method of entrainment. RT-qPCR experiments in Results 1 (section 3.2.6) did not identify rhythms in *nlp-36* or *sur-5* under the entrainment conditions in which they were previously identified as rhythmic, but did identify significant *nlp-36* rhythms using collective light and temperature cycles (Figure 3.8A). Further, these conditions provided the basis for detecting transcriptome-wide circadian expression patterns in Results 2. However, using this approach for RNA-seq ultimately did yield relatively few rhythmic genes (fewer than prior *C. elegans* work and other organisms; as discussed in section 5.3.1). While not methodologically comparable approaches, the fact that using dual zeitgebers in this work did not identify a substantially greater number of rhythmic genes than prior work using light or temperature alone (van der Linden et al., 2010) does not support the idea that entrainment approaches could be improved by using two environmental signals in phase. Further, given that prior work reported rhythms most effectively under antiphasic entrainment conditions (Goya et al., 2016), a small number of rhythmic genes identified in RNA-seq may even suggest interference of light and temperature if *C. elegans* preferentially entrains to antiphasic cycles. Entrainment therefore remains a question to be addressed in *C. elegans* chronobiology. Whether it forms the primary experimental question or not, further work would likely be most informative in exploring multiple entrainment protocols.

5.5.3 The reproductive system of *C. elegans* creates difficulties for time series RNA sampling

A final potential barrier to detecting circadian rhythms effectively in *C. elegans* are issues surrounding its basic biology, specifically those of reproduction and larval development. Most chronobiology experiments (including all those described in this thesis) are carried out on adult nematodes (reviewed in Introduction chapter, section 1.3) for the reason that, while circadian clocks do emerge through the course of

development (Vallone et al., 2007), they are a facet of fully-developed animals. Using adults in *C. elegans* also enables the avoidance of the potential confounding variable of the gene changes and oscillations that occur throughout development (discussed in Introduction chapter, section 1.2.3). Using adults does pose substantial difficulties for circadian experiments however; as noted throughout this work, *C. elegans* differs from conventional circadian models in being able to self-fertilise, producing up to 300 progeny from the start of adulthood (Corsi et al., 2015). In addition, the generation time of *C. elegans* is approximately three days (depending on temperature), meaning that only a very narrow window of time exists before developing larvae are capable of reproduction themselves. The work in Results 1 and Results 2 took two different approaches to mitigating these issues. The experiments in Results 2 took a conventional approach in treating nematodes with the DNA synthesis inhibitor FUDR to prevent reproduction, while the work in Results 1 took a less conventional approach in allowing nematodes to reproduce and manually selecting adults for RNA extraction. Both of these methods have respective advantages, but also have substantial disadvantages that create difficulties for the accurate study of molecular rhythms.

The primary advantages of using FUDR are that it enables the collection of a large number of animals efficiently and easily without contamination by an abundance of larvae. The disadvantages of FUDR use lie in its unknown potential effects that could affect wild-type gene expression. In using FUDR, nematodes could be sampled as whole petri plate populations for RNA-seq, meaning 100-200 adults could be collected per sample in the complete absence of larvae. This enabled the collection of sufficient material for effective RNA extraction and sequencing and contributed to a sequencing read-depth that would be effective for rhythm detection in mice and *Drosophila* (Li et al., 2015). Considering the variability in expression observed between replicates in these experiments (Figure 4.2), FUDR would also enable much greater sample numbers to be efficiently collected if needed for effective reporting of rhythms. However, FUDR does effectively halt development at the late L4 stage

(see Figure 1.3, Introduction chapter). As such, the nematodes being examined may not truly reflect adult gene expression as would be seen in wild type, normally reproducing hermaphrodites. These nematodes may effectively sit in a permanent state of trying to develop and may consequently show unknown gene expression abnormalities. FUDR therefore represents a trade-off between experimental ease and accuracy. Such a trade-off would perhaps be justified if rhythms were abundant and robust. However, given that this has not been found to be the case here and in other work, and because the FUDR approach has generally been applied in most other *C. elegans* chronobiology literature (van der Linden et al., 2010; Goya et al., 2016), it is a possibility that FUDR could be limiting positive results in *C. elegans* chronobiology. Methods for FUDR avoidance should therefore be thoroughly investigated.

An alternative strategy to FUDR use was applied in RT-qPCR experiments in Results 1, in which FUDR use was circumvented by manually picking adult nematodes into TRIzol and freezing prior to RNA extraction, avoiding developing larvae. This strategy would have provided results that are more closely representative of fully-developed adult expression than if FUDR was used. This approach is also bound by a number of limitations however. The primary difficulty for circadian experiments in not using FUDR, is the need to work within the limitation of a five to six-day experimental window; freely reproducing adult nematodes must be entrained (for three days in this work) and samples must be collected before the subsequent generation reaches adulthood. A further limitation of this approach is that sample size is capped by the number of nematodes that can be manually collected (quickly, and without excessive light exposure for dark timepoints), such to represent a single timepoint (in this case, n = 50; Results 1). Issues that arose from using this sample size included a small amount of RNA, limiting the possibilities for sample purification (spin column protocols could not be used) as well as data that are presumably more prone to variation and less accurate estimates of a population mean. This could explain some of the variation in RT-qPCR data, and possibly why prior data could not be replicated

in places. The FUDR-free approach could be quite simply improved however, by increasing the number of nematodes picked; 200 nematodes could be quickly isolated to produce samples less prone to stochastic variation and to increase the amount of RNA isolated. With some protocol alterations, RNA experiments could provide reliable and informative results without using FUDR treatment.

A final point of note on sampling is that neither method discussed here fully controls for developmental effects, and this cannot be achieved using whole adult hermaphrodite *C. elegans* for RNA extraction. Because freely reproducing hermaphrodite adults are gravid throughout adulthood, gene expression measurements will also likely capture expression in developing embryos. FUDR does not fully mitigate this issue, in that while development is halted, embryos are also present in developing adults. As such, both FUDR and non-FUDR approaches could be measuring embryonic gene expression. Given that tissue-specific expression is a considerable factor in circadian experiments in mammals (Zhang et al., 2014; discussed above in section 5.3.1), embryonic expression could certainly be limiting in detecting rhythmic genes in *C. elegans* experiments. To circumvent this issue, either tissue-specific methods must be developed or, as discussed in with regards to optimising reporters (section 5.2) and used in some Results 1 experiments, circadian biology could benefit from the use of males. Using males requires a manual approach to collection, but is the only other way to avoid potential larval or embryonic interference with gene expression measurements.

Overall, *C. elegans* chronobiology experiments are subject to a considerable number of issues in reporting circadian expression that could be a result of a non-robust clock, suboptimal entrainment strategies and *C. elegans*-specific sampling difficulties. These factors individually, or in conjunction, may be limiting the understanding of the *C. elegans* molecular clock. Some of these issues may be resolvable however by further study, as described in the paragraphs above. The most effective strategy for

the advancement of *C. elegans* chronobiology could be with experiments targeted at solving these questions.

5.6 Final conclusions

C. elegans remains a substantial challenge in chronobiology. It is fundamentally unlikely to ever be a model of circadian rhythms based on the methodological complications exemplified by experiments in this thesis, including its reproductive capacity and potential non-robustness or poor entrainment of its clock. These challenges make performing circadian assays, consistent reporting and consequently, generating insights into the *C. elegans* molecular clock, tremendously difficult. Nevertheless, the work presented here and a varied body of past literature do suggest the presence of statistically significant rhythms in gene expression and other aspects of its biology.

The presence of harder-to-detect or less robust rhythms and fewer oscillating genes than other species makes the question of the *C. elegans* clock all the more interesting from an evolutionary perspective; either the rhythms identified here and elsewhere are false positives, and *C. elegans* is an arrhythmic animal, or *C. elegans* is rhythmic with a clock that diverges from other models. This could be a result of a decaying or residual clock, one that is poorly, or differently, entrained to those in other animals, and possibly even one with a novel genetic basis.

The work in this thesis did not uncover the molecular basis of circadian rhythms in *C. elegans*, but the results described here provide a range of insights that will hopefully be used to direct future work. In particular, a number of new target genes, like *acl-12*, were highlighted by the convergence of MetaCycle RNA-seq results and prior expression studies in *C. elegans* and mammals. The success of dual, in-phase entrainment, along with the profound between-replicate variation seen in constant dark/15°C conditions will hopefully inform future experimental design. Finally, the presence of 16-hour oscillations and diurnal variation identified by WGCNA offers

new directions for *C. elegans* research, further links between circadian clocks and development and a mass of candidate genes for further study.

Appendices

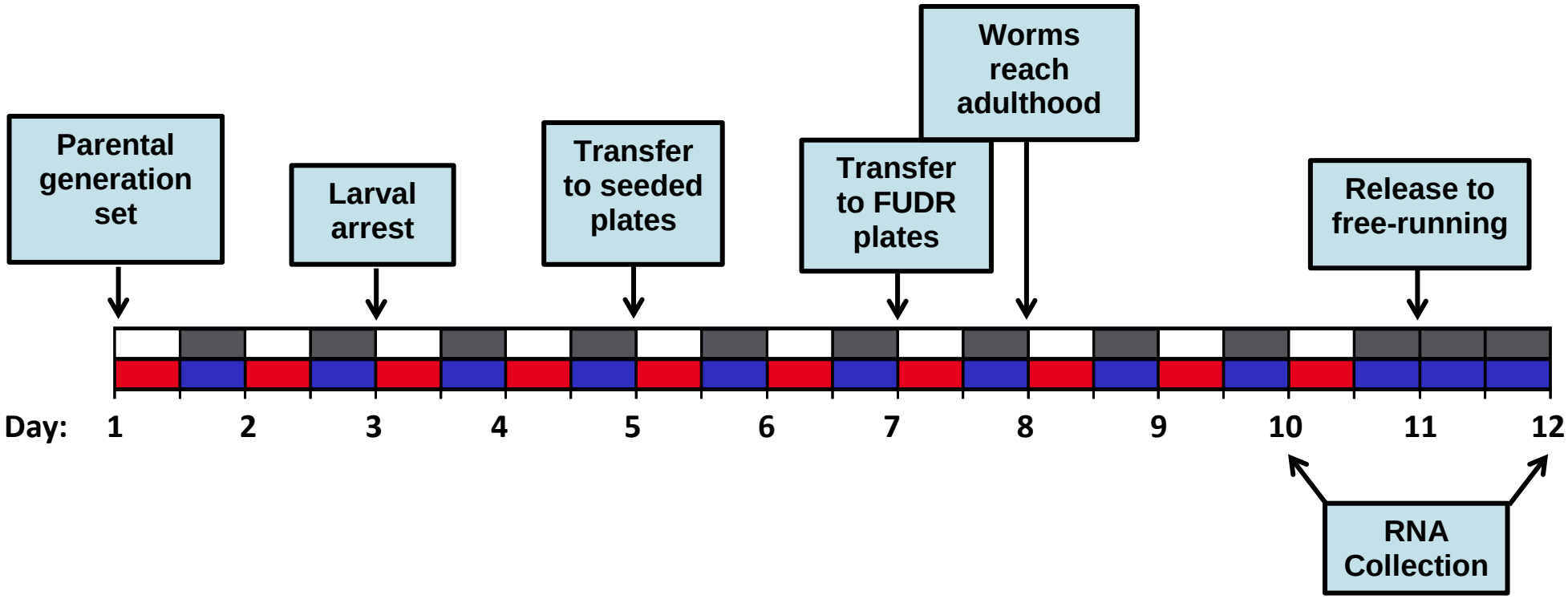


Figure 6.1: Schematic of protocol used for RNA-seq Experiment.

Nematodes were exposed to 12-hour cycles of light at 20°C (white and red boxes) and dark at 15°C (grey and blue boxes). F0 individuals were entrained until gravid. Eggs were isolated by bleaching adults (see Chapter 2, Methods) Nematodes were maintained in starvation conditions to generate synchronised aged worms before being moved to plates seeded with OP50 strain *Escherichia coli*. Once at L4 larval stage, nematodes were transferred to plates containing (25 µM) 5-fluorodeoxyuridine (FUDR) to prevent reproduction. All manipulations took place around dawn (lights on). F1 nematodes were entrained for three days as adults and then released into constant dark/cold conditions. Collections took place every four hours on the final day of entrainment and first day in constant conditions. Each sample comprised an individual petri plate containing 100-200 nematodes. Sampling was performed over two identical time series, with three biological replicate samples collected in each (six replicates overall).

Table 6.1: Summary of transcriptome-wide circadian gene expression time series performed in mammals and mammalian cells and tissues.

Reference	Organism/ Tissues	Transcriptomic Method	Sampling Frequency	Entrainment Cycles	Analysis Period Length	No. Rhythmic Transcripts/Genes (entrainment)	No. Rhythmic Transcripts/Genes (free-running)	Other Notes (of interest to thesis text)
Zhang et al., 2014	Mouse (12 tissues)	Microarray + RNA-seq	2 hrs (microarray) 6 hrs (RNA-seq)	Light:Dark (12:12 hr)	48 hrs	-	8504 (of 19788, 43% across 12 different tissues)	Analysis by JTK_Cycle, FDR < 0.05 Maximum (liver) still had < 3500 genes individually. 6 tissues had < 1000 genes Only 10 genes oscillated in all 12 tissues, 7 of which were core clock genes.
Suzuki et al., 2014	Mouse liver	RNA-seq	3 hrs	Light:Dark (12:12 hr)	24 hrs	-	1126 (of 11,926, 9.4% of expressed genes)	Used own curve fitting approach.
Hughes et al., 2012b	Mouse liver Wild type, *Clock* abolished and *Clock*-rescue mice	Microarray	2 hrs	Light:Dark (12:12 hr)	48 hrs	-	576 in wild type 253 in *Clock*-rescue 67 in *Clock* abolished	Analysis by JTK_Cycle, FDR < 0.05 > 50% of 576 overlapped with previous study Hughes et al., 2009.
Hughes et al., 2009	Mouse liver, NIH3T3 and U2OS cells	Microarray	1 hr	Light:Dark (12:12 hr)	48 hrs	-	by COSOPT: 5282 by Fisher's G: 4148 Both: 3667 (FDR < 0.05) (of 18,581 transcripts)	COSOPT is a cosine curve fitting approach (Straume, 2004) Fisher's G is a periodogram-based method (Wichert et al., 2004) Also reported 11 circadian genes in NIH3T3 cells, 7 in U2OS cells (COSOPT 5% FDR) FDR < 0.01, COSOPT: 3371, Fisher's G: 2914, Both: 2412.
Mockler et al., 2011	MMH-D3 murine hepatocytes	Microarray	2 hrs	Synchronisation by serum shock	48 hrs	-	1130	Analysis by Fisher's G q < 0.05.

Hughes et al., 2007	Mouse pituitary gland	Microarray	1 hr	Light:Dark (12:12 hr)	48 hrs	-	by COSOPT: 334 by Fisher's G: 1152 Both: 274 (FDR < 0.05)	Also present FDR < 0.01 values: COSOPT: 131, Fisher's G: 316, Both: 120.
Menger et al., 2007	NIH/3T3 fibroblasts	Microarray	6 hrs	Forskolin application	48 hrs	-	323	Analysis by curve fitting approach. Reports overlap with other mammalian cell studies; 7 genes with Rat SCN tissue, 8 with SCN2.2 cells (Menger et al., 2005) 5 genes with Rat-1 (Duffield et al., 2002) 8 with Rat 3Y1 (Grundschober et al., 2001).
Menger et al., 2005	Rat SCN tissue and immortalised SCN2.2 cells	Microarray	6 hrs	Light:Dark (12:12 hr) (SCN tissue) Forskolin application (cells)	48 hrs	-	SCN tissue: 301 (3.38%) SCN2.2 cells: 162 (1.82%) Both: 57 (0.64%)	Analysis by curve fitting approach.
Duffield et al., 2002	rat-1 fibroblasts	Microarray (not whole transcriptome)	4 hrs	Serum shock	48 hrs	-	44 genes (2% of a subset of genes sampled)	Analysed by CORRCOS algorithm, on which COSOPT is based.
Grundschober et al., 2001	rat-1 fibroblasts	Microarray	4 hrs	Serum shock	76 hrs	-	85 genes (of 9957)	Analysed by own spectral analysis approach.

Table 6.2: Summary of transcriptome-wide circadian gene expression time series performed in *Drosophila melanogaster*.

Reference	Organism/ Tissues	Transcriptomic Method	Sampling Frequency	Entrainment Cycles	Analysis Period Length	No. Rhythmic Transcripts/Genes (entrainment)	No. Rhythmic Transcripts/Genes (free-running)	Other Notes (of interest to thesis text)
Hughes et al., 2012a	*D. melanogaster* brains	RNA-seq	6 hrs (LD) 4 hrs (DD)	Light:Dark (12:12 hr)	42 hrs (LD) 48 hrs (DD)	JTK_Cycle: 870 Fisher's G: 782 (FDR < 0.05)	JTK_Cycle: 771 Fisher's G: 1214 (FDR < 0.05)	Also report at 1% and 0.1% FDR, and find significant rhythms in *per^0* mutants: *per^0* FDR < 0.05 in LD, JTK_Cycle: 262, Fisher's G: 302 in DD, JTK_Cycle: 332, Fisher's G: 486 9 of 159 cycling genes (5.66%) in DD were also found to cycle in LD at FDR < 0.01, both tests. Included most cycling core clock genes.
Rodriguez et al., 2013	*D. melanogaster* heads	Nascent-seq and mRNA-seq	4 hrs	Light:Dark (12:12 hr)	24 hrs	Nascent-seq: 136 mRNA: 237 Both: 44	-	Analysed by Fourier analysis approach.
Abruzzi et al., 2017	*D. melanogaster* brains Subsets of neurons	Rna-seq	4 hrs	Light:Dark (12:12 hr)	24 hrs	LNvs: 249 LNds: 303 DN1s: 185 TH neurons: 31 All sets: 4	-	Rhythmic genes given from both JTK_Cycle and Fourier-based approach, p < 0.05) Stringent approach, additional caveats include 2-fold change requirement. Some results omitted here.
Keegan et al., 2007	*D. melanogaster* Reanalysed data from the 5 studies below	Microarray	Artificial 6-point time series.	Light:Dark	24 hrs (but analysed by appending data)	214	-	Analysis by appending single day data and Fourier-based method. Only 81 genes found overlapped with prior studies. Reported overlapping gene no. between studies: 27.8% (highest between 2 reports). 17.4% (between 3), 10.4% (4) and 9.7% (5).

								7 genes found in all most clock genes included.
Claridge-Chang et al., 2001	*Drosophila melanogaster* heads	Microarray	4 hrs	Light:Dark (12:12 hr)	48 hrs	Combined entrainment and free-running days: 293 ($p < 0.05$)	Combined entrainment and free-running days: 293 ($p < 0.05$)	Analysed by using Fourier-based method. 158 genes at $p < 0.02$ Also report on per^0, tim^{01}, and Clk^{Jrk} clock mutants.
McDonald and Rosbash, 2001	*Drosophila melanogaster* heads	Microarray	4 hrs	Light:Dark (12:12 hr)	24 hrs	-	134	Stringent screening approach: Including low expression, those with no significant expression differences, then cosine fitting method. Core clock genes (*per*, *tim*, *clk*, *vri*) are among 10 most significant.
Ueda et al., 2002	*Drosophila melanogaster* heads	Microarray	4 hrs	Light:Dark (12:12 hr)	24 hrs	712 (5.3%)	456	Cosine curve fitting approach. 115 genes identified in LD were identified in DD (25.22%). Identified the core clock genes. Also used clk^{jrk} arrhythmic mutants.
Ceriani et al., 2002	*Drosophila melanogaster* heads and bodies	Microarray	4 hrs	Light:Dark (12:12 hr)	48 hrs (both entrainment and free-running)	Head: 1206 Body: 1144 $P < 0.05$	-	Analysis by COSOPT. Also report 0.01 and 0.025 significance levels. Failed to identify per and clock oscillations at 0.01, does at 0.05. Overlap: find half genes overlap in dd, 4/120 in mutants.

								Also looked at clk^{Jrk} arrhythmic mutants.
Lin et al., 2002	*D. melanogaster* heads	Microarray	4 hrs	Light:Dark (12:12 hr)	44 hrs	377	447	Analysed by autocorrelation approach. 46 genes identified in LD were identified in DD (10.29%). Also included per^{01} clock mutants.

Table 6.3: Select examples of transcriptome-wide circadian gene expression time series performed in animals other than mammals and *Drosophila melanogaster*.

Reference	Organism/ Tissues	Transcriptomic Method	Sampling Frequency	Entrainment Cycles	Analysis Period Length	No. Rhythmic Transcripts/Genes (entrainment)	No. Rhythmic Transcripts/Genes (free-running)	Other Notes (of interest to thesis text)
Van der Linden et al., 2010	*Caenorhabditis elegans*, whole organism	Microarray	4 hrs	Light:Dark (12:12 hrs) Warm:Cold (12:12 hrs)	44 hrs (but analysed by appending data)	Light: 775 Temperature: 1817 Both: 107	Free-run only: Light: 111 Temperature 198 Both: 2 Entrainment and Free-run combined: Light: 292 Temp.: 88 Both: 0	Analysis by appending single day data and Fourier based method. Also used by Keegan et al., 2007 below.
Matsumae et al., 2015	*Ciona intestinalis* (sea squirt), whole organism	Microarray	6 hrs	Unclear	42 hrs	817 (of 21,938)	-	Analysis by COSOPT method. One replicate. No p values given.
Oren et al., 2015	*Nematostella vectensis* (sea anemone), whole organism	RNA-seq	4 hrs	Light:Dark (12:12 hrs)	48 hrs	180	-	Fourier/periodogram analysis approach. One replicate.
Leming et al., 2014	*Aedes aegypti* (mosquito) heads	Microarray	4 hrs	Light:Dark (12:12 hrs)	44 hrs	1035 (of 15202)	887 (of 15202)	JTK_Cycle 5% 13,528 non rhythmic 248 genes identified in LD were identified in DD (27.96%).
Rund et al., 2013 Rund et al., 2011	*Anopheles gambiae* (mosquito) heads and bodies	Microarray	4 hrs	Light:Dark (12:12 hrs)	48 hrs (LD and DD)	JTK_Cycle: Heads: 1943 Bodies: 1216 (FDR < 0.05)	JTK_Cycle: Heads: 913 Bodies: 458 (FDR < 0.05)	Also report at different statistical thresholds and using COSOPT and DFT algorithms. Also detail overlapping genes between tests.

Ptitsyn et al., 2011	*Aedes aegypti* (mosquito) heads	Microarray	4 hrs	Light:Dark (12:12 hrs)	32 hrs	Fisher's G: 8445(19%) Pt-test: 19067 (42%) Autocorrelation: 6058 (13%) (of 45220 probes, 5% significance)	-	Analysed using own Pt-test. No FDR Correction.
Rodriguez-Zas et al., 2012	*Apis mellifera* (honey bee) brains Nurses and forager bees	Microarray	4 hrs	Light:Dark (12:12 hrs)	24 hrs	-	Nurses: 160 Foragers: 541	Used own cosine fitting approach. Did not detect some core clock genes in foragers- notably *per*.

Table 6.4 All significant genes (q < 0.05) identified by meta2d analysis over the full two-day time series, ordered by q-value.

q-value Rank	Gene	meta2d q-value
1	*fat-6*	0
2	*acl-12*	1.06E-10
3	*R07E5-4*	3.52E-08
4	*maoc-1*	9.39E-08
5	*col-93*	1.72E-07
6	*lea-1*	2.36E-06
7	*ruvb-1*	5.02E-06
8	*alh-9*	5.26E-06
9	*F46F2-3*	6.79E-06
10	*Y82E9BL-18*	6.79E-06
11	*cnc-4*	8.21E-06
12	*eif-3-E*	8.21E-06
13	*T10G3-3*	9.84E-06
14	*F10D11-6*	9.84E-06
15	*F54B8-4*	1.84E-05
16	*Y46D2A-2*	2.00E-05
17	*F08F3-4*	3.97E-05
18	*F35D11-4*	4.32E-05
19	*col-80*	6.13E-05
20	*B0507-3*	6.13E-05
21	*col-160*	6.17E-05
22	*ugt-12*	6.64E-05
23	*xrn-2*	8.63E-05
24	*Y47G6A-19*	0.00011727
25	*ctl-2*	0.00011787
26	*C23H3-2*	0.00013142
27	*T04A11-1*	0.00013142
28	*B0379-1*	0.00013142
29	*Y37H2A-14*	0.00013588
30	*gstk-1*	0.00013794
31	*abce-1*	0.00017617
32	*T25B9-1*	0.00017708
33	*nol-58*	0.00019565
34	*Y80D3A-9*	0.00019593
35	*C06B3-6*	0.00019593
36	*col-94*	0.00020103
37	*Y17G7B-21*	0.00020292
38	*mct-6*	0.00020292
39	*F13H8-3*	0.00022971
40	*C30H6-12*	0.00026084
41	*Y73B6BL-29*	0.0002953
42	*K11H12-9*	0.00029549
43	*clec-1*	0.00029549
44	*prg-1*	0.00030188
45	*K02D7-1*	0.00030188
46	*F53F1-4*	0.00034435
47	*hsp-16-1*	0.00034435
48	*dnj-12*	0.00046565
49	*F31D4-8*	0.000471
50	*acp-6*	0.00049576
51	*W04C9-4*	0.00054429
52	*F40F12-7*	0.00068703
53	*C48D1-5*	0.00068785
54	*acs-7*	0.0007091
55	*F17C8-9*	0.00074124
56	*dhs-18*	0.00083112
57	*aqp-4*	0.00087358
58	*nlp-33*	0.00103085
59	*C08G5-7*	0.00107615
60	*F42A8-1*	0.00117465
61	*fbxa-163*	0.00127737
62	*dhs-28*	0.0013014
63	*C44B11-4*	0.00132596
64	*T05E12-10*	0.00132917
65	*dct-11*	0.00172211
66	*B0302-4*	0.00178176
67	*col-124*	0.00205887
68	*K08C7-1*	0.00205887
69	*xbp-1*	0.0023149
70	*C05C8-7*	0.00248885
71	*syf-2*	0.00249652
72	*clec-186*	0.00249652

73	*C27B7-9*	0.00251376
74	*fat-5*	0.00251376
75	*fat-7*	0.00251376
76	*F10D2-10*	0.00251376
77	*ugt-26*	0.00275377
78	*fars-1*	0.00275377
79	*hpo-6*	0.0028984
80	*pbo-1*	0.0028984
81	*dhs-17*	0.00294864
82	*K01C8-1*	0.00297857
83	*C17G1-2*	0.00304681
84	*F45D3-4*	0.00349526
85	*ppw-2*	0.00349526
86	*ctl-1*	0.0037553
87	*F40H3-2*	0.00385236
88	*Y54G2A-73*	0.00385236
89	*C18D11-6*	0.00394133
90	*fust-1*	0.00394133
91	*ZK1127-5*	0.00394809
92	*T06A1-5*	0.00397952
93	*F10A3-1*	0.00414355
94	*C56C10-15*	0.00421469
95	*dim-1*	0.00421469
96	*Y43F8A-5*	0.00429323
97	*did-2*	0.00429323
98	*K08D12-6*	0.00429323
99	*dod-20*	0.0043219
100	*ZK1055-7*	0.0043219
101	*sma-2*	0.0043219
102	*C01B7-7*	0.00451201
103	*T01D1-3*	0.00456766
104	*sph-1*	0.00459826
105	*ZK822-8*	0.00459979
106	*mct-4*	0.00464792
107	*T23E7-2*	0.00473218
108	*acox-1-3*	0.0047947
109	*K01G5-5*	0.00504469
110	*Y45F10C-6*	0.00504469
111	*ZK1058-9*	0.00504469
112	*C53A3-2*	0.00508791
113	*F55F3-2*	0.00514216
114	*Y73B6BL-31*	0.00514216
115	*F17C11-6*	0.00514216
116	*cyp-33C9*	0.00514216
117	*swt-7*	0.00520611
118	*T19B10-2*	0.00526839
119	*rab-21*	0.00526839
120	*bath-19*	0.00609074
121	*F21A3-3*	0.00620072
122	*lmn-1*	0.00622026
123	*F21C10-10*	0.0062565
124	*W06A7-2*	0.00653142
125	*fbxa-73*	0.00669017
126	*VB0395L-1*	0.00669017
127	*Y48G1A-2*	0.00675868
128	*C38H2-2*	0.00724177
129	*R09E12-9*	0.00724177
130	*ttr-30*	0.00724177
131	*T24D1-3*	0.00724177
132	*C18B2-4*	0.00755993
133	*kynu-1*	0.00772159
134	*ZK1320-3*	0.00772159
135	*C25F9-12*	0.00776908
136	*egl-1*	0.0078253
137	*smd-1*	0.00813882
138	*T23E7-6*	0.00817126
139	*gst-35*	0.00875287
140	*pkg-2*	0.0089791
141	*F55C12-4*	0.00897938
142	*T01G5-1*	0.00897938
143	*oac-20*	0.00913943
144	*cor-1*	0.00932533
145	*pgp-1*	0.00956616
146	*arf-1-1*	0.00959856
147	*C34C12-4*	0.00995566
148	*F53F10-2*	0.01008282
149	*unc-97*	0.01012906
150	*F10E9-11*	0.01020382
151	*gmps-1*	0.01029129
152	*cut-4*	0.01040694
153	*C48B6-2*	0.01056793

154	*Y43C5A-3*	0.01056793
155	*cgh-1*	0.01056793
156	*pgp-9*	0.01056793
157	*T22F7-1*	0.01060433
158	*Y22D7AL-11*	0.01062967
159	*anmt-2*	0.01114019
160	*sqst-5*	0.01160952
161	*linc-81*	0.01160952
162	*T05E7-1*	0.01160952
163	*acox-3*	0.01163776
164	*twk-22*	0.01163776
165	*F53A9-6*	0.01176422
166	*anmt-3*	0.01225552
167	*F53F1-2*	0.01250722
168	*T20D4-3*	0.01261538
169	*dct-17*	0.0127692
170	*F43C9-1*	0.01285495
171	*spp-23*	0.01325938
172	*R03D7-5*	0.01351091
173	*Y57A10B-6*	0.01360608
174	*pcp-3*	0.01415385
175	*K08E4-2*	0.01502959
176	*F42G4-7*	0.0154938
177	*mlt-10*	0.0156086
178	*W01A8-6*	0.01636478
179	*prx-11*	0.01649883
180	*F39G3-3*	0.01761849
181	*oac-32*	0.01775559
182	*Y67D2-2*	0.01833951
183	*szy-2*	0.0183707
184	*bed-1*	0.0183707
185	*T11B7-2*	0.01867522
186	*Y47G6A-14*	0.01867522
187	*C29F5-1*	0.01867522
188	*F18F11-1*	0.01887239
189	*T28D6-7*	0.01887239
190	*Y51A2D-13*	0.01887239
191	*C56A3-6*	0.01887239
192	*C16C10-2*	0.01907302
193	*sbds-1*	0.01907302
194	*ZC434-9*	0.01907302
195	*ugt-63*	0.01907707
196	*F58A6-1*	0.01907707
197	*ddx-17*	0.01909699
198	*Y104H12D-2*	0.01929169
199	*Y48E1B-8*	0.01938611
200	*lst-1*	0.01986271
201	*cam-1*	0.02004853
202	*lin-18*	0.02004853
203	*nas-3*	0.02022857
204	*hil-2*	0.02026976
205	*ugt-32*	0.02098123
206	*F10E9-7*	0.02098123
207	*ZC247-2*	0.02098123
208	*alh-12*	0.02319826
209	*acox-1-5*	0.02323362
210	*bed-2*	0.02349164
211	*F30A10-9*	0.02372836
212	*ZK1321-4*	0.02421246
213	*C39H7-4*	0.02492339
214	*K10C2-12*	0.02497432
215	*ugt-13*	0.02512664
216	*R09A1-3*	0.02513948
217	*C06B3-7*	0.02514693
218	*pah-1*	0.02667193
219	*H06I04-6*	0.02722394
220	*ent-4*	0.02836192
221	*F30F8-9*	0.02911894
222	*hsp-16-48*	0.02911894
223	*col-19*	0.02946669
224	*Y54G11A-7*	0.02946669
225	*C49G7-12*	0.02946669
226	*nlp-24*	0.03037844
227	*F57B9-3*	0.03188764
228	*pot-3*	0.03257114
229	*col-142*	0.03257411
230	*slc-25A21*	0.03379218
231	*Y11D7A-7*	0.03379218
232	*vha-5*	0.03402242
233	*pqn-73*	0.03622522
234	*Y39B6A-1*	0.03727788

235	*C08F11-3*	0.03756952
236	*mrpl-24*	0.03756952
237	*clec-8*	0.03789896
238	*Y39B6A-41*	0.03811079
239	*ugt-11*	0.03822286
240	*ZK512-7*	0.03864519
241	*abf-2*	0.03882845
242	*vglu-2*	0.03882845
243	*F33C8-4*	0.03976525
244	*sdz-12*	0.03998241

245	*gst-26*	0.04064026
246	*oac-57*	0.04110849
247	*aqp-8*	0.04237729
248	*ZK813-4*	0.04263086
249	*tdo-2*	0.0427572
250	*akir-1*	0.04340187
251	*Y45G5AM-3*	0.04365392
252	*elpc-3*	0.04374061
253	*C25F9-16*	0.04384049
254	*cct-2*	0.04392696

255	*C37C3-2*	0.04527322
256	*fkh-6*	0.04739331
257	*T04F8-8*	0.04781395
258	*hda-11*	0.04784267
259	*tre-2*	0.04866781
260	*grd-14*	0.04896542
261	*E03H4-8*	0.04962503
262	*cyb-2-2*	0.04962503
263	*Y58A7A-5*	0.0499648

Table 6.5 All significant genes (q < 0.05) identified by meta2d analysis from the free-running day alone, ordered by q-value.

q-value Rank	*Gene*	meta2d q-value
1	*fat-6*	0.00104648
2	*sws-1*	0.00262845
3	*B0302-4*	0.00263666
4	*snf-7*	0.01259639
5	*F25B4-7*	0.01263205
6	*gstk-1*	0.01263205
7	*C39B5-14*	0.01637129
8	*pqn-73*	0.01947605
9	*F46G11-1*	0.01947605
10	*F07G6-10*	0.01947605
11	*Y7A9A-1*	0.01947605
12	*xbx-5*	0.02092428
13	*F28C6-5*	0.02223191
14	*ZC487-1*	0.02223191
15	*W01C9-2*	0.02651184
16	*C27B7-9*	0.03465754
17	*eif-2A*	0.03541046
18	*mls-1*	0.04211201
19	*eif-3-E*	0.04258919
20	*F56A4-10*	0.04258919
21	*F54F7-2*	0.04258919
22	*grd-5*	0.04493954
23	*K09H9-7*	0.04530596
24	*ZK185-5*	0.04530596
25	*F22E5-21*	0.04530596
26	*pbo-1*	0.04848048

Table 6.6: Functional analysis results from Brown module. All categories for which > 50% of terms had FDR < 0.1.

Annotation Cluster	Category	Functional Term/Decriptor	No. genes	Fold Enrichment	FDR (corrected)	Cluster Enrichment Score
1	GOTERM_BP_DIRECT	GO:0006470~protein dephosphorylation	56	8.00757576	1.0525E-33	17.1743379
	INTERPRO	IPR000242:Protein-tyrosine phosphatase, receptor/non-receptor type	39	9.45573309	2.8793E-26	
	GOTERM_BP_DIRECT	GO:0035335~peptidyl-tyrosine dephosphorylation	39	8.84234234	7.8819E-25	
	SMART	SM00194:PTPc	38	8.57782364	6.371E-25	
	INTERPRO	IPR003595:Protein-tyrosine phosphatase, catalytic	38	8.92536345	1.7225E-24	
	GOTERM_MF_DIRECT	GO:0004725~protein tyrosine phosphatase activity	39	7.92133965	3.3237E-23	
	SMART	SM00404:PTPc_motif	38	7.68430034	7.1756E-23	
	GOTERM_MF_DIRECT	GO:0004721~phosphoprotein phosphatase activity	35	7.65052462	3.7628E-20	
	INTERPRO	IPR000387:Protein-tyrosine/Dual specificity phosphatase	33	8.26770509	6.2122E-20	
	INTERPRO	IPR006186:Serine/threonine-specific protein phosphatase/bis(5-nucleosyl)-tetraphosphatase	22	10.7839632	1.5543E-15	
	SMART	SM00156:PP2Ac	22	9.28446357	9.4369E-15	
	INTERPRO	IPR004843:Metallophosphoesterase domain	21	7.39865654	2.773E-11	
	UP_KEYWORDS	Protein phosphatase	24	5.37700274	4.2186E-10	
	INTERPRO	IPR016130:Protein-tyrosine phosphatase, active site	17	6.84501558	2.0345E-08	
	GOTERM_MF_DIRECT	GO:0016787~hydrolase activity	57	1.64973482	0.00213043	
	UP_KEYWORDS	Hydrolase	55	1.26587976	0.49742858	
	GOTERM_BP_DIRECT	GO:0008152~metabolic process	51	1.2074318	0.80732286	
2	GOTERM_BP_DIRECT	GO:0008360~regulation of cell shape	51	13.0969388	3.3552E-43	14.8458839
	GOTERM_BP_DIRECT	GO:0018105~peptidyl-serine phosphorylation	51	9.79770992	1.9455E-35	
	GOTERM_MF_DIRECT	GO:0004672~protein kinase activity	77	4.33154703	1.1624E-27	
	GOTERM_BP_DIRECT	GO:0006468~protein phosphorylation	76	4.41724403	2.1994E-27	
	INTERPRO	IPR000719:Protein kinase, catalytic domain	78	4.24822791	1.1965E-26	
	INTERPRO	IPR011009:Protein kinase-like domain	81	3.79711271	2.3664E-24	
	GOTERM_MF_DIRECT	GO:0004674~protein serine/threonine kinase activity	58	4.57454049	1.2628E-21	

	INTERPRO	IPR017441:Protein kinase, ATP binding site	52	4.52706913	1.1235E-18	
	SMART	SM00220:S_TKc	54	3.74392979	1.0511E-16	
	GOTERM_MF_DIRECT	GO:0005524~ATP binding	92	2.22267873	1.4454E-12	
	INTERPRO	IPR008271:Serine/threonine-protein kinase, active site	30	3.26786762	5.1129E-07	
	UP_KEYWORDS	Kinase	43	2.09661138	9.724E-05	
	GOTERM_MF_DIRECT	GO:0000166~nucleotide binding	69	1.56333524	0.00187536	
	GOTERM_CC_DIRECT	GO:0005737~cytoplasm	90	1.459499	0.00173575	
	UP_KEYWORDS	Serine/threonine-protein kinase	17	1.60902734	0.53716574	
	UP_KEYWORDS	ATP-binding	48	1.2680096	0.58859454	
	UP_KEYWORDS	Transferase	61	1.13171629	0.93377877	
	UP_KEYWORDS	Nucleotide-binding	50	1.07529166	0.99561617	
	GOTERM_CC_DIRECT	GO:0005634~nucleus	80	1.04308827	0.99592972	
3	INTERPRO	IPR016187:C-type lectin fold	51	4.21231728	6.4011E-17	14.0900903
	GOTERM_MF_DIRECT	GO:0030246~carbohydrate binding	52	3.9782728	1.5703E-16	
	INTERPRO	IPR016186:C-type lectin-like	48	4.29491173	3.1279E-16	
	UP_KEYWORDS	Lectin	52	3.24071703	1.8154E-12	
	INTERPRO	IPR001304:C-type lectin	38	3.79130483	6.3601E-11	
	SMART	SM00034:CLECT	37	3.25013513	3.3689E-09	
4	INTERPRO	IPR000535:Major sperm protein	40	7.98169437	1.0164E-23	11.8237746
	INTERPRO	IPR008962:PapD-like	40	7.8428823	2.1458E-23	
	GOTERM_CC_DIRECT	GO:0005856~cytoskeleton	29	3.14488815	1.5157E-06	
	UP_KEYWORDS	Cytoskeleton	28	2.33764349	0.00076067	
	UP_KEYWORDS	Cytoplasm	36	0.93154214	0.99999991	
5	INTERPRO	IPR017441:Protein kinase, ATP binding site	52	4.52706913	1.1235E-18	8.20335975
	GOTERM_BP_DIRECT	GO:0038083~peptidyl-tyrosine autophosphorylation	20	11.7054264	7.3275E-15	
	GOTERM_CC_DIRECT	GO:0031234~extrinsic component of cytoplasmic side of plasma membrane	20	11.1894923	1.5654E-14	

	GOTERM_MF_DIRECT	GO:0004715~non-membrane spanning protein tyrosine kinase activity	21	9.83638879	4.1078E-14	
	GOTERM_BP_DIRECT	GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway	20	9.49685535	7.3208E-13	
	INTERPRO	IPR000980:SH2 domain	22	7.87400485	1.8188E-12	
	SMART	SM00252:SH2	21	7.27986348	1.3325E-11	
	INTERPRO	IPR020635:Tyrosine-protein kinase, catalytic domain	21	6.57658359	3.1597E-10	
	GOTERM_MF_DIRECT	GO:0005102~receptor binding	20	6.7504629	5.1201E-10	
	UP_KEYWORDS	Tyrosine-protein kinase	21	5.8997669	1.4033E-09	
	GOTERM_MF_DIRECT	GO:0004713~protein tyrosine kinase activity	21	5.87784208	2.3331E-09	
	SMART	SM00219:TyrKc	21	5.66211604	2.4998E-09	
	INTERPRO	IPR008266:Tyrosine-protein kinase, active site	17	5.39888552	8.4132E-07	
	GOTERM_BP_DIRECT	GO:0030154~cell differentiation	25	3.53464419	1.8791E-06	
	INTERPRO	IPR001245:Serine-threonine/tyrosine-protein kinase catalytic domain	21	3.913339	4.6504E-06	
	GOTERM_BP_DIRECT	GO:0042127~regulation of cell proliferation	21	3.67013889	1.2927E-05	
	UP_KEYWORDS	Kinase	43	2.09661138	9.724E-05	
	GOTERM_BP_DIRECT	GO:0016310~phosphorylation	30	2.16954023	0.00168329	
	GOTERM_MF_DIRECT	GO:0016301~kinase activity	29	1.99280132	0.00790327	
	GOTERM_BP_DIRECT	GO:0045087~innate immune response	26	2.06414301	0.01103985	
	UP_KEYWORDS	ATP-binding	48	1.2680096	0.58859454	
	GOTERM_MF_DIRECT	GO:0016740~transferase activity	43	1.18052354	0.91139335	
	UP_KEYWORDS	Transferase	61	1.13171629	0.93377877	
	UP_KEYWORDS	Nucleotide-binding	50	1.07529166	0.99561617	
6	GOTERM_BP_DIRECT	GO:0007060~male meiosis chromosome segregation	4	25.1666667	0.00319172	3.18243676
	GOTERM_CC_DIRECT	GO:0000785~chromatin	7	6.89272727	0.00425754	
	GOTERM_CC_DIRECT	GO:0000790~nuclear chromatin	8	4.10281385	0.03196519	

Table 6.7: Functional analysis results from Yellow module. All categories for which > 50% of terms had FDR < 0.1.

Annotation Cluster	Category	Functional Term/Descriptor	No. genes	Fold Enrichment	FDR (corrected)	Cluster Enrichment Score
1	GOTERM_CC_DIRECT	GO:0016021~integral component of membrane	241	1.34795996	1.2226E-09	6.14687936
	GOTERM_CC_DIRECT	GO:0016020~membrane	242	1.29626338	9.888E-08	
	UP_KEYWORDS	Transmembrane helix	238	1.26292493	0.00027195	
	UP_KEYWORDS	Transmembrane	238	1.26195295	0.0002921	
	UP_KEYWORDS	Membrane	238	1.21159917	0.00449929	
2	SMART	SM00051:DSL	5	12.564300 4	0.0044374	2.04899704
	INTERPRO	IPR001774:Delta/Serrate/lag-2 (DSL) protein	5	12.130350 2	0.00660282	
	UP_KEYWORDS	EGF-like domain	12	3.13951236	0.01690292	
	INTERPRO	IPR000742:Epidermal growth factor-like domain	13	2.83850195	0.02840443	
	GOTERM_BP_DIRECT	GO:0007154~cell communication	5	8.37047101	0.03344772	
	INTERPRO	IPR013032:EGF-like, conserved site	13	2.72932879	0.03943941	
	GOTERM_BP_DIRECT	GO:0001708~cell fate specification	7	3.58734472	0.16522858	
	SMART	SM00181:EGF	9	2.78824201	0.15122748	
	UP_KEYWORDS	Disulfide bond	30	1.25010825	0.89373622	
	GOTERM_BP_DIRECT	GO:0040020~regulation of meiotic nuclear division	5	1.55008722	0.99940039	
	GOTERM_BP_DIRECT	GO:0042127~regulation of cell proliferation	5	0.94403808	1	

List of Abbreviations

BP: Base pair

BH: Benjamini-Hochberg (procedure or correction)

CPS: Counts per second

DAVID: Database for Annotation, Visualization and Integrated Discovery

FDR: False discovery rate

GO: Gene Ontology

ipRGC: Intrinsically photosensitive retinal ganglion cell

LD: Light:Dark

KEGG: Kyoto Encyclopedia of Genes and Genomes

RNA-seq: RNA sequencing

RT-qPCR: Quantitative reverse transcription polymerase chain reaction

TPM: Transcripts per Million

TTFL: Transcription-translation feedback loop

WC: Warm:Cold

WGCNA: Weighted gene co-expression network analysis

Glossary

Acrophase: The highest point of a circadian cycle

Antiphase/antiphasic entrainment: Term used here to refer to the application of light and temperature simultaneously as zeitgebers, pairing a light phase with a cold/cryophase and a dark phase with a warm/thermophase

Bathyphase: The lowest point of a circadian cycle

Circadian: A cycle lasting approximately one day, taken from the Latin *circa* (around) and *dies* (day)

Counts per second/CPS: A measure of luminescence (photon count)

Cryophase: The cold phase of a warm:cold cycle

Ectotherm: An organism that cannot regulate its own body temperature

Endotherm: An organism that can regulates its own body temperature

Entrainment/entrainment phase: Exposure to diurnal cycles to synchronise and set the phase of a circadian clock

Free-running/free-running phase: Constant conditions in which observed circadian rhythms reflect endogenous activity of the circadian clock

In-phase entrainment: Term used here to refer to the application of light and temperature simultaneously as zeitgebers, pairing a light phase with warm/thermophase a and a dark phase with a cold/cryophase

Period: The frequency of a rhythmic pattern

Poikilotherm: An organism whose body temperature varies due to environmental change

Subjective day/subjective night: In constant conditions following diurnal cycles, the phase of a 24-hour cycle in which day/night conditions would have occurred

Thermophase: The warm phase of a warm:cold cycle

Transcripts per Million/TPM: A measure of relative gene expression in RNA-seq data

Ultradian: A rhythm with a duration of less than one day

Zeitgeber: Translated from ‘time-giver’, an environmental signal by which the phase of a circadian clock is set.

References

Adams, M., D. et al. The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185-2196 (2000).

Alexander, A. G., Marfil, V. & Li, C. Use of *Caenorhabditis elegans* as a model to study Alzheimer's disease and other neurodegenerative diseases. *Front. Genet.* **5,** 279 (2014).

Androwski, R. J., Flatt, K. M. & Schroeder, N. E. Phenotypic plasticity and remodeling in the stress-induced *Caenorhabditis elegans* dauer. *Wiley Interdiscip. Rev. Dev. Biol.* **6,** e278 (2017).

Banerjee, D., Kwok, A., Lin, S-Y. & Slack, F. J. Developmental timing in *C. elegans* is regulated by *kin-20* and *tim-1*, homologs of core circadian clock genes. *Dev. Cell* **8,** 287-295 (2005).

Barrière, A. & Félix, M. A. High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Curr. Biol.* **15,** 1176-1184 (2005).

Ben-Moshe Livne, Z. et al. Genetically blocking the zebrafish pineal clock affects circadian behavior. *PLoS Genet.* **12,** e1006445 (2016).

Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Series B (Methodological)* **57,** 289-300 (1995).

Benson, D. A. et al. GenBank. *Nucleic Acids Res.* **41,** D36-D42 (2013).

Berardi, S. et al. The Period protein homolog LIN-42 regulates germline development in *C. elegans*. *Mech. Dev.* **153,** 42-53 (2018).

Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli. J. Bacteriol*. **62,** 293-300 (1951).

Borbély, A. A., Daan, S., Wirz-Justice, A. & Deboer, T. The two-process model of sleep regulation: A reappraisal. *J. Sleep Res*. **25,** 131-143 (2016).

Borner, J., Rehm, P., Schill, R. O., Ebersberger, I. & Burmester, T. A transcriptome approach to ecdysozoan phylogeny. *Mol. Phylogenet. Evol.* **80,** 79-87 (2014).

Bourlat, S. J. et al. Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* **444,** 85-88 (2006).

Bozek K, et al. Regulation of clock-controlled genes in mammals. *PLoS One* **4,** e4882 (2009).

Brock, T. J., Browse, J. & Watts, J. L. Genetic regulation of unsaturated fatty acid composition in *C. elegans*. *PLoS Genet.* **2,** e108 (2006).

Brown, L. A., Fisk, A. S., Pothecary, C. A. & Peirson, S. N. Telling the time with a broken clock: Quantifying circadian disruption in animal models. *Biology* **8,** 18 (2019).

Brown, L. A. et al. Meta-analysis of transcriptomic datasets identifies genes enriched in the mammalian circadian pacemaker. *Nucleic Acids Res.* **45,** 9860-9873 (2017).

Buhr, E. D., Yoo, S-H. & Takahashi, J. S. Temperature as a universal resetting cue for mammalian circadian oscillators. *Science* **330,** 379-385 (2010).

Burr, A. H. The photomovement of *Caenorhabditis elegans*, a nematode which lacks ocelli. Proof that the response is to light not radiant heating. *Photochem. Photobiol.* **41,** 577-582 (1985).

Byerly, L., Cassada, R. C. & Russell, R. L. The life cycle of the nematode *Caenorhabditis elegans*. I. Wild-type growth and reproduction. *Dev. Biol.* **51,** 23-33 (1976).

The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282,** 2012-2018 (1998).

Cavallari, N. et al. A blind circadian clock in cavefish reveals that opsins mediate peripheral clock photoreception. *PLoS Biol.* **9,** e1001142 (2011).

Ceriani, M. F. et al. Genome-wide expression analysis in *Drosophila* reveals genes controlling circadian behavior. *J. Neurosci.* **22,** 9305-9319 (2002).

Chasnov, J. R. & Chow, K. L. Why are there males in the hermaphroditic species *Caenorhabditis elegans*? *Genetics* **160,** 983-994 (2002).

Chen, L. et al. Correlation between RNA-Seq and microarrays results using TCGA data. *Gene* **628,** 200-204 (2017).

Chen, Y. C., Liu, T., Yu, C. H., Chiang, T. Y., & Hwang, C. C. Effects of GC bias in Next-Generation-Sequencing data on de novo genome assembly. *PLoS ONE* **8,** e62856 (2013).

Cheng, H., Govindan, J. A. & Greenstein, D. Regulated trafficking of the MSP/Eph receptor during oocyte meiotic maturation in *C. elegans*. *Curr. Biol.* **18,** 705-714 (2008).

Chhangawala, S., Rudy, G., Mason, C. E., & Rosenfeld, J. A. The impact of read length on quantification of differentially expressed genes and splice junction detection. *Genome Biol.* **16,** 131 (2015).

Choi, S., Chatzigeorgiou, M., Taylor, K. P., Schafer, W. R. & Kaplan, J. M. Analysis of NPR-1 reveals a circuit mechanism for behavioral quiescence in *C. elegans*. *Neuron* **78,** 869-880 (2013).

Coburn, C. M. & Bargmann, C. I. A putative cyclic nucleotide-gated channel is required for sensory development and function in *C. elegans*. *Neuron* **17,** 695-706 (1996).

COMBINE-lab. wasabi. (2018). at https://github.com/COMBINE-lab/wasabi

Cormack, B. P., Valdivia, R. H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173,** 33-38 (1996).

Corsi, A. K., Wightman, B. & Chalfie, M. A transparent window into biology: A primer on *Caenorhabditis elegans*. *Genetics* **200,** 387-407 (2015).

Čopič, A. et al. A giant amphipathic helix from a perilipin that is adapted for coating lipid droplets. *Nat. Commun.* **9,** 1332 (2018).

Currie, J., Goda, T. & Wijnen, H. Selective entrainment of the *Drosophila* circadian clock to daily gradients in environmental temperature. *BMC Biol.* **7,** 49 (2009).

de Bekker, C., Will, I., Hughes, D. P., Brachmann, A. & Merrow, M. Daily rhythms and enrichment patterns in the transcriptome of the behavior-manipulating parasite *Ophiocordyceps kimflemingiae*. *PLoS One* **12**, e0187170 (2017).

Denzel, M. S., Lapierre, L. R. & Mack, H. I. D. Emerging topics in *C. elegans* aging research: Transcriptional regulation, stress response and epigenetics. *Mech. Ageing Dev.* **177,** 4-21 (2019).

Dusik, V. et al. The MAP Kinase p38 is part of *Drosophila melanogaster's* circadian clock. *PLoS Genet.* **10,** e1004565 (2014).

Edelman, T. L. B. et al. Analysis of a *lin-42*/*period* null allele Implicates all three isoforms in regulation of *Caenorhabditis elegans* molting and developmental timing. *G3- Genes. Genom. Genet.* **6,** 4077-4086 (2016).

Edgar, R. S. et al. Peroxiredoxins are conserved markers of circadian rhythms. *Nature* **485,** 459-464 (2012).

Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32,** 3047-3048 (2016).

Feeney, K. A., Putker, M., Brancaccio, M. & O'Neill, J. S. In-depth characterization of firefly luciferase as a reporter of circadian gene expression in mammalian cells. *J. Biol. Rhythms* **31,** 540-550 (2016).

Félix, M. A. & Duveau, F. Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biol.* **10,** 59 (2012).

Fogg, P. C. M. et al. Class IIa histone deacetylases are conserved regulators of circadian function. *J. Biol. Chem.* **289,** 34341-34348 (2014).

Frøkjær-Jensen, C. et al. Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat. Genet.* **40,** 1375-1383 (2008).

Fuchikawa, T., Eban-Rothschild, A., Nagari, M., Shemesh, Y. & Bloch, G. Potent social synchronization can override photic entrainment of circadian rhythms. *Nat. Commun.* **7,** 11662 (2016).

Gal, T. Z., Glazer, I. & Koltai, H. An LEA group 3 family member is involved in survival of *C. elegans* during exposure to stress. *FEBS Lett.* **577,** 21-26 (2004).

The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25,** 25-29 (2000).

The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47,** D330-D338 (2019).

Glaser, F. T. & Stanewsky, R. Synchronization of the *Drosophila* circadian clock by temperature cycles. *Cold Spring Harb. Symp. Quant. Biol*. **72,** 233-242 (2007).

Glaser, F. T. & Stanewsky, R. Temperature synchronization of the *Drosophila* circadian clock. *Curr. Biol.* **15,** 1352-1363 (2005).

Glynn, E. F., Chen, J. & Mushegian, A. R. Detecting periodic patterns in unevenly spaced gene expression time series using Lomb–Scargle periodograms. *Bioinformatics* **22,** 310-316 (2006).

Goda, T., Umezaki, Y., Alwattari, F., Seo, H. W. & Hamada, F. N. Neuropeptides PDF and DH31 hierarchically regulate free-running rhythmicity in *Drosophila* circadian locomotor activity. *Sci. Rep.* **9,** 838 (2019).

Golden, J. W. & Riddle, D. L. The *Caenorhabditis elegans* dauer larva: Developmental effects of pheromone, food, and temperature. *Dev. Biol.* **102,** 368-378 (1984).

Gong, J. et al. The *C. elegans* taste receptor homolog LITE-1 is a photoreceptor. *Cell* **167,** 1252-1263 (2016).

Goya, M. E., Romanowski, A., Caldart, C. S., Benard, C. Y., Golombek, D. A. Circadian rhythms identified in *Caenorhabditis elegans* by in vivo long-term monitoring of a bioluminescent reporter. *Proc. Natl. Acad. Sci. U. S. A*. **113,** E7837-E7845 (2016).

Han, S. M., Cottee, P. A. & Miller, M. A. Sperm and oocyte communication mechanisms controlling *C. elegans* fertility. *Dev. Dyn.* **239,** 1265-1281 (2010).

Hardin, P. E. The circadian timekeeping system of *Drosophila*. *Curr. Biol.* **15,** R714-R722 (2005).

Hargreaves, J. K., et al. Wavelet spectral testing: Application to nonstationary circadian rhythms. *Ann. Appl. Stat.* (2019).

Harmer, S. L. The circadian system in higher plants*. Annu. Rev. Plant Biol.* **60,** 357-377 (2009).

Harper, R. E. F., Dayan, P., Albert, J. T. & Stanewsky, R. Sensory conflict disrupts activity of the *Drosophila* circadian network. *Cell Rep.* **17,** 1711-1718 (2016).

Harper, R. E. F., Ogueta, M., Dayan, P., Stanewsky, R. & Albert, J. T. Light dominates peripheral circadian oscillations in *Drosophila melanogaster* during sensory conflict. *J. Biol. Rhythms* **32,** 423-432 (2017).

Hayashi Y, Sanada K, Hirota T, Shimizu F & Fukada Y. p38 Mitogen-activated Protein Kinase regulates oscillation of chick pineal circadian clock. *J. Biol. Chem.* **278,** 25166-25171 (2003).

Hedgecock, E. M. & Russell, R. L. Normal and mutant thermotaxis in the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* **72,** 4061-4065 (1975).

Hendriks, G. J., Gaidatzis, D., Aeschimann, F. & Großhans, H. Extensive oscillatory gene expression during *C. elegans* larval development. *Mol. Cell* **53,** 380-392 (2014).

Herrero, A. et al. Pigment-dispersing factor signaling in the circadian system of *Caenorhabditis elegans*. *Genes. Brain. Behav.* **2,** 493-501 (2015).

Hoogewijs, D., Houthoofd, K., Matthijssens, F., Vandesompele, J. & Vanfleteren, J. R. Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in *C. elegans*. *BMC Mol. Biol.* **9,** 9 (2008).

Huang DW, Sherman BT & Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protoc.* **4,** 44-57 (2009a).

Huang DW, Sherman BT & Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37,** 1-13 (2009b).

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

Higgins, B. J. & Hirsh, D. Roller mutants of the nematode *Caenorhabditis elegans*. *MGG Mol. Gen. Genet.* **150,** 63-72 (1977).

Hughes, M. E. et al. Guidelines for genome-scale analysis of biological rhythms. *J. Biol. Rhythms* **32,** 380-393 (2017).

Hughes, M. E., Grant, G. R., Paquin, C., Qian, J. & Nitabach, M. N. Deep sequencing the circadian and diurnal transcriptome of *Drosophila* brain. *Genome Res.* **22,** 1266-1281 (2012a).

Hughes, M. E. et al. High resolution time course analysis of gene expression from the liver and pituitary. *Cold. Spring. Harb. Symp. Quant. Biol.* **72,** 381-386 (2007).

Hughes, M. E. et al. Harmonics of circadian gene transcription in mammals. *PLoS Genet.* **5,** e1000442 (2009).

Hughes, M. E. et al. Brain-specific rescue of *Clock* reveals system-driven transcriptional rhythms in peripheral tissue. *PLoS Genet.* **8,** e1002835 (2012b).

Hurley, J. M. et al. Analysis of clock-regulated genes in *Neurospora* reveals widespread posttranscriptional control of metabolic potential. *Proc. Natl. Acad. Sci.* **111,** 16995-17002 (2014).

Hurley, J., Loros, J. J. & Dunlap, J. C. Dissecting the mechanisms of the clock in *Neurospora*. *Methods Enzymol.* **551,** 29-52 (2015).

Ignowski, J. M. & Schaffer, D. V. Kinetic analysis and modeling of firefly luciferase as a quantitative reporter gene in live mammalian cells. *Biotechnol. Bioeng.* **86,** 827-834 (2004).

Imanikia S., Sheng M., Castro C., Griffin J. L., and Taylor R. C. XBP-1 remodels lipid metabolism to extend longevity. *Cell Rep.* **28,** 581-589 (2019).

Isojima, Y. et al. CKIε/δ-dependent phosphorylation is a temperature-insensitive, period-determining process in the mammalian circadian clock. *Proc. Natl. Acad. Sci.* **106,** 15744-15749 (2009).

Ito, C. & Tomioka, K. Heterogeneity of the peripheral circadian systems in *Drosophila melanogaster*: A review. *Front. Physiol.* **7,** 8 (2016).

Jagannath, A., Taylor, L., Wakaf, Z., Vasudevan, S. R., & Foster, R. G. The genetics of circadian rhythms, sleep and health. *Hum Mol Genet.* **26,** R128-R138 (2017).

Jeon, M., Gardner, H. F., Miller, E. A., Deshler, J & Rougvie, A. E. Similarity of the *C. elegans* developmental timing protein LIN-42 to circadian rhythm proteins. *Science* **286,** 1141-1146 (1999).

Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8,** 118-127 (2007).

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44,** D457-D462 (2016).

Kasukawa, T. et al. Quantitative expression profile of distinct functional regions in the adult mouse brain. *PLoS One* **6,** e23228 (2011).

Keegan, K. P., Pradhan, S., Wang, J-P. & Allada, R. Meta-analysis of *Drosophila* circadian microarray studies identifies a novel set of rhythmically expressed genes. *PLoS Comput. Biol.* **3,** e208 (2007).

Kenyon, C. Ageing in *C. elegans*, *C. elegans II*, ed. D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1997).

Kippert, F., Saunders, D. S.**,** Blaxter, M. L. *Caenorhabditis elegans* has a circadian clock. *Curr. Biol.* **12,** R47-R49 (2002).

Klass, M., R. & Hirsh, D. Sperm isolation and biochemical analysis of the major sperm protein from *C. elegans*. *Dev. Biol.* **84,** 299-312 (1981).

Ko, H. W., Jiang, J. & Edery, I. Role for Slimb in the degradation of *Drosophila* Period protein phosphorylated by Doubletime. *Nature* **420,** 673-678 (2002).

Kon, N., Sugiyama, Y., Yoshitane, H., Kameshita, I. & Fukada, Y. Cell-based inhibitor screening identifies multiple protein kinases important for circadian clock oscillations. *Commun. Integr. Biol.* **8,** e982405 (2015).

Konopka, R. J. & Benzer, S. Clock mutants of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **68,** 2112-2116 (1971).

Koreth, J. & van den Heuvel, S. Cell-cycle control in *Caenorhabditis elegans*: How the worm moves from G1 to S. *Oncogene* **24,** 2756-2764 (2005).

Kristensen, A. K. J. Temperature and heat balance of soil. *Wiley on behalf of Nordic Society Oikos*. **10,** 103-120 (1959).

Krupp, J. J., Billeter, J., Wong, A., Choi, C., Nitabach, M. N. & Levine, J. D. Pigment-Dispersing Factor modulates pheromone production in clock cells that influence mating in *Drosophila*. *Neuron.* **79,** 54-68 (2013).

Kurosawa, G. & Iwasa, Y. Temperature compensation in circadian clock models. *J. Theor. Biol.* **233,** 453-468 (2005).

Kyriacou, C. P., Oldroyd, M., Wood, J., Sharp, M. & Hill, M. Clock mutations alter developmental timing in *Drosophila*. *Heredity* **64,** 395-401 (1990).

Lagido, C., Pettitt, J., Flett, A. & Glover, L. A. Bridging the phenotypic gap: real-time assessment of mitochondrial function and metabolism of the nematode *Caenorhabditis elegans*. *BMC Physiol.* **8,** 7 (2008).

Lahiri, K., Froehlich, N., Heyd, A., Foulkes, N. S. & Vallone, D. Developmental stage-specific regulation of the circadian clock by temperature in zebrafish. *Biomed Res. Int.* **2014,** 930308 (2014).

Lahiri, K. et al. Temperature regulates transcription in the zebrafish circadian clock. *PLoS Biol.* **3,** e351 (2005).

Langfelder, P. & Horvath, S. Tutorials for the WGCNA package. (2016). at https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/

Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9,** 559 (2008).

Lažetić, V. & Fay, D. S. Molting in *C. elegans*. *Worm* **6,** e1330246 (2017).

Lee, R. Y. N. et al. WormBase 2017: Molting into a new stage. *Nucleic Acids Res.* **46,** D869-D874 (2018).

Leek, J. T. et al. sva: Surrogate Variable Analysis. R package version 3.30.0. (2019). at https://bioconductor.org/packages/release/bioc/html/sva.html

Lehmann, R. et al. Assembly of a comprehensive regulatory network for the mammalian circadian clock: a bioinformatics approach. *PLoS One* **10,** e0126283 (2015).

Leming, M. T., Rund, S. S. C., Behura, S. K., Duffield, G. E. & O'Tousa, J. E. A database of circadian and diel rhythmic gene expression in the yellow fever mosquito *Aedes aegypti*. *BMC Genomics* **15,** 1128 (2014).

Leyva-Díaz, E. et al. Silencing of repetitive DNA is controlled by a member of an unusual *Caenorhabditis elegans* gene family. *Genetics* **207,** 529-545 (2017).

Li, J., Grant, G. R., Hogenesch, J. B. & Hughes, M. E. Considerations for RNA-seq analysis of circadian rhythms. *Methods Enzymol.* **551,** 349-367 (2015).

Li, S. et al. CGDB: A database of circadian genes in eukaryotes. *Nucleic Acids Res.* **45,** D397-D403 (2017).

Liu, Y., Tsinoremas, N. F., Johnson, C. H., Lebedeva, N. V., Golden S.S., Ishiura, M. & Kondo, T. Circadian orchestration of gene expression in cyanobacteria. *Genes & Dev.* **9,** 1469-1478 (1995).

Lin, Y. et al. Influence of the *period*-dependent circadian clock on diurnal, circadian, and aperiodic gene expression in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* **99,** 9562-9567 (2002).

Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* **25,** 402-408 (2001).

López-Olmeda, J. F., Madrid, J. A. & Sánchez-Vázquez, F. J. Light and temperature cycles as zeitgebers of zebrafish (*Danio rerio*) circadian activity rhythms. *Chronobiol. Int.* **23,** 537-550 (2006).

López-Olmeda, J. F. & Sánchez-Vázquez, F. J. Zebrafish temperature selection and synchronization of locomotor activity circadian rhythm to ahemeral cycles of light and temperature. *Chronobiol. Int.* **26,** 200-218 (2009).

Lowrey, P. L. et al. Positional syntenic cloning and functional characterization of the mammalian circadian mutation *tau*. *Science* **288,** 483-491 (2000).

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

Matsumae, H. et al. Detection of periodic patterns in microarray data reveals novel oscillating transcripts of biological rhythms in *Ciona intestinalis*. *Artif. Life Robot.* **20,** 347-352 (2015).

McCarter, J., Bartlett, B., Dang, T. & Schedl, T. On the control of oocyte meiotic maturation and ovulation in *Caenorhabditis elegans*. *Dev. Biol.* **205,** 111-128 (1999).

McClung, C. R. Plant circadian rhythms. *Plant Cell* **18,** 792-803 (2006).

McCulloch, K. A. & Rougvie, A. E. *Caenorhabditis elegans period* homolog *lin-42* regulates the timing of heterochronic miRNA expression. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 15450-15455 (2014).

McDonald, M. J. & Rosbash, M. Microarray analysis and organization of circadian gene expression in *Drosophila*. *Cell* **107,** 567-578 (2001).

Meireles-Filho, A. C. A., Bardet, A. F., Yáñez-Cuna, J. O., Stampfel, G. & Stark, A. Cis-regulatory requirements for tissue-specific programs of the circadian clock. *Curr. Biol.* **24,** 1-10 (2014).

Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10,** 3959-3970 (1991).

Möller, G. et al. Characterization of the HSD17B4 gene: D-specific multifunctional protein 2/17beta-hydroxysteroid dehydrogenase IV. *J. Steroid Biochem. Mol. Biol.* **69,** 441-446 (1999).

Nelson, M. D. et al. The neuropeptide NLP-22 regulates a sleep-like state in *Caenorhabditis elegans*. *Nat. Commun.* **4,** 2846 (2013).

Nicholas, W. L., Dougherty, E. C. & Hansen, E. L. Axenic cultivation of *Caenorhabditis briggsae* (Nematoda: Rhabditidae) with chemically undefined supplements; comparative studies with related nematodes. *Ann NY Acad Sci.* **77,** 218-236 (1959).

Ntambi, J. M. & Miyazaki, M. Regulation of stearoyl-CoA desaturases and role in metabolism. *Prog Lipid Res.* **43,** 91-104 (2004).

O'Neill, J.S., Reddy, A.B., Circadian clocks in human red blood cells. *Nature* **469,** 498-503 (2011).

Olmedo, M., Geibel, M., Artal-Sanz, M. & Merrow, M. A high-throughput method for the analysis of larval developmental phenotypes in *Caenorhabditis elegans*. *Genetics* **201,** 443-448 (2015).

Olmedo, M. et al. Circadian regulation of olfaction and an evolutionarily conserved, nontranscriptional marker in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **109,** 20479-20484 (2012).

Oren, M. et al. Profiling molecular and behavioral circadian rhythms in the non-symbiotic sea anemone *Nematostella vectensis*. *Sci. Rep*. **5,** 11418 (2015).

Parton, W. J. & Logan, J. A. A model for diurnal variation in soil and air temperature. *Agric. Meteorol.* **23,** 205-216 (1981).

Pasqualone, D. et al. Chromosome cohesion is regulated by a clock gene paralogue TIM-1. *Nature* **423,** 1002-1009 (2003).

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14,** 417-419 (2017).

Peirson, S. N., Halford, S. & Foster, R. G. The evolution of irradiance detection: melanopsin and the non-visual opsins. *Philos. Trans. R. Soc. B* **364,** 2849-2865 (2009).

Pembroke, W. G., Babbs, A., Davies, K. E., Ponting, C. P. & Oliver, P. L. Temporal transcriptomics suggest that twin-peaking genes reset the clock. *Elife* **4,** e10518 (2015).

Peng et al. Inhibition of fat accumulation by Hesperidin in *Caenorhabditis elegans*. *J. Agric. Food Chem.* **64,** 5207-5214 (2016).

Perales, R., King, D. M., Aguirre-chen, C. & Hammell, C. M. LIN-42, the *Caenorhabditis elegans* PERIOD homolog, negatively regulates microRNA transcription. *PLoS Genet.* **10,** e1004486 (2014).

Petersen, C. et al. Travelling at a slug's pace: possible invertebrate vectors of *Caenorhabditis* nematodes. *BMC Ecol.* **15,** 19 (2015).

Philip, N. S., Escobedo, F., Bahr, L. L., Berry, B. J. & Wojtovich, A. P. Mos1 element-mediated CRISPR integration of transgenes in *Caenorhabditis elegans*. *G3- Genes. Genom. Genet.* early online, (2019).

Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* **14,** 687-690 (2017).

Pizarro, A., Hayer, K., Lahens, N. F. & Hogenesch, J. B. CircaDB: A database of mammalian circadian gene expression profiles. *Nucleic Acids Res.* **41,** D1009-D1013 (2013).

Ptitsyn, A. A. et al. Rhythms and synchronization patterns in gene expression in the *Aedes aegypti* mosquito. *BMC Genomics* **12,** 153 (2011).

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2018). at https://www.R-project.org/

Raizen, D. M. et al. Lethargus is a *Caenorhabditis elegans* sleep-like state. *Nature* **451,** 569-572 (2008).

Ralph, M. & Menaker, M. A mutation of the circadian system in golden hamsters. *Science* **241,** 1225-1127 (1988).

Refinetti, R. Entrainment of circadian rhythm by ambient temperature cycles in mice. *J. Biol. Rhythms* **25,** 247-256 (2010).

Renn, S. C. P., Park, J. H., Rosbash, M., Hall, J. C. & Taghert, P. H. A *pdf* neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in *Drosophila*. *Cell* **99,** 791-802 (1999).

Reuter, J. A., Spacek, D. V., & Snyder, M. P. High-throughput sequencing technologies. *Mol. Cell* **58,** 586-597 (2015).

Rhodehouse, K. et al. The Doubletime homolog KIN-20 mainly regulates *let-7* independently of its effects on the Period homolog LIN-42 in *Caenorhabditis elegans*. *G3- Genes. Genom. Genet.* **8,** 2617-2629 (2018).

Rodriguez, J. et al. Nascent-Seq analysis of *Drosophila* cycling gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **110,** E275-E284 (2013).

Rodriguez-zas, S. L. et al. Microarray analysis of natural socially regulated plasticity in circadian rhythms of honey bees. *J. Biol. Rhythms* **27,** 12-24 (2012).

Romanowski, A., Garavaglia, M. J., Goya, M. E., Ghiringhelli, P. D. & Golombek, D. A. Potential conservation of circadian clock proteins in the phylum Nematoda as revealed by bioinformatic searches. *PLoS One* **9,** e112871 (2014).

Ruben M. D., et al., A database of tissue-specific rhythmically expressed human genes has potential applications in circadian medicine. *Sci. Transl. Med*. **10,** eaat8806 (2018).

Rund, S. S. C., Gentile J. E. & Duffield, G. E. Extensive circadian and light regulation of the transcriptome in the malaria mosquito *Anopheles gambiae*. *BMC Genomics* **14,** 218 (2013).

Rund, S. S. C., Hou, T. Y., Ward, S. M., Collins, F. H. & Duffield, G. E. Genome-wide profiling of diel and circadian gene expression in the malaria vector *Anopheles gambiae*. *Proc. Natl. Acad. Sci.* **108,** E421-E430 (2011).

Sagalovich, V. N., Fal'kov, E. Y. & Tsareva, T. I. Determination of diurnal soil-temperature cycles using remote sensing data. *Mapp. Sci. Remote Sens.* **39,** 46-55 (2002).

Sahar, S. & Sassone-Corsi, P. Metabolism and cancer: the circadian clock connection. *Nat. Rev. Cancer* **9,** 886-896 (2009).

Saigusa, T. et al. Circadian behavioural rhythm in *Caenorhabditis elegans*. *Curr. Biol.* **12,** R46-R47 (2002).

Salomé, P. A. & McClung, C. R. PSEUDO-RESPONSE REGULATOR 7 and 9 are partially redundant genes essential for the temperature responsiveness of the Arabidopsis circadian clock. *Plant Cell* **17,** 791-803 (2005).

Scheer, F. A. J. L., Hilton, M. F., Mantzoros, C. S. & Shea, S. A. Adverse metabolic and cardiovascular consequences of circadian misalignment. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 4453-4458 (2009).

Schmid, B., Helfrich-Förster, C. & Yoshii, T. A new ImageJ plug-in 'ActogramJ' for chronobiological analyses. *J. Biol. Rhythms* **26,** 464-467 (2011).

Schroeder, A. et al. The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* **7,** 3 (2006).

Schurch, N. J. et al. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA* **22,** 839-851 (2016).

Sehadova, H. et al. Temperature entrainment of *Drosophila'*s circadian clock involves the gene *nocte* and signaling from peripheral sensory tissues to the brain. *Neuron* **64,** 251-266 (2009).

Siepka, S. M. & Takahashi, J. S. Methods to record circadian rhythm wheel running activity in mice. *Methods Enzymol.* **393,** 230-239 (2005).

Simonetta, S. H. & Golombek, D. A. An automated tracking system for *Caenorhabditis elegans* locomotor behavior and circadian studies application. J*. Neurosci. Methods* **161,** 273-280 (2007).

Simonetta, S. H., Migliori, M. L., Romanowski, A. & Golombek, D. A. Timing of locomotor activity circadian rhythms in *Caenorhabditis elegans*. *PLoS One* **4,** e7571 (2009).

Simoni, A., Wolfgang, W., Topping, M. P., Kavlie, R. G., Stanewsky, R. & Albert, J. T. A mechanosensory pathway to the *Drosophila* circadian clock. *Science* **343,** 525-528 (2014).

Sinner, D. I., Kim, G. J., Henderson, G. C. & Igal, R. A. StearoylCoA Desaturase-5: a novel regulator of neuronal cell proliferation and differentiation. *PLoS One.* **7,** e39787 (2012).

Sokolove, P. G. & Bushell, W. N. The chi square periodogram: Its utility for analysis of circadian rhythms. *J. Theor. Biol.* **72,** 131-160 (1978).

Southern, M. M. & Millar, A. J. Circadian genetics in the model higher plant, *Arabidopsis thaliana*. Methods Enzymol. **393,** 23-35 (2005).

Stanewsky, R., Jamison, C. F., Plautz, J. D., Kay, S. A. & Hall, J. C. Multiple circadian-regulated elements contribute to cycling *period* gene expression in *Drosophila*. *EMBO J.* **16,** 5006-5018 (1997).

Sterken, M. G., Snoek, L. B., Kammenga, J. E. & Andersen, E. C. The laboratory domestication of *Caenorhabditis elegans*. *Trends Genet.* **31,** 224-231 (2015).

Stiernagle, T. Maintenance of *C. elegans*, *WormBook*, ed. The *C. elegans* Research Community, WormBook (2006). at: http://www.wormbook.org.

Stokkan, K-A., Yamazaki, S., Tei, H., Sakaki, Y. & Menaker, M. Entrainment of the circadian clock in the liver by feeding. *Science* **291,** 490-493 (2001).

Straume M. DNA microarray time series analysis: automated statistical assessment of circadian rhythms in gene expression patterning. *Methods Enzymol.* **383,** 149-166 (2004).

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

Tennessen, J. M., Gardner, H. F., Volk, M. L. & Rougvie, A. E. Novel heterochronic functions of the *Caenorhabditis elegans* period-related protein LIN-42. *Dev. Biol.* **289,** 30-43 (2006).

Tennessen, J. M., Opperman, K. J. & Rougvie, A. E. The *C. elegans* developmental timing protein LIN-42 regulates diapause in response to environmental cues. *Development* **137,** 3501-3511 (2010).

Top, D. et al. CK1/Doubletime activity delays transcription activation in the circadian clock. *Elife* **7,** e32679 (2018).

Tyson, J. J. & Novak, B. Temporal organization of the cell cycle. *Curr. Biol.* **18,** 759-768 (2008).

Ueda, H. R. et al. Genome-wide transcriptional orchestration of circadian rhythms in *Drosophila*. *J. Biol. Chem.* **277,** 14048-14052 (2002).

Uriu, K. Genetic oscillators in development. *Dev. Growth Differ.* **58,** 16-30 (2016).

Vallone, D., Gondi, S. B., Whitmore, D. & Foulkes, N. S. E-box function in a *period* gene repressed by light. *Proc. Natl. Acad. Sci.* **101,** 4106-4111 (2004).

Vallone, D., Lahiri, K., Dickmeis, T. & Foulkes, N. S. Start the clock! Circadian rhythms and development. *Developmental Dynamics* **236,** 142-155 (2007).

van der Linden, A. M. et al. Genome-Wide Analysis of light- and temperature-entrained circadian transcripts in *Caenorhabditis elegans*. *PLoS Biol* **8,** e1000503 (2010).

VanVickle-Chavez, S. J. & Van Gelder, R. N. Action spectrum of *Drosophila* cryptochrome. *J. Biol. Chem.* **282,** 10561-10566 (2007).

Vatine, G., Vallone, D., Gothilf, Y. & Foulkes, N. S. It's time to swim! Zebrafish and the circadian clock. *FEBS Lett.* **585,** 1485-1494 (2011).

Vaze, K. M. & Sharma, V. K. On the adaptive significance of circadian clocks for their owners. *Chronobiol. Int.* **30,** 413-433 (2013).

Vitaterna, M. H. et al. Mutagenesis and mapping of a mouse gene, clock, essential for circadian behavior. *Science* **264,** 719-725 (1994).

Wang, R., Kniazeva, M. & Han, M. Peroxisome protein transportation affects metabolism of branched-chain fatty acids that critically impact growth and development of *C. elegans*. *PLoS One* **8,** e76270 (2013).

Ward, A., Liu, J., Feng, Z. & Xu, X. Z. S. Light-sensitive neurons and channels mediate phototaxis in *C. elegans*. *Nat. Neurosci.* **11,** 916-922 (2008).

Weger, M. et al. Real-time in vivo monitoring of circadian E-box enhancer activity: A robust and sensitive zebrafish reporter line for developmental, chemical and neural biology of the circadian clock. *Dev. Biol.* **380,** 259-273 (2013).

Welsh, D. K., Yoo, S-H., Liu, A. C., Takahashi, J. S., Kay, S. A. Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression. *Curr. Biol.* **14,** 2289-2295 (2004).

Whitmore, D., Foulkes, N. S. & Sassone-Corsi, P. Light acts directly on organs and cells in culture to set the vertebrate circadian clock. *Nature* **404,** 87-91 (2000).

Wichert, S., Fokianos, K. & Strimmer, K. Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics* **20,** 5-20 (2004).

Wickham, H. ggplot2: *Elegant Graphics for Data Analysis*. (Springer-Verlag, New York, 2016).

Winbush, A., Gruner, M., Hennig, G. W. & van der Linden, A. M. Long-term imaging of circadian locomotor rhythms of a freely crawling *C. elegans* population. *J. Neurosci. Methods* **249,** 66-74 (2015).

Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K. & Hogenesch, J. B. MetaCycle: an integrated R package to evaluate periodicity in large scale data. *Bioinformatics* **32,** 3351-3353 (2016).

Xu, K., DiAngelo, J. R., Hughes, M. E., Hogenesch, J. B & Sehgal, A. Interaction between circadian clocks and metabolic physiology: implications for reproductive fitness. *Cell Metab.* **13,** 639-654 (2011).

Yang, M., Lee, J-E., Padgett, R. W. & Edery, I. Circadian regulation of a limited set of conserved microRNAs in *Drosophila*. *BMC Genomics* **9,** 83 (2008).

Yang, Y Cao, J. & Shi, Y. Identification and characterization of a gene encoding human lysophosphatidylglycerol acyltransferase. *J. Biol. Chem.* **279,** 55866-55874 (2004).

Yoo, S-H. et al. PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 5339-5346 (2004).

Yoshii, T., Hermann, C. & Helfrich-Förster, C. Cryptochrome-positive and -negative clock neurons in *Drosophila* entrain differentially to light and temperature. *J. Biol. Rhythms* **25,** 387-398 (2010).

Yoshii, T., Heshiki, Y., Ibuki-Ishibashi, T., Matsumoto, A., Tanimura, T. & Tomioka, K. Temperature cycles drive *Drosophila* circadian oscillation in constant light that otherwise induces behavioural arrhythmicity. *Eur. J. Neurosci.*, **22,** 1176-1184 (2005).

Yu, W. & Hardin, P. E. Circadian oscillators of *Drosophila* and mammals. *J. Cell Sci.* **119,** 4793-4795 (2006).

Yuen, K. W. Y., Nabeshima, K., Oegema, K. & Desai, A. Rapid de novo centromere formation occurs independently of heterochromatin protein 1 in *C. elegans* embryos. *Curr. Biol.* **21,** 1800-1807 (2011).

Zee, P. C., Attarian, H. & Videnovic, A. Circadian rhythm abnormalities. *Contin. Lifelong Learn. Neurol.* **19,** 132-147 (2013).

Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: Implications for biology and medicine. *Proc. Natl. Acad. Sci.* **111,** 16219-16224 (2014).

Zhang, S. O. et al. Genetic and dietary regulation of lipid droplet expansion in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 4640-4645 (2010).

Zhao, S., Fung-Leung, W-P., Bittner, A., Ngo, K. & Liu, X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One* **9,** e78644 (2014).