Functional analysis of a *Drosophila* clock gene

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

J. Michael Hennessy
Department of Genetics, University of Leicester

March 1999

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGMENTS

First I wish to thank a rather unique supervisor and good friend, Bambos Kyriacou. To my good fortune Bambos introduced me to Alex Peixoto whose infectious enthusiasm helped me to plan the early work, and has sustained my enthusiasm over many years. The work was subsequently completed in two laboratories in different countries, funded by CEC and BBSRC grants to Bambos.

In Italy where my education was immeasurably boosted by 'la dolce vita', I thrived with Rudi Costa's supervision, and lessons in the arts of life. Close friendship with Ezio Rosato and Mauro Zordan and the incredible cooking of Rudi's wife Luciana helped me feel rather 'at home' despite my very poor Italian. Special thanks are due to many students who worked with me at Padova: Denis Dujardin for help with cloning, Raphaella de Lisi with P-element injections and Chris Wood and Alexandra Hatzopoulou for *in situ* Hybridization. Thanks also to all my other Italian friends; Mario, Daniella , Modesto, Federica, Ricardo, Sylvana and many others for two wonderful and unforgettable years.

The second phase of my work, undertaken at Leicester, was mainly the enormous task of behavioural testing of the transgenic flies which simply could not have been undertaken without computer programming provided by Martin Couchman. As a late starter with protein work, I was also grateful to Alberto Piccin for help with Western blotting. I am indebted to Lesley Sawyer for work carried out on natural variants, which have complemented transgenic results contained in this work. Also to Helen Parkinson initially for providing the $(Thr\text{-}Gly)_{17}$ fragment and subsequently for extending my work with (*Thr-Gly*) deletions. Similarly to Ian Townson for extending the work with the chimaeric transgenes and helping me with southern blots. A special thanks also to Karen Garner for all the locomotor activity testing of combined temperature/light cycles.

I wish to express my gratitude to the whole staff and student body of the Genetics department at Leicester, past and present, collectively for their help and support. During the last few years I have ceaslessly travelled from Wales to Italy to Leicester and back again, somehow fathered two children, and attempted to complete the work presented here. I have always found friends willing to give their time to help this chaotic wanderer with techniques, coffee and chat, or emergency accomodation if needed. Nobody can ask for more.

I of course owe an enormous debt to my parents Rona and Evan, who did all, perhaps more than they should have, for my early education, and to Irene whose enthusiasm helped me to start my higher education.

Finally I am indebted to Caroline who has lovingly supported me in this mad pursuit, and especially for what she has done over the last five years since we have had Alana and Sean.

ABBREVIATIONS

aa	Amino acid
bp	Base pair
^{0}C	Degrees Centigrade
CT	Circadian time
g, mg, µg,	Grams, milligrams, micrograms
h/hr	hour
Kb	Kilobase pairs
LD	Light/dark
LL	Light/light
Min	Minutes
M	Molar
µl	microliters
ZT	Zeitgeber time

Abstract

The *period* (*per*)gene of *Drosophila* encodes a fundamental part of the circadian clock. The PER protein interacts with other proteins in a negative feedback loop which induces cycling of PER and other clock proteins. These proteins subsequently act as transcriptional regulators of other proteins in the loop. It is assumed therefore that they also act at the transcriptional level to affect the production of downstream genes which then control rhythmic phenotypes such as circadian activity patterns, pupal eclosion and lovesong cycle.

The work presented here concentrates on a polymorphic '(*Thr-Gly*)' region of *per*, part of which encodes for a series of threonine-glycine (Thr-Gly) repeats. Clock genes from orders as diverse as mammals, fungi and cyanobacteria have identifiable (*Thr-Gly*) regions albeit encoding for smaller (Thr-Gly) repeats, therefore the possibility exists that the region has some fundamental importance to clock function.

The role of the the repeat and its interaction with sequences immediately 5′ to it, is investigated with *D.pseudoobscura*/*D.melanogaster* interspecific chimaeric *per* transgenes. These molecular manipulations reveal a 'coevolved' functional unit which consists of the repeat and the approximately 60 amino acid upstream flanking sequence. Using these interspecific chimaeric per genes also allowed the subsequent mapping of species-specific behavioural differences between *D.pseudoobscura* and *D.melanogaster* to discrete regions of the *per* gene. The role of the polymorphic (Thr-Gly) repeat in temperature compensation (the process whereby clock function is maintained over a range of temperatures) is investigated with a series of (*Thr-Gly*) *invitro* mutated transgenes, and the results suggest a selective explanation for the latitudinal cline in (*Thr-Gly*) length variations in European populations of *D.melanogaster*. These (*Thr-Gly*) variants are also shown to affect the behavioural patterns displayed in light/dark cycles. Finally, novel responses to combined light/dark and temperature cycles in *per* mutants suggest that *per* may have pleiotropic functions

CONTENTS

AKNOWLEDGEMENTS i
ABREVIATIONS ii
ABSTRACT iii
CONTENTS vi

Chapter 1. Introduction

Biologcal clocks 1
Problems with observations of rhythmycity 2
The negative feedback model 4
How Clocks in diverse species conform to the negative feedack model 6
Unicells (*Synechococcus*) 5
Neurospora crassa 7
Drosophila 8
The PER/TIM complex in *Drosophila* 9
Mammals 12
The Threonine-Glycine (Thr-Gly) region of PER and its relation to the work presented in chapters 3-7 13

Chapter 2. Materials and methods

2.1 Molecular techniques 16
Plasmids used 16
Preparation of plasmid DNA 16
Screening of *E.coli* colonies for transformants 17
PCR For amplification of genomic DNA 17
PCR generation of recombinant DNA 17
Sequencing 18

2.2 Subcloning and construction of transgenes 18
PCR amplification and subcloning of the $(Thr\text{-}Gly)_1$ gene fragment 18
PCR amplification of and subcloning of the $(Thr\text{-}Gly)_{17}$ gene fragment 19
Subcloning the $(Thr\text{-}Gly)_{20}$ 13.2 kb *per* gene into a new P-element vector 20
PCR amplification and subcloning of *D. melanogaster / D. pseudoobscura* chimaeric gene fragment used in *mps4* 21

2.3 P-element mediated transformation 22

2.4 Verification of unique Chromosomal insertion 24

In situ hybridization to transgenic polytene chromosomes 24
preparation of polytene chromosome 25
preparation of slides and coverslips 25
denaturation of the chromosome 26
preparation of the probe 26
Hybridization 26
Antibody binding (Transgenic insert positions) 27

Southern blotting **27**
Fly genomic DNA extraction for southern blot 27
Preparation of probes for southern blotting 28

2.5 Locomotor activity experiments **28**
Protocols for DD experiments 29
Protocols for LD experiments 31
'wrapping' of data 31

2.6 Western Blots **32**
Protein extraction 32
Polyacrylamide gel electrophoresis of proteins 32
Electroblotting of proteins 32
Immunoblotting 33

2.7 Fly stocks **33**

Chapter 3. Molecular coevolution within *per* **34**

3.1 Introduction **34**
(The *mel* transgene and *mps1,2,3* and *4* chimaeric transgenes)

3.2 Methods **36**
Transgene construction (also see Chapter 2 for further details) 36
Locomotor activity 37
Western blots 37

3.3 Results **38**
Overall mean periods 38
Statistical analysis of the intragenotype variation of the transgenes: 39
mel 39
mps1 40
mps2 41
mps3 42
mps4 43
Testing the effects between genotypes 44
Western blots 45
Discussion 46

Chapter 4.: The *per* gene and species-specific behaviour **53**

4.1 Introduction 53

4.2 Methods 54
Lines and strains 54
Activity analysis 54

4.3 Results **56**
DD analysis 56
D.melanogaster and *D.pseudoobscura* (LD at 18^0C) 60
Testing the *D.pseudoobscura/D.melanogaster* transgenes for LD effects 63
Planned comparisons 66

4.4 Discussion 69

Chapter 5 : Temperature compensation effects of the Thr-Gly repeat **74**

5.1 Introduction 74

5.2 Methods 77
Transgene construction (also see Chapter 2 for further details)

5.3 Results **77**
Overall mean periods 77

Statistical analysis of the intragenotype variation of the transgenes: 78
Δ (*Thr-Gly*) 78

(*Thr-Gly*) $_{79}$ 89

(*Thr-Gly*)$_{17}$ 80

(*Thr-Gly*)$_{20}$ 81

Testing the effects between genotypes 81

5.4 Discussion 84

Chapter 6. How temperature and *per* interact to affect behaviour patterns in *Drosophila* ***89***

6.1 Introduction 89

6.2 Methods 92

6.3 Results 93

6.4 Discussion 97

Chapter 7: Investigating the potential pleiotropic effects of *per* **101**

7.1 Introduction **101**

7.2 Materials and methods **104**
Fly stocks 104
Activity analysis of combined light/dark and temperature cycles 104
Number of flies tested at each condition 105

7.3 Results **106**

7.4 Discussion **116**

Chapter 8: Conclusions **124**

Appendix **129**
Appendicees for Chapter 3 **130**
Appendicees for Chapter 4 **132**
Appendicees for Chapter 5 **138**
Appendicees for Chapter 6 **140**
Appendicees for Chapter 7 **144**

Bibliography **151-163**

Chapter 1 Introduction

Biological clocks

Rhythmic processes have been observed across the plant and animal kingdoms, and until recently the origin of these rhythms have been regarded as one of the great mysteries of nature. The vast majority of organisms are subject to a cyclically changing environment. These cycles exist in many different time scales; some of the most prominent cycles are circannual (once a year), circadian (once a day) and circatidal (twice daily), and are related to the movement of our planet within the solar system. For any organism the major implication of these rhythms is to predict and prepare for changes in the external and internal environment, which will impinge on, for example, methods of feeding, reproductive strategies and avoiding predation. This anticipation requires the organism to be equipped with an internal method of keeping time, an endogenous clock.

How the vast array of living things from single celled organisms to man can be controlled by rhythmic forces that can powerfully influence behaviour, has in the past been mainly studied by descriptive ananlysis of these rhythms. While much has been learned, for example, about the physiology of sleep disorders in humans, the biological processes underlying the behaviours have remained elusive. However in the last three decades some spectacular advances have been made in cyanobacteria (Ishiura *et al.*, 1988), fungi (Crosthwaite *et al.*, 1997), insects (Rosbash *et al.*,1996) and mammals (Ralph, 1996), which are providing some consensus as to the nature of this phenomenon. Although it is still unsure whether the oscillators in these diverse organisms have evolved from a common ancestor, they nevertheless appear to have mechanisms in common.

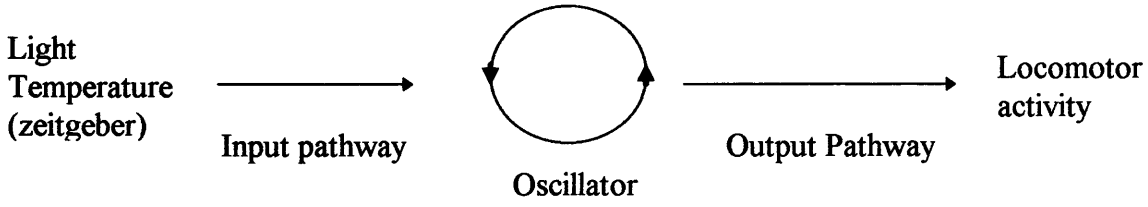
The very ubiquity of clocks in nature and their connection to apparently complex behaviour, led to the assumption that they must have a similarly complex polygenic determination, where removal of one or some 'clock' genes would only have minor effects on the overall pattern of behaviour. Therefore it came as some surprise when the period of circadian rhythmicity in fruit flies was shown to be changed by as much as 20% through mutations within a single gene (*period*, Konopka and Benzer, 1971). Other mutations in fruit flies have since shown that behaviour, from simple stimuli-induced responses, to the complexity of memory, can be drastically affected by single mutations. In *Drosophila* these include *sevenless* (lacking phototaxis to UV light, Tomlinson and

Ready 1986), *fruitless* (alterations in mating behaviour, Gailey and Hall, 1989) and the memory mutants *dunce*, *rutabaga* and *amnesiac* (Quinn and Greenspan, 1984).

To investigate behaviour, simple models are required where discrete changes in an otherwise constant genetic background can be shown to have a measurable and replicable result. The marriage of the study of rhythmicity (which lends itself to automation of observation, see Chapter 2), fruitfly genetics and modern molecular techniques have allowed spectacular advances in our understanding of biological clocks (Young, 1996; Dunlap, 1998). The initial step is the design of experiments where the effect of temperature and light (the most common zeitgebers [time givers]) on mutants can be observed.

To design meaningful experiments, a hypothesis of how a circadian oscillator may function is needed before it's existence can be tested. A simple clock model consists of three major components (Hardin *et al.*, 1992b); An input pathway from the environment, through which external messages (zeitgebers) can reset the endogenous clock, the clock itself which we imagine to be some kind of oscillator that can sustain it's oscillations for a number of circadian cycles, and the output pathways which give external manifestation of the clock's oscillation (see Fig 1.1). Behavioural studies have shown that the rhythmicity in the absence of zeitgebers, rarely follows an exact 24hr cycle (Saunders 1982).

Fig 1.1 A simple model of a biological clock



Problems with observations of rhythmicity

In the observation of rhythmic phenotypes (the output from the oscillator) we meet a paradox. In order to identify a putative clock we need to eliminate zeitgebers to determine if the putative oscillation is autonomous, yet in complex organisms it is not always known if the output is a true reflection of the internal oscillator, locomotor rhythms for example are directly affected by zeitgebers and in animal orders by other

factors as well. How can we be sure that in the absence of the zeitgeber they truly reflect the state of the internal oscillator? Our preconceptions about clocks based on the ones that we see and use in our daily lives can beguile us into thinking that what we measure (the output) is a true representation of the phase of the clock, whereas in reality the physiology of the organism may serve to dampen or elaborate the message from the oscillator.

Only recently (Edery *et al.*, 1994b) has direct measurement of a putative oscillating clock component (western blot of oscillating PER concentration from *Drosophila* heads) been accomplished. Yet even this direct approach can be misleading, because it has been shown, under some circumstances, that individual clock components can be rhythmic within an otherwise non-functioning clock (eg TIM cycling in per^0 flies in LD, Price *et al.*, 1995). Furthermore it is known that clocks can work even in the absence of an identifiably rhythmic phenotype, for example, mutations have been identified where one output phenotype is affected but another is not (Newby and Jackson 1996).

Clock components can also be involved in input pathways; TIM (Myers *et al*, 1996) and WC-2 (Crosthawite *et al*, 1997) are photosensitive clock molecules that help to connect the clock to the environment and alter clock function as a result. To sense the input the molecule has to change and this may lead to an altered competence for that component in the feedback loop. This raises the possibility that the connection of clock components to output pathways could similarly alter oscillator function. At the present time, it is assumed that components of the clock promote the transcription of other clock controlled genes (Van Gelder *et al.*, 1995; Van Gelder and Krasnow 1996) with the attendant delays that this implies. Another theoretical explanation is that clock proteins may act directly on the physiology of the cell, perhaps in the manner of DCO (Majercak *et al.*, 1997), which with other proteins, forms a cAMP dependent kinase that acts in the locomotor output pathway. The biological complexity inherent in these processes therefore suggests that output cannot always be assumed to be an accurate reflection of an internal clock.

To help the researcher identify *bona fide* clock function some criteria have been established (Saunders, 1982). The first of these criteria is that a clock can sustain rhythmic output without external cues. The second is that it is not affected by normal daily fluctuations in temperature. In all clock models the components are ultimately cellular factors, and their interactions must follow set laws which state that as temperature increases, so does the rate of reaction. This rate normally approximately doubles with

every 10^0C rise (Q_{10} = 2). How this property is controlled by the clock machinery is a poorly understood process known as 'temperature compensation', that results in biological clocks having a Q_{10} of approximately 1 (Pittendrigh, 1954). Temperature compensation is vital because in the presence of daily temperature fluctuations, correct time must be maintained if the organism is to be able to predict change in its environment. The third property of a clock is that although it must sustain its oscillation in the absence of a zeitgeber, it must also retain the ability to reset.

As the temperature compensation mechanism is seldom perfect, and in addition day-length changes with the seasons, so the fundamental zeitgeber of light coming and going during the morning and night must reset the endogenous clock. This ability of the input pathway to modify the phase of the oscillator can be represented by the phase response curve (PRC) (Saunders *et al.*, 1994). The organism is entrained to a light/dark (LD) cycle and then released into constant darkness (DD) during which the organism is given pulses of light at different times of the subjective day or night. The effect of the pulses on the phase of the clock as determined by the state of external manifestation of rhythmicity (such as locomotor activity), indicates which parts of the cycle are responsive to light and which are not. For *Drosophila* light pulses during the subjective day have little effect, whereas a light pulse in the early part of the night resets the clock to subjective lights off, and in the late part of the night to subjective lights on (Saunders, 1982). Nature has therefore devised a way of preventing incongruous resetting of the oscillator. This is important because if an animal moves away from midday light, by going underground for example, it would not make sense for its clock to reset to a new dawn on re-emergence into light.

Until recently light was believed to be the most important factor in resetting the *Drosophila* clock (Sidote *et al.*, 1998). However recent work with *frq* gene in *Neurospora* indicates that in certain circumstances, temperature could in fact dominate the light dark cycle in resetting the clock (Liu *et al.*, 1998). Intuitive reasoning however suggests that the dominant zeitgeber is likely to be whichever is most ecologically relevant to the organism under study.

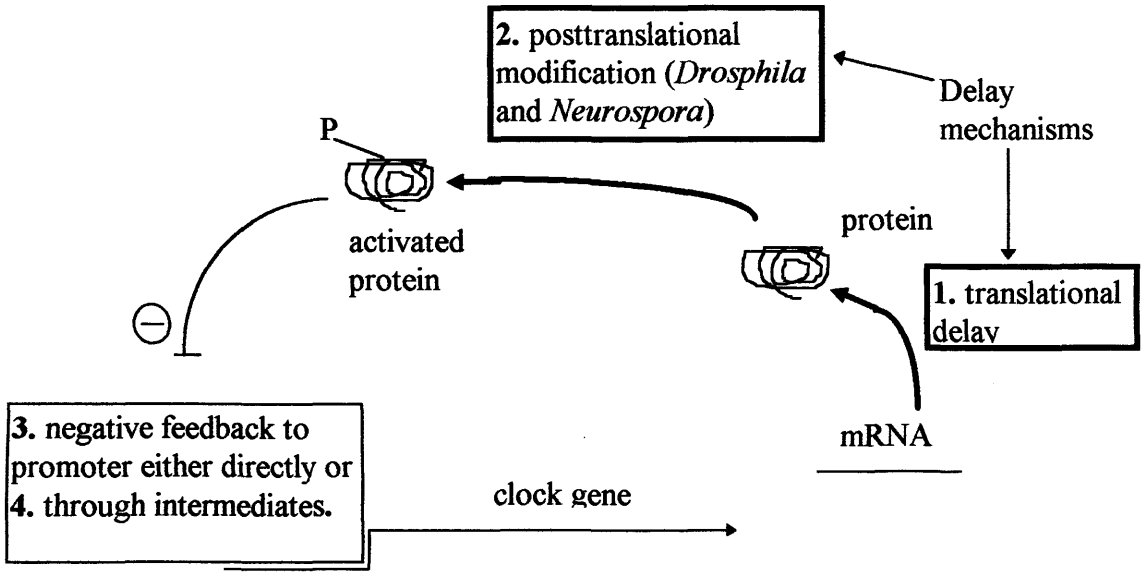
The negative feedback model

The cloning of clock genes from *D.melanogaster* and observation of their mRNA cycling resulted in the formulation of a clock model that can be termed the transcriptional negative feedback loop (Hardin *et al.*, 1990). The model proposes that the production of

a clock protein, and its subsequent negative feedback on its own transcription, after delays at translational and nuclear entry stages, is sufficient for the generation of an oscillating transcriptional loop that may form the basis of the clock. Whilst fruit fly researchers have now shown that a cohort of proteins are involved in the accumulation and feedback processes, this model still forms the basis of our understanding of how clocks work in a wide range of species (Ishiura *et al.*, 1998; Dunlap, 1998).

All clocks so far identified at the molecular level appear to have this negative feedback as a common factor. All are transcriptional loops in which direct or indirect negative feedback of the translated protein on it's own production causes oscillation in that protein's concentration. Theoretical models have also indicated that a lag is needed between transcription and the presence of an active molecule available for negative feedback (Goldbeter, 1995). Without this lag the oscillation cannot occur as competition between positive and negative feedback would bring the system to a steady state (see Fig 1.2).

Fig 1.2 A skeleton model of a molecular negative feedback oscillator which illustrates the transcriptional feedback loop in terms that can be generally applied to all the systems uncovered to date. Thick arrows indicate where delays in production of active protein are known to exist in *Drosophila*. Although the intermediates which relay the negative feedback are known from insect and mammal models the simplicity of the overall scheme illustrates what is known for examples from cyanobacteria, fungi, insects and mammals



To help identify putative components of the clock involved in negative feedback, a set of criteria was established (Aronsen *et al.*, 1994a): 1. mutations in any component

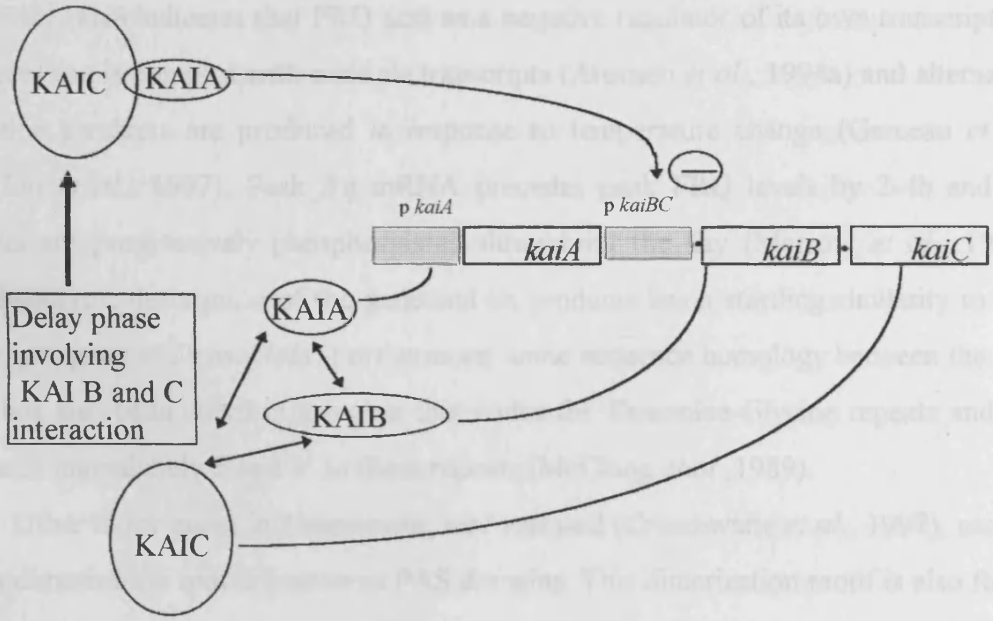
must affect the properties of the clock; 2. the non mutant form of the component should oscillate at the normal periodicity for that organism; 3. alterations in concentration (or activity, for example by post-translational modification) will via negative feedback alter the level of it's own production; 4. it's oscillatory phase is reset by zeitgebers; 5. blocking oscillation of this component results in cessation of the measured rhythmicity.

How Clocks in diverse species conform to the negative feedback model as outlined above:

Unicells (*Synechococcus*)

Prokaryotic cells have some particular advantages over more complex systems as output cannot be attenuated by multicellular pathways. Rhythmicity has been shown in the daily nitrogen fixation cycles in the Cyanobacteria *Synechococcus* (Grobbelar, 1986), the simplest organism known to have a biological clock. Free running rhythmicity, entrainability and temperature compensation have also been shown (Kondo *et al.*, 1993, Kondo and Ishiura, 1994).

Fig 1.3 Outline of molecular interactions believed to represent the negative feedback loop in Cyanobacteria



A recent report shows that the fundamental principles of a molecular negative feedback loop (as outlined in Fig 1.2) are present in the *Synochoccus* oscillator (Ishiura *et al.*, 1998). However the sequence of the genes and the mechanism of the feedback loop indicates that they may have evolved seperately from the *Neurospora* and *Drosophila*/mammalian systems (Ishiura *et al.*, 1998; Dunlap, 1998). The clock is based on the expression of a cluster of three genes known as *kai*A, B and C. Constitutive overexpression of the the *kaiC* gene causes repression at it's promotor which is shared with *kaiB*, and inactivation of any of the three genes abolish rhythmicity. *kaiC* genes appear to fulfill the role outlined in points 1 and 3 from Fig 1.2., with the important delay phase (Goldbeter, 1995) allowing the system to oscillate instead of reaching a steady state. Although different to the fungal and insect/mammalian clocks, the functional similarities shown by the delay and negative feedback phases indicate that the system developed for *Drosophila* will have general applications as a model system for clocks.

Neurospora crassa

The most studied clock gene in *Neuospora* is *frequency* (*frq*) (Feldman and Hoyle 1973). The *frq* locus has semidominant short (frq^1), long (frq^3, frq^7) and null period mutants (frq^9, frq^{10}) which alter the period between conidiation events (Dunlap, 1996). Cycling *frq* mRNA is required for phenotypic rhythmicity with constitutive *frq* expression stopping rhythmicity, and low levels of the *frq* transcript restarting the clock (Aronsen *et al.*, 1994a). This indicates that FRQ acts as a negative regulator of its own transcription. *frq* expression is complex with multiple transcripts (Aronsen *et al.*, 1994a) and alternative translation products are produced in response to temperature change (Garceau *et al.*, 1997; Liu *et al.*, 1997). Peak *frq* mRNA precedes peak FRQ levels by 2-4h and the products are progressively phosphorylated throughout the day (Merrow *et al.*, 1997). This phenotypic description of the gene and its products has a startling similarity to that for the *per* gene of *Drosophila*. Furthermore, some sequence homology between the two genes has also been noted in a region that codes for Threonine-Glycine repeats and the sequences immediately 5'and 3' to these repeats (McClung *et al.*,1989).

Other clock genes in *Neurospora*, *wc1* and *wc2* (Crosthwaite *et al.*, 1997), encode protein dimerization motifs known as PAS domains. This dimerization motif is also found in PER *and* dCLOCK (Huang *et al.*, 1993; Allada *et al.*, 1988). WC-2 is thought to be the clock component that promotes *frq* gene expression and is itself repressed by FRQ (Crosthwaite *et al.*, 1997; Dunlap, 1998), (see point 4. Fig 1.2). WC-1, but not WC-2, is

required for photoinduction of the *frq* gene (Crosthwaite *et al.*, 1997). Therefore WC-1 appears to be part of the input pathway to the oscillator, but it acts very near to the autonomous clock mechanism and may have a support role, as the clock operates poorly in it's absence (Crosthwaite *et al.*, 1997). Interestingly *wc2* has been shown to be pleiotropic as it is involved in the regulation of a clock control gene *ccg-2* in the absence of a functional clock (Arpaia *et al.*, 1993). This is of particular interest as in Chapter 6 and 7 I present evidence to suggest that *per* may also have pleiotropic fuctions. *wc-2* like mouse *mper* (Albrecht *et al.*, 1997), *Drosophila dper* (Colot *et al.*, 1988), *frq* (Mclung *et al.*, 1989) and *kaiC* (Ishiura *et al.*, 1998) also encodes a region of Thr-Gly repeats. The (*Thr-Gly*) repeat region of *per* has specific clock functions (Sawyer *et al.*, 1997; Peixoto *et al.*, 1998) and a clock role has also been suggested for the *kaiC* repeat (Ishiura *et al.*, 1998). It remains to be seen if the repeats in *frq* and *wc-2* have functional significance, or if the occurrence of these repeats in clock genes from four different orders is coincidental.

Drosophila

The most intensively studied gene in *Drosophila* is *per*, and as for *frq* in *Neurospora* a negative feedback loop is at the core of PER cyclicity (Hardin *et al.*, 1990). The three original per mutants per^{S}, per^{L} and per^{01} (Konopka and Benzer 1971) produced short, long and arrhythmic circadian phenotypes in locomotor activity and pupal eclosion, In addition the mutants gave corresponding changes in the male courtship song rhythm (Kyriacou and Hall 1980). The mutations are semidominant like the period-changing *frq* mutants. When the mutations were identified at the sequence level (Yu *et al.*, 1987a; Baylies *et al.*, 1987b), little could be inferred of PER's function from their coding changes. One feature however that did stand out in the putative protein was the run of Thr-Gly and Ser-Gly encoding repeats (Citri *et al.*, 1987). On the basis of similarity with these repeats it was proposed that PER may be a proteoglycan affecting behavioural rhythms by altering gap junctional communication in the *Drosophila* nervous system (Bargiello *et al.*, 1987). The link to proteoglycans however turned out to be spurious (Siwicki *et al.*, 1992).

Like FRQ, PER is now known to be progressively modified by phosphorylation throughout its cycle and not extensively glycosylated (Edery *et al.*, 1994b). Interestingly an in-frame deletion of the entire (*Thr-Gly*) repeat including some sequence from the flanking regions [*Δ(Thr-Gly)*] gave a reduction in the ultradian song cycle period from

~60s to ~40s (Yu *et al.*,1987b) and a decrease in temperature compensation of per^{01} transformants (Ewer *et al.*, 1990). Three *frq* mutants displaying a range of temperature sensitive and free-running circadian period changes, remarkably map to the immediate 5' (frq^{3} temperature sensitive) and 3'($frq^{7,8}$ long period and frq^{1} short period) regions that flank the *frq* (*Thr-Gly*) repeat (Aronsen *et al.*, 1994b).

per is widely expressed in both central and peripheral nervous systems (CNS and PNS) and other tissues such as the malphigian tubules, gut and ovaries (Siwicki *et al.*, 1988; Saez and Young, 1988; Liu *et al.*, 1992). *per* expression in the CNS is concentrated in a small group of protocerebral neurons, the glia of the brain, the optic lobes and the thoracic ganglion. Sites of PNS expression include the photoreceptors, the ocelli, the ring gland and some peripheral nerves (Siwicki *et al.*, 1988; Saez and Young, 1988; Liu *et al.*, 1992). The adult brain has been indicated as the site of the circadian pacemaker, and mosaic analysis has shown that robust rhythms are produced in flies that only express *per* in a subset of neurons termed the lateral neurons (Ewer *et al.*, 1992). A transformed strain expressing PER only in the lateral neurons has also found to be robustly rhythmic (Frisch *et al.*, 1994). However the adult brain is not the only site of pacemaker function as decapitated flies still have circadian activity in their malphigian tubules (Giebultowicz and Hege, 1997) and autonomous photoreceptive clocks are also present throughout the body of the fly (Plautz *et al.*, 1997). The functions of *per* in many tissues is simply unknown, but regional pacemakers appear to be driving specialized rhythmic activity in separate organs, with the locomotor activity cycle orchestrated from the central nervous system.

The PER/TIM complex of *Drosophila*

Most work has focused on PER protein and mRNA cycling. However another cycling protein TIM has been identified (Vosshall *et al.*, 1994) and mutants of *tim* behave in a similar way to *per*, abolishing rhythmic eclosion and locomotor activity (Sehgal, *et al.*, 1994). The *tim* gene has also been cloned (Gekakis *et al.*, 1995; Myers *et al.*, 1995) and it has become clear that PER and TIM interact with each other and respond to light input, to provide a self regulated oscillating feedback system. (Young *et al.*, 1996).

These investigations have been aided by immunohistochemistry using clock gene/reporter gene fusions, An example is the PER-β-gal fusion protein which shows predominantly nuclear signals (Liu *et al.*, 1992). PER immunoreactivity also fluctuates in

a circadian manner in the lateral neurons and the optic lobes in both LD and DD (Siwicki *et al.*, 1988; Zerr *et al.*, 1990; Curtin *et al.*, 1995). Peak immunoreactivity in wild type flies occurs late at night (ZT21) and almost completely dissapears by the end of the day (ZT12, Zerr *et al.*, 1990). When *per* mutants were examined in LD, per^{S} flies were found to have their protein peak earlier than wild type flies. However, per^{L1} flies had no obvious peak in immunoreactivity and cycling was therefore difficult to assess (Zerr *et al.*, 1990). These results have since been confirmed with more sensitive western blots of head extracts (Rutila *et al.*,1996)

The tim^{0} mutant abolished the nuclear localisation of PER-β-gal fusion proteins indicating that TIM was required for PER to become nuclear (Vosshall *et al.*, 1994). When western blot analysis of PER in tim^{+} and tim^{0} backgrounds was performed there were very low levels of PER present in the tim^{0} flies, suggesting that PER may be unstable in the absence of TIM (Price *et al.*, 1995). TIM's physical interaction with PER was subsequently confirmed by the use of PER as a bait in the yeast two-hybrid screen (Gekakis *et al.*, 1995). Surprisingly however TIM did not encode the PAS protein interaction domain (Gekakis *et al.*, 1995).

The importance of the TIM/PER interaction is further illustrated by the poor interaction between TIM and the PER^{L1} (the long period mutant with poor temperature compensation Rutila *et al* 1996) which results in delayed nuclear entry (Gekakis *et al.*, 1995; Curtin *et al.*, 1995). In the total absence of TIM, PER cannot become nuclear (Price *et al.*, 1995). Therefore correct timing and nuclear entry is shown to be essential for temperature compensation and negative feedback. Both *per* and *tim* mRNAs cycle in phase with each other but *per* mRNA cycles with a peak at ZT15, ~6 h before the PER protein peak (Hardin *et al.*, 1990; Zerr *et al.*, 1990). per^{S} mRNA cycling is advanced and per^{L1} mRNA shows a delay (Hardin *et al.*, 1990). The fact that when PER levels are high, *per* mRNA levels are low suggests that PER may negatively regulate its own transcription. This was confirmed when PER was overexpressed in the photoreceptors and endogenous *per* mRNA cycling was dampened (Zeng *et al.*, 1994). The lack of mRNA cycling in per^{01} flies, which have no functional PER protein, is also consistent with this hypothesis. However even when *per* mRNA is under the constitutive control of a heterologous non-cycling promoter (Vosshall and Young, 1995), it cannot be taken for granted that the per mRNA under this foreign promoter will not cycle because the cycling

of *per* mRNA is known to be due to a combination of transcriptional and postranscriptional processes (So and Rosbash, 1997; Stanewsky *et al.*, 1997a, 1997b).

As stated above, *tim* mRNA cycles with the same phase as *per* mRNA (Sehgal *et al.*, 1995). Expression of *per* and *tim* is linked, because peak *tim* mRNA levels in a per^S background are advanced like per^S mRNA, and *tim* mRNA cycling is abolished in per^{01}as is *per* mRNA in tim^0 (Sehgal *et al.*, 1995). Therefore *tim* and *per* are mutually involved in regulating their own expression. Furthermore, TIM concentration, but not *tim* mRNA level, is reduced by light suggesting that TIM probably has a role in light induced resetting of the clock (Hunter-Ensor, 1996; Lee *et al.*, 1996; Zeng *et al.*, 1996; Myers *et al.*, 1996). The involvement of TIM in light induced phase shifts and behavioural entrainment has recently been confirmed (Suri *et al.*, 1998; Yang *et al.*, 1998). Available *tim* mRNA is unaffected by the light and so translation of TIM can recommence after lights off and the dark phase 'stable' TIM is therefore available to interact with PER, in order to mediate nuclear entry (Vosshall *et al.*, 1994; Price *et al.*, 1995).

The *tim* gene is therefore now known to produce a protein essential for *per* cycling via heterodimerization with PER, nuclear entry and subsequent light induced degradation (Young *et al.*, 1996). A model proposed by Young *et al.* (1996), essentially explains the patterns of *tim* and *per* expression noted in LD. TIM and PER are the primary clock products which cause negative feedback (see Fig 1.2). It is also now that this feedback occurs via their interaction with the products of two recently cloned genes, *cycle* (Rutila *et al.*, 1998) and *Jerk* (the *Drosophila* homologue of m*Clock*)(Allada *et al.*, 1998; Gekakis *et al.*, 1998). These genes code for bHLH binding proteins which like WC-1 and WC-2 from *Neurospora* also contain PAS dimerization motifs. The CYCLE-JERK heterodimer binds to specific sites in the *per* and *tim* promoters known as E boxes (Hao *et al.*, 1997) and drives circadian regulated expression of these genes (Rutila *et al.*, 1998; Allada *et al.*, 1998; Darlington *et al.*, 1998; Gekakis *et al.*, 1998)

Transcription of *per* and *tim* begins during the light phase. PER levels are kept low during this early phase by DOUBLETIME [DBT] which phosphorylates PER and earmarks it for degradation (Price *et al.*, 1998, and see below). Levels of *per* and *tim* RNA reaching a peak after lights off (ZT 12) at which point heterodimerization of PER and TIM occur in the early dark phase and PER-TIM complexes translocate to the nucleus at ~ZT20. Peak nuclear accumulation of PER-TIM complexes in the late night leads to suppression of the positive effect of factors JERK and CYC which in turn reduces *per* and *tim* mRNA levels. As PER and TIM degrade, the transcriptional block

on per and tim is lifted, after which the cycle begins again. [see Fig 1.4 for illustration of PER amd TIM cycling after Young *et al.* (1996) modified to take into account the new information on *jrk, cyc* and *dbt*].

Efficient degradation is also a prerequisite for functional cycling proteins, and as PER and TIM are progressively phosphorylated (Edery et al., 1994b; Rutila *et al.*, 1996), this may play a role in preparing the protein for degradation. Experiments performed on cycling of β-galactosidase-*per* (β-gal-*per*) reporter genes suggest that a specific region of PER that encompasses the Thr-Gly repeat is required for protein cycling. Two β-gal-*per* reporter genes were constructed, one in which PER was truncated before the Thr-Gly region (SG), and the other, downstream of the Thr-Gly region (BG). Both the corresponding mRNA's cycled in LD in a per^{+} background but only the longer BG protein cycled, suggesting that this region is required for autonomous post-translational control of cycling, and may represent a site for targeted degradation (Dembinska *et al.*, 1997).

Recent work has uncovered the *doubletime* (*dbt*) locus (Price *et al.*, 1998) which encodes a protein closely related to the human casein kinase Iε (Kloss *et al.*, 1998). Different mutations of this gene shorten and lengthen the free-running period (Kloss *et al.*, 1998; Price *et al.*, 1998). DBT reduces the stability and hence accumulation of monomeric PER and so contributes to the delay between the appearance of *per* mRNA and protein. It is thought to do this by phosphorylating the PER molecule and so making it a target for degradation (Price *et al.*, 1998). Larval PER in a dbt^{-} background (created by P-element insertion into the locus and referred to as dbt^{P}) does not cycle and so behaves like the 'Thr-Gly deleted' SG-β-gal fusion (Price *et al.*, 1998). Therefore it is possible that the site of DBT induced degradation within PER is in the Thr-Gly containing fragment. The free-running periods of short (dbt^{S} ~ 18h) and long (dbt^{L} ~ 27 h) mutations are close to the periods of original per^{S} and per^{L} mutations (Konopka and Benzer 1971). DBT^{S} and PER^{S} are phosphorylated earlier in their circadian cycles than PER^{+} while DBT^{L} and PER^{L} products are phosphorylated later (Edery *et al.*, 1994; Rutila *et al.*, 1996; Price *et al.*, 1998). This is evidence for a common factor in these phenotypes. Interestingly dbt^{P}/dbt^{P} is lethal in the adult (Price *et al.*, 1998) and as all other identified fly clock genes are non-vital, this indicates that *dbt* has other important functions in the organism.

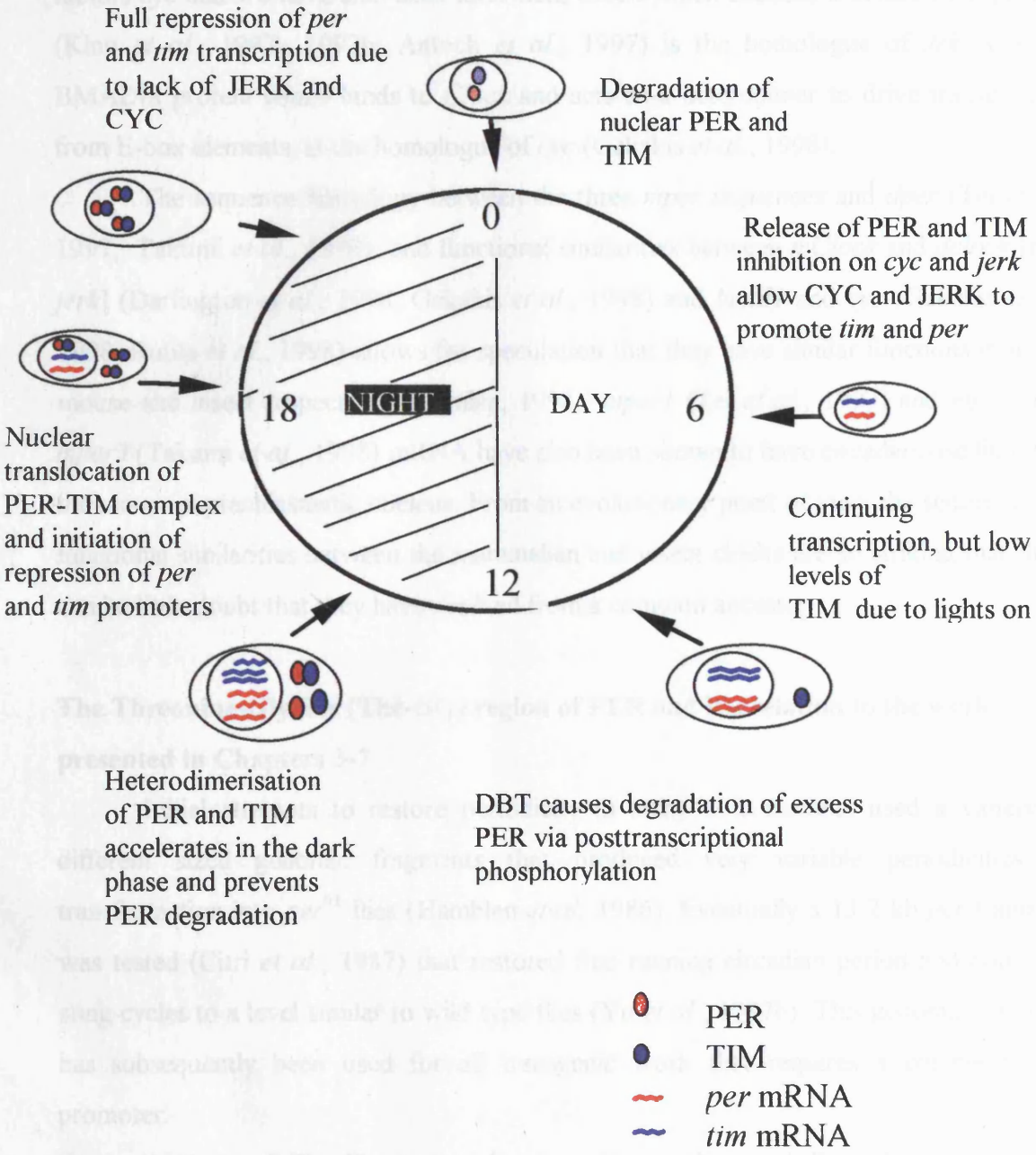


Fig 1.4 A schematic representation of the current model of *per* and *tim* mRNA and protein cycling in LD (Young *et al.*, 1996, Rosbash *et al.*, 1996, Dunlap, 1988)

Mammals

Mammalian *per* homologues; *mper1* (Tei *et al.*, 1997), *mper2* (Albrecht *et al.*, 1997) *and mper3* (Takumi *et al.*, 1998) have been identified and sequenced. All have PAS dimerization domains and like *dper* (Hao *et al.*, 1997) *mper1* has an E box enhancer in it's promoter (Gekakis *et al.*, 1998). Mammalian homologues of the positive *Drosophila* factors *cyc* and *Jrk* have also been identified; *Clock* which encodes a bHLH-PAS protein (King *et al.*, 1997a 1997b; Antoch *et al.*, 1997) is the homologue of *Jrk*, and the BMAL1a protein which binds to *Clock* and acts as a heterodimer to drive transcription from E-box elements, is the homologue of *cyc* (Gekakis *et al.*, 1998).

The sequence homology between the three *mper sequences* and *dper* (Tei *et al.*, 1997; Takumi *et al.*, 1998), and functional similarities between *mClock* and *dclock* [née *jerk*] (Darlington *et al.*, 1998; Gekakis *et al.*, 1998) and *bmal1* and *cyc* (Gekakis *et al.*, 1998; Rutila *et al.*, 1998) allows for speculation that they have similar functions in in the mouse and insect respectively (Dunlap, 1998). *mper1* (Tei *et al.*, 1997) and *mper2* and *mper3* (Takumi *et al.*, 1998) mRNA have also been shown to have circadian oscillation in the mouse suprachiasmatic nucleus. From an evolutionary point of view the sequence and functional similarities between the mammalian and insect clocks are so striking that there can be little doubt that they have evolved from a common ancestor.

The Threonine-Glycine (Thr-Gly) region of PER and it's relation to the work presented in Chapters 3-7

Initial attempts to restore periodicity in arrhythmic mutants used a variety of different sized genomic fragments that produced very variable periodicities on transformation into per^{01} flies (Hamblen *et al*, 1986). Eventually a 13.2 kb *per* fragment was tested (Citri *et al.*, 1987) that restored free running circadian period and courtship song cycles to a level similar to wild type flies (Yu *et al.*, 1987b). This genomic fragment has subsequently been used for all transgenic work that requires a competent *per* promoter.

Analysis of Thr-Gly regions between *D. simulans* and *D. melanogaster* using transformants based on the 13.2 fragment but chimaeric for the two species in this region, showed it to control species-specific differences in the song cycle (Wheeler *et al.*, 1991). In addition locomotor activity changes in the temperature compensation characteristics of the Δ(*Thr-Gly*) transformants were also observed (Yu *et al.*, 1987b). At lower

temperatures wild type periods of ~24h were observed, but at higher temperatures lack of temperature compensation lead to periods 1-2h longer (Ewer *et al.*, 1990). A number of studies in our laboratory have subsequently focused on the length of the (*Thr-Gly*) repeat region length and differences in behaviour associated with the naturally occuring length variation (Costa *et al.*, 1991, 1992; Nielsen *et al.*, 1994; Sawyer *et al.*, 1997).

The (*Thr-Gly*) region is conspicuous in the less conserved middle part of the gene (Colot *et al.*, 1988), and appears to have evolved by expansion (Costa *et al.*, 1992). Theoeretical models indicate that expansion of a repetitive region requires at least two tandem repeats (Gray and Jeffreys, 1991), and after the initial expansion a consolidation occurs where a number of repeats are maintained at a fixed level. Diptera outside the *Drosophila* have two pairs of Thr-Gly's and this appears to be the ancestral state (Nielsen *et al.*, 1994). Structural predictions (Chou and Fasman, 1974) suggest that the Thr-Gly repeat represents a series of flexible turns, (Costa *et al.*, 1992; Parkinson 1997). Experiments with synthetic Thr-Gly polypeptides also suggest that the $(Thr\text{-}Gly)_3$ motif is a structural monomer which generates a type-2 β–turn (Castiglione-Morelli *et al.*, 1995), and that different Thr-Gly lengths can adopt alternative conformations which can also change as a function of temperature (Castiglione-Morelli *et al.*, 1995). Although Yu *et al.* (1987b) showed that deletion of this region resulted in a temperature sensitive phenotype, the function of (*Thr-Gly*) repeat has largly been ignored in more recent investigations of temperature compensation (Huang *et al.*, 1995; Gekakis *et al.*,1995; Saez and Young, 1996).

However a large body of work is now accumulating (Yu *et al.*, 1987b; Sawyer *et al.*, 1997; Parkinson, 1997; Peixoto *et al.*, 1998) which shows that the Thr-Gly repeat is intimately involved in temperature compensation. These studies have used *in-vitro* deletions and chimaeric constructs of this region, in addition to locomotor activity observations of natural (*Thr-Gly*) length variants from European *D. melanogaster* populations. It is worth remembering however that insect species eg. *Musca Domestica* without this extended (*Thr-Gly*) repeat also show good temperature compensation (Nielsen *et al.*, 1994). The extreme expansion displayed by some of the Drosophilids (e.g *D. pseudoobscura*) must therefore be a specialized adaptation. However investigation of this adaptation and the mutant phenotypes caused by it's selective deletion should illuminate our understanding of the phenomenon of temperature compensation in general.

Sequence analysis of the (*Thr-Gly*) region in flies from European populations have shown that there is a robust clinal distribution of the two most common allelic variants, the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ (Costa *et al.*, 1992). The longer alleles are found more frequently in the North and the shorter in the South. One explanation for the cline may be that thermal selection on the *per* gene favours the $(Thr\text{-}Gly)_{17}$ variant in the warmer southern latitudes and $(Thr\text{-}Gly)_{20}$ allele in the colder North. This explanation is supported by a similar Australian clinal pattern with the $(Thr\text{-}Gly)_{20}$ allele predominating in the cooler south (Sawyer 1996). Consistent with this natural selection scenario Sawyer *et al.* (1997) found that the $(Thr\text{-}Gly)_{20}$ allele in natural populations when tested under laboratory conditions gave more robust temperature compensation than the $(Thr\text{-}Gly)_{17}$ allele.

The length of the repeat varies widely between dipteran species (Costa *et al.*, 1991, 1992, Peixoto *et al.*, 1993, Nielsen *et al.*, 1994). In particular Peixoto *et al.* (1993) and Nielsen *et al.* (1994) suggested that the repeat length coevolves with the immediate 5′ and 3′ flanking region regions. If this is indeed correct, then a chimaeric transgene where the species-specific junction divides flanking sequence derived from a species with a short (*Thr-Gly*) region and a long (*Thr-Gly*) repeat from another species, should reveal some kind of defective clock phenotype. **Chapter 3** investigates the above scenario with a range of chimaeric *per* gene constructs combining *D.pseudoobscura* with *D.melanogaster* sequences, with junctions positioned to test the hypothesis that the repeat and flanking regions for each species have coevolved.

Chapter 4 demonstrates that different rhythmic behaviour patterns observed in two *Drosophila* species, can be assigned to the *per* genes from those species. Analysis of *per* transgenes (using the chimaeric transgenes from Chapter 3) for these two species further maps these behaviours to subregions within *per*.

In **Chapter 5** I show that *in vitro* deletions of the (*Thr-Gly*) region which artificially recreate as transgenes the naturally occuring $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ alleles, give temperature compensation profiles similar to the corresponding natural variants. Larger deletions $(Thr\text{-}Gly)_1$ and the previously constructed Δ(*Thr-Gly*) transgene (Yu *et al.*, 1987) are also tested and shown to give results consistent with the hypothesis that this region is involved in temperature compensation.

D.melanogaster like many insects has a bimodal activity profile in LD (Saunders, 1982). It is often assumed that the peaks in activity are simply responses to changes in the zeitgeber. In **Chapter 6** I investigate the nature of this bimodality via variation in

temperature on the range of *in vitro* generated (*Thr-Gly*) length transformants used in Chapter 3.

The effect of temperature and the light dark cycle on the fly, and it's response to these challenges, is a constant theme throughout this thesis. In nature light and temperature cycling are superimposed by virtue of the lighting and heating effect of the sun on the earth's surface. There is however little in the literature about the specific effects of temperature variation on the activity of insects within the framework of an LD activity cycle. **Chapter 7** attempts to address this and reveals novel behavioural effects due to combinations of light and temperature zeitgebers.

Chapter 2 Materials and Methods

2.1 Molecular techniques:

The basic molecular techniques used in this work, such as agarose gel electrophoresis, restriction enzyme digestions, ligations, etc., followed Sambrook *et al.* (1989). Transformation of bacterial cells was carried out by a simplified Hanahan method (Sambrook *et al.*, 1989) or by electroporation with a Biorad gene pulser. The DH5 alpha *E. coli* strain (Stratagene), or XL1- blue was used.

The restriction enzymes, buffers, phosphatase and ligase used in the cloning were supplied by a variety of manufacturers mainly Gibco Brl and Promega.

Plasmids used:

pTZ18 phagemid (Pharmacia P-L Biochemicals), used here as a simple cloning vector has BamHI, SacI and KpnI restriction sitesremoved.

pW8, an 8kb P-element *Drosophila* transformation vector (Klemenz et al., 1987).

cp20.1, a P-element *Drosophila* transformation vector, with the 13,2kb fragment) containing the *D. melanogaster per* gene (Citri et al., 1987).

p Bluescript sk a cloning vector containing the *D. pseudoobcura per* gene (Citri et al., 1987).

Preparation of plasmid DNA

Large scale preparations were made using Tip-100 Qiagen DNA preparation kits, according to the manufacturers' instructions (DIAGEN Qiagen Inc.). These yielded up to 100 μg of supercoiled plasmid DNA. Small scale quick plasmid minipreps were prepared according to Serghini *et al.* (1989). DNA fragments used in the cloning were recovered from normal agarose gels using the QUIEX gel extraction kit (DIAGEN) or from low melting point agarose (high purity NuSieve GTG, FMC BioProducts) gels by phenol and subsequent double phenol-chloroform extractions.

Screening *E.coli* colonies for transformants

This was carried out either by miniprepping (Serghini *et al.*, 1989), or by direct PCR screening of bacterial colonies (D.Gussow and T.Clackson 1989). The PCR

reactions were carried out according to Jeffreys *et al.* (1988). Taq polymerase from a variety of manufacturers were used. The cycle conditions used were 95°C for 1 min (denaturation), 65°C for 1 min (annealing) and 70°C for 1 min (elongation).

PCR

For amplification of *Drosophila* genomic DNA

Fly genomic DNA used in PCR was prepared according Gloor and Engels (1990). 50 µl of 'squishing' buffer (10 mM Tris-Hcl pH 8.2 , 1mM EDTA, 25mM Nacl and 200µg/ml Proteinase K) is drawn into a 200µl yellow tip, and the tip used to grind a single male fly in an eppendorf tube. The buffer is then used to suspend tissue which is incubated at 37^0C for 30 min and then subsequently heated to 95^0C for 2 min to inactivate the proteinase K. 1 µl of this DNA preparation was used for each 10 µl PCR reaction.

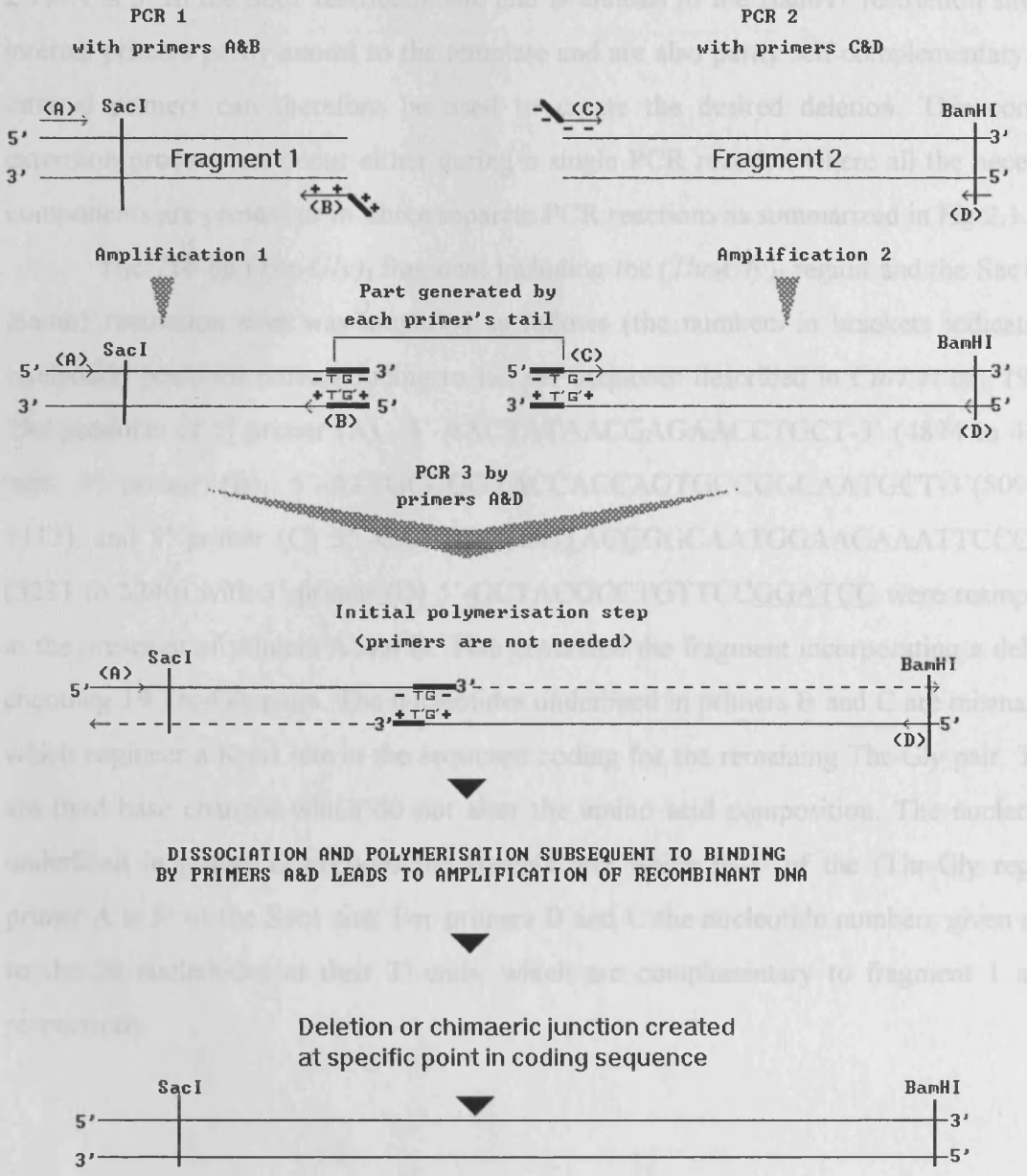
PCR generation of recombinant DNA

PCR reactions can induce base substitutions with an estimated substitution rate of 1 in 400 during a 30 cycle amplification (Sambrook *et al.*, 1989). To reduce the possibility of generating fragments with the incorrect sequence, a polymerase isolated from *Pyrococcus furiosus* (*Pfu*, STRATAGENE) a hyperthermophilic marine archaebacterium was used. This thermostable DNA polymerase possesses both 5' to 3' DNA polymerase activity and 3' to 5' proofreading exonuclease activity. It has a temperature optimum near 75°C and remains greater than 95% active following one hour incubation at 95°C. The 3' to 5' proofreading exonuclease activity results in a reported 12 fold increase in fidelity of DNA synthesis in comparison with Taq DNA polymerase (STRATAGENE). To join two recombinant DNA fragments with a novel juction containing deletions or insertions and for the generation of chimaeric fragments, a three step amplification strategy was used, this is procedure is outlined in Fig 2.1. The procedure follows that of Erlich (1989)

Sequencing:

All PCR generated fragments were sequenced to check for errors, before reinsertion into the *per* gene fragment. The DNA sequence analysis was carried out by the dideoxy chain-termination method, using the Sequenase version 2.0 kit from United States Biochemicals

Fig 2.1 The use of a three step amplification system to generate alterations in DNA sequence or for the generation of chimaeric junctions. In this report these alterations were always within the (*Thr-Gly*) repetitive region of the *per* gene



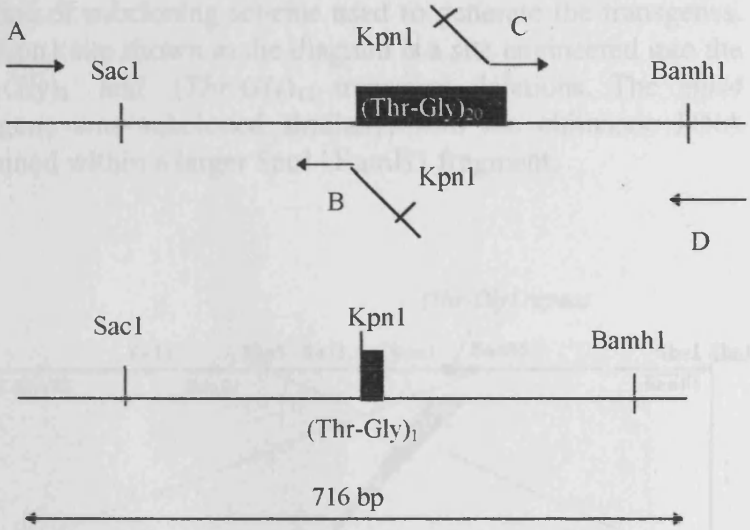
2.2 Subcloning and Construction of transgenes

PCR amplification and subcloning of the (*Thr-Gly*)$_1$ gene fragment

The (*Thr-Gly*)$_1$ transgene was made by amplifying a new Sac1 to BamH1 fragment (see Figs 2.1 and 2.2) which includes the (*Thr-Gly*) region. The generation of the recombinant DNA is made by using two pairs of primers (primers A-B and C-D) (see Fig 2.1). A is 5' to the Sac1 restriction site and D anneals to the BamH1 restriction site, the internal primers partly anneal to the template and are also partly self complementary. The internal primers can therefore be used to create the desired deletion. This complex extension process can occur either during a single PCR reaction where all the necessary components are present or in three separate PCR reactions as summarized in Fig 2.1.

The 716 bp (*Thr-Gly*)$_1$ fragment including the (*Thr-Gly*)$_1$ region and the Sac1 and Bamh1 restriction sites was amplified as follows (the numbers in brackets indicate the nucleotide positions corresponding to the *per* sequence described in *Citri et al.*, 1987,): The products of 5' primer (A), 5'-AACTATAACGAGAACCTGCT-3' (4874 to 4893), with 3' primer (B), 5'-ATTGCCGGTACCACCAGTGCCGGCAATGCT-3'(5094 to 5113), and 5' primer (C) 5' -CACTGGTGGTACCGGCAATGGAACAAATTCCGC-3' (5231 to 5240) with 3' primer (D) 5'-GCTACGCCTGTTCCGGATCC were reamplified in the presence of primers A and D. This generated the fragment incorporating a deletion encoding 19 Thr-Gly pairs. The nucleotides underlined in primers B and C are mismatches which engineer a Kpn1 site in the sequence coding for the remaining Thr-Gly pair. These are third base changes which do not alter the amino acid composition. The nucleotides underlined in primer D encodes the BamH1 site which is 3' of the (Thr-Gly region), primer A is 5' to the Sac1 site. For primers B and C the nucleotide numbers given relate to the 20 nucleotides at their 3' ends, which are complimentary to fragment 1 and 2 respectively.

Fig 2.2. Generation of the (*Thr-Gly*)$_1$ Sac1-BamH1fragment



Restriction of the external primer annealing sites are used after production of the recombinant fragment for replacement of the (*Thr-Gly*)$_{20}$ Sac1 to BamH1 fragment via a series of subcloning steps (see Fig 2.3). Restriction analysis of the complete recombinant (*Thr-Gly*)$_1$ transgene in the pW8 *Drosophila* transformation vector is shown in Fig 2.4.

PCR amplification and subcloning of the (*Thr-Gly*)$_{17}$ gene fragment

The (*Thr-Gly*)$_{17}$ gene fragment was generated by amplifying a 364-bp Thr-Gly fragment using a natural (*Thr-Gly*)$_{17}$ variant as a template. The relevant fragment of this gene had previously been sequenced and other than the Thr-Gly repeat the coding sequence was identical to that of the (*Thr-Gly*)$_{20}$ and (*Thr-Gly*)$_1$ sequences. The primers used were the 5' primer A (see above) and a 3' primer (E) 5'-CATTGCCGGTACCAGTGCCT-3' (5217-5236) which again carried mis-match bases (underlined) forming a Kpn1 site. The amplified product was digested with Sac1 and Kpn1 and now carrying the coding sequence for 16 Thr-Gly repeats, used to replace the Sac1-Kpn1 fragment from the (*Thr-Gly*)$_1$ gene fragment.

Fig 2.3 Outline of subcloning scheme used to generate the transgenes. The Kpn1 site shown in the diagram is a site engineered into the $(Thr\text{-}Gly)_1$ and $(Thr\text{-}Gly)_{17}$ transgene deletions. The *mps4* transgene was subcloned similarly, with the chimaeric DNA contained within a larger Sac1- BamH1 fragment.

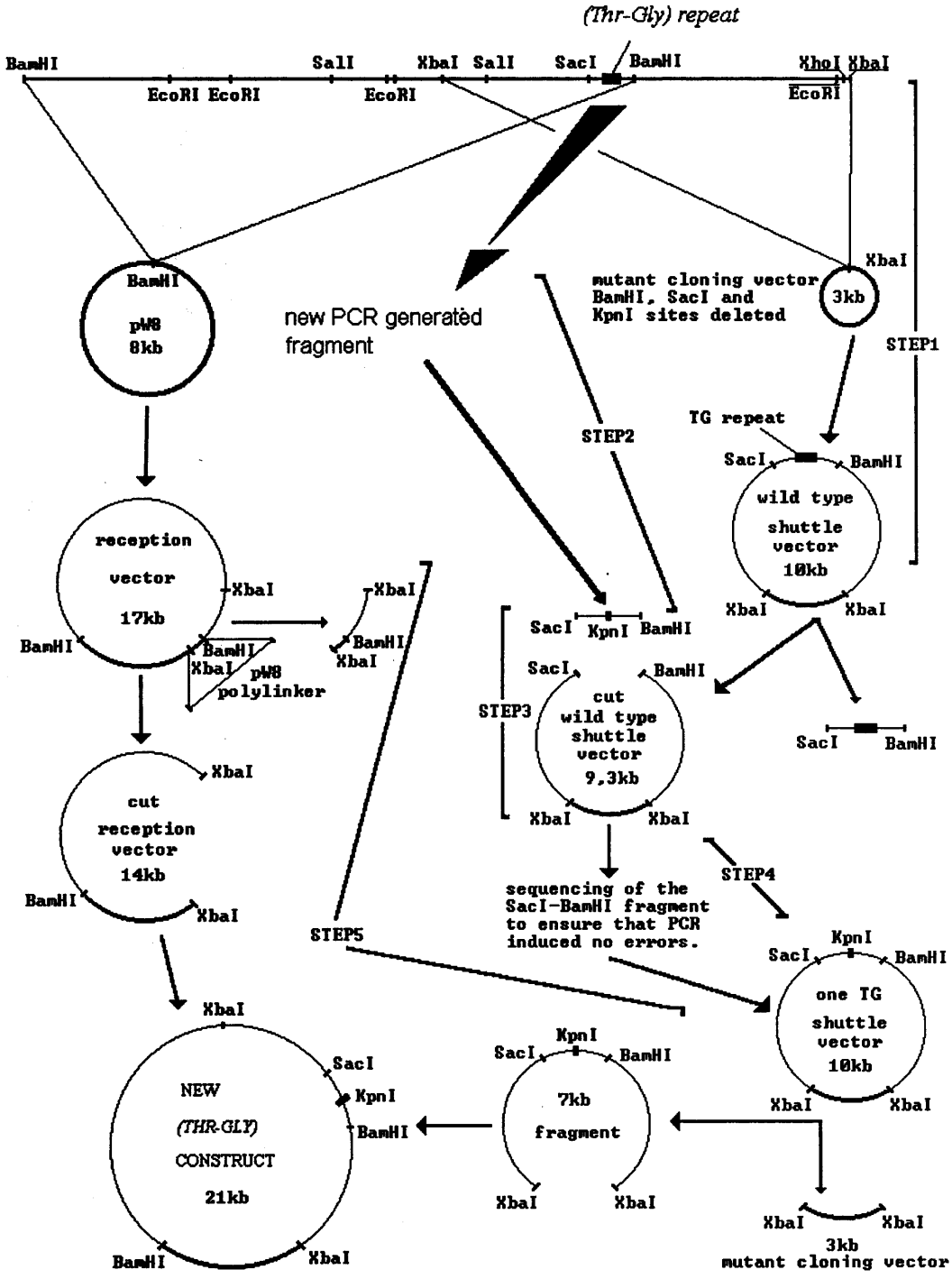
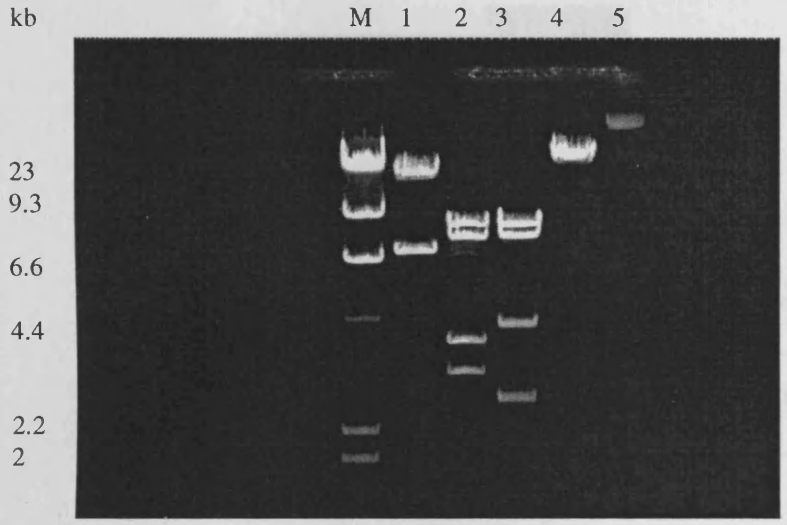


Fig 2.4 Restriction digests of the $(Thr\text{-}Gly)_1$ transgene (which is illustrated below). M; molecular wt marker λ–Hind 111. **lane 1;** Xba1. **lane 2;** Xba1/BamH1 double digest. **lane 3**; Xba1/Kpn1 double digest. **lane 4;** plasmid linearized with Xho1, and **lane 5** uncut plasmid.



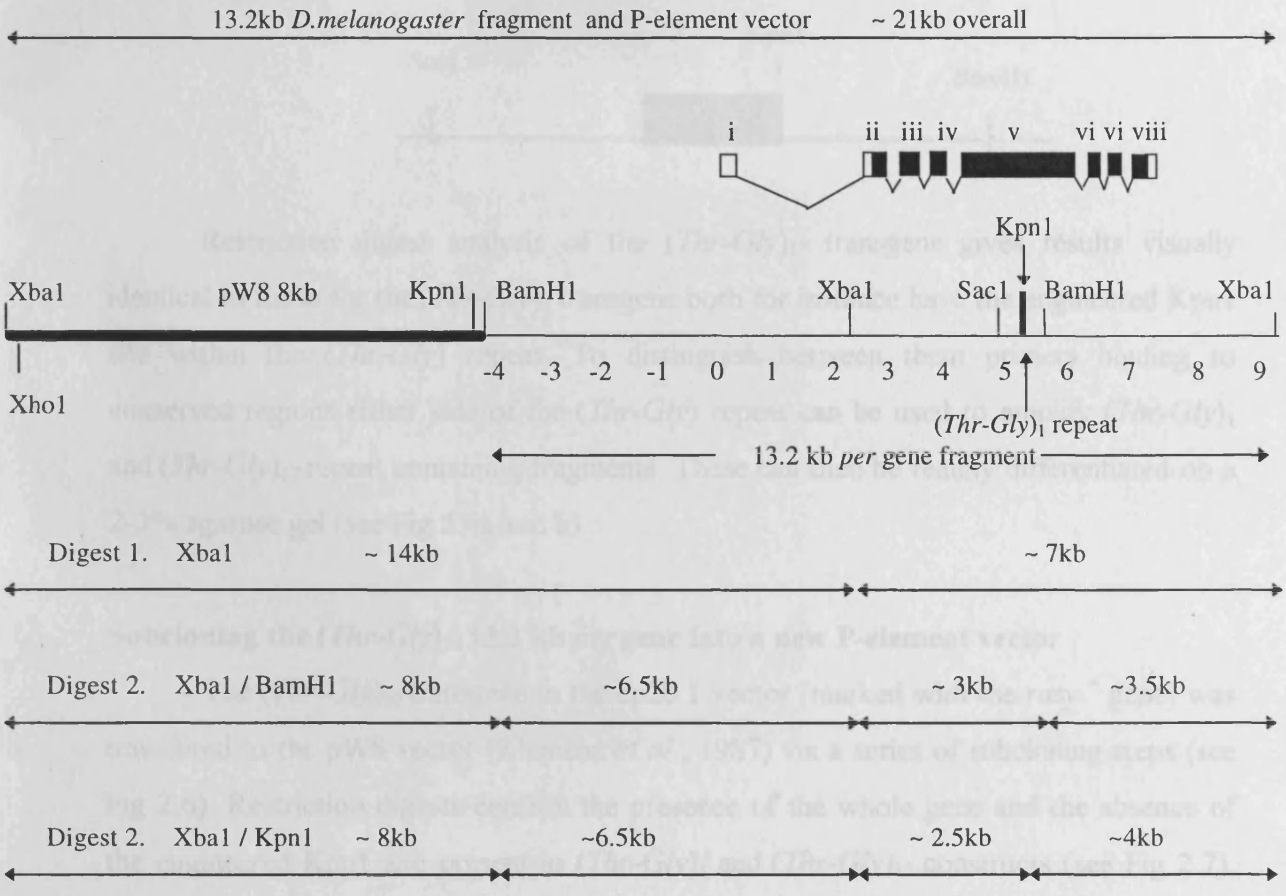
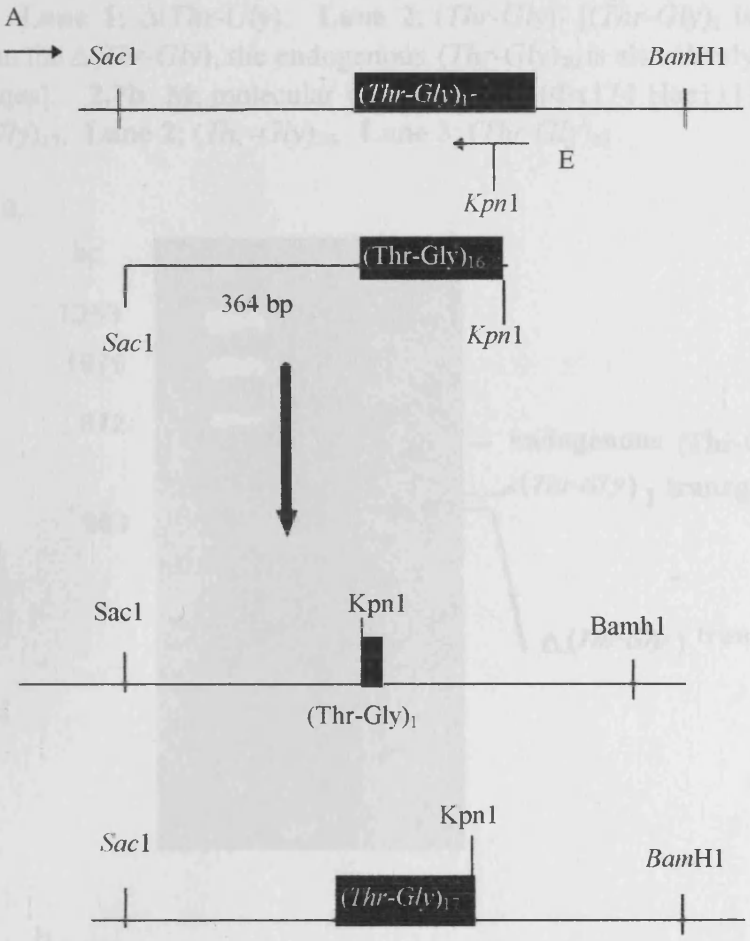


Fig 2.5 Generation of the $(Thr\text{-}Gly)_{17}$ fragment



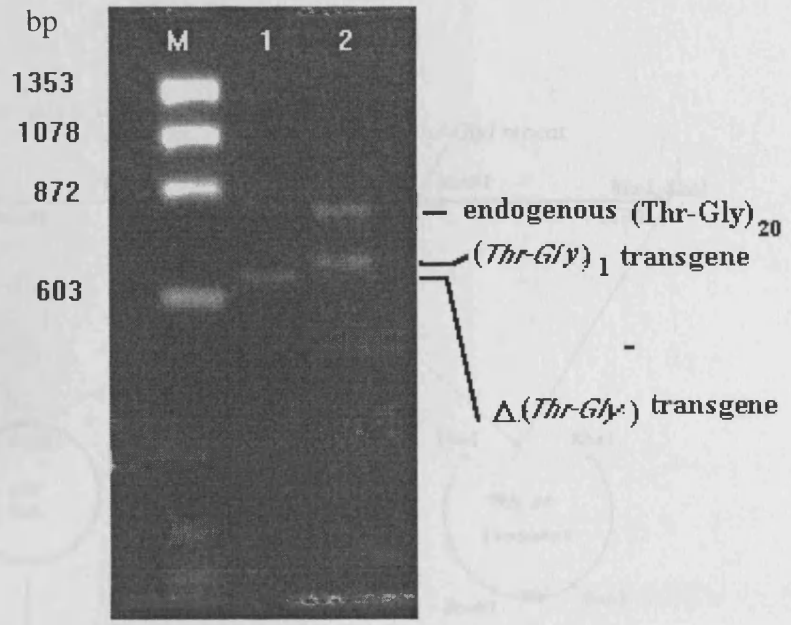
Restriction digest analysis of the $(Thr\text{-}Gly)_{17}$ transgene gives results visually identical to those for the $(Thr\text{-}Gly)_1$ transgene both for instance have the engineered Kpn1 site within the (*Thr-Gly*) repeat. To distinguish between them primers binding to conserved regions either side of the (*Thr-Gly*) repeat can be used to amplify $(Thr\text{-}Gly)_1$ and $(Thr\text{-}Gly)_{17}$ repeat containing fragments. These can then be readily differentiated on a 2-3% agarose gel (see Fig 2.6a and b)

Subcloning the $(Thr\text{-}Gly)_{20}$ 13.2 kb *per* gene into a new P-element vector

The $(Thr\text{-}Gly)_{20}$ transgene in the cp20.1 vector (marked with the *rosy* $^+$ gene) was transfered to the pW8 vector (Klemenz *et al.*, 1987) via a series of subcloning steps (see Fig 2.6). Restriction digests confirm the presence of the whole gene and the absence of the engineered Kpn1 site present in $(Thr\text{-}Gly)_1$ and $(Thr\text{-}Gly)_{17}$ constructs (see Fig 2.7). The $(Thr\text{-}Gly)_{20}$ repeat can also be readily distinguished from the $(Thr\text{-}Gly)_{17}$ by amplification and agarose gel electrophoresis (see Fig 2.6b).

Fig 2.6a Agarose gel electrophoresis of DNA generated by PCR using primers 5' and 3' to the (*Thr-Gly*) region. **M**; molecular weight marker (Φx174 Hae111). **Lane 1**; Δ(*Thr-Gly*). **Lane 2**; $(Thr\text{-}Gly)_1$ [$(Thr\text{-}Gly)_1$ is 27 bp longer than the Δ(*Thr-Gly*), the endogenous $(Thr\text{-}Gly)_{20}$ is also clearly visible in both lanes]. **2.7b** M; molecular weight marker (Φx174 Hae111). **Lane 1**; $(Thr\text{-}Gly)_{17}$. **Lane 2**; $(Thr\text{-}Gly)_{20}$. **Lane 3**; $(Thr\text{-}Gly)_{23}$

a.



b.

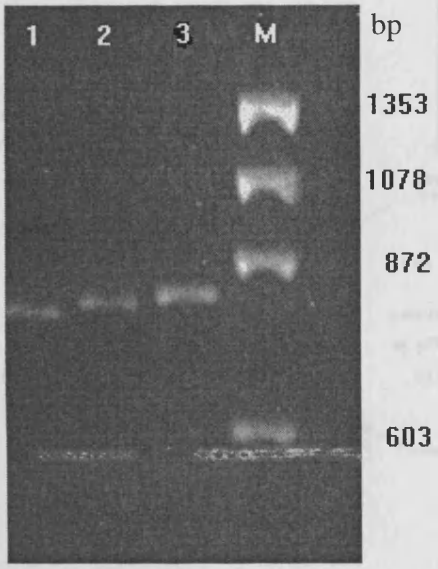


Fig 2.7 Subcloning of $(Thr\text{-}Gly)_{20}$ transgene from *rosy* vector (cp20.1) to *white* vector (pW8)

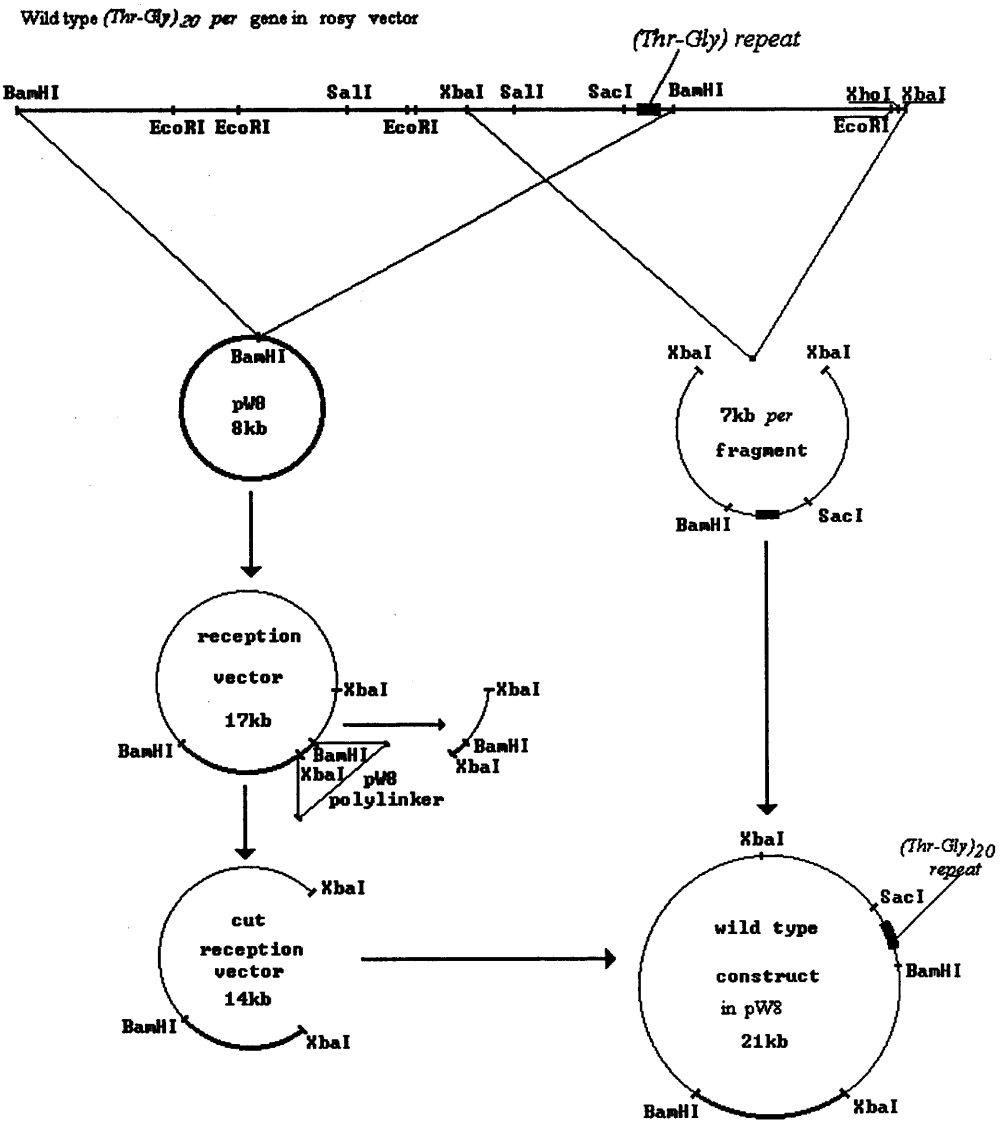
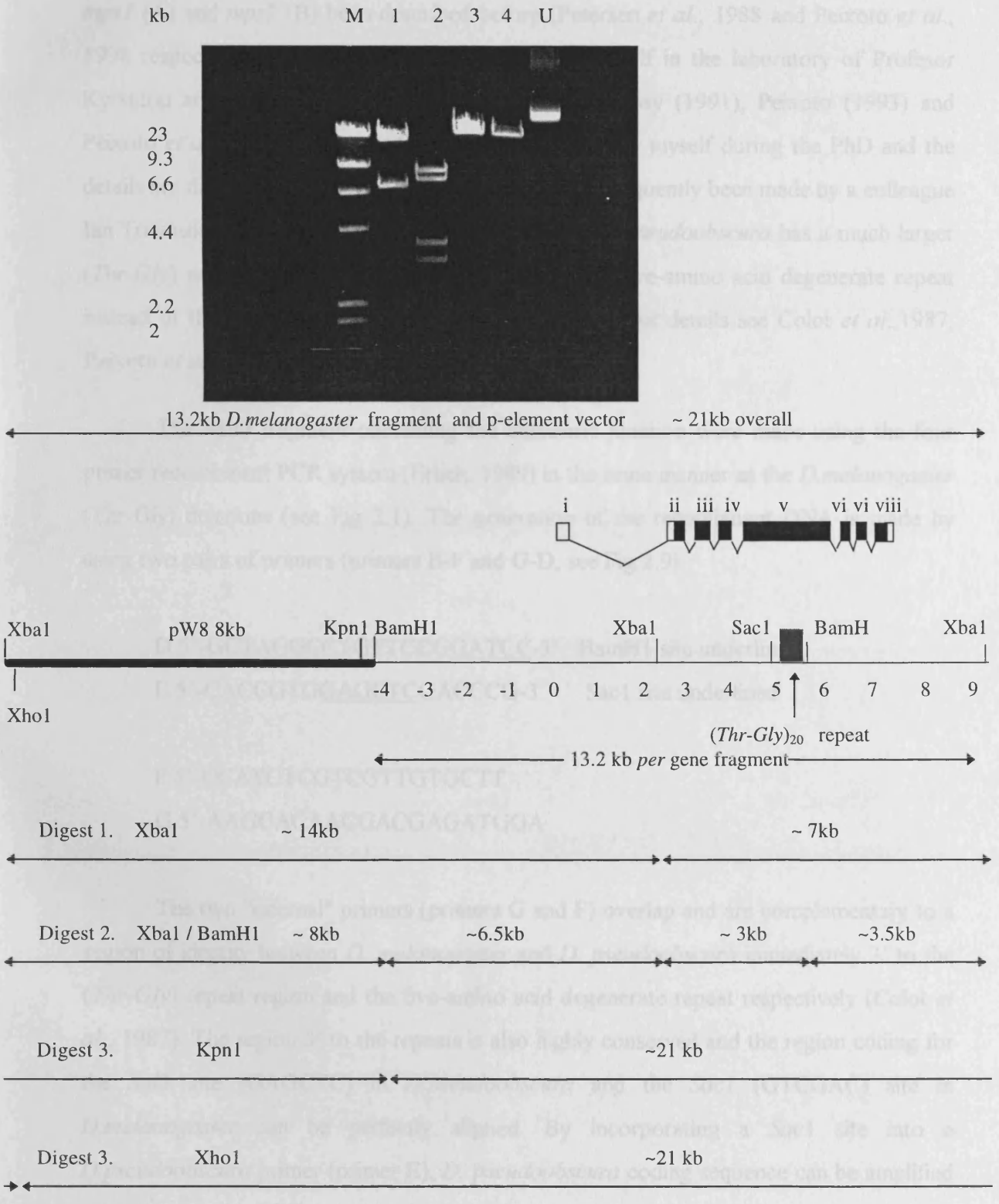


Fig 2.8 Restriction digests of the $(Thr\text{-}Gly)_{20}$ transgene (which is illustrated below). **M**; molecular wt marker λ–Hind 111. **Lane 1**; Xba1digest. **Lane 2**; Xba1/BamH1 double digest. **Lane 3**; Kpn1 digest and **Lane 4** plasmid linearized with Xho1. **U**; uncut plasmid.



PCR generation of the *D. melanogaster / D. pseudoobscura* chimaeric gene fragment used in *mps4*

Fig 2.9 briefly illustrates all of the *D. melanogaster / D. pseudoobscura* chimaeric transgenes genes that have been subjected to behavioural and molecular tests in this work. *mps1* (A) and *mps2* (B) been described before (Petersen *et al.*, 1988 and Peixoto *et al.*, 1998 respectively). *mps3* (C) was constructed by myself in the laboratory of Profesor Kyriacou at Leicester and is fully described in Hennessy (1991), Peixoto (1993) and Peixoto *et al.*(1998). Only *mps4* (D) was constructed by myself during the PhD and the details are dicussed below. An *Mps5* transgene has subsequently been made by a colleague Ian Townson and is briefly mentioned in Chapter 4. *D.pseudoobscura* has a much larger (*Thr-Gly*) region than *D.melanogaster* and encodes a five-amino acid degenerate repeat instead of the Thr-Gly repeat seen in *D.melanogaster* (for details see Colot *et al.*,1987, Peixoto *et al.*,1992 and Peixoto *et al.*, 1998).

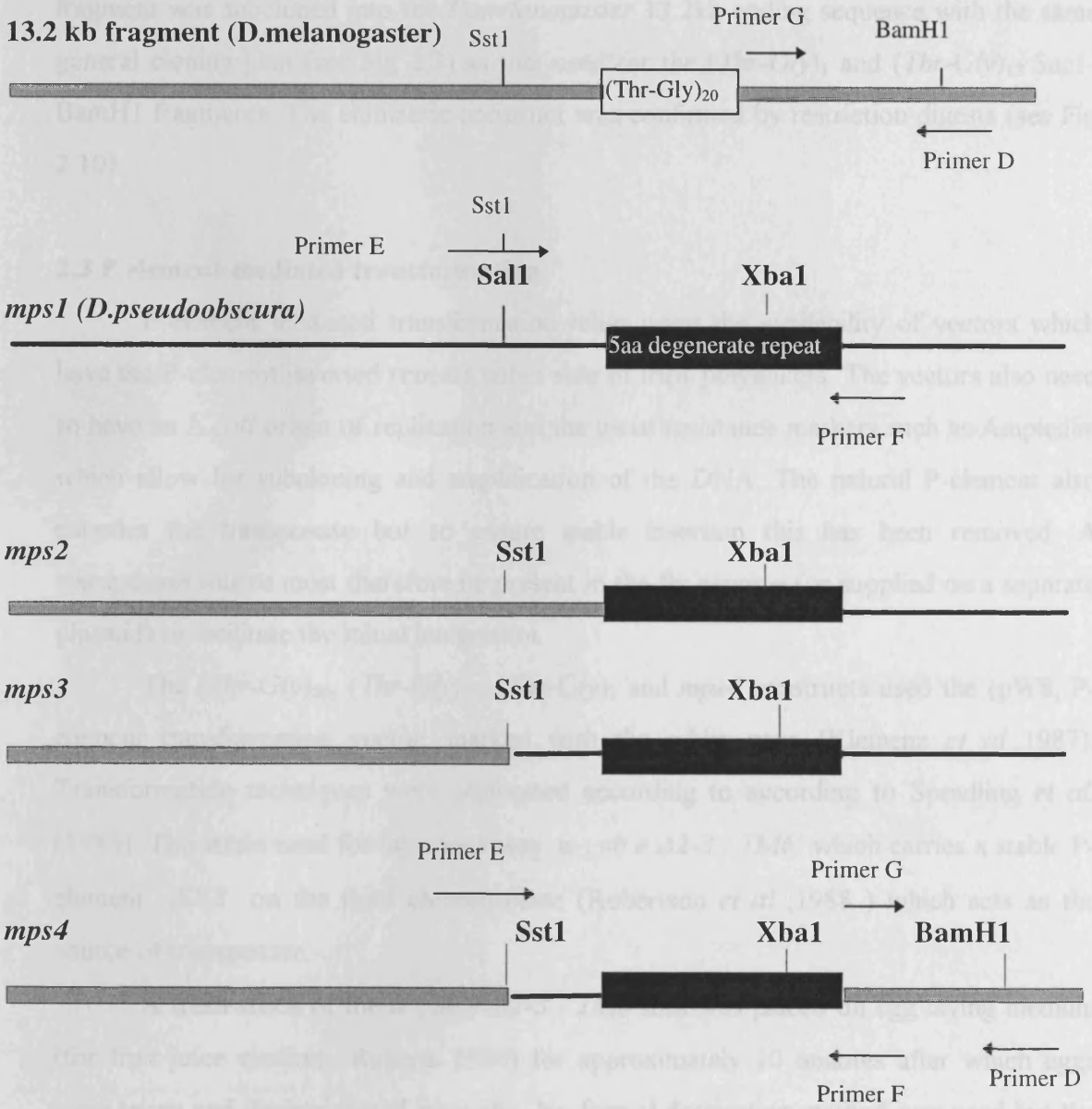
The *mps4* fragment containing the chimaeric junction were made using the four primer recombinant PCR system (Erlich, 1989) in the same manner as the *D.melanogaster* (Thr-Gly) deletions (see Fig 2.1). The generation of the recombinant DNA is made by using two pairs of primers (primers E-F and G-D, see Fig 2.9).

D 5'-GCTACGCCTGTTCCGGATCC-3' BamH1 site underlined
E 5'-CACCGTGGAGCTCGACCCG-3' Sac1 site underlined

F 5'-CCATCTCGTCGTTGTGCTT
G 5'-AAGCACAACGACGAGATGGA

The two "internal" primers (primers G and F) overlap and are complementary to a region of identity between *D. melanogaster* and *D. pseudoobscura* immediately 3' to the (*Thr-Gly*) repeat region and the five-amino acid degenerate repeat respectively (Colot *et al.*, 1987). The region 5' to the repeats is also highly conserved and the region coding for the *Sal*1 site (GAGCTC) in *D.pseudoobscura* and the *Sac*1 (GTCGAC) site in *D.melanogaster* can be perfectly aligned. By incorporating a *Sac*1 site into a *D.pseudoobscura* primer (primer E), *D. pseudoobscura* coding sequence can be amplified using primers E and F that contains only the five -amino acid degenerate repeat and its

Fig 2.9 Various *D.melanogaster/D.pseudoobscura* chimaeric (*Thr-Gly*) region constructs. Grey lines show *D.melanogaster* and black *D.pseudoobscura* coding material.



immediate 5' flanking sequences but incorporating a *Sac*1 site for ligation into *D.melanogaster* sequence. Primer G and D are used to amplify the *D.melanogaster* sequences downstream of the repeat region including the *Bam*H1 site. these two fragments can then be coamplified in the presence of primers E and D to produce a chimaeric *D.pseudoobscura/D.melanogaster* Sac1-BamH1fragment (see Fig 2.1). This fragment was subcloned into the *D.melanogaster* 13.2kb coding sequence with the same general cloning plan (see Fig 2.3) as that used for the $(Thr\text{-}Gly)_1$ and $(Thr\text{-}Gly)_{17}$ Sac1-BamH1 fragments. The chimaeric construct was confirmed by restriction digests (see Fig 2.10)

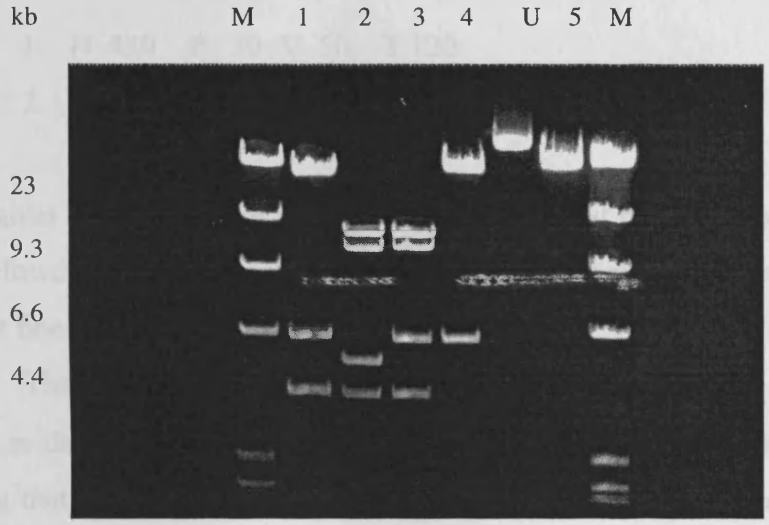
2.3 P element-mediated transformation

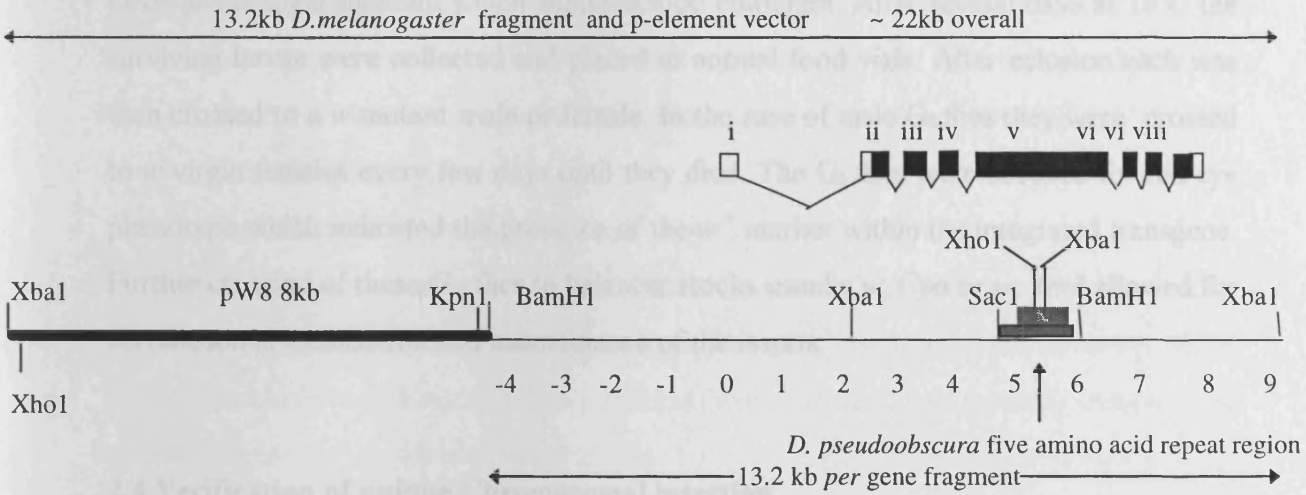
P-element mediated transformation relies upon the availability of vectors which have the P-element inverted repeats either side of their polylinkers. The vectors also need to have an *E.coli* origin of replication and the usual resistance markers such as Ampicillin which allow for subcloning and amplification of the DNA. The natural P-element also encodes the transposase but to ensure stable insertion this has been removed. A transposase source must therefore be present in the fly genome (or supplied on a separate plasmid) to facilitate the initial integration.

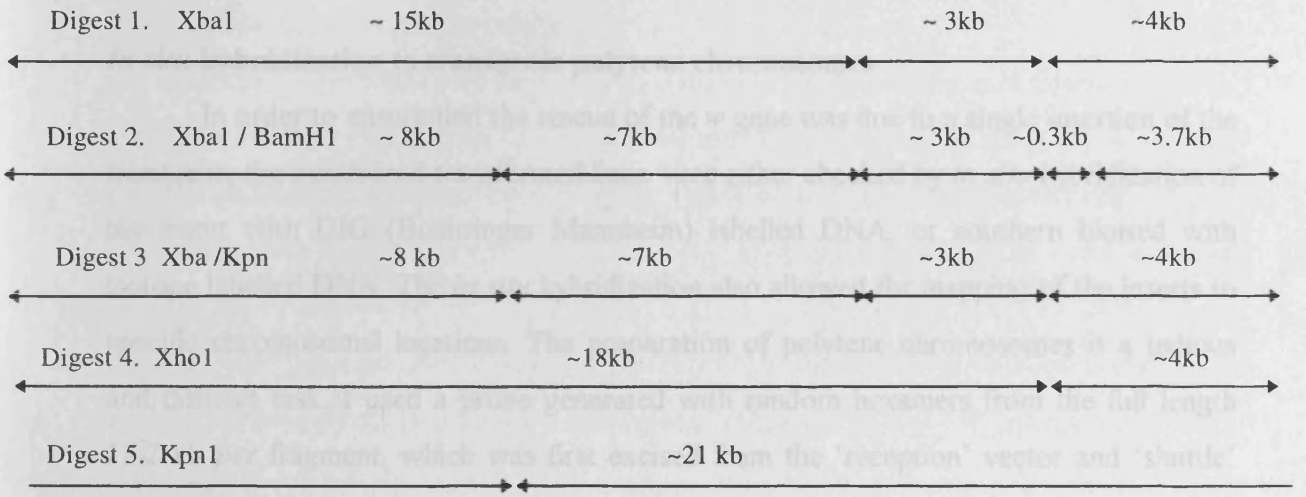
The $(Thr\text{-}Gly)_{20}$, $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_1$ and *mps4* constructs used the (pW8, P-element transformation vector, marked with the *white* gene (Klemenz *et al.*,1987). Transformation techniques were performed according to according to Spradling *et al.* (1986). The strain used for injections was *w* ; *sb e Δ2-3 / TM6* which carries a stable P-element *Δ2-3* on the third chromososme (Robertson *et al* ,1988) which acts as the source of transposase.

A fresh stock of the *w* ; *sb e Δ2-3 / TM6* flies was placed on egg laying medium (for fruit juice medium, Roberts 1986) for approximately 10 minutes after which eggs were taken and dechorionated manually. No formal dessication method was used but the maximum number of eggs were dechorionated for 10-15 minutes, and were then placed on the edge of double sided tape adhesive tape (after many trials *Scotch Tape* was found to be optimal) which itself was mounted on a cover slip. The eggs were layed so that the pole cells slightly protuded over the edge of the double sided tape. The cover slip was then mounted on a microscope slide using a drop of water to fix the cover slip and was placed on an inverted microscope at 100x magnification. The eggs were then covered

Fig 2.10 Restriction digests of the *mps4* transgene. **M**; molecular wt marker λ–Hind 111. **Lane 1**; Xba1 digest showing D.pseudoobscura repeat Xba1 site. **Lane 2**; Xba1/BamH1 double digest (0.3kb fragment not shown). **Lane 3**; Xba1/Kpn1 double digest digest. **Lane 4**; plasmid digested with Xho1 showing D.pseudoobscura repeat Xho1 site. **U**; u ncut plasmid. **Lane 5**; Kpn1 digest.







with Voltalef mineral oil (grade 10s). DNA was injected into the pole cell region either by direct injection from a pneumatic pump (eg PV830 from World Presicion Instruments) or by continuous flow of DNA from the needle tip. Microinjection needles were pulled on a Flaming Brown needle puller set to the following program:

1. H 830 P 30 V 50 T 120
2. H 960 P 100 V 220 T 55

This gives a strong needle with a short parallel sided but fine secondary tip with closed end. However the uniform nature of the secondary tip allows it to be cut several times and still be fine enough for egg injection.

The constructs were all approximately 21-22kb and were injected at 150 - 250 ng/µl in the injection buffer [5mM KCL; 0.1 mM NaH_2P04 (pH 6.8)]. It is worth noting that this buffer is extremely weak and it's pH maintenance is probably a critical factor for injection success. After injection the cover slip was placed on normal *Drosophila* sugar medium which included food colourant. After several days at 18^0C the surviving larvae were collected and placed in normal food vials. After eclosion each was then crossed to a *w* mutant male or female. In the case of male G_0 flies they were crossed to *w* virgin females every few days until they died. The G_1 flies were checked for red eye phenotype which indicated the presence of the w^+ marker within the integrated transgene. Further crossing of these G_1 flies to balancer stocks usually *w*; *Cyo* or *w*; *Tm3* allowed for chromosomal localisation and maintainance of the inserts.

2.4 Verification of unique Chromosomal insertion

In situ hybridization to transgenic polytene chromosomes

In order to ensure that the rescue of the *w* gene was due to a single insertion of the transgene, the maintained transformed lines were either checked by *in situ* hybridization of the insert with DIG (Boehringer Mannheim) labelled DNA, or southern blotted with isotope labelled DNA. The *in situ* hybridization also allowed for mapping of the inserts to specific chromosomal locations. The preparation of polytene chromosomes is a tedious and difficult task. I used a probe generated with random hexamers from the full length 13.2 kb *per* fragment, which was first excised from the 'reception' vector and 'shuttle'

vector (see Fig 2.3) using the BamH1 and Xba1 sites respectively, and then cleaning the relevant fragments by phenol extraction from low melting point gel. Approximately 0.5μg of this was then labelled with a mixture of nucleotides containing some labelled with the DIG epitope (for details see DIG labelling kit, Boehringer Mannheim).

Preparation of the polytene chromosome

1. One large larva was placed in 100ul of freshly prepared 45% acetic acid and the salivary glands dissected out.
2. The salivary gland was transferred to 8-10ul of 45% acetic acid on a siliconized coverslip.
3. The coverslip was inverted and pressed onto a gelatinized microscope slide. sharp tapping on the side of the slide then caused rupture of the salivary glands and dispersion of the polytene chromosomes which at this stage are in a globular state. This process can be easily followed on a normal light microscope fitted with phase contrast at 400x
4. Very gentle tapping with a rubber pencil using a circular pattern of points spreads the polytene chromosomes until the three visible chromosomes have clearly separated.
5. Using absobent paper eg Whatman 3 MM, folded so that it is above and below the slide, the inverted slide is then pressed with the thumb to make the spread chromosome fix to the gelatinized slide.
6. The slide is then dried on a heating block at 40-50^{0}C for 5-10 minutes.
7. The slide is then put into a -80^{0}C freezer for 5 minutes on a previously cooled aluminium block. On removal the cover slip is flicked off using a mounted needle. After dehydration in 95% alcohol, the slide can either be viewed and the successful chromosome preparations subsequently probed, or stored.

Preparation of slides and coverslips

Cover slips were first cleaned very carefully using ethanol and Kleenex tissues. Any dust left after the alcohol has dried is removed using a photographic air brush. The cover slips were then passed through silane (sigmacote, Sigma) and individually airdried by holding them with a forceps in a fume hood. They are then stored in a dry dust free box.

Slides were carefully washed in detergent and then rinsed in distilled water. They were gelatinized by dipping (using a slide rack), into a solution of 1% gelatine and 0.01%

chromealum ($CrK[SO_4]_2 .12\ H_2O$) at room temperature, and subsequent drying in a dust free environment.

Denaturation of the chromosome

1. Using slide baths, the slides with fixed polytene chromososmes were passed into 95% ethanol for 5min, then incubated in SSC 2X at 75^{0}C for 30 min
2. Dehydration was then repeated using first 70% and then 95% ethanol both for 5min.
3. The DNA was denatured with freshly prepared NaOH 0.07N (pH 12.5) for exactly three minutes.
4. Dehydration was repeated as in 2.

Preparation of probe

1. 0.5ug of DNA in 10ul of ddH_2O was denatured for 10 min at 95^{0}C and placed on ice. To the cooled denatured DNA was added; 2ul of random hexanucleotide primers, 2ul of DIG DNA labelling mixture (1mMdATP; 1mM dCTP; 1mM dGTP; 0.65mM dTTP 0.35mM DIGdUTP supplied in the DIG labelling kit) and ddH_2O was added to a final volume of 20µl. the labelling reaction was allowed to proceed overnight at 37^{0}C. The labelled probe was then boiled, 20X SSC, SDS and H_2O were then added to a final volume of 200µl. Final concentrations of SSC and SDS were 5X and 0.1% respectively.

Hybridization

1. The probe was denatured at 90^{0}C for 5 min and then immediately placed on ice
2. 7ul of probe were then put on top of the area known to contain the chromosomes, and covered with an 18 x 18mm cover slip to spead the probe. The coverslip was then seale with glue (UHU-extra) to prevent dehydration of the probe
3. The slide was then put into a sealed plastic box which was kept humid with paper soaked in SSC 5x and incubated at 68^{0}C for four hours
4. The cover slip was removed with tweezers and the slide put into 100mM maleic acid 150mM NaCL pH7 for one minute, and then 1% Boehringer Mannheim blocking buffer for 30min

Antibody binding

Commercially prepared pimary antibody raised to the DIG epitope was then used according to Boehringer Mannheim protocols and subsequent to this step a variety of commercially prepared secondary antibodies labelled with alkaline phosphatase or fluorescent tags could be used to visualize the insert.

Initially I used alkaline phosphatase labelled secondary antibody (Boehringer Mannheim) was used. This technique was used to probe the $(Thr\text{-}Gly)_1$ line 88a (see Fig 2.11). Subsequently all the other transgenic lines created in Padova and Leicester were analysed by fluorescent *in situ* hybridization by Hatzopoulou (1994) [This technique proved to generate more consistent results]. Polytene map positions were designated where possible (see Fig 2.12a for an example of the hybridization of a transgenic insert with a fluorescent [FITC] secondary antibody). The DIG system allowed for the same probe to be used in both protocols. Fig 2.12b gives the chromosomal location and position (if known) for the $(Thr\text{-}Gly)_1$, $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_{20}$ and *mps4* lines used in this work. The $(Thr\text{-}Gly)_{17}$ lines 11a and 41a were generated some time after the other lines and single insertion in these two lines was verified by the considerably quicker technique of southern blot (see Fig 2.13).

Southern Blotting

Fly genomic DNA extraction for Southern blotting

This was performed as follows. Fifty males were collected, anaesthetised with CO_2 and placed in eppendorf tubes on ice. The flies were ground in 50 µl Solution A (0.1 M Tris-HCl pH 9.0, 1 mM EDTA, 1% SDS and 1% DEPC) and incubated at 70 °C for 30 min. Then 70 µl of 14 M Potassium Acetate was added and the tubes were incubated on ice for 30 min. The tubes were spun for 30 min at 13,000 rpm in a microfuge at 4 °C. The supernatant was transferred to a fresh tube and 5 µl of 10 mg/ml RNAse was added, the tubes were incubated at 37 °C for 30 min, and the solution was then extracted once with phenol:chloroform. The solution was precipitated on dry ice with 2 volumes of isopropanol and 1/40 volume of 4M NaCl. The tube was spun at 13,000 rpm in a microfuge and the resulting pellet was washed in 70% ethanol and dried at room

Fig 2.11. *In situ* hybridization of polytenic chromosome from (*Thr-Gly*)$_1$ line 88a with *per* DNA. **White arrow** shows location of endogenous *per* gene close to the end of the X chromosome. **Green arrow** shows transgenic insert

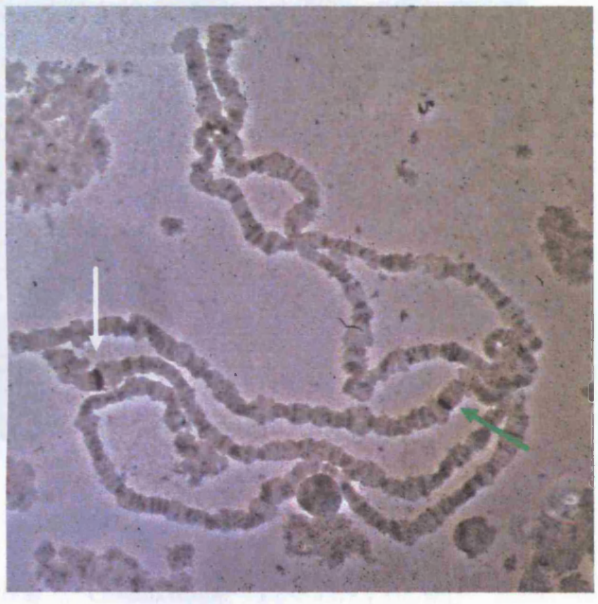
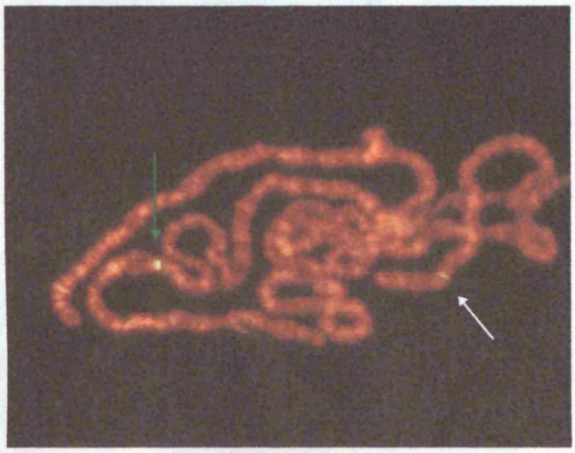


Fig 2.12a. Fluorescent *in situ* hybridization with FITC secondary antibody on *per* transgene (background staining of chromosome with ethidium bromide). **White arrow** endogenous *per* at end of X chromosome, **green arrow** transgenic *per* gene. **2.12b** Table of identified single transgenic insertions with chromosomal locations (polytenic banding positions given if known)

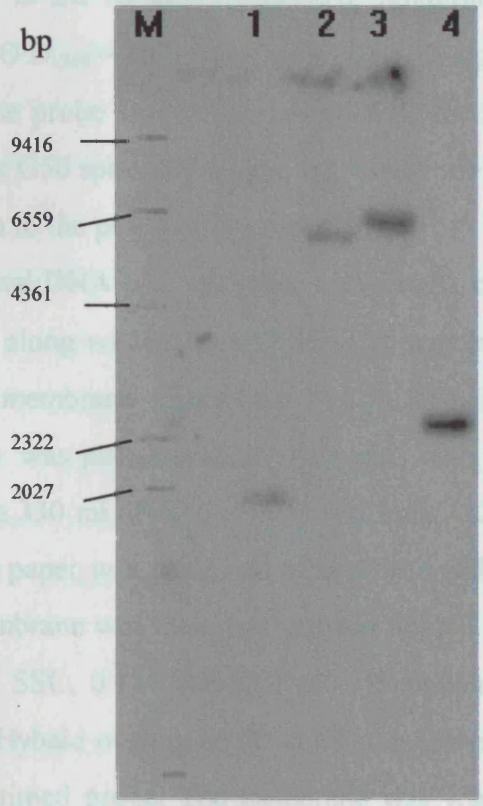
a.



b.

Genotype	line	Chromosome	Position (if known)	Insert No	Balancer
(Thr-Gly)$_1$	44a	2	2R 51F-52A	1	CyO
	57a	2	2R 42F-43A	1	CyO
	88a	2	2R 43ABC	1	CyO
(Thr-Gly)$_{17}$	11A	3			TM6
	41A	2			CyO
(Thr-Gly)$_{20}$	17a	3	3L 75DEF	1	TM6
	116a	2	2R 48F	1	CyO
mps4	6f	2	2R 55E-56A	1	CyO
	16b	3		1	TM3

Fig 2.13 Southern blot of $(Thr\text{-}Gly)_{17}$ transgenes showing single ectopic insertion into genome [Genomic DNA cut with EcoR1 which cuts in the pW8 polylinker immediately 3' to the 5'P-element end and outside the transposed sequences in the genome, probed with a 580 bp Hind 111 fragment generated from the 5' P-element end (Klemenz *et al.*,1987)]. **M**; molecular weight marker λ Hind 111. **Lane 1**; $(Thr\text{-}Gly)_{17}$ line 11a. **Lane 2**; $(Thr\text{-}Gly)_{17}$ line 41a. **Lanes 3 and 4**; genomic DNA from other transgenic flies known to carry single insertions and used here as controls.



temperature. The dry pellet was resuspended in 50 µl TE (1 mM Tris-HCl pH 8, 0.1 mM EDTA). DNA concentration was estimated by gel electrophoresis (0.8% HGT agarose gel) against a known standard

Preparation of Probes for Southern Blotting

Probes were labelled by a random priming method as follows. 10-50 ng of probe DNA was suspended in 25 µl water, boiled for 5 min and incubated on ice for 2 min. 10 µl of OLB (OLB consists of solutions A, B and C in the ratio 100:250:150. Solution A is 1.25 M Tris-HCl pH8.0, 0.125 M $MgCl_2$, 0.5 mM dATP, 0.5 mM TTP and 0.5 mM dGTP. Solution B is 2.0 M HEPES pH 6.6. Solution C is hexadeoxyribonucleotides dissolved in TE at $O.D._{260}$ = 90 U/ml.) 2 µl of BSA, 2.5 µl ^{32}P dCTP, and 1 µl of Klenow were added and the probe was incubated at 37 °C for 2 h. The probe was recovered through a sephadex G50 spin column and then boiled for 5 min and incubated on ice for 2 min before addition to the pre-hybridisation buffer.

Chromosomal DNA was restricted with EcoRI or KpnI and run on 0.8% HGT-TAE agarose gels, along with 10 ng of λ HindIII marker DNA. These were then blotted onto Hybond N+ membrane (Amersham Ltd., UK) using an alkaline transfer blotting procedure. Dry N+ was placed on top of the gel, with its wells removed on a blotting platform containing 330 ml of ATS (1.5 M NaCl and 0.25 M NaOH). A 10 cm stack of quickdraw blotting paper was placed on top and secured with a glass plate, this was left overnight. The membrane was then removed and floated in 0.4 M NaOH for 10 min each side, rinsed in 2 x SSC, 0.1M Tris-HCl pH 7.5 and finally in 2x SSC. The filter was prehybridised in a Hybaid oven at 65 °C in Church buffer for 15 min and then hybridised with a randomly primed probe. The membrane was washed, while monitoring, 3 times with, 5 x SSC 0.1 x SDS, 4 x SSC 0.1 x SDS, and 3 x SSC 0.1% SDS. The membrane was then autoradiographed for 2 days and developed in an automatic developer. EcoRI digested DNA was probed with a 580 bp probe derived from a Hind111 restrction fragment of the 5' P-element end contained in the pW8 plasmid (Klemenz *et al.*, 1987, shown in Figure 2.13).

2.5 Locomotor activity experiments

Homozygous or heterozygous balanced transgenic flies were crossed to per^{01} virgin female flies approximately 14-18 days before activity analysis was carried out. The

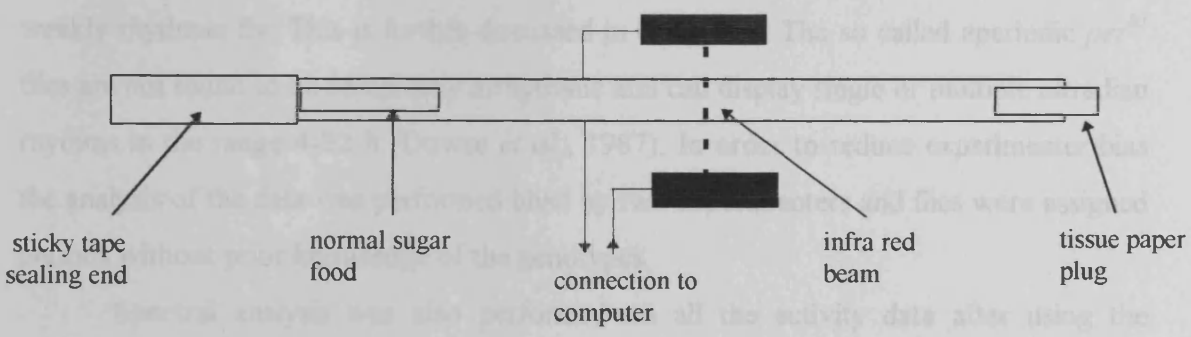
flies were then kept at 25^{0}C LD 12:12 with lights on/off at 09:00/21:00 h. At this temperature the progeny start to eclose about 12 days after crossing. Only male offspring between 1 and 7 days old, carrying a *per* ^{01} bearing X chromosome and one copy of the transgene, were used.

The locomotor monitor was produced by Biodata Ltd (Manchester, UK). and consists of rows of the activity collection units depicted in Fig 2.14. Each fly was loaded into a glass tube (8 x 0.3 cm) with 3 cm of sugar food in one end secured with tape to prevent desiccation and a tissue bung at the other. When loaded, each tube was situated between an infra-red light emitter and a detector which records an event every time the light beam is broken by the fly. Infra-red light beams are used because flies are insensitive to this part of the spectrum (Bertholf, 1932; Kyriacou and Burnet, 1979). The apparatus has the ability to measure the number of times the beam is broken within a given time span, which in this work was always 30 minutes (one activity bin)

Protocols for DD experiments

Flies two to seven days old were loaded into tubes and placed in an incubator at either 18, 25 or 29^{0}C in the same LD 12:12 conditions in which they were raised on day one. On day two the incubators started DD at ZT 12. Data collection started on day three and the first data were collected at C.T 30, 18h after the last light to dark transition. Data is collected in 48, 30 minute bins per 24 h and writen directly to an EXCEL spreadsheet

Fig 2.14 Apparatus for measuring activity of *Drosophila*



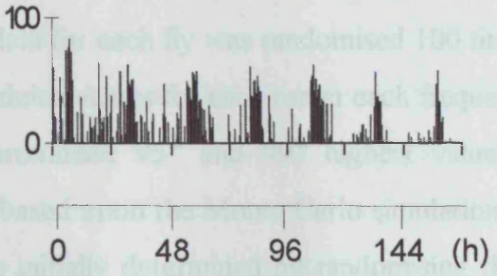
Raw activity data were edited before analysis as follows. Dead flies were removed from the data set. Flies which died part-way through the experiment were removed if they

had generated less than four and a half days of data, and hyperactive flies were removed (those that consistently produced bin scores of more than 200 events per 30 minute bin). Some unusual counts usually with the format 301 or 701 were occasionally seen and were probably caused by a software malfunction. These random errors were dealt with by deleting them and substituting an average of the figures either side. Other activity recorders have shown similar problems (Hamblen-Coyle *et al.*, 1992).

In the work presented here all free-running (DD) periodicity was determined by autocorrelation analysis (Diggle, 1990 available in the SPSS/PC+ Version 2.0 software package) and Spectral analysis (Roberts *et al.*, 1987: Kyriacou and Hall, 1989). Autocorrelation analysis works thus: the number of events separated by a certain time lag are correlated. Therefore with a time lag of one time bin (30 min), bin 1 data is correlated with bin 2, bin 2 with bin 3, bin 3 with bin 4..., etc. For example, if the data shows a 24 h period with bins of 30 min, then bin 1 and bin 49, bin 2 and bin 50..., etc. will produce the highest correlation coefficient. These correlation coefficients are plotted against each lag (1, 2, 3, etc.) in a correlogram which extends to 116 lags or 58 h. Figure 2.15b shows an example of such a correlogram for a fly with a period of 24.5h. The peak is at lag 49 and a second peak is near its harmonic at lag 98. The confidence limits indicated by the horizontal lines are set at 95% ($2/\sqrt{N}$, where N is the number of bins). Rhythmicity was determined for each fly by taking the highest peak above the 95% confidence limit in the range 15- 40h in the correlogram. This peak represents the freerunning period of activity of the fly. The overall pattern of each correlogram was also considered. A robust significant period should be followed by another significant peak with twice the period of the first peak. There is a degree of subjectivity in that there is 'noise' present on some correlograms which can make it difficult to distinguish between an arrhythmic fly and a weakly rhythmic fly. This is further discussed in Chapter 3. The so called aperiodic per^{01} flies are not found to be completely arrhythmic and can display single or multiple ultradian rhythms in the range 4-22 h (Dowse *et al.*, 1987). In order to reduce experimenter bias the analysis of the data was performed blind by two experimenters and flies were assigned periods without prior knowledge of the genotypes.

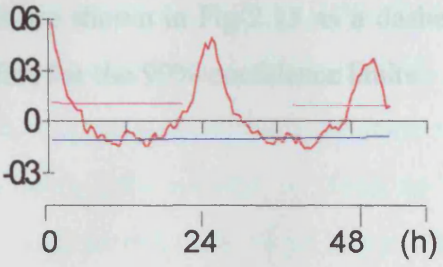
Spectral analysis was also performed on all the activity data after using the CLEAN algorithm (Roberts *et al.*, 1987: Kyriacou and Hall, 1989). Spectral analysis works by breaking down a signal into its sine and cosine waves shown in Figure 2.1c where the spectrograms show the frequencies that best represent the signal. The highest relevant peak is taken as the period. However, if the data was bimodal, i.e. there were

a.



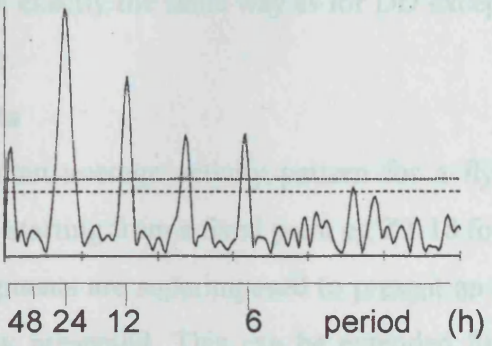
HISTOGRAM

b.



AUTOCORRELATION

c.



SPECTROGRAM

Fig 2.15 Typical free-running locomotor activity histogram, autocorrelation and spectral analysis of an individual *D.melanogaster* fly showing period *circa* 24.5h.

significant ~12 h and ~24 h peaks, and the ~ 12 h peak was much more robust then the this peak was doubled to obtain the circadian period. This doubling of the strong ~12 h spectral period also usually agreed better with the circadian correlogram value for those flies. Flies with 12 h peaks were seen most often when experiments were performed at higher temperatures in DD, where most of the activity is limited to subjective dawn and dusk.

Activity data for each fly was randomised 100 times and the CLEAN analysis was repeated. The modulus values for each run at each frequency were ordered from lowest to highest. The approximate 95th and 99th highest values represent the 95% and 99% confidence limits based upon the Monte Carlo simulation of the random data. The 99 and 95 % limits were initially determined by randomising the data 1000 times for each fly. However, this was expensive in computer time and as values for 1000 and 100 randomisations were very similar, 100 randomisations were subsequently used in all Monte Carlo simulations. To be determined significant, a peak had to be above the 99% limit. These limits are shown in Fig 2.15 as a dashed horizontal line for the 95% and a dotted horizontal line for the 99% confidence limits.

Protocols for LD experiments

Data is collected in exactly the same way as for DD except that the incubators continue in the LD 12:12 cycle.

'wrapping' of data

To present an average activity pattern for a fly the 7-10 days of data can be 'wrapped'. That is starting from a fixed point eg ZT 18 for a fly in LD. The data for all the succesive 24 h segments are superimposed to present an average activity pattern that can then be graphically presented. This can be extended further, by the collation of many wraps the activity patterns of a population of flies can be visualised and compared statistically with other populations. Although in DD the clock free-runs and the wrapping of individuals with different circadian periods is problematic, it has been attempted in this work and is explained in the materials section of the relevant chapters.

with another three pieces of Whatman paper. Proteins were then blotted at 400mA for 4hrs (for a 16x18 cm gel).

Immunoblotting

[Rabbit anti-PER antibody (gift of J.Hall and R.Stanewsky) was used at a concentration of 1:10,000, and rat anti-TIM antibody (gift of M. Myers) was used at 1:1000]

After electroblotting the membrane with the bound proteins was then visualised with Ponceau S to ensure that the transfer had taken place before going on to antibody binding. If transfer was sufficient and no major areas were without transfer then the membrane was incubated for one hour in blocking solution (2% w/v non fat dried milk. 0.05 % w/v Tween 20 in TBS [10mM Tris-Cl pH 7.5, 150mM NaCl] and then for 2 h in blocking solution containing the primary rabbit or rat antibody (which was diluted by trial and error until a satisfactory binding was achieved). The filter was then subjected to three 5 min washes in TBST (TBS containing 0.05% Tween 20). The horse radish peroxidase cojugated secondary (antirat or antirabbit) antibody, diluted according to manufacturers suggestion, was then incubated in blocking solution with the membrane for 1hr. The membrane was subsequently washed 3x 5min in TBST. Sufficient chemiluminescent solution [6.25uM Luminol and 6.38uM p-Coumaric acid 0.1M Tris-Cl pH 8.5 in H_20] to cover the membrane was then freshly prepared and poured over it taking care that the distribution was even. The chemiluminescent signal was then detected by exposing the membrane to autoradiography film for between 1-5 min to achieve the desired exposure.

2.7 *Fly stocks*

Fly stocks were maintained on sugar/agar medium (4.63 g of sucrose, 4.63 g of dried brewers' yeast, 0.71 g of agar and 0.2 g of Nipagin in 100 ml of water), in either glass vials (10 x 2.2 cm) or one third pint milk bottles. The stocks were kept in temperature controlled rooms at 18 or 25 °C and in light dark cycles of 12 hours (LD12:12). Lights on/off were maintained at 09.00/21.00 h BST year-round. Stocks used and transformant strains generated are described in the relevant chapters.

Chapter 3 Molecular coevolution within *per*

3.1 Introduction

The Thr-Gly region of *per* has been shown to control species-specific lovesong rhythms using chimaeric *per* constructs from *D.melanogaster* and *D.simulans* (Wheeler *et al.*, 1991). The Thr-Gly encoding regions of the two species varied by only a few amino acids, but the differences were enough to switch species-specific behaviour from one species type to another. Although species-specific traits in circadian locomotor behaviour have been suggested to be under the control of *per* transgenes for *D.pseudoobscura* and *D.melanogaster* (Petersen *et al.*, 1988), mapping of these traits to regions of the *per* gene has not been investigated. To begin to address this problem a series of chimaeric constructs using previously cloned fragments containing the *D. melanogaster per* gene (Citri *et al.*, 1987) and the *D. pseudoobscura per* gene (Colot *et al.*, 1988) was planned.

The constructs were also created so that they would contain either the alternating Thr-Gly encoding hexamers of *D. melanogaster*, or the much larger *D. pseudoobscura* repeat in which the Thr-Gly hexamers are followed by many copies of a degenerate pentapeptide sequence. In creating a chimaeric protein there is always the risk that the protein will not function *in vivo*, so for the first construct a junction site identical to one in the successful *D.melanogaster/D.simulans* constructs of Wheeler *et al.* (1991) was chosen. The construct had the putative 5' regulatory region and the upstream part of the *D.melanogaster* coding sequence fused to the 3' half of *D.pseudoobscura per*, which included the *D.pseudoobscura* repeat region (see Materials Methods). The junction of this construct is 177 bases 5' to the beginning of the *D. melanogaster* Thr-Gly perfect repeat, so that the *D.pseudoobscura* repeat is conspecific with it's immediate 5' flanking sequence (Hennessy, 1991; Peixoto, 1993). This transgene is referred to as *mps3* (see Fig 3.1). At the same time a similar construct was being made at Brandeis University by Dr. Michael Rosbash's group in which the *D. melanogaster* coding sequence continued up to the first Thr-Gly pair, after which it is replaced by the *D.pseudoobscura* repeat and the rest of the *D.pseudoobscura* 3' sequence. In this second construct therefore the *D.pseudoobcura* repeat is heterospecific with it's 5' flanking sequence. This construct is referred to as *mps2* (see Fig 3.1). The intact *D.pseudoobscura per* gene, the Brandeis construct and the first Leicester construct are called *mps1*, *mps2* and *mps3* respectively.

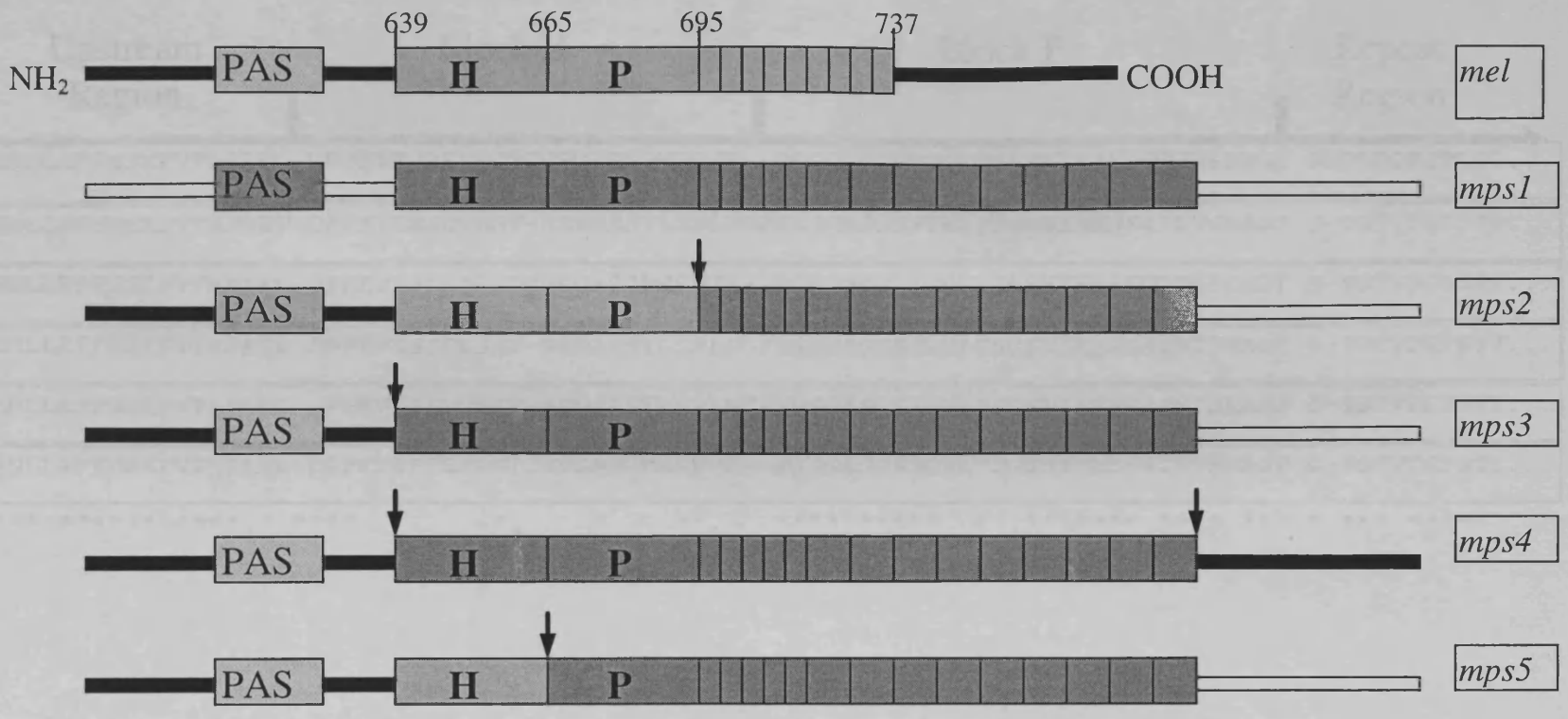


Fig 3.1a The coding regions that were manipulated to create the three chimaeric *per* genes (*mps2-4*) between *D.melanogaster* (light grey) and *D. pseudoobscura* (dark grey) coding sequences are illustrated, with the arrows representing the position of the chimaeric junctions. *mps5* is a transgene made by Ian Townson (see Peixoto *et al.*, 1998).

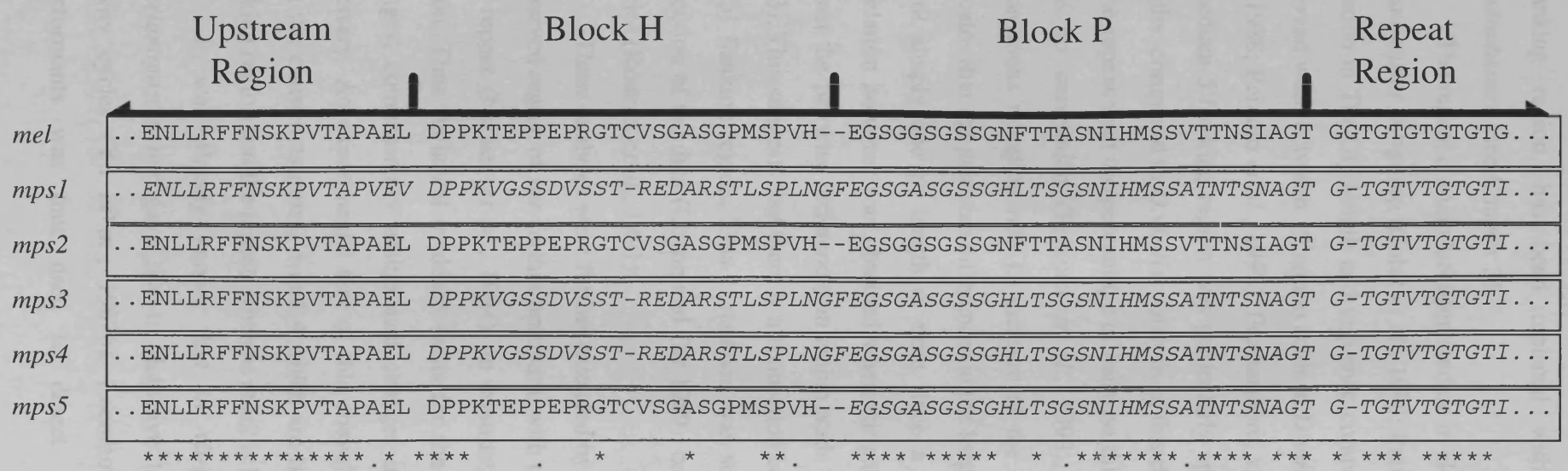


Fig 3.1b Amino acid sequences immediately upstream of the two species repeats in the chimaeric constructs. The upstream block labelled P may coevolve with interspecific repeat length (Peixoto *et al.* 1993, Nielsen *et al.* 1994) whereas block H cannot be aligned, but may nevertheless represent additional sequences that could be involved in the putative interaction between the repeat and the immediately adjacent region. Asterisks and points denote identical and similar amino acids respectively.

In addition I have made another transgene (*mps4*), which encodes the complete *D.melanogaster per* gene in which a fragment encoding the repeat and conspecific 5'flanking region, has been replaced with the corresponding fragment of *D. pseudoobscura* (see Chapter 2)

The repeat domain has been studied in eleven *Drosophila* species from outside the *melanogaster* subgroup (Colot *et al.*, 1998; Peixoto *et al.*, 1993) and shows interspecific variation in Thr-Gly length and sequence composition. The length of the repeat regions observed were between 19 amino acids in *D virilis* to 209 in *D pseudoobscura* (Colot *et al.*, 1988; Peixoto *et al.*, 1993). Furthermore, variation in amino acid composition in the immediate 5'flanking region was particularly apparent in species which had long repeat lengths, compared to *D.virilis*. Analysis of this data, comparing pairwise length difference of the repeat with the percentage of pairwise third base divergence of the flanking region gave no correlation (Peixoto *et al.*, 1993). This absence of correlation between synonymous mutation rate (a measure of the molecular clock), and length of repeat, indicates that the pairwise differences in the length of the repeat between any two species is not simply due to length of time since a common ancestor. A highly significant correlation however was obtained when the pairwise differences in length were plotted against the pairwise differences in amino acids in the 5' flanking region (Peixoto *et al.*, 1993). This suggests that there is a connection between repeat length and the evolution of the 5' flanking region. This correlation was still significant after using a phylogenetic correction of the data (Coyne and Orr 1989), based on an rRNA phylogeny for the same species (Rousset *et al.*, 1991).

These analyses were further extended by comparing protein divergence of another conserved region of *per* further upstream, with the pairwise length differences of the Thr-Gly repeat (Nielsen *et al.*, 1994). In contrast, no correlation was obtained with this region. Thus statistical evidence exists for the view that as interspecific repeat length changes, compensatory amino acid changes in the immediate 5' flanking region are necessary. An experimental test of this coevolutionary scenario now became available with the constructs *mps1, 3* and *4* which are homospecific for the *D pseudoobscura* 5' flanking region and the repeat, whereas *mps2* is heterospecific for these two regions.

It was already known that *in vitro* deletions of this repeat region in *D.melanogaster per* gave rise to defective temperature compensation in locomotor activity cycles (Yu *et al.*, 1987b). Locomotor activity of the corresponding *mps* transformants was thus used to detect whether disruption of the proposed

coevolution/interaction between the repeat and adjacent upstream sequences would yield any phenotypic effects on circadian behaviour. In addition, western blots were performed with head extracts from these transformants over the circadian cycle using anti-PER and anti-TIM antibodies.

The data presented in this Chapter has been published in *Proc. Natl. Acad. Sci. USA* Vol.95, pp. 4475-4480, April 1988. Title: " Molecular coevolution within a *Drosophila* clock gene". A reprint is enclosed in Appendix 3.

3.2 Methods

Transgene construction

(see Chapter 2 for further details)

The *mel* transgene contains a 13.2 kb fragment of the *D. melanogaster per*$^{+}$ gene and has been described previously (Citri *et al.*, 1987). It contains all the coding information plus approximately 5kb of upstream sequence and rescues rhythmicity in *per*^{01} flies. Two lines generated at Brandeis University (2a and 34a) with the 13.2 kb fragment inserted into the cp.20.1 *rosy* vector (Citri *et al.*, 1987) were tested (see table 3.1), as were an additional two 13.2 lines (17a and 116a) generated by myself in which the 13.2 kb fragment was cut out of the cp.20.1 vector and ligated into the pW8 vector (see Fig 3.1 and Chapter 2).

As the Brandeis and Leicester chimaeric trangenes used cp.20.1 and pW8 respectively, testing these four *mel* transgenes provided a control for any unforseen behavioural effects due to differences in the vectors used.

mps1 has also been described previously (Petersen *et al.*, 1988). It contains the complete coding sequence of the *D.pseudoobscura per* gene fused to the 5' upstream region of *D.melanogaster* at a position close to the 3' end of the large first intron.

mps2 - 4 The cloning steps for generating chimaeric transgenes *mps2-4* are described in Chapter 2. The brief details below are illustrated in Fig 3.1 a and b:-

mps2 has not been described previously and was made by G. Hasan at Brandeis. It contains *D.melanogaster per* material up to the end of region P (see Fig 3.1), followed by 3′ *D.pseudoobscura per* sequence. ***mps3*** was made by myself at Leicester and consists of *D.melanogaster per* up to the beginning of region H ligated to 3' *D.pseudoobscura per*

material. ***mps4*** was also made by myself at Leicester and Padova, and is a *D.melanogaster per* gene in which regions H, P and the Thr-Gly repeat have been replaced by the equivalent regions from *D.pseudoobscura* (for all details see Chapter 2).

Locomotor activity. Male per^{01} transformants carrying one autosomal copy of the transgene were examined with respect to their circadian locomotor activity under free-running conditions (DD, constant darkness), at 18, 25 and 29^{0}C as described in Chapter 2. Each fly's activity was analysed using autocorrelation (Dowse *et al.*, 1989). and a high resolution spectral analysis (Roberts *et al.*, 1987, Kyriacou *et al.*, 1989). If a record gave a significant period with autocorrelation, but not spectral analysis, or *vice-versa,* then the record was judged 'arrhythmic' (see Fig 3.2). Various ANOVAS were performed followed by the Newman Keuls *a posteriori* procedure (Winer, 1971).

Western Blots. Heads were collected at 2h intervals in LD 12:12 and snap frozen in liquid nitrogen. Extracts were prepared and Western blotting was performed as in Edery *et al.*, (1994b) with minor modifications (see Chapter 2). Equal amounts of proteins (~50 mg of total protein) were used for each time point. Rabbit anti-PER antibody (gift of J. Hall and R. Stanewsky), was used at a concentration of 1:10000, and rat anti-TIM antibody (gift of M. Myers) was used at a concentration of 1:1000. Quantification of the bands was performed using a densitometer.

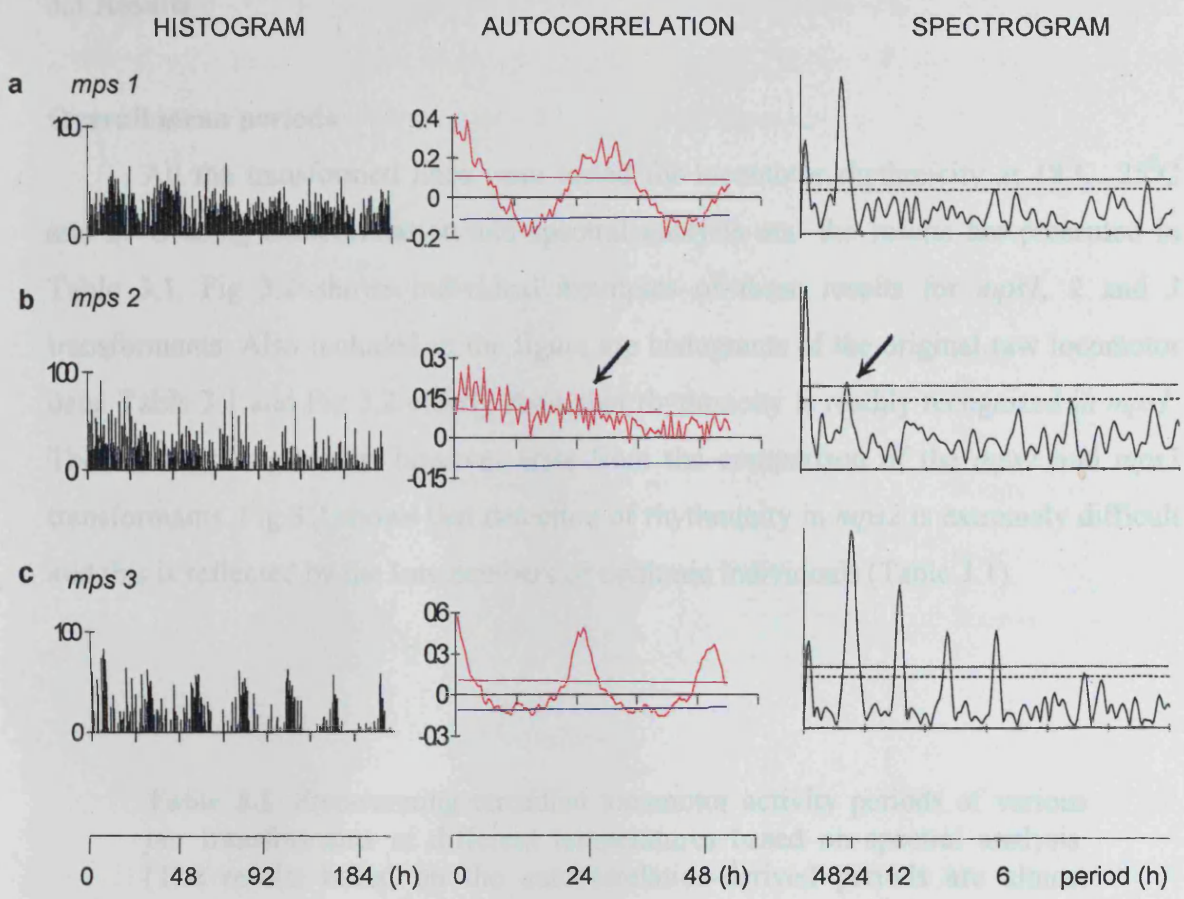


Fig 3.2 Free-running locomotor activity histograms, autocorrelation and spectral analysis for three male *per*^{01} transformants monitored at 29°C carrying one copy of the *mps1, mps2* and *mps3* transgenes. Histograms give raw activity events in 30 min bins. The correlograms show r and 95% confidence limits either side of r = 0 . Spectrograms show 95% (dashed) and 99% (lined) confidence limits.

a. *mps1* male showing 28h cycle. **b.** *mps2* record illustrated is one of the few rhythmic by our criteria (26 h arrowed) **c.** *mps3* profile also typical of *mps4* and *mel* transformants (not shown) showing period of *ca* 24.5h.

3.3 Results

Overall mean periods

All the transformed lines were tested for locomotor rhythmicity at 18^{0}C, 25^{0}C and 29^{0}C using autocorrelation and spectral analysis and the results are presented in Table 3.1. Fig 3.2 shows individual examples of these results for *mps1, 2* and *3* transformants. Also included in the figure are histograms of the original raw locomotor data. Table 3.1 and Fig 3.2 clearly show that rhythmicity is readily recognized in *mps1*. The most striking results however arise from the comparison of the *mps2* and *mps3* transformants. Fig 3.2 shows that detection of rhythmicity in *mps2* is extremely difficult and this is reflected by the low numbers of rhythmic individuals (Table 3.1).

Table 3.1. Free-running circadian locomotor activity periods of various *per* transformants at different temperatures based on spectral analysis. (The results based on the autocorrelation-derived periods are almost identical - data not shown)

Construct	line	18^{0}C				25^{0}C				29^{0}C			
		N	n	period	± sem	N	n	period	± sem	N	n	period	± sem
mel	2a	61	41	24.61	0.11	25	23	25.02	0.11	53	45	24.80	0.09
	17a	59	31	24.07	0.18	24	20	24.34	0.66	46	26	24.18	0.19
	34a	51	30	24.21	0.16	47	39	25.02	0.13	32	32	24.92	0.16
	116a	50	41	23.51	0.07	33	30	25.52	0.11	33	31	25.43	0.16
pooled		221	143	24.09	0.07	129	112	25.03	0.13	164	134	24.85	0.08
mps1	I20	26	17	29.37	0.83	31	18	28.34	1.24	61	13	29.02	1.24
	I26	25	21	27.37	0.65	37	17	28.88	0.76	82	8	29.50	0.98
pooled		51	38	28.26	0.53	68	35	28.60	0.59	143	21	29.21	0.84
mps2	6	31	4	22.73	4.12	61	5	32.98	1.61	65	5	29.77	1.62
	9	28	6	24.66	2.10	29	8	28.64	1.96	39	7	28.56	2.73
	22	28	11	27.48	2.18	68	8	28.12	2.66	34	5	28.80	2.42
pooled		87	21	25.76	1.54	158	21	29.48	1.33	138	17	28.99	1.34
mps3	65c	19	15	24.14	0.14	44	42	25.71	0.11	15	13	25.62	0.18
	67a	22	15	23.33	0.18	32	32	24.37	0.08	12	12	25.01	0.24
pooled		41	30	23.74	0.14	76	74	25.13	0.11	27	25	25.33	0.16
mps4	6f	17	15	24.27	0.14	23	20	24.82	0.11	66	62	24.84	0.10
	16b	27	16	23.95	0.11	20	17	25.15	0.18	24	21	24.72	0.14
pooled		44	31	24.11	0.09	43	37	24.97	0.10	90	83	24.81	0.08
per^{01}		31	5	22.58	1.06	32	3	19.26	1.86	57	4	21.53	2.13

Fig 3.3 illustrates the the results of the more sensitive spectral analysis only. The *mps2* transformants had very poor recovery of arrhythmicity (Table 3.1, and see Fig 3.2) and are left out of this comparison as they are dealt with separately in Fig 3.6.

Fig 3.3 Mean periods of free-running circadian locomotor analysis for all transgenic lines at different temperatures.

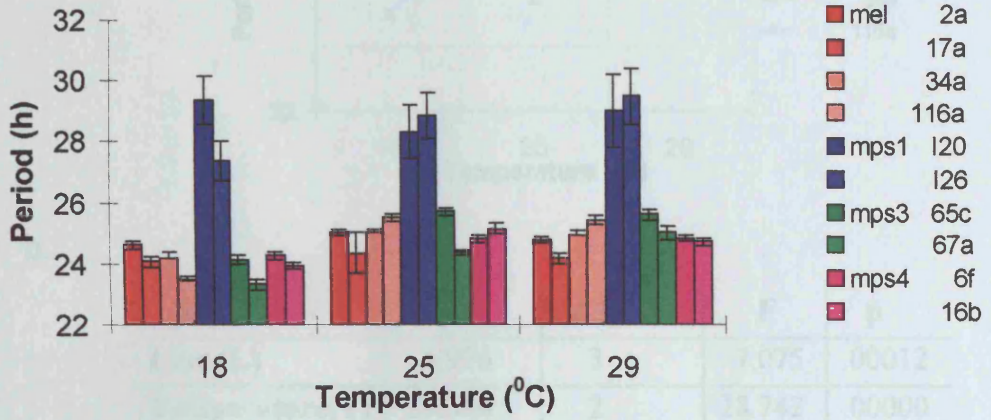


Fig 3.3 shows the results for all lines of all genotypes with high levels of rhythmicity (see Table 3.1). It can plainly be seen that the *mps1* lines I20 and I26 have significantly longer periods (in the range 28-29 h) than the rest of the lines which are more tightly grouped in the 23.5-25.5h range. In order to examine the data more closely individual ANOVAS were perfomed for the lines within each genotype

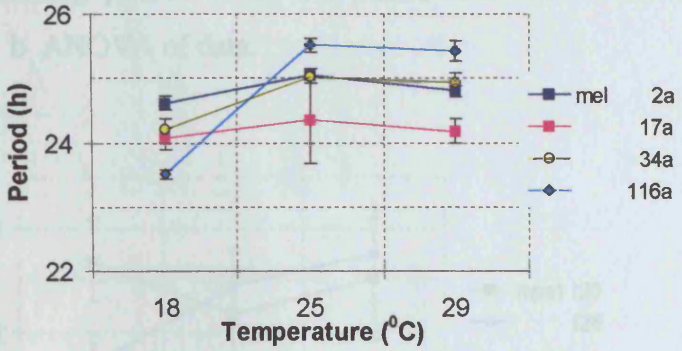
Statistical analysis of the intragenotype variation of the transgenes

mel

The four lines of the *D. melanogaster per* transgene (*mel)* rescue rhythms in a high percentage of *per* ^{01} hosts with periods in the circadian 24-25h range at all three temperatures (Table 3.1 and Fig 3.4). All lines appear to be temperature compensated across the three temperatures with the exception of line 116a between 18^{0}C and 25^{0}C. The results of a two way ANOVA (Fig 3.4) show there are highly significant line (L), temperature (T) and L * T interactions.

Fig 3.4 a. Free-running periods of the different transformant lines of the *mel* transgene. **b.** ANOVA table of data.

a.



b.

Source	MS	d.f.	F	p
Line (L)	6.976	3	7.075	.00012
Temperature(T)	28.340	2	28.742	.00000
L*T	7.767	6	7.877	.00000
Error	.986	375		

Newman-Keuls *post hoc* comparisons (Appendix 3.4) reveal that at 18^{0}C, lines 2a 17a and 34a are not significantly different, but line 116a is significantly different, to the other three ($p< 0.03$). At 25^{0}C, lines 2a, 17a, and 34a, are statistically inseparable. 116a is significantly different to line 17a ($p< 0.0001$). At 29^{0}C there is a detectable difference between 17a and 34a ($p< 0.04$) and a highly significant difference between 17a and 116a ($p< 0.00004$). Line 2a and 17a do not vary significantly across the temperature range. The period of line 116a increases between 18^{0}C and 25^{0}C ($p< 0.00002$) as does that of line 34a ($p< 0.02$)

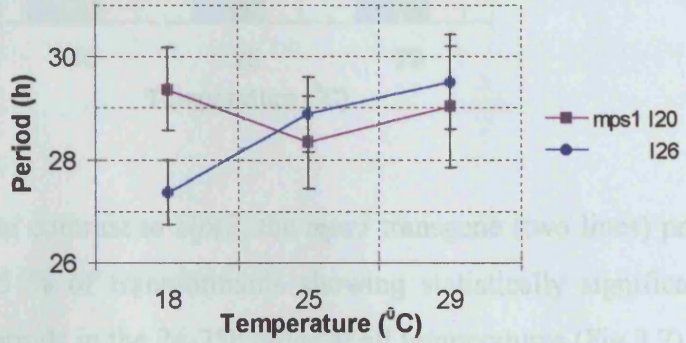
mps1

The *D. pseudoobscura per* coding region in *mps1* is not as effective in rescuing *per*^{01} arrhythmicity as the *D.melanogaster per* gene, yet approximately 50-75% of the transformants show significant rhythmicity at 18 and 25^{0}C (see Table 3.1 & Fig 3.2) with longer-than-normal periods of ~28h. These periods are stable, even at the highest temperature of 29^{0}C, when only 15% of the transformants are rhythmic (Table 3.1). Although visual inspection of the means of the two lines suggest they are different at

18^0C the results of a two way ANOVA show there are no significant Line (L) or Temperature (T) effects and no significant L * T interaction (Fig 3.5).

Fig 3.5 a. Free running periods of the two transformant lines of the *mps1* transgene. **b** ANOVA of data.

a.



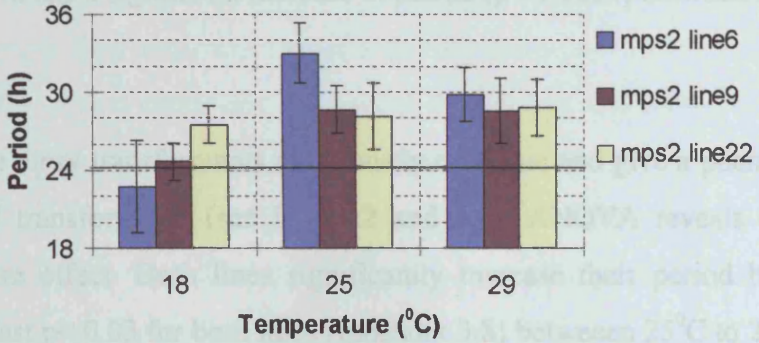
b.

Source	MS	d.f.	F	p
Line (L)	2.287	1	.188	.6654
Temperature(T)	5.221	2	.429	.6520
L*T	17.670	2	1.453	.2391
Error	12.153	88		

mps2

The *mps2* construct (three transformed lines) generated a very small percentage of weakly rhythmic flies, 15% across all three temperatures, and barely above the level of rhythmicity found for per^{01} mutants (Table 3.1). This rhythmicity was detected in spectral analyses and correlograms and was qualitatively much poorer than the analyses associated with *mps1* (see Fig 3.2). The low numbers of rhythmic individuals per line, their weak rhythmicity and the very large error associated with the estimation of their periods (see Table 3.1 and Fig 3.2), resulted in a decision to exclude them from statistical comparison with the other transformants presented here. The mean periods for the few rhythmic flies however are in the 28-29h range at 25^0C and 29^0C and about 24h at 18^0C (Fig 3.6).

Fig 3.6 Mean free-running periods (+/- sem) for the three *mps2* transgene lines at all temperatures.

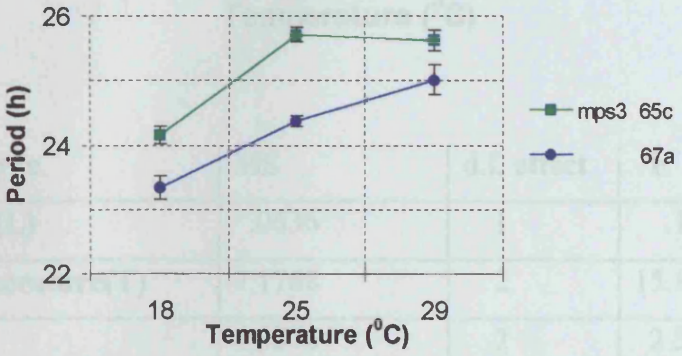


mps3

In complete contrast to *mps2*, the *mps3* transgene (two lines) produced excellent rescue with 75-95 % of transformants showing statistically significant rhythms with mean circadian periods in the 24-25h range at all temperatures (Fig 3.7). Both lines show an ~1.5h increase in period from 18°C to 29°C although line 65c seems more temperature compensated from 25°C to 29°C. The results of the ANOVA (Fig 3.7) reveal significant Line and Temperature effects, and significant L * T interaction.

Fig 3.7a. Mean free-running periods of the two *mps3* transgene lines. **b.** ANOVA of the data.

a.



b.

Source	MS	d.f.	F	p
Line (L)	21.92	1	51.684	.0000
Temperature (T)	22.36	2	52.720	.0000
L*T		2	3.645	.0289
Error		123		

Newman-Keuls *post hoc* comparisons (Appendix 3.7) confirm that the two lines are significantly different at all three temperatures ($p < 0.006$ at least). Across temperatures only line 67a has a significant increase in period ($p < 0.0059$) between 25^0C and 29^0C.

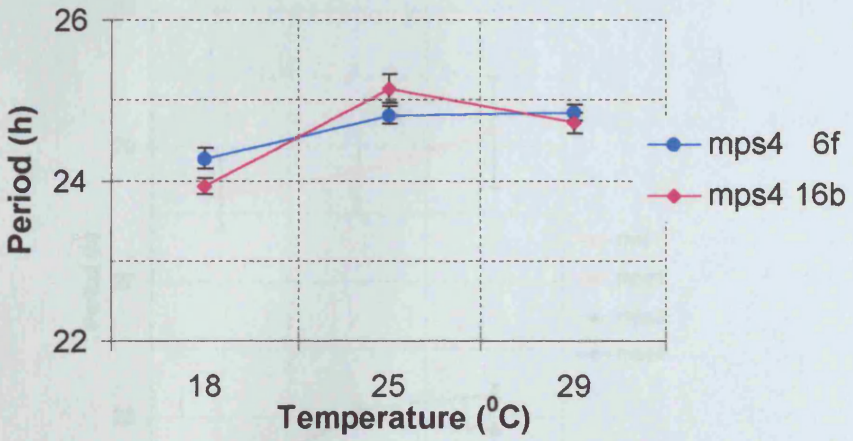
mps4

The *mps4* transformants are robustly rhythmic and give a phenotype very similar to the *mel* transformants (see Figs 3.2 and 3.4). ANOVA reveals only a significant Temperature effect. Both lines significantly increase their period between 18^0C and 25^0C (at least $p < 0.03$ for both lines Appendix 3.8) betweeen 25^0C to 29^0C no significant differences are seen.

Fig 3.8 a. Mean free-running periods of the two *mps4* transgene lines

b. ANOVA of data.

a.



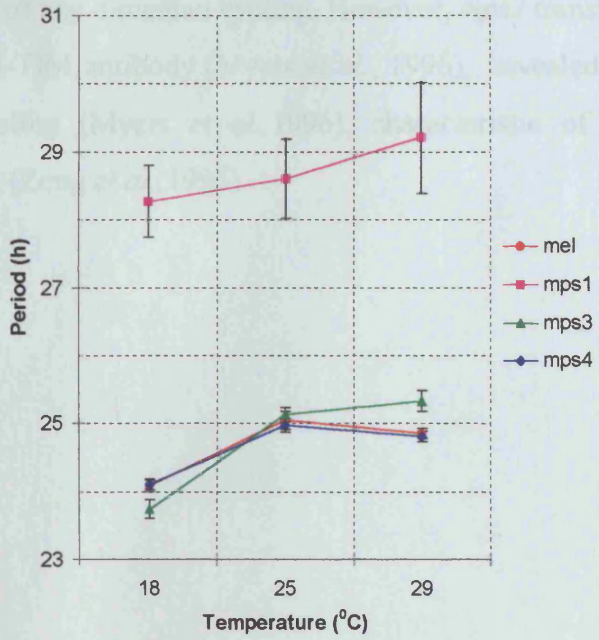
b.

Source	MS	d.f. effect	F	p
Line (L)	.0636	1	.140	.7081
Temperature(T)	7.1768	2	15.874	.0000
L*T	1.0576	2	2.339	.1000
Error	.4521	145		

Testing the effects betweeen genotypes

Although there were significant L * T interactions observed with *mel* , *mps3* and *mps4*, these were very small and in many cases the profiles within a genotype were very similar eg *mps4* (Fig 3.8). I therefore pooled the data within genotypes, given that a nested ANOVA could not be applied because there were different numbers of lines in each genotype. Fig 3.9 reveals significant Genotype (G) x Temperature (T) effects. Inspection of the results shows slight increases in period with temperature in all genotypes and large increases in period at all temperatures for *mps1* (see Appendix 3.9). There were no significant period differences at any of the temperatures between *mel*, *mps3* and *mps4* (see Appendix 3.9 a).

Fig 3.9 The free-running period of the transformants (lines pooled)



Source	MS	d.f.	F	p
Genotype (G)	414.680	3	182.246	.0000
Temperature(T)	49.205	2	21.625	.0000
G*T	3.002	6	1.319	.2458
Error	2.275	749		

Further pairwise comparisons of *mel/mps3* and *mel/mps4* were made by performing ANOVA with the pooled data. There were no significant genotype effects

between either pair of genotypes. There were minor Genotype x Temperature interactions in the *mel/mps3* comparison ($p< 0.021$) (see Appendix 3.9b). For the *mel/mps4* comparison no significant effects were seen (see Appendix 3.9c). In conclusion *mel*, *mps3* and *mps4* show almost identical profiles.

Western blots

Western blots were performed on *mps2* and *mps3* transformants (see Chapter 2 for details) Flies carrying one copy of the transgene were maintained in L.D.12:12 and the heads were harvested in 2h intervals throughout the circadian cycle. per^{+} (Canton-S) and per^{01} flies were used as controls. At ZT18 Fig 3.10 reveals that the *mps3* transformants show the characteristic wild-type cycle of PER abundance (Edery *et al.*,1994b), with a peak in intensity at ZT18-22 (Edery *et al.*,1994b). In contrast, the *mps2* PER product appears to be expressed at a lower level than the *mps3* product and with little evidence of any circadian cycling. However, *mps2* transformant head extracts, blotted with an anti-TIM antibody (Myers *et al.*, 1996), revealed a low amplitude (3X peak-to-trough) cycling (Myers *et al.*,1996), characteristic of per^{01} mutants under LD12:12 conditions (Zeng *et al.*,1996).

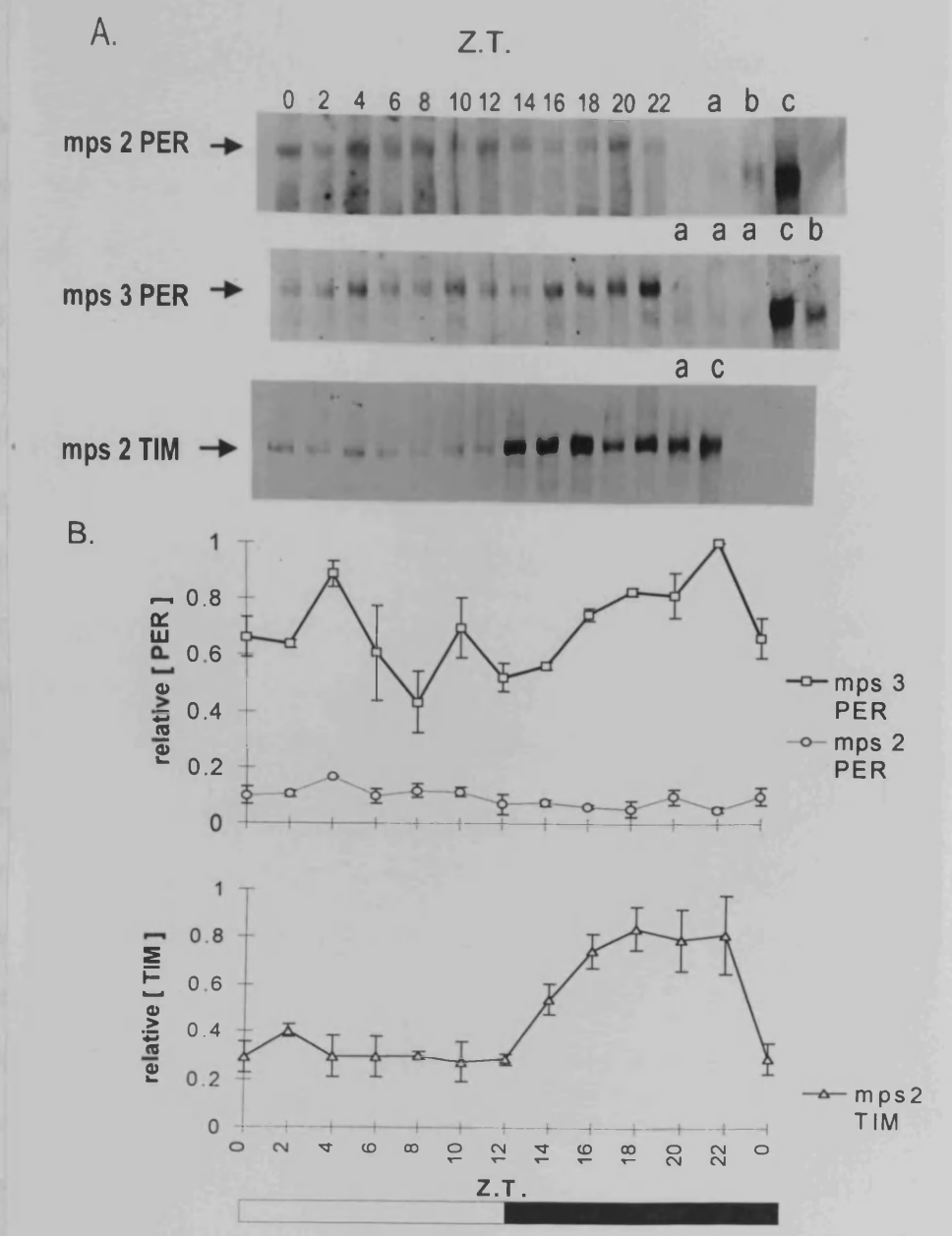


Fig 3.10A. Western blots of *mps2* and *mps3* transformant heads collected at different Zeitgeber times (ZT 0, lights on; ZT12, lights off) in LD 12:12 at 25^{0}C. Lanes on the right show *per* ^{01} (a), *tim* ^{01} (b) and *per* $^+$ Canton-S (c) controls at ZT 0, except for the TIM blot where the controls were taken at ZT 18. From top to bottom are shown : *mps2* blotted with anti-PER; *mps3* blotted with anti-PER; *mps2* blotted with anti-TIM. The predicted pimary M_r of the *mps2-5* PER product is 136,000 , heavier than either *D.melanogaster* (128,000)or *D.pseudoobscura* (132,000) proteins. However, *D.melanogaster* PER runs at about 180,000 (Edery *et al.*, 1984), and the chimaeric proteins run a little higher.

B. Densitometry analysis for the three separate anti-PER and anti-TIM blots. The highest intensity *PER* or TIM band within each blot was given a value of unity. Means +/- SEM are shown. Control experiments in which serial amounts of total protein were loaded for mps3PER, wild typePER and mps2PER at ZT0, followed by western blotting with anti-PER, revealed an approximate 5-7-fold reduction in mps2PER intensity for the same amount of total protein (data not shown). No characteristic wild type *PER* cycling (Edery *et al.*, 1984), was ever observed in *mps2 blots,* although all three blots showed the highest protein levels at ZT4.

Discusssion

The results presented in this chapter have revealed a remarkable difference in the function of the *mps2* and *mps3* transgenes. The *mps2* construct, in which more of the *melanogaster* coding sequence is present, generated a very small percentage of weakly rhythmic flies, barely above the level of rhythmicity found for per^{01} mutants (Table 3.1, Fig 3.2). This result was surprising as an initial assumption may be that any increase in the proportion of the *melanogaster* species-specific amino acid content of the protein within a *melanogaster* host should enhance *per* function. If this were the case *mps3* should give slightly worse rescue than *mps2*, given that *mps3* has more *D.pseudoobscura* sequence than *mps2* (see Table 3.1, Fig 3.6). In dramatic contrast, the *mps3* gene produced excellent rescue with 75-95 % of transformants showing statistically significant rhythms. The mean period of these rhythms are virtually indistinguishable from those produced by the control *mel* transformants at the three temperatures tested (Figs 3.4 and 3.7). Thus in *mps2* transformants, where the repeat of *D. pseudoobscura* lies directly adjacent to 5' flanking material of *D.melanogaster* (Fig 3.1), *per* gene function is severely disrupted (Table 3.1, Figs 3.2 and 3.6). In contrast, *mps3,* in which the repeat of *D.pseudoobscura* lies next to its conspecific 5' flanking sequences (blocks H and P in Fig 3.1), generates an ostensibly wild-type rescue of the per^{01} arrhythmic phenotype (Fig 3.7). This dramatic result strongly supports the idea of an intragenic coevolution between the length of the repeat and the immediate flanking upstream amino acids (~60) in blocks A and B (Peixoto *et al.*, 1993, Nielsen *et al.*, 1994). In fact without the evolutionary framework the results would be incomprehensible.

The *mps4* transgene essentially reproduced the *mps3* results, but seems to mimic the *mel* transformants behavioural phenotype even more faithfully, (Table 1, Appendix 3.8 and 3.9). The *mps3* and *mps4* results (Fig 3.7 and 3.8) therefore show that the *pseudoobscura* repeat plus its flanking sequences represents a 'coevolved' module which can function autonomously within an otherwise *melanogaster* PER protein.

Previous results with the same *D. pseudoobscura mps1* transformants (Petersen *et al.*, 1988), found fewer rhythmic individuals at 25^0C, but this was due to the use of a less sensitive statistical measure of rhythmicity (Dowse and Ringo, 1989; Kyriacou *et al.*,

1989). The examples show (Fig 3.2) that rhythms can be readily identified in these *mps1* flies. However the enhanced rescue of *mps3* over *mps1* in terms of period and temperature compensation reveals that the *melanogaster* N-terminal region is required for robust behavioural rhythms in addition to a functional/coevolved Thr-Gly region. This may reflect the fact that the N-terminal PAS dimerization domain of PER encoded in *mps3* is conspecific with the TIMELESS (TIM, Myers *et al.*, 1996) partner molecule of the hosts. Support for this view comes from observations that the *Musca domestica* PAS domain within PER is more similar to *D.melanogaster* than is *D. pseudoobscura* PAS (Piccin, 1988). This is surprising because the estimated divergence time for *D. melanogaster* and *M. domestica* is 100 Myr whereas for *D.melanogaster* and *D. pseudoobscura* it is 25 Myr (Hennig, 1981, Russo *et al.*, 1994). Interestingly the *Musca per* gene rescues *melanogaster* per^{01} arryhmicity to wild type levels in contrast to *D.pseudoobscura mps1*, suggesting that the strength of rescue in D.melanogaster hosts is correlated with the interspecific PAS similarity. By the same argument the *D.melanogaster* PAS domain in *mps3* may interact with the *mel* host TIM (as in *M. domestica*) more efficiently than the more diverged *D.pseudoobscura* PAS encoded in *mps1*.

The sequences flanking the repeat in blocks H and P (Fig 3.1) contain 27 amino acid substitutions between the two parental species, most of which are distributed in block H. Comparing this flanking region to sequences in the databases did not reveal any significant similarities to motifs which might illuminate the function of this region. *D.pseudoobscura per* does however encode possible PEST sequences (Rogers *et al.*,1986; Dembinska *et al.*, 1997) within the body of the repeat, which are characteristic of proteins which have a rapid turnover, and recent work with PER-reporter fusions has suggested that a region of PER encoded by an ~700bp fragment which contains the *melanogaster* Thr-Gly region, may be a target for degradation (Dembinska *et al.*, 1997).

While not explaining the results presented, any alteration in PER stability might be expected to affect the negative feedback of PER protein on its own mRNA (Hardin *et al.*, 1990, Goldbeter *et al.*, 1995) and so alter some aspect of it's cycling. In the western blots (Fig 3.10a and b), the *mps3* transformants show PER protein cycling with low levels at

ZT 8 (late day) and a peak intensity at ZT 22 (late night) which is comparable in phase to the cycling seen in wild type PER (Edery *et al.*, 1994b). In contrast, mps2 PER is expressed at a level 5-7 times lower than that of mps3 PER (Fig 3.10a and b), with little evidence of any circadian cycling, This failure of mps2 PER to cycle may suggest an altered stability of the chimaeric protein or some other factor which alters the kinetics of its negative feedback loop, such as degradation (Dembinska *et al.*, 1997). Perhaps the interspecific coevolutionary dynamics uncovered between the repeats and their conspecific flanking regions represent the signature of natural selection as it attempts to maintain the appropriate degradation kinetics for PER in the presence of a relatively high mutation rate in the repetitive region (Rosato *et al.*, 1997).

It is not clear whether mutations in the flanking region may have imposed selection pressure on the length of the repeat, or *vice-versa,* but because mutations in the repeat and flanking region are tightly linked, they conform to models where compensatory neutral mutations can become fixed even when the individual mutations are deleterious (Kimura, 1991). This process would be enhanced in a region associated with a high mutation rate (Rosato *et al.* 1997), and if the domain determined an adaptive character.

It is assumed that TIM and mps2 PER will interact (Hunter-Ensor *et al.*, 1996, Zeng *et al.*, 1996) via the mps2 PAS domain (Fig 3.1). However, *mps2* transformant head extracts, blotted with an anti-TIM antibody (Myers *et al.*, 1996), revealed a low amplitude cycling which is characteristic of per^{01} mutants under LD12:12 conditions (Zeng *et al.*, 1996). The mps2 PER is therefore either not binding to TIM or is binding inappropriately and affecting normal TIM cycling.

A further construct *mps5* (see Fig 3.1) has been made by Ian Townson at Leicester in which the 5' conspecific flanking region H and P of *mps3* has been changed so that it has been made heterospecific (H-region from *D.melanogaster* and P-region from *D.pseudoobscura*). The *mps3 pseudoobscura* content has thus been reduced, and as with the change from *mps1* to *mps2*, the change from *mps3* to *mps5* increases the species-specific amino acid content. Again contrary to the simplistic view that an increase in the *mel* content of the coding sequence would improve interactions *in vivo*, significant and informative phenotypic changes are seen. The *mps5* transformants produce a phenotype

intermediate between *mps2* and *mps3,* in that locomotor rhythms are statistically robust, but the periods are extremely temperature-sensitive, 24, 28 and 36h at 18^0C 25^0C and 29^0C respectively (Peixoto *et al.*, 1998). This indicates that block H may be directly involved in controlling temperature compensation, with block P more intimately involved with basic rhythmicity. This idea is of course very speculative and would require a reciprocal transgenic control of *mps5*, with a *pseudoobscura* block H and a *melanogaster* block P to check its validity. The *a priori* hypothesis would be that the construct would be temperature compensated like *mps3* (Fig 3.6) but would have poor rhythmicity like *mps2*.

The *mps5* data suggests that the unalignable amino acids in block H (Fig 3.2) also play a role in the coevolutionary interaction between block P and the repeat, and also implies that these residues are under positive natural selection rather than drift. Coevolution within molecules are rarely studied, some notable examples being the *Drosophila rRNA* genes, where compensatory mutations appear to maintain the secondary structure of the molecules (Hancock *et al.*, 1990; Rousset *et al.*,1991), and the *Drosophila Adh* gene where compensatory mutations appear to maintain interactions between the 5' and 3' regions of the transcript (Parsch *et al.*, 1997). Overall, these results suggest that the coevolutionary interaction between the repeat and its flanking sequences may serve a function related to the clock's cardinal property of temperature compensation (Pittendrigh 1954).

Like the *mps1, 2* and *5* transformants, the poorly temperature compensated per^{L1} mutant also shows significant increases in period with higher temperatures (Konopka *et al.*, 1989) and has been used as a model in most studies describing temperature compensation (Curtin *et al.*,1995; Gekakis *et al.*, 1995; Rutila *et al.*, 1996). The mutation maps to the PAS region of the PER protein which is involved in protein-protein interactions (Huang *et al.*, 1993), and which may mediate the negative feedback mechanism believed to be fundamental to *per* function (Hardin *et al.*,1990). Nuclear translocation of PER requires a functional *tim* gene (Vosshall., 1994), and a model of intramolecular interactions involving PAS and a domain C-terminal to PAS, and intermolecular dimerisation between PER and TIM has been proposed to underlie

temperature compensation (Huang *et al.*, 1995). The C-domain identified in these studies, includes PER amino acids 524-685, which overlaps with blocks H and P (aa 639-695 Fig 3.1).

This work (Huang *et al.*, 1995) seemed to offer supporting evidence from the yeast two hybrid system, for a link between the Thr-Gly region and temperature compensation effects shown in this chapter. The model proposes that PER intramolecular attractions sequester PER in monomeric form which reduces the amount of PER available for dimer formation with TIM, for nuclear entry (see Chapter 1). The PER-L fragment (aa 233-390) is shown to have a stronger self-affinity with the PER C-domain (aa 524-685, which overlaps the H P region [aa 639-695]), than the corresponding wild type fragment. The model has however become neglected since Gekakis *et al.* (1995), using yeast assays, demonstrated temperature sensitivity in TIM/PER ^{L1} interactions (with decreased affinity at higher temperatures) for the PAS only fragment (aa 233-390) and PER ^{L1} PAS to C-domain fragment (aa 233-685). PER and TIM interaction is needed for nuclear entry (Vosshall *et al.*, 1994) and PER^{L} has been shown to have delayed nuclear entry (Curtin *et al.*, 1995). The Gekakis *et al.* results therefore suggest that the temperature insensitivity of wild type PER/TIM interaction is the basis of temperature compensation and that the C-domain is not required. Taken with the Vosshall *et al.* and Curtin *et al.* results, it suggests that poor TIM/PER ^{L1} interaction is the basis of the long period and poor temperature compensation phenotypes.

The hypothesis that temperature compensation is a function of the strength of binding of PER and TIM (Gekakis *et al.*, 1995), has however been shown to be too simplistic. While disruption of the PER-TIM complex in the cytoplasm delays PER phosphorylation and nuclear entry (Lee *et al.*, 1996), the mechanism of temperature compensation in the presence of a functional complex is more complicated (Rutila *et al.*, 1996). Investigation of *tim* ^{SL}, the suppressor of *per* ^{L} (Rutila *et al.*, 1996), has in yeast studies revealed no improvement in the binding efficiency of PER ^{L} /TIM ^{SL} in comparison with PER ^{L}/TIM at 30^{0}C, even though the period of *per* ^{L}/*tim* ^{SL} double mutants is 5h shorter than *per* ^{L} at this temperature (Rutila *et al.*, 1996). This result suggests that even though the binding of the complex is still weak, temperature compensation is robust. In

addition Gekakis *et al.* (1995) also demonstrated that the TIM/PER ^{L1} interaction for the aa 233-685 PER ^{L1} fragment was stronger at 22^{0}C than the corresponding interaction between TIM and wild type PER (at 22^{0}C, *per* ^{L} has a period of 27h). The simple prediction that low affinity PER-TIM binding will lead to period lengthening (Gekakis *et al.*,1995), is therefore not necessarily true. A hyperphosphorylated state of the TIM ^{SL} protein seems to compensate for an initial lack of *per*^{L} phosphorylation (Rutila *et al.*, 1996). Although reinforcing the importance of the PER/TIM complex in temperature compensation, this suggests that the physical interaction of the proteins is secondary to other important molecular events involved in temperature compensation.

To summarize, the temperature compensation mechanism may be disrupted at several different sequential levels: initially, as shown by elements of the Huang *et al.* model (1995), intramolecular effects may slow down complex formation, secondly binding efficiency is changed with temperature (Gekakis *et al.*, 1995), thirdly postranslational modifications of the complex may alter PER and TIM dynamics (Rutila *et al.*, 1996). It is conceivable that delay in functional PER/TIM dimer formation required for nuclear entry (which would cause free-running period increases) in response to temperature challenge, could occur at any or all of these levels. The molecular basis for the poor *mps 2* and *5* temperature compensation, may therefore act at a different level to that of *per* ^{L}. The picture is certainly too complicated to assume that failure of temperature compensation is directly proportional to defective binding of PER to TIM (see Chapter 5 for similar argument regarding temperature compensation and Thr-Gly deletions).

In terms of the importance of the Thr-Gly region, examples of clock genes with similar regions of homology are now available across the animal and plant kingdoms. The *frequency (frq)* gene in *Neurospora crassa*, which can be mutated to alter free running period also includes a run of Thr-Gly and Ser-Gly pairs (McClung *et al.*, 1989). It may be more than coincidence that two mutations, *frq3* and *frq7*, which disrupt the temperature compensation of the circadian period (Gardner and Feldman, 1981), have been mapped to within a 40 amino acid region either side of the repetitive domain (Aronson *et al.*, 1994b). Perhaps these repetitive regions within the two clock genes serve a common function

related to temperature compensation. Further emphasis on the importance of the repeat and asociated flanking regions will come from the recently identified mouse *per* homologues *mper1* and *mper2* (Tei *et al.*, 1997, Shearman *et al.*, 1997). Sequence comparisons for the Thr-Gly region are quite striking with *mper1* having a 37% similarity to *D. melanogaster per* while *mper1* and *mper2* have 58% similarity to each other (Shearman *et al.*,1997).

In conclusion, the most remarkable feature of the results is that the interaction between subregions of the Thr-Gly 'domain', was predicted by relatively simple statistical analyses of the relevant DNA sequences from a number of Dipteran species (Peixoto *et al.*, 1993; Nielsen *et al.*, 1994). As can be seen, a slight difference in the position of the chimaeric junction can produce major phenotypic effects. The repeat and the flanking regions appear to interact to maintain the clock's temperature compensation, and this is further developed in Chapter 5.

Chapter 4 The *per* gene and species-specific behaviour

4.1 Introduction

The *per* locus has previously been associated with certain patterns of species-specific behavioural activity (Petersen *et al.*, 1988; Wheeler *et al.*, 1991). Wheeler *et al.* (1991) convincingly demonstrated species-specific transfer of the 40s cycle of the male *D. simulans* courtship song (Kyriacou and Hall, 1986) to *D. melanogaster* hosts using reciprocal chimaeric *per* gene constructs. These species-specific courtship song cycles mapped to a ~600 base fragment spanning the Thr-Gly repeat. Because of the similarity of *per* sequence in these two closely related species, these differences could be further mapped to interspecific substitutions in the Thr-Gly repeat and/or to a small number of amino acid changes in the immediate flanking sequences.

Petersen *et al.* (1988) reported that *D.pseudoobscura* had shorter and 'stronger' free running rhythms than *D.melanogaster* and that the primary amino acid sequence of the *D.pseudoobscura per* gene is able to transfer this species-specific behaviour on transformation into per^{+} hosts. Essentially the timing and relative amounts of peak locomotor activity during the day was altered between the two species. [Differences in patterns of locomotor actvity have also been noted for a variety of *per* mutations, for example per^{S} and per^{L}, where evening peak activity occurs earlier and later respectively in LD 12:12 conditions (Hamblen Coyle *et al.*, 1992)]. Because of these demonstrated species-specific influences in behaviour which may have implications for reproductive isolation, the *per* locus has been proposed to be a 'speciation gene' (Coyne, 1992).

Initially I chose to reinvestigate the reported free-running locomotor activity differences in *D.pseudoobscura* and *D.melanogaster* (Petersen *et al.*, 1988) free from the influence of a zeitgeber, so that any inherent species-specific activity might be more visible. The problem with this approach is that flies with different free running periods (normal for any population in free-run) are difficult to compare. To solve this problem I developed a novel technique to standardize activity patterns between individual flies (see Materials and methods 4.2). A population of flies could then be analysed by superimposing all their activity patterns to generate an average profile. As well as attempting to confirm Petersen *et al.'s* results, an additional question was whether the two species would show similar activity profile responses to temperature?

Therefore behavioural analysis in DD was undertaken with *D.melanogaster* and *D.pseudoobscura* at a range of temperatures. The *mps1* (all *per* coding sequence from *D.pseudoobscura*) and *mps3* transformants (3' half of *per* including the Thr-Gly region from *D.pseudoobscura*) were then analysed in DD to map any species-specific activity to sub-regions of the *per* gene. After identification of differences in DD, the analysis was extended to LD. For the LD analysis the *mps4* transformant (essentially *D.melanogaster per* with just the Thr-Gly region of *D.pseudoobscura*) was included.

4.2 Methods

Lines and strains

The chimaeric *D.melanogaster/ D.pseudoobscura* transformants *mps1*, *3*, *4*, and the *13.2 D.melanogaster* transformants (lines 2a and 17a) were used, and are described in Chapter 2 and 3. per^{+} Brandeis (Br) [a Canton S wild type strain used to isolate the $(Thr\text{-}Gly)_{20}$ carrying 13.2 kb fragment] (Citri *et al.*, 1987) and *17m* a $(Thr\text{-}Gly)_{17}$ homozygous line isolated in Northern Italy were also used. Two strains of *D.pseudoobscura* used were, Ayala (Ay) [the strain used to isolate the *D.pseudoobscura per* gene fragment (Petersen *et al.*, 1988) and Pachuca (Pa) (see Colot *et al.*, 1988). The data for the *D.melanogaster* wild type controls in LD 12:12 was generated by a colleague L.Sawyer, using two lines isolated from Rethimnon in Greece, Ret.4 and Ret.9, which represent $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ variants respectively. Prior to activity analysis, transgenes were placed on a per^{01} arrhythmic background, by crossing male transformants to per^{01} virgin females. The F1 were collected 1-7 days after eclosion.

Activity analysis

The flies were entrained in LD 12:12 at 25°C, followed by activity analysis in either LD 12:12 at 18°C or DD at 18°C, 25°C and 29°C. In both cases the flies were given a complete circadian cycle in the new conditions before activity recording began. Recording lasted for between 7 and 10 days. For both LD and DD recordings the data was processed by 'wrapping' (also see Chapter 2.5) to achieve an average mean activity pattern for each fly.

For LD data, a mean activity cycle for each of 48, 30 min bins across 7-10 days was calculated by superimposing the data for each of the days. This results in an average mean activity per bin for each fly. The genotype means were then calculated per bin and for statistical analysis the 48 bins were subsequently divided into twelve 2h windows,

starting from ZT 18. A mean activity for each 2h window was calculated. This had to be done simply because the ANOVA statistics package (STATISTICA) could not run with the amount of data carried in the 48 bins.

For DD data, the activity is wrapped using the spectrally derived period for each individual fly. This means that the range of periodicities observed in free run, will also result in patterns representing the range of circadian values. To compare these patterns the value of the X-axis was standardized as follows: For each fly the number of sample points (30 min bins of activity) that are closest to the circadian period (for example periods of 24.2 and 24.3 h are assigned 48 and 49 bins respectively) are connected by a computer program. Standardized activity patterns are then created by sampling the Y-value of the regression lines at 48 equally spaced X-axis values. The activity for each of the newly generated 48 bins for a population of free-running flies can then be averaged either using raw activity levels, or proportions of total activity where the value per bin is divided by the sum of the 48 bins and transformed into arcsine. Given the nature of these data manipulations ANOVA on the complete data set is inappropriate as the data are interdependent, so to examine intergenotype activity differences, the bin numbers carrying the peak activity values for each individual were compared.

For historical reasons data collection started at CT30 after the last lights-off signal and so a correction factor was needed for any fly with a period departing from 24 hours. To do this the assumption was made that the underlying process of speeding up or slowing down of the clock in DD would be effective immediately from the lights-off transition and would proceed at a constant level throughout the circadian cycle. To correct for the amount of time the fly had already been in free-run the deviation of the spectrally derived period from 24h was multiplied by 18/24 (the fraction of a circadian day the fly had already been in free-running conditions) and the value obtained used to advance or retard the individual activity patterns accordingly. In other words , the locomotor profile was readjusted to CT 12 (lights off). To find the 'peak bin' the highest value was then read manually from a corrected 'Microsoft Excel' graph of the data representing each fly.

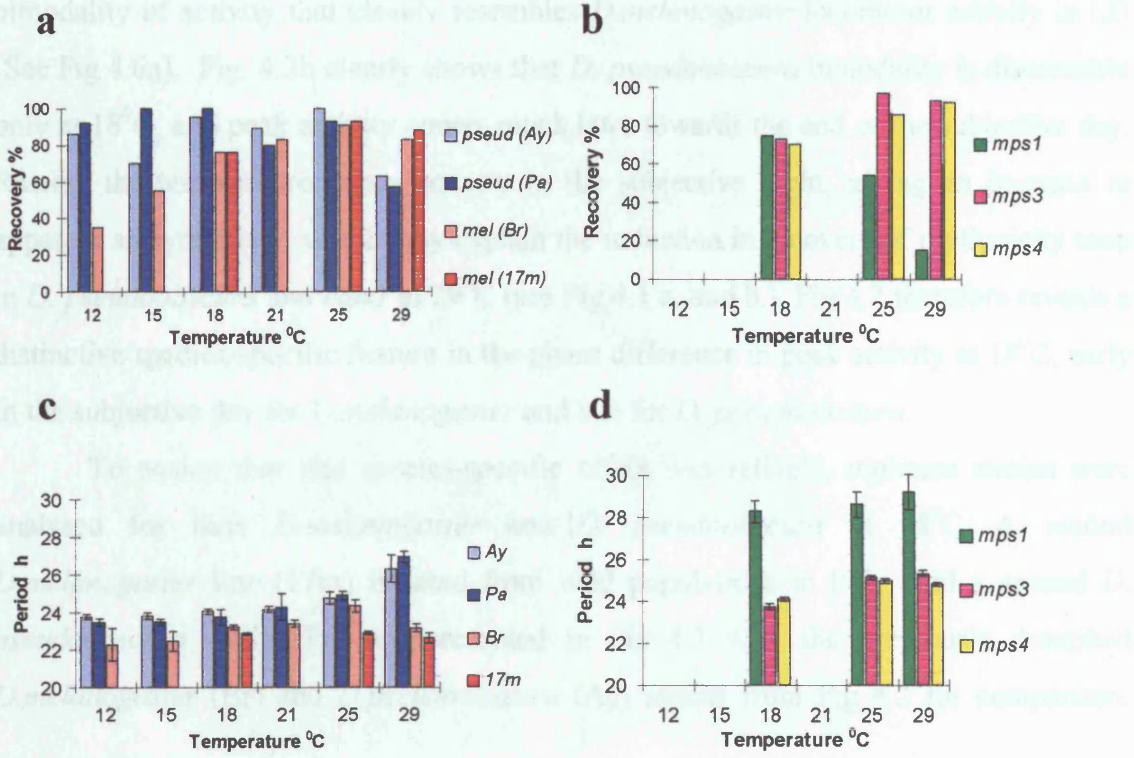
For DD analysis, only rhythmic flies could be used as a spectrally derived period was needed to wrap the data for analysis. Flies analysed in LD were given no preselection other than for hyper or hypo activity and early death.

4.3 Results

DD analysis

Initial investigation of *D. pseudoobscura* (Ay and Pa) and *D.melanogaster* (Br and 17m) free-running locomotor activity profiles indicated different, and possibly species-specific, responses to a range of temperature challenges. Fig 4.1a and b (data is in Appendix 4.1) shows the variation in the recovery of rhythmicity and the free-running period of two strains each of *D.pseudoobscura*, *D.melanogaster* and the chimaeric *D.melanogaster/D.pseudoobscura* transformants. Only one of the *D.melanogaster* strains (Br) has been tested at all temperatures and the data for the chimaeric transformants is reproduced from Chapter 3. Fig 4.1a clearly shows that the recovery of rhythmicity for the two *D.pseudoobscura* strains is very good (~70-90%) at temperatures as low as 12^0C but drops considerably for both lines at 29^0C (~60%).

Fig 4.1 a. and **c.** Recovery and free-running periods of *D.pseudoobscura* (strains Ay and PA) and *D.melanogaster* (strains Br and 17m). **b.** and **d.** Recovery and free-running periods of *mps1*, *mps3* and *mps4* in DD. [see Appendix 4.1 and Fig. 3.1 for data]



D.melanogaster on the other hand appears to be quite the reverse with low recovery at 12-15^{0}C (~30-55%), but very good recovery at 29^{0}C (~80-90%). These low temperature results must be treated with some caution as only one strain of *D.melanogaster* was tested at 12^{0}C and 15^{0}C. Results for the transformants tested at 18^{0}C, 25^{0}C and 29^{0}C (see Fig 4.1b) reveal that the *D.pseudoobscura mps1* transgene has reduced levels of recovery from 18-29^{0}C but the chimaeric transgenes give the '*melanogaster*' pattern of recovery, increasing over this temperature range. This '*D.pseudoobscura*' phenotype (low recovery of rhythmicity at 29^{0}C) is also accompanied by a raised free-running period for both the *D.pseudoobscura* strains and two lines of the *mps1* transformant (See Fig 4.1c and d respectively).

In order to investigate these observed differences in more detail the activity profiles (overall activity throughout the subjective day) of *D.pseudoobscura* (Ay) and *D.melanogaster* (Br) and their respective transformants (*mps1* line I20 and *13.2* line 2a) were examined at a range of temperatures. Fig 4.2a shows that at 18^{0}C, *D.melanogaster* and the *D.melanogaster per* transformant show an early major peak of activity close to subjective lights on, between the adjusted CT0 and CT6. When the temperature is raised to 25^{0}C the major peak moves towards subjective lights off. At 29^{0}C, the bimodality apparent at 25^{0}C becomes more pronounced with a decrease in activity in the middle of the subjective day. This activity phase shift is associated with the development of a bimodality of activity that closely resembles *D.melanogaster* locomotor activity in LD (See Fig 4.6a). Fig. 4.2b clearly shows that *D. pseudoobscura* bimodality is discernable only at 18^{0}C, and peak activity comes much later towards the end of the subjective day. Raising the temperature raises activity in the subjective night, giving an increase in apparent arrhythmicity, which may explain the reduction in recovery of rhythmicity seen in *D. pseudoobscura* and *mps1* at 29^{0}C (see Fig 4.1 a. and b.). Fig 4.2 therefore reveals a distinctive species-specific feature in the phase difference in peak activity at 18^{0}C, early in the subjective day for *D.melanogaster* and late for *D. pseudoobscura.*

To ensure that this species-specific effect was reliable, replicate strains were analysed for both *D.melanogaster* and *D. pseudoobscura* at 18^{0}C. A second *D.melanogaster* line (17m) isolated from wild populations in Italy, and a second *D. pseudoobscura* strain (Pa) are presented in Fig 4.3 with the previously described *D.melanogaster* (Br) and *D.pseudoobscura* (Ay) strains from Fig 4.2 for comparison.

Both *D.melanogaster* lines show early peaks and both *D.pseudoobscura* strains show late peaks, confirming the species specific patterns at this temperature.

Fig 4.2 a. *D.melanogaster* (Br) and *13.2 (line* 2a) transformant free-running locomotor activity (Y-axis = transformed proportions, X-axis = adjusted bin number) at 18^{0}C, 25^{0}C and 29^{0}C in DD. **b.** *D.pseudoobscura* (Ay) and *mps1* (line I20) transformant. [beginning of bin 13 and 37 = CT0 and CT12 respectively]

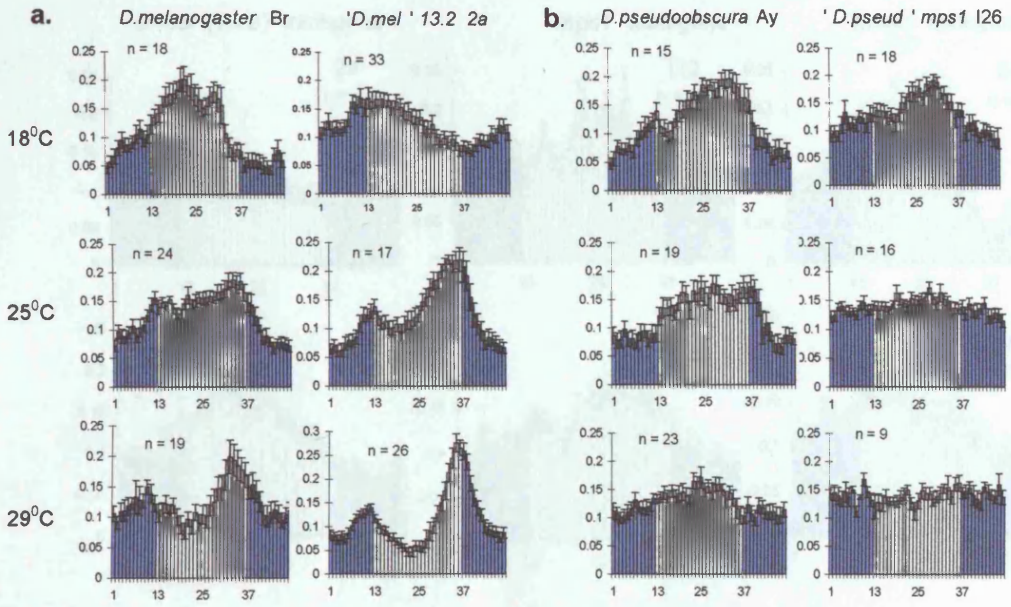
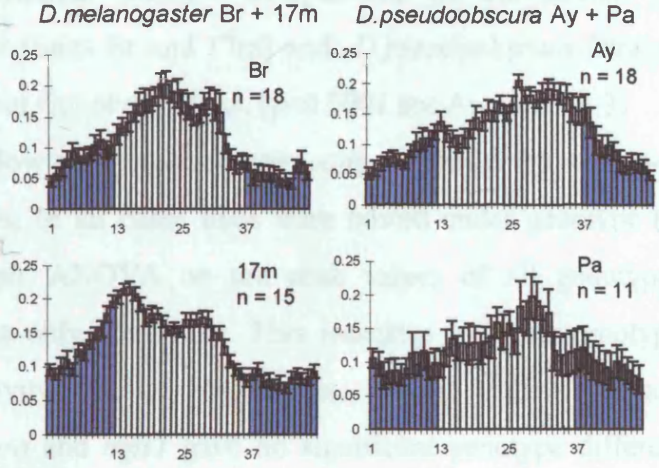
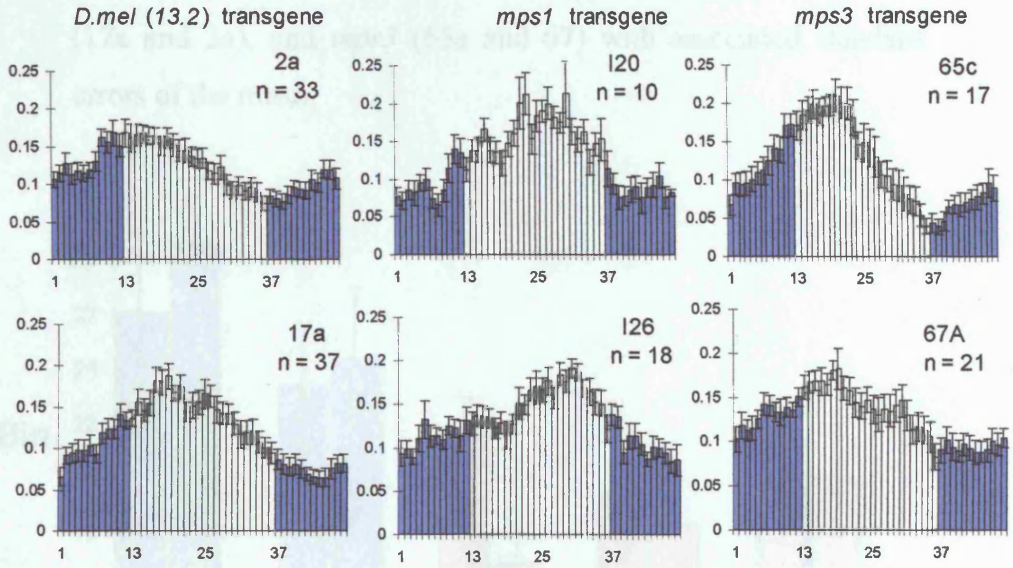


Fig 4.3 Locomotor activity (transformed proportions, Y-axis) at 18^{0}C in DD for *D.melanogaster* (lines Br and 17m) and *D.pseudoobscura* (strains Ay and Pa).



The next step involved comparison of *mps1* with *13.2* and *mps3* transformants to see whether *mps3* gave a *D.melanogaster or D.pseudoobscura* profile. Inspection of the activity profiles presented in Fig 4.4 indicates that the *13.2* (2a and 17a) and *mps3* (65c and 67a) transformants have an early morning '*melanogaster*' peak while the *mps1* (I20 and I26) has a late afternoon '*pseudoobscura*' peak.

Fig 4.4 Free-running locomotor activity (transformed proportions) at 18^0C in DD for*13.2* (17a and 2a), *mps1*(I20 and I26) and *mps3* (65c and 67a) transformants

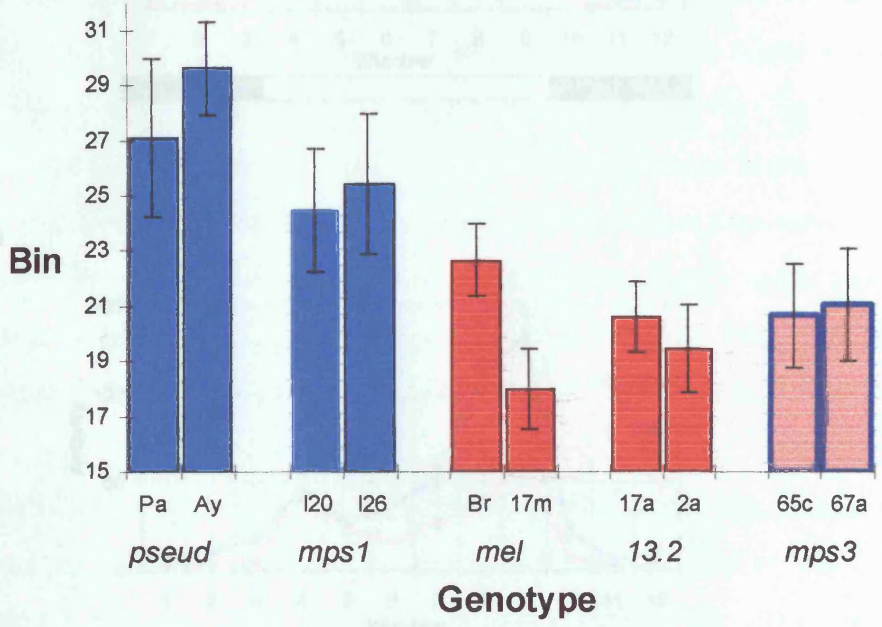


To test this observation statistically the bins carrying the peak values of all the individual fly data presented in Fig 4.3 and Fig 4.4 were calculated and one way ANOVA performed on the data. The mean peak values and standard errors for each line are presented in Fig 4.5. Although there is some heterogeneity in the peak activity of the *melanogaster* strains, planned comparisons of the pooled peak bin values for *D.melanogaster* (lines Br and 17m) and *D.pseudoobscura* (strains Ay and Pa), gave a highly significant Genotype effect. ($p<0.0001$ see Appendix 4.3)

The following planned comparisons confirmed the species-specific differences in peak bin values. In all cases lines were nested under genotype (see Appendix 4.4 for ANOVA tables). ANOVA on the peak values of all genotypes gave a significant genotype effect only ($p<0.021$). This indicates that the genotypes have significantly different activity peaks (Appendix 4.4a). Pairwise comparisons between *D.pseudoobscura* and *mps1* gave no significant genotype differences ($p<0.13$) or line

effect ($p<0.73$) (Appendix 4.4b). Comparison between *D.melanogaster*, *13.2* and *mps3* also gave no significant genotype differences ($p< 0.90$) or line effect ($p<0.34$) (Appendix 4.4c). However, direct comparison of *mps1* and *13.2* gave a significant genotype effect ($p<0.028$) with no line effect ($p<0.84$). The conclusion is therefore that *mps3* has a *D.melanogaster* like activity for this particular aspect of behaviour, suggesting that in DD the N-terminal half of the protein controls this species-specific behaviour pattern.

Fig 4.5 Bin number containing peak activity for *D.pseudoobscura* (Ay and Pa), *mps1* (I20 and I26), *D.melanogaster* (Br and 17m), *13.2* (17a and 2a), and *mps3* (65a and 67) with associated standard errors of the mean.



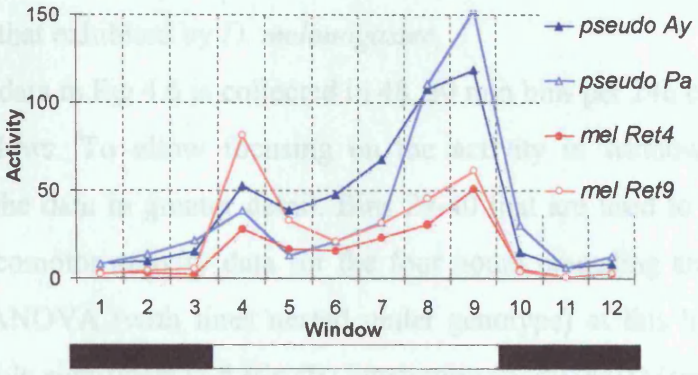
The above results were obtained in DD where the flies exhibit rhythms without the imposition of an external zeitgeber. Would species-specific behaviour also be visible when driven by an LD cycle?

D.melanogaster* and *D.pseudoobscura (LD at 18^0C)

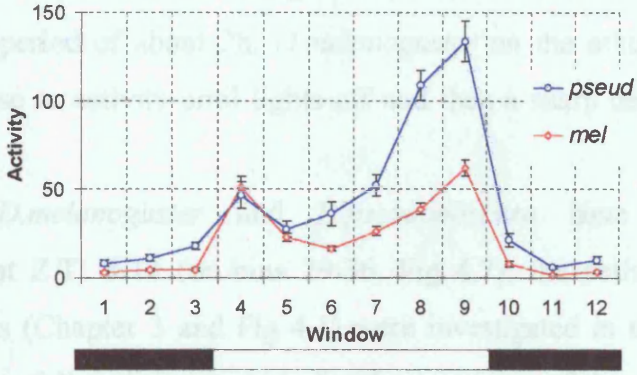
I chose to perform the LD experiments at 18^0C as in the latter DD experiments. The activity profiles are initially plotted with average activity against 12, 2h windows. This format gave the maximum amount of data that could be analysed by ANOVA with the statistical package.

Fig 4.6 a. Locomotor activity at 18^{0}C in LD 12:12 of *D. pseudoobscura* (AY and PA) and *D. melanogaster* (Ret.4 and Ret.9). **b.** Strains pooled into genotypes, (+/-sem). **c.** ANOVA table of *D.pseudoobscura* versus *D.melanogaster* (lines and strains as random factors nested under Genotype)

a



b



c

melanogaster (Ret 4 + Ret 9) v *pseudoobscura* (Ay + Pa)						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	118835.7	2	6813.273	17.44179	0.05283
2	2	6813.273	1188	1187.5	5.737493	0.00331
3	11	88954.86	22	3996.99	22.25546	0.00000
12	--	--	--	--	--	--
13	11	18663	22	3996.99	4.669263	0.00104
23	22	3996.99	1188	1187.5	3.365886	0.00000
123	--	--	--	--	--	--

Fig 4.6a shows two replicate lines of *D.melanogaster* (Ret.4 and Ret.9) and the two strains of *D.pseudoobscura* (Ay and Pa) used previously in DD. Fig 4.6b shows lines pooled into genotypes. ANOVA (with lines nested within genotype) indicates significant Genotype (G) * Window (W) interactions ($p<0.001$ Fig 4.6c). The critical species-specific activity causing the interaction appears to originate from activity variation in windows 8 and 9 where the activity of the *D.pseudoobscura* lines appears to be 2-3 times higher than that exhibited by *D. melanogaster*.

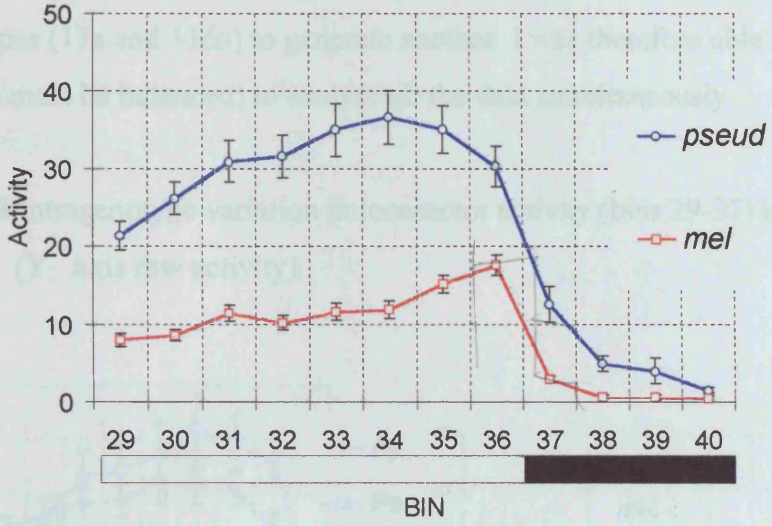
The data in Fig 4.6 is collected in 48, 30 min bins per 24h cycle but is analysed as 12 2h windows. To allow focusing on the activity in windows 8-10 of Fig 4.6, I reanalysed the data in greater detail. Bins 29-40 that are used to create windows 8-10, have the locomotor activity data for the four hours preceding and the two hours after lights off. ANOVA (with lines nested under genotype) at this higher resolution again shows a highly significant G * Bin (B) interaction ($p<<0.0001$) (see Fig 4.7)

As in Fig 4.6, *D.pseudoobscura* is 2-3 times more active in peak activity values. However this higher resolution analysis reveals another effect; *D.pseudoobscura* activity peaks in bin 34 (an hour before lights off) after which activity slowly declines to near zero over a period of about 2h. *D.melanogaster* on the other hand appears to have a linear increase in activity until lights off and then a sharp decrease to zero in the dark phase.

As *D.melanogaster* and *D.pseudoobscura* have significant behavioural differences at Z.T. 8-12 (or bins 29-36, Fig 4.7), the activity of the chimaeric *per* transformants (Chapter 3 and Fig 4.1) were investigated in the same way. Transgenes containing the full coding sequence for the *per* genes of the two parental species (*13.2* and *mps1*), and two *D.melanogaster/D.pseudoobscura* chimaeric transgenes (*mps3* and *mps4*, see Chapter 3 for details), were analysed to see if the species-specific effects shown in LD could first be mapped to the *per* gene and second to subregions within the gene.

Fig 4.7 **a.** Higher resolution of *D.pseudoobscura* and *D.melanogaster* locomotor activity in LD 12:12 from Fig 4.6 (Bins 29-40 = windows 8-10). **b.** ANOVA table

a.



b.

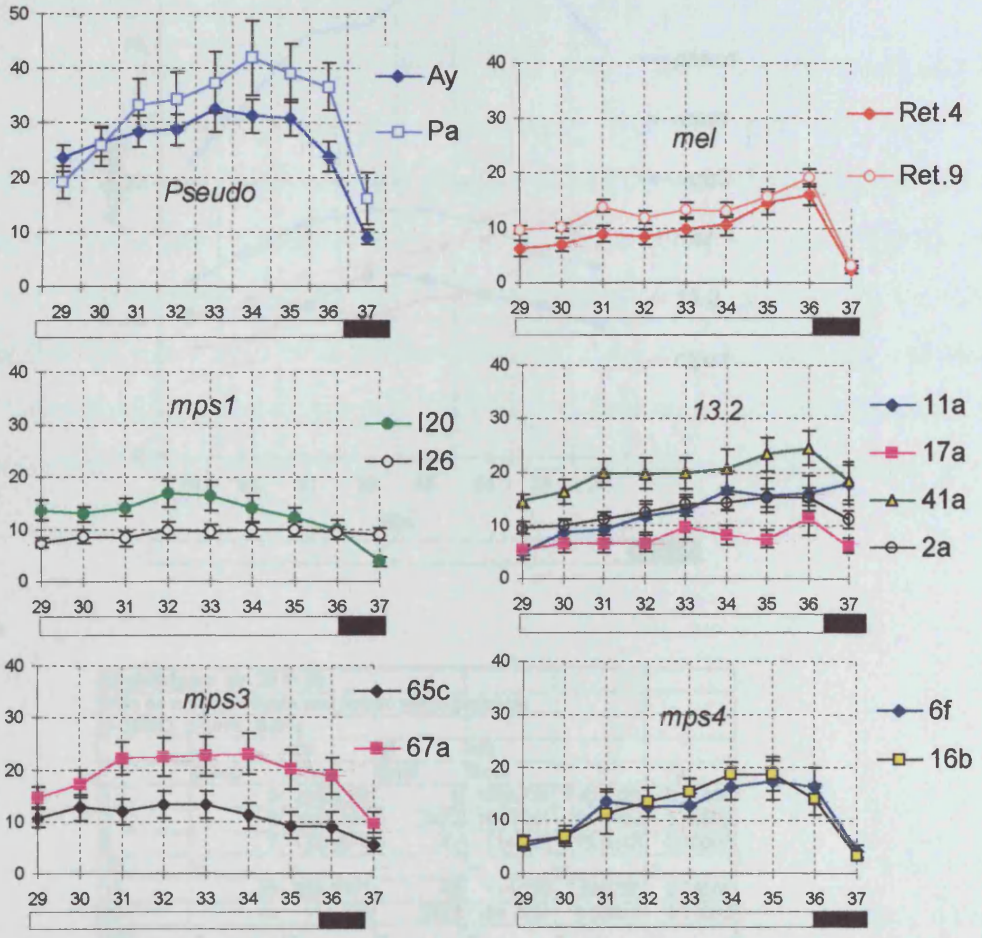
melanogaster (Ret 4 + Ret 9) v *pseudoobscura* (Pa + Ay) bin 29 to 40						
lines as random effects and nested within genotype						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	61133.57	2	2273.613	26.88829	0.035237
2	2	2273.613	1188	208.6547	10.89653	2.04E-05
3	11	8876.936	22	183.017	48.50335	7.9E-13
12	--	--	--	--	--	--
13	11	1669.399	22	183.017	9.121552	7.04E-06
23	22	183.017	1188	208.6547	0.877129	0.626586
123	--	--	--	--	--	--

Testing the *D.pseudoobscura/D.melanogaster* transgenes for LD effects

Initially all the lines available were tested for intragenotype variation with 2-way ANOVA (see Fig 4.8). Bins 29-37 were chosen for the analysis because I wished to concentrate on the observed differences immediately prior to lights off (see Fig 4.6 and Fig 4.7). All genotypes had at least two available lines. Two way ANOVA was performed within each genotype, and only the *mps1* Line * Bin interaction was significant (see Appendix 4.8a-f). However even in *mps1* line I26 peak activity occurs in the same bin as line I20, and so inspite of the interaction the two lines are quite similar (see Fig 4.8) Also if bins 29-36 are analysed no interaction is seen (see Appendix 4.8b),

revealing the differences are due to the lights off transition. As the primary *D.pseudoobscura/D.melanogaster* species-specific difference in locomotor behaviour has been identified as being in the bins leading up to lights off (see Fig 4.7), bin 37 was excluded from susequent analyses, which were designed to identify intergenotype heterogeneity. The 13.2 data was used to create two lines for this nested analysis by pooling the two cp20.1 vector lines (2a and 34a) to generate one new line, and the two pW8 vector lines (17a and 116a) to generate another. I was therefore able to use a nested design (which must be balanced) to analyse all the data simultaneously

Fig 4.8 Intragenotype variation in locomotor activity (bins 29-37) in LD 12:12. (Y- axis raw activity)



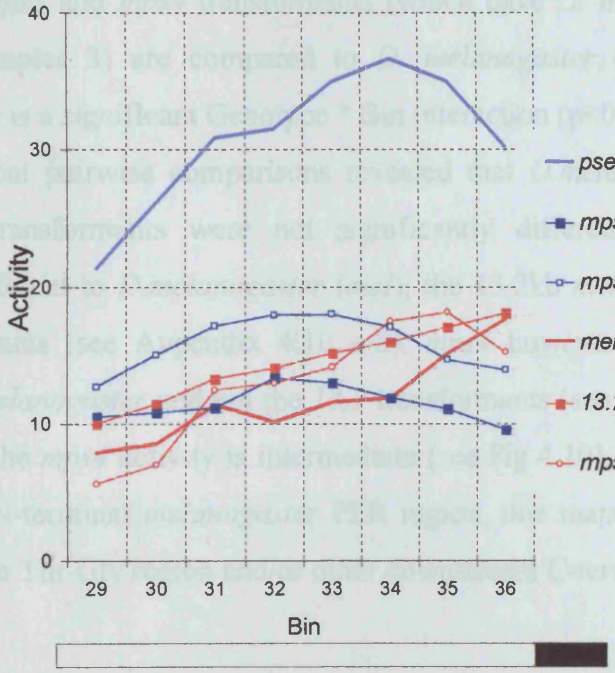
Preliminary pairwise comparisons of *D.pseudoobscura* v *mps1* and *D.melanogaster* v *13.2* with lines nested within genotypes revealed that both pairs gave a significant Genotype * Bin interaction over bins bin 29-37 but not bins 29-36 (see

Appendix 4.9a and b). This indicates that the lights-off transition is affecting the locomotor activity differently for natural variants and the corresponding transformed lines (This provides the rationale for excluding bin 37). Fig 4.9 has a deliberately elongated Y-axis to emphasize the differences in activity patterns between the genotypes. Profile differences are immediately apparent, and standard error bars are ommitted in Fig 4.9 for clarity, but are included in further analyses of these data in Figs 4.10-11

Fig 4.9 a Locomotor activity of all genotypes in LD 12:12 (bin 29-36).

b ANOVA table

a.



b.

All genotypes bin 29 to 36						
lines as random effects and nested within genotype						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p
1	5	23525.46	6	3085.781	7.623826	0.01405
2	6	3085.781	2472	164.7961	18.72484	0.00000
3	7	1750.573	42	114.235	15.3243	0.00000
12	--	--	--	--	--	--
13	35	259.1691	42	114.235	2.268735	0.00585
23	42	114.235	2472	164.7961	0.69319	0.93306
123	--	--	--	--	--	--

The overall ANOVA with lines nested within genotypes, shows a highly significant Genotype * Bin interaction ($p<0.006$). This indicates that the genotypes are exhibiting significantly different activity profiles. The *mps1* and *mps3* transformants

peak before lights out at bin 32-33, anticipating lights-off (the end of bin 36) by 1-2h, and are clearly similar in profile to *D.pseudoobscura* which peak at bin 34. *D.melanogaster* and *13.2* have an approximately linear increase in activity until bin 36, their activity peak clearly not anticipating lights-off. *mps4* however appears to have an intermediate pattern, peaking at bin 35 just before the lights/dark transition.

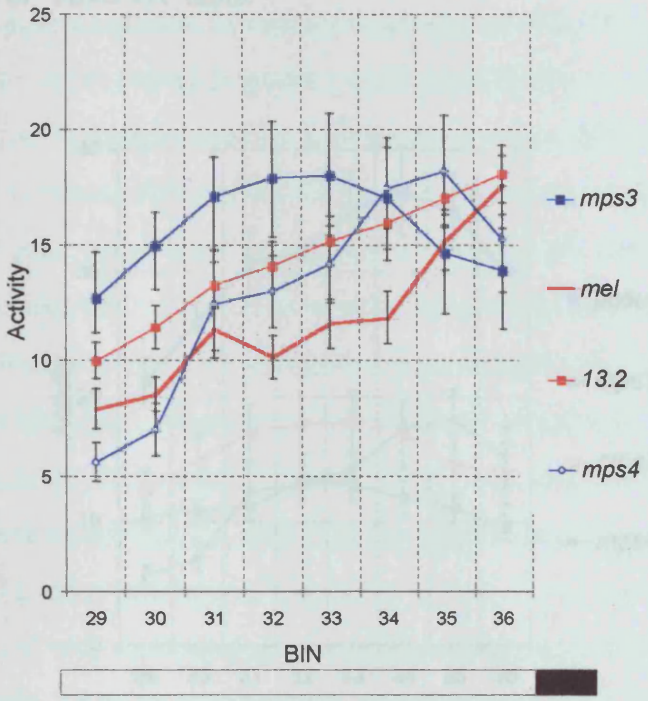
Planned comparisons

Again I used the nested design to compare the different groups. The first *a priori* assumption from the DD results (Fig 4.5) suggests that the N-terminal region of *per* will generate *melanogaster*-type behaviour (see Fig 4.4 and 4.5). However Fig 4.10 reveals that when the *mps3* and *mps4* transformants (which have *D. melanogaster* N-terminal regions, see Chapter 3) are compared to *D. melanogaster* (*mel*) and the 13.2 kb transgenes, there is a significant Genotype * Bin interaction ($p<0.001$).

Subsequent pairwise comparisons revealed that *D.melanogaster* (*mel*) and the 13.2kb *mel* transformants were not significantly different, but that *mps3* was significantly different to *D.melanogaster* (*mel*), the 13.2kb *mel* transformants and the mps4 transformants (see Appendix 4.10 a-d). *mps4* however was only significantly different to *D.melanogaster* and not the *13.2* transformants (see Appendix 4.10 e and f). suggesting that the *mps4* activity is intermediate (see Fig 4.10). As both *mps* transgenes have the same N-terminal *melanogaster* PER region, this maps any differences to the *D.pseudoobscura* Thr-Gly region and/or other downstream C-terminal sequences.

Fig 4.10 a. Locomotor activity in LD 12:12 of transformants carrying *D.melanogaster* N-terminal PER (bins 29-36). **b.** ANOVA table

a.



b.

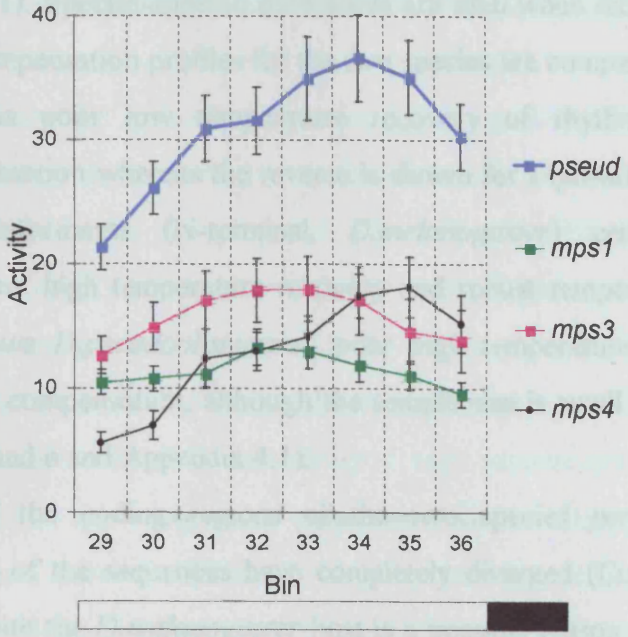
melanogaster, *132*, *mps3* and *mps4* (bin 29 to 36) lines as random effects and nested within genotype 1-GENO, 2-LINE, 3-Bin						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	3	979.5793	4	3405.923	0.287611	0.833041
2	4	3405.923	1688	109.971	30.9711	6.44E-25
3	7	1137.859	28	43.37032	26.2359	1.07E-10
12	--	--	--	--	--	--
13	21	160.7705	28	43.37032	3.706923	0.000724
23	28	43.37032	1688	109.971	0.39438	0.998185
123	--	--	--	--	--	--

If this is the case, then *D.pseudoobscura* and any transformant carrying *D.pseudoobscura* C-terminal material should show a similar profile. A second 'planned' comparison was made between *D.pseudoobscura* and the transgenes which contain *D.pseudoobscura per* coding sequence *mps1*, *mps3* and *mps4*. No significant Genotype * Bin interaction is observed (Fig 4.11).

Fig 4.11 a. Locomotor activity in *D.pseudoobscura* and transformants carrying *D.pseudoobscura per* sequences at bin 29-36 in LD

b. ANOVA table.

a.



b.

Pseudoobscura, mps1, mps3 and mps4						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	3	31272.54	4	2388.046	13.09545	0.015509
2	4	2388.046	1232	230.1999	10.37379	2.89E-08
3	7	1172.937	28	150.8087	7.777647	3.1E-05
12	--	--	--	--	--	--
13	21	265.0441	28	150.8087	1.757485	0.081345
23	28	150.8087	1232	230.1999	0.655121	0.915633
123	--	--	--	--	--	--

Thus, surprisingly, the species-specific locomotor behaviour in LD cycles between *D.melanogaster* and *D.pseudoobscura* maps to the 3' half of the *per* gene, including sequences C-terminal to the beginning of the Thr-Gly repeat.

Discussion

The results illustrate a number of novel findings concerning *D.pseudoobscura and D.melanogaster* species-specific locomotor behaviour in both (DD) and LD cycles. In DD *D.pseudoobscura* and *D.melanogaster* exhibit different locomotor responses to temperature (Fig 4.1). Species-specific differences are seen when recovery of rhythmicity and temperature compensation profiles for the two species are compared (Fig 4.1b and d). *D.melanogaster* has poor low temperature recovery of rhythmicity but excellent temperature compensation whereas the reverse is shown for *D.pseudoobscura*. The *mps3* and *mps 4* transformants (N-terminal, *D.melanogaster*) certainly show good '*D.melanogaster* like' high temperature recovery and robust temperature compensation. *mps1* however shows *D.pseudoobscura*'s poor high temperature recovery, and also poorer temperature compensation, although the sample size is small at high temperatures for *mps1* (Fig 4.1c and e and Appendix 4.1).

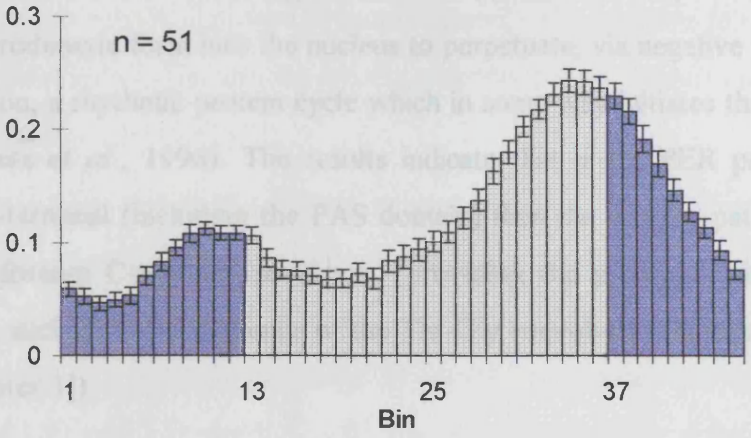
Analysis of the coding regions of the two species *per* genes shows that approximately 30% of the sequences have completely diverged (Colot *et al.*, 1988), so non-compatibility with the *D.melanogaster* host is a possible reason for the low recovery of rhythmicity of the *mps1* transgene in a per^{01} background at higher temperatures (see Fig 4.1). This lack of rhythmicity is however also seen in *D.pseudoobscura* at 29^0C and so *mps1*'s poor recovery may be due to a combination of species-specific differences in the protein and not necessarily just poor interaction with other host proteins, such as TIM (Gekakis *et al.*, 1995; Saez and Young, 1996).

(Petersen *et al.*, 1988) suggested that the *D.pseudoobscura* gene product is essentially 'weak'. This is not supported by the low temperature (18^0C) recovery of *mps1* (74% for two *mps1* lines) and *D.pseudoobscura* strains (nearly 100% for both strains, see Fig 4.1). Inspection of the *D.pseudoobscura* free-running activity profile (Fig 4.2) shows that at 18^0C *D.pseudoobscura* and *mps1* exhibit a profile qualitatively as rhythmic and robust as that of the *D.melanogaster* (Br) and *13.2* (line 2a) transformant. So the *pseudoobscura* PER protein cannot simply be 'weak', but it appears that the two species have different temperature ranges for optimum PER function. This may correspond to the more restricted geographical range of *D.pseudoobscura* (British Colombia to Mexico as opposed to *D.melanogaster's* cosmopolitan distribution, Wang and Hey 1996).

To investigate the possible species-specific effects outlined above, *D.melanogaster* and *D.pseudoobscura* free running daily activity profiles, were generated for a range of temperatures, but the largest differences were found at 18^0C. At this temperature *D.melanogaster* activity is bimodal with a large peak in the subjective morning and a small peak in the subjective evening while at this temperature *D.pseudoobscura* is also bimodal but with a small 'morning' peak and large 'evening' peak. The finding of a late day major peak of activity in *D.pseudoobscura* in DD, confirms the findings of Petersen *et al.* (1988) who used a different type of analysis and different entrainment regimes to those used in this study.

The lack of bimodal activity at higher temperature in *D.pseudoobscura* (see Fig 4.2) can be considered to be a factor which masks the rhyhmicity and temperature compensation of the species. I compared this with the DD activity profile of a *D.melanogaster per* Thr-Gly deletion transformant which also has poor recovery and poor temperature compensation (see 5.3 table 1). The *a priori* assumption was that if poor bimodality was linked to lack of temperature compensation, then this transformant, like *D.pseudoobscura,* would show little bimodality at high temperature. Fig 4.13 shows the locomotor activity of a Δ(Thr-Gly) transformant at 29^0C in DD (at this temperature it's period is ~27h, [see Chapter 5, Table 1] and the data has been plotted using X-axis standardization described in materials and methods). Bimodality is strong and clearly comparable to the *D.melanogaster* profile at 29^0C (see Fig 4.2).

Fig 4.13 Locomotor activity in DD of Δ(*Thr-Gly*) line 52c at 29^0C (Y-axis = transformed proportions, X-axis = adjusted bin number)



The maintenance of this distinctive feature, even for a transformant with poor rescue and temperature compensation, shows that the presence of bimodality is not the primary factor ensuring robust rhythmicity and temperature compensation at 29^{0}C in *D.melanogaster*. Conversely the absence of this pronounced bimodality in *D.pseudoobscura* is therefore not likely to be the result of a poorly functioning PER protein, as is the case for the product of the Δ(*Thr-Gly*) transgene. This does however reinforce the concept that the species-specific differences shown in Fig 4.2 are produced or facilitated by the specific activity of the protein. On the basis of the above arguments, the activity differences did indeed appear to be species-specific, so the analysis was continued.

18^{0}C was chosen as the most suitable temperature for further analysis as *D.melanogaster* and *D.pseudoobscura* are clearly statistically distinguishable in DD by early and late peaks respectively in the subjective day (Figs 4.3 and 4.5). Data for the *D.pseudoobscura* transformant *mps1* supports this species-specific scenario as the transgene confers the 'late day' '*pseudoobscura*' activity patterns to *per* ^{01} hosts and comparison of the peak bin values of *D.pseudoobscura* and *mps1* show no significant differences (Figs 4.4 and 4.5).

Figs 4.4 and 4.5 show that the N-terminal (*D.melanogaster*) part of the *per* gene, which includes the PAS region, is sufficient for the maintainance of robust *D.melanogaster* patterns, and comparison of *D.melanogaster*, the *mel 13.2*, and *mps3* transformants show no significant differences in the bin numbers which carry the peak activity values (Appendix 4.4c). This may indicate that the interaction of the transgenes' PAS region with TIM (Gekakis *et al.*, 1995) is involved in maintainance of species-specific activity profiles, at least in DD. The DD results therefore fit the current ideas of PER forming an interaction through binding via the PER PAS region to TIM (Gekakis *et al.*, 1995) and then progressing in heterodimeric form into the nucleus to perpetuate, via negative feedback on it's own transcription, a rhythmic protein cycle which in some way initiates the locomotor activity cycle (Myers *et al.*, 1996). The results indicate that if the PER protein has a *D.melanogaster* N-terminal (including the PAS domain) then the activity patterns in DD are unaffected by foreign C-terminal sequences (providing the protein retains a certain degree of integrity, such as in the presence of the Thr-Gly coevolved region irespective of species origin[Chapter 3]).

Fig 4.6 shows that in LD 12:12, *D.pseudoobscura* has a much higher (2-3x) level of activity than *D.melanogaster* in the evening peak, and there is a highly significant difference in the activity profile of the two species (Fig 4.6). Subsequent analysis indicated that *D.pseudoobscura* has different activity in the two hours prior to lights off when compared to *D.melanogaster* (Fig 4.7). This takes the form of higher activity levels than *D.melanogaster* and an anticipation of lights-off, in terms of a reduction in activity levels just before lights off, not seen in *D.melanogaster* or the *13.2. mel* transformants (Fig 4.8). *D.pseudoobscura* and the corresponding *mps1* transformants show similar early peaks of activity in LD (Fig 4.9). However in contrast to the DD results, this '*D.pseudoobscura* like' behaviour is also shown by the *mps3* and *mps4* transgenes, which have *D.pseudoobscura per* coding material in the C-terminal and Thr-Gly encoding regions of the gene (Fig 4.11). Furthermore, *mps4* was different both from *D.melanogaster* and *mps3*, posssibly reflecting the lowered *D.pseudoobscura per* coding content (Appendix 4.10).

The limit of resolution for this experiment is determined by the error in synchronising the light cycle of the incubators with the data collection equipment. I estimate this error to be in the order of 10 minutes, and certainly less than one 30 minute bin. Thus even though the LD differences between species are subtle the similarity between each species and it's respective transformants adds confidence to my interpretation.

It is known that light resets the clock and that it has measurable effects on at least one of the major clock components TIM. Constant light (LL) has inhibitory effects on the accumulation and phosphorylation of TIM (Marrus *et al.*, 1996) and light causes rapid proteolysis of phosphorylated TIM (Zeng *et al.*, 1996). The LD cycle is also known to cause fluctuations in *per* mRNA that cannot be explained by the negative feedback model (Qiu *et al.*, 1996). How light influences the activity patterns of *per* variants and mutants is further investigated in Chapters 5 and 6.

In conclusion, the results show that species-specific elements of locomotor behaviour in DD have been mapped to the N-terminal conserved part of the *per* coding sequence (Fig 4.5). In LD the C-terminal diverged region containing the Thr-Gly repeat region and 3' sequences are involved in generating species-specific activity (Figs 4.8-11). The comparison of DD and LD results could also be interpreted as showing that the imposition of a zeitgeber at 18^0C has a much larger effect on *D.melanogaster* behaviour,

than on *D.pseudoobscura*. Nevertheless in the appropriate regime, species-specificity can be assigned to various *per* gene fragments.

Chapter 5 : Temperature compensation effects of the Thr-Gly repeat

5.1 Introduction

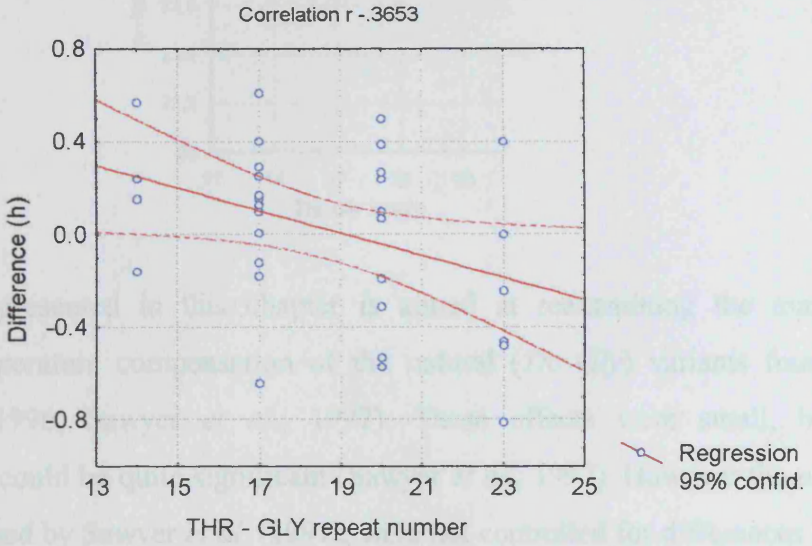
The Thr-Gly repeat in *D. melanogaster* is known to be polymorphic in length (Yu *et al.*, 1997b, Costa *et al.*, 1991, Peixoto *et al.*, 1992, Rosato *et al.*, 1994, Sawyer, 1996) with variation in natural populations ranging between 14 and 24 repeats. Alleles encoding 14, 17, 20 and 23 (Thr-Gly) pairs, account for 99% of the length variation seen in European populations (Costa *et al.* 1992). Other rare alles such as (Thr-Gly) $_{15,21 \text{ and } 24}$ make up the rest of the natural population. Variants differ largely in the length of the repeat, with no significant amino acid alteration in the immediate flanking region (Peixoto *et al.*, 1992, Rosato *et al.*, 1996). An S to F replacement is the only amino acid polymorphism that has been encountered in the immediate flanking regions surrounding the repeat in European (Costa *et al.* 1991, 1992 and Rosato *et al.* 1996) and other populations (Kliman and Hey, 1993). This suggests that the region may be under some kind of constraint. The two major variants in European populations $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ which account for ~90% of the total, are distributed as a highly significant latitudinal cline in Europe and North Africa (Costa *et al.* 1992), with high levels of the former observed in the southern Mediterranean, and the latter predominating in northern Europe. This latitudinal comparison of the two alleles has also been replicated in Australian natural populations (Sawyer, 1996). Some selective advantage may therefore be held by the different length alleles that predominate at the different latitudinal extremes. Average daily temperature seemed the most likely environmental difference due to latitude but there are of course many others, for example variation in photo period.

In an initial attempt to relate some of these effects to Thr-Gly variation, the free-running circadian locomotor activity rhythms of males from a large number of different attached-*X* lines whose *per* carrying *X* chromosomes originated from eight European and North African localities, were examined at 18°C and 29°C (see Fig 5.1 taken from Sawyer, 1996)

The natural Thr-Gly length variants $(Thr\text{-}Gly)_{14,17,20,23,}$ show slightly different patterns of circadian temperature compensation (see Fig. 5.1). The data clearly show that there is a linear relationship betweeen Thr-Gly length and temperature compensation.

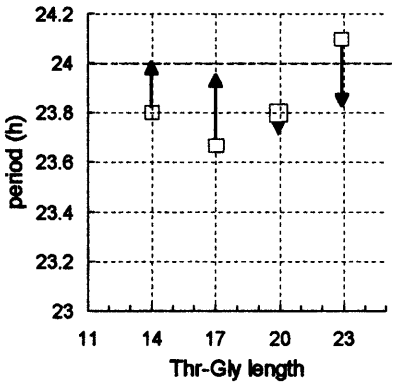
One noticeable feature seen when the mean periods at the two temperatures of the data from Fig 5.1 are plotted (see Fig 5.2), is that the periods of the $(Thr\text{-}Gly)_{14}$ and $(Thr\text{-}Gly)_{17}$ flies become longer, and $(Thr\text{-}Gly)_{23}$ become shorter, on increasing the temperature (the direction of the arrow shows the change in period at the warmer temperature). The $(Thr\text{-}Gly)_{20}$ flies on the other hand are virtually perfectly temperature compensated.

Fig 5.1 Temperature compensation of 33 European populations of *D.melanogaster* with varying Thr-Gly repeat lengths (data from Sawyer, 1996). y-axis is the free running period at 29^{0}C minus that at 18^{0}C.



Structural investigation of the Thr-Gly motif has suggested that the region is composed of a series of turns based on a $(Thr\text{-}Gly)_3$ peptide which represents a conformational monomer, and generates a type II or III β-turn (Castiglione-Morreli *et al.*, 1995). The approximately linear relationship between the 14-17-20-23 allelic series and temperature compensation may therefore involve a structural component related to the dynamic properties of the $(Thr\text{-}Gly)_3$ motif, and illustrates a correlation between natural variation in protein conformation and a behavioural phenotype. This is supported by the behavioral analysis of flies carrying the rare length alleles $(Thr\text{-}Gly)_{15,\ 21,\ 24}$ which represent less than 1% of European populations (Costa *et al.*, 1991). These flies do not appear to follow the temperature compensation trends presented in Fig 5.1 and Fig 5.2 (Sawyer,1996).

Fig 5.2 Period change from 18^{0}C (open squares) to 29^{0}C (arrow heads) for the pooled means of 33 European populations of *D.melanogaster* with varying Thr-Gly repeat lengths (after Sawyer, 1996)



The work presented in this chapter is aimed at reexamining the marginal differences in temperature compensation of the natural (*Thr-Gly*) variants found by Sawyer (Sawyer 1996, Sawyer *et al.*, 1997). These effects were small, but in evolutionary terms could be quite significant (Sawyer *et al.*, 1997). However the natural *per* variants examined by Sawyer *et al.* (1997), were not controlled for differences in the X chromosomes, which were taken intact from each European population. Consequently the effect seen by Sawyer *et al.* (1997) might be due to other genes on the X chromosome. The only way to see whether these effects were due to the Thr-Gly repeat itself was to make internal deletions of the repeat from a $(Thr\text{-}Gly)_{20}$ transgene. In this way the effect of any sequence in linkage disequilibrium with the repeat (Rosato *et al.*, 1997) or more distantly located on the X, could be disregarded. I therefore generated *per* transgenes in which internal deletions of the repetitive tract from a cloned $(Thr\text{-}Gly)_{20}$ *per* gene were made. $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_1$ transgenes were thus constructed (see Chapter 2), and a $\Delta(Thr\text{-}Gly)$ transgene (Yu *et al.*, 1987b, Ewer, *et al.*, 1990) containing a slightly more extensive deletion than the perfect (*Thr-Gly*) repeat, was included for comparison and analysis. All the constructs were made from the original cloned 13.2 kb per^+ fragment thus avoiding any possibility of polymorphism within the transgenes. The transformants were tested for their free-running locomotor periods at various temperatures.

Part of the work presented in this Chapter has been published in Sawyer *et al.*, 1997, *Science* Vol 278, pp. 2117-2120. Title: 'Natural variation in a *Drosophila* clock gene and temperature compensation'. A reprint is enclosed in Appendix 5.

5.2 Materials and methods

Four $(Thr\text{-}Gly)_{20}$, two $(Thr\text{-}Gly)_{17}$, three $(Thr\text{-}Gly)_{1}$ and two $\Delta(Thr\text{-}Gly)$ independently generated transformant lines were tested for free-running locomotor activity at 18°C, 25°C and 29°C. Two of the $(Thr\text{-}Gly)_{20}$ transformants (2a and 34a) and the $\Delta(Thr\text{-}Gly)$ transformants were generated with the transformation vector cp 20.1. All the other transgenics were transformed using pW8 (Chapter2). Locomotor activity recording and statistical analysis was identical to that previously described (see methods Chapter 4).

5.3 Results

Overall mean periods

Table 1 shows the results of spectral analysis on all lines with associated standard errors.

Table 1 Free-running periods of *Thr-Gly* transgenes at 18°C, 25°C and 29°C. N = number of flies tested, n = number rythmic

Thr-Gly construct	line	N 18°C	n	N 25°C	n	N 29°C	n	period (h) ± sem 18°C		period (h) ± sem 25°C		period (h) ± sem 29°C	
Δ(Thr-Gly)	18A	50	27	39	23	55	50	25.02	0.28	25.67	0.42	26.16	0.13
	52C	46	32	27	33	58	55	24.58	0.22	26.31	0.16	27.38	0.17
pooled		96	59	66	56	113	105	24.78	0.18	26.01	0.25	26.80	0.12
(Thr-Gly) 1	44A	44	34	30	27	37	35	24.76	0.10	26.70	0.17	26.05	0.13
	57A	23	19	14	11	17	16	24.34	0.12	26.83	0.19	26.25	0.20
	88A	33	26	48	41	51	49	24.93	0.18	26.78	0.11	26.58	0.10
pooled		100	79	92	79	105	100	24.71	0.08	26.76	0.09	26.34	0.08
(Thr-Gly) 17	11A	35	30	18	13	35	24	23.40	0.12	23.77	0.55	24.27	0.15
	41A	73	57	54	50	88	83	24.30	0.11	25.22	0.10	25.48	0.10
pooled		108	87	72	63	123	107	23.99	0.10	24.93	0.15	25.20	0.10
(Thr-Gly) 20	2A	61	41	25	23	53	45	24.61	0.11	25.02	0.1	24.80	0.09
	17A	59	31	24	20	46	26	24.07	0.18	24.34	0.66	24.18	0.19
	34A	51	30	47	39	32	32	24.21	0.16	25.02	0.13	24.92	0.16
	116A	50	41	33	30	33	31	23.51	0.07	25.52	0.11	25.43	0.16
pooled		221	143	129	112	164	134	24.09	0.07	25.03	0.13	24.85	0.08

Statistical analysis of the intragenotype variation of the transgenes

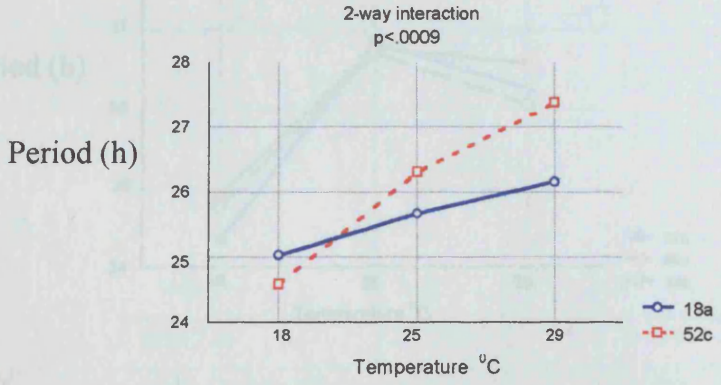
The lines were initially analysed independently within genotype.:-

Δ (*Thr-Gly*)

Fig 5.3 shows the two lines of the *Δ(Thr-Gly)* transgene which rescue rhythms in 80% of per^{01} hosts (rescue pooled across temperatures and lines). Periods are in the 25 to 27h range at all three temperatures (Fig 5.3) and there are significant line (L), temperature (T) and L * T interactions. However it appears that both lines increase their periods across the temperatures range in an approximately linear manner varying in period from 24.5 to 27.5h.

Fig 5.3a Mean free running periods of Δ(*Thr-Gly*) transgene lines. **b**. ANOVA of the data. [y-axis Period (h)]

a.



b.

Source	MS	d.f.	F	p
Line (L)	10.785	1	6.0024	**.01511**
Temperature(T)	72.757	2	40.4920	**.00000**
L*T	13.021	2	7.2469	**.00090**
Error	1.796	208		

Newman-Keuls *post hoc* comparisons (Appendix 5.3) show that within temperatures the lines only differ significantly at 29°C ($p<.001$). Within lines both show significant increases in period between 18 and 29°C (at least $p<.002$, see Appendix 5.3). Line 52c is also significantly different between 18 and 25°C ($p<.001$) whereas line 18a marginally misses significance between these temperatures ($p<.053$).

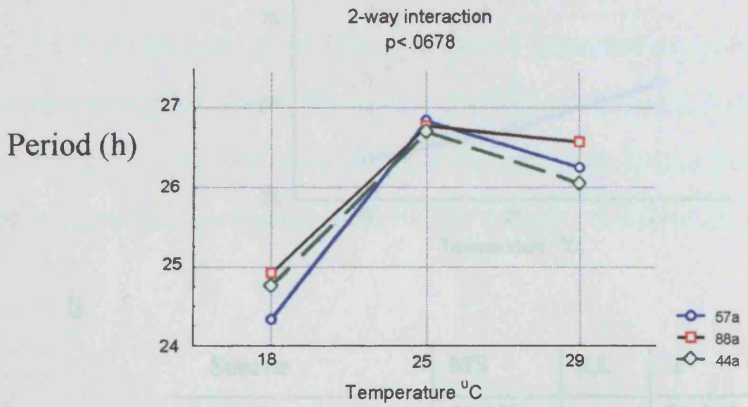
(*Thr-Gly*)$_1$

The three lines of the (*Thr-Gly*)$_1$ transgene rescue rhythms in 86% of *per*^{01} hosts (data pooled across temperatures and lines) with periods in the circadian 25-27h range at all three temperatures (Fig 5.4). The (*Thr-Gly*)$_1$ transformants therefore have a period range similar to (Fig 5.3). However unlike Δ(*Thr-Gly*) the longest periods are found at 25^{0}C in all three lines.

Fig 5.4 shows the mean periods of the (*Thr-Gly*)$_1$ transformants and reveals significant Line and Temperature effects in the ANOVA. All lines show an increase of ~2h between 18 and 25^{0}C and a small decrease from 25 to 29^{0}C.

Fig 5.4a Mean free running periods of (*Thr-Gly*)$_1$transgene lines.**b** ANOVA table of the data. [y-axis Period (h)]

a.



b.

Source	MS	d.f.	F	p
Line (L)	2.212	2	4.1873	**0.0162**
Temperature(T)	84.277	2	159.4931	**0.0000**
L*T	1.170	4	2.2159	0.0678
Error	0.528	248		

Newman-Keuls *post hoc* comparisons (Appendix 5.4) confirm the changes in period between 18-25^{0}C. Between 25 and 29^{0}C a marginally significant decrease in period of 30-40 minutes are seen for lines are 57 and 44a ($p<.047$, $p<.012$ respectively). The three lines are therefore remarkably uniform in their profiles.

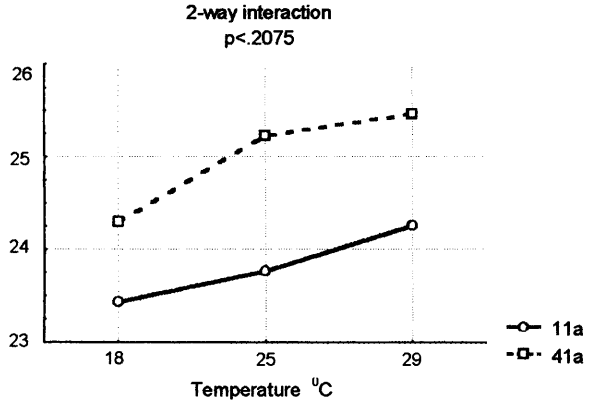
The (*Thr-Gly*)$_1$ and Δ(*Thr-Gly*) genotypes therefore have different behavioural responses to the temperature challenges, which is important as the Δ(*Thr-Gly*) deletion removes some sequence flanking the uninterrupted repeat.

***(Thr-Gly)*$_{17}$**

The two lines of the *(Thr-Gly)*$_{17}$ transformants rescue rhythms in 86% of *per*^{01} hosts (data pooled across temperatures and lines). The *(Thr-Gly)* $_{17}$ transformants (Fig 5.5) show significant Line and Temperature effects but no interactions revealing that they are behaving in a similar manner. Both lines appear to show an increase in period of ~1 h between 18^0-29^{0}C, and are separated by ~1-1.5h in period at all three temperatures.

Fig 5.5a Mean free running periods of *(Thr-gly)*$_{17}$ transgene lines. **b.** ANOVA table of the data. [y-axis Period (h)]

a.



b.

Source	MS	d.f.	F	p
Line (L)	61.608		76.464	**.0000**
Temperature(T)	19.531	2	24.119	**.0000**
L*T	1.280	2	1.607	.2075
Error	.809	250		

Newman-Keuls *post hoc* comparisons (Appendix 5.5) confirm that the two lines are significantly different from each other at all three temperatures. Within lines there are significant increases in period for one line between 18-25^{0}C (p<.0001, line 41a) and for the other line between 25-29^{0}C (p<.03 line 11a). Overall, between 18-29^{0}C both lines show highly significant increases of ~ 90 minutes (p<.0001 for line 41a and p<.005 for line 11a)

(*Thr-Gly*)$_{20}$

This data has been presented in Fig 3.4 (Chapter 3). To reiterate the four lines of the (*Thr-Gly*)$_{20}$ transgene rescue rhythms in 76% of *per*^{01} hosts (data pooled across temperatures and lines) with periods in the circadian 24-25h range at all three temperatures.

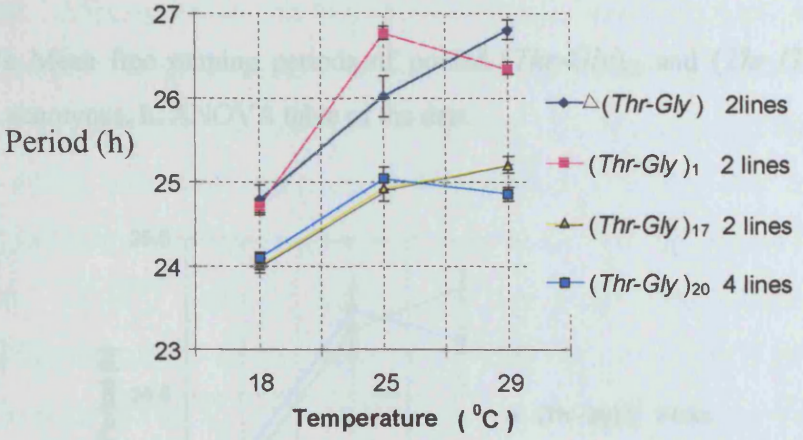
Testing the effects between genotypes

Pooled genotypes

Although there were some L * T interactions, the differences were small and so I pooled the lines to give an overall comparison of the of the four genotypes. It can be seen from Fig 5.6 that the genotypes fall into two groups. The first [(*Thr-Gly*)$_{20}$ and (*Thr-Gly*)$_{17}$] have periods close to 24h at all three temperatures and follow the temperature compensation profiles shown by natural populations (Fig 5.2 Sawyer, 1996 Sawyer *et al.*, 1997). The Δ(*Thr-Gly*) and (*Thr-Gly*)$_1$ have generally longer periods and are clearly less well temperature compensated with ~1-2hr longer periods at all temperatures

Fig 5.6a. Mean free running periods of pooled (Thr-Gly) transformants b. ANOVA table of the data

a.



b.

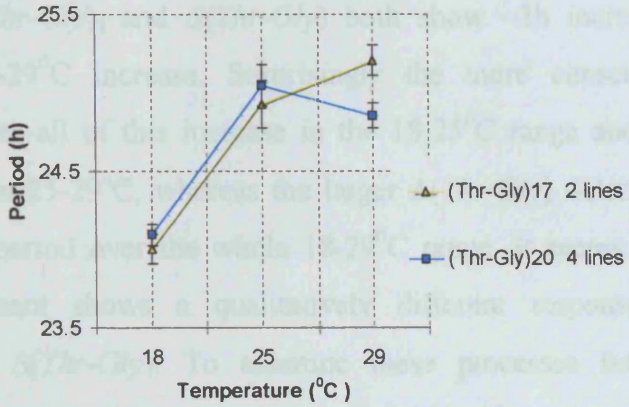
Source	MS	d.f.	F	p
Genotype (G)	131.076	3	114.4906	**0.0000**
Temperature(T)	209.908	2	183.3478	**0.0000**
G * T	12.953	6	11.3140	**0.0000**
Error	1.144	1102		

ANOVA shows all effects to be significant and *post hoc* comparisons (Appendix 5.6) within the temperatures confirm that the Δ(*Thr-Gly*) and $(Thr\text{-}Gly)_1$ are statistically different at 25°C and 29°C ($p<.0001$ and $p<.02$ respectively) from each other, and both are highly statistically different at all three temperatures from the $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$. Differences between Δ(*Thr-Gly*) and $(Thr\text{-}Gly)_1$ are also confirmed by Newman-Keuls tests (Appendix 5.6). In comparison there are no significant differences between $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ at any temperature (Fig 5.6). However because of the importance of the $(Thr\text{-}Gly)_{20}$ to $(Thr\text{-}Gly)_{17}$ comparison, in relation to the differences found by Sawyer *et. al.*, (1997) between these two variants in natural populations, these were examined separately.

Fig 5.7 shows a significant G * T interaction (p<.033) and a Temperature effect. Clearly the interaction is due to the change in period of the two genotypes between 25 and 29^{0}C, so a subtle difference can be detected between the two genotypes.

Fig 5.7a Mean free running periods of pooled $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ genotypes. **b.** ANOVA table of the data

a.



b.

Source	MS	d.f.	F	p
Genotype (G)	.321	1	.29238	.5888
Temperature(T)	64.894	2	59.08729	**.0000**
G*T	3.767	2	3.42994	**.0329**
	1.098	637		

Discussion

These results clearly reveal that changes in period and temperature compensation are associated with a disruption of the (*Thr-Gly*) region. Δ(*Thr-Gly*) transformants are not temperature compensated at this temperature range, confirming an earlier report (Ewer *et.al.*, 1992). However my results, taken from a variety of deletions now show that not all deletions in this region produce similar effects. This, in turn suggests that it may be possible to assign particular functions to the different 'domains' that comprise the (*Thr-Gly*) region. $(Thr\text{-}Gly)_1$ and Δ(*Thr-Gly*) both show ~2h increases in period in response to the 18-29^{0}C increase. Surprisingly the more conservative $(Thr\text{-}Gly)_1$ transformant generates all of this increase in the 18-25^{0}C range and is conspicuously compensated between 25-29^{0}C, whereas the larger Δ(*Thr-Gly*) deletion shows a more gradual increase in period over the whole 18-29^{0}C range. It seems therefore that the $(Thr\text{-}Gly)_1$ transformant shows a qualitatively different response to temperature challenge than the Δ(*Thr-Gly*). To examine these processes futher, the deletion constructs presented in this chapter were compared with three other (*Thr-Gly*) deletions constructs generated by my colleague Helen Parkinson. They are described below and diagramatically compared with previously discussed deletions in Fig 5.8.

(a) The *L3* transgene retains 5' flanking sequences and the $(Thr\text{-}Gly)_{17}$ array but 75 base pairs from the the 3'sequences have been deleted

(b) In the *L4* transgene this 3' flanking material and the (*Thr-Gly*) array has been deleted [but leaving a single pair of Thr-Gly pair from the uninterrupted repeat].

(c) The *L2* deletion has the 5'flanking material, the Thr-Gly repeat and the 3' flanking sequences removed.

By combining the data derived from the transgenes *L3*, *L4* and *L2* (Parkinson, 1997) with the ones described in this chapter into a coherent series, we can gain an informative insight into the functional subdomains which comprise the repetitive region and which influence temperature compensation (see Fig 5.8).

Fig 5.8 Diagramatic comparison of all the (*Thr-Gly*) deletion constructs. Drawings not to scale

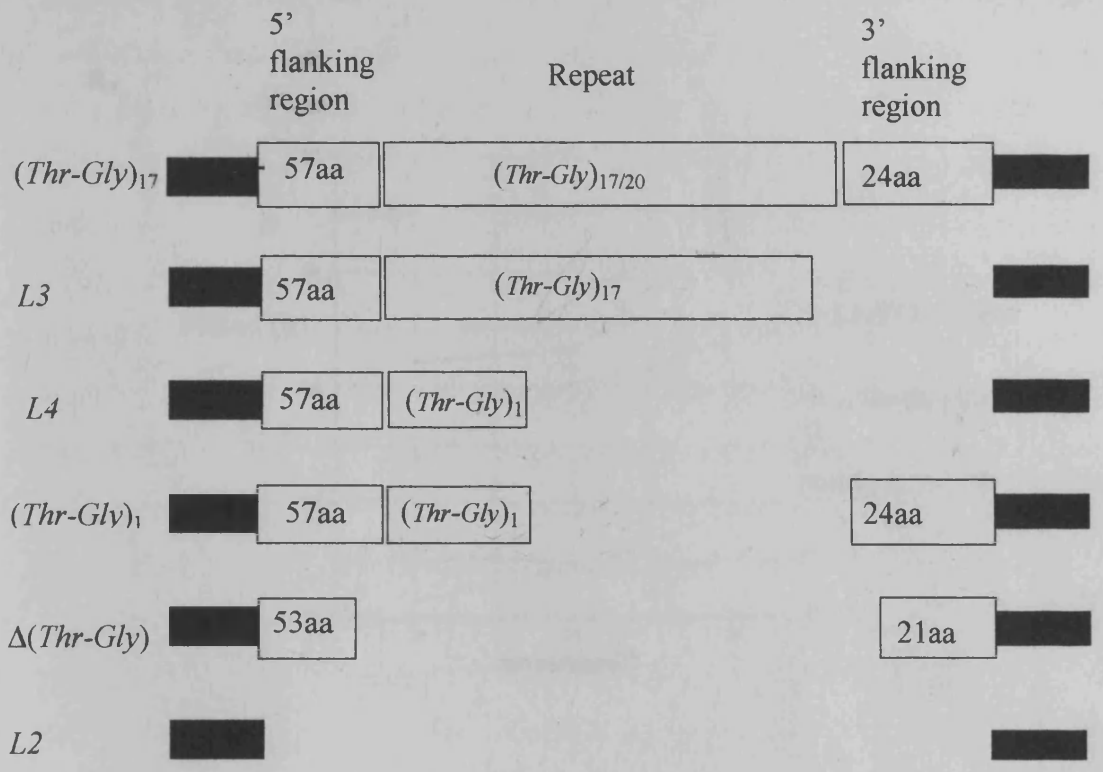
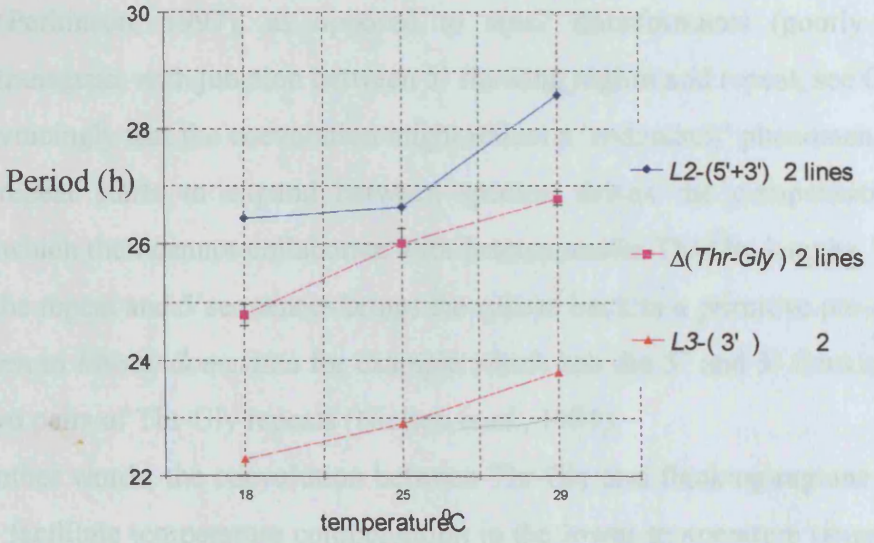


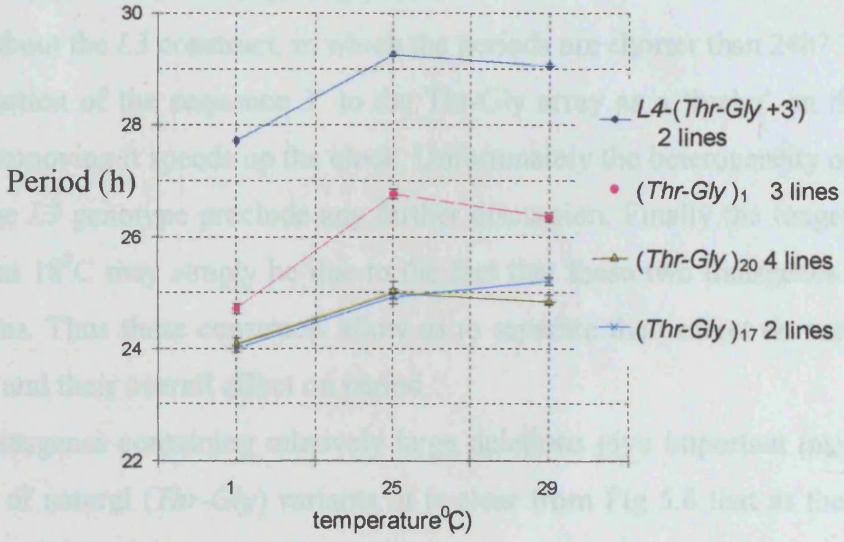
Fig 5.9 divides the constructs into those which show a significant increase in period between 25-29^{0}C (Fig 5.9a) and those that do not (Fig5.9b) [sem for the data included where available]. The *L3* genotype in Fig 5.9a consists of two independently transformed lines which were heterogeneous in their temperature compensation profiles so I cannot comment securely on these results, other than noting the overall shorter period of 22-23h in this genotype. All the other genotypes were homogeneous in terms of their line effects (Parkinson 1997) and the Δ(*Thr-Gly*) and *L2* show clear temperature rises in the 25-29^{0}C temperature range. The constructs in Fig 5.9b have either an extended repeat $(Thr\text{-}Gly)_{17}$ or $(Thr\text{-}Gly)_{20}$ or at least one Thr-Gly pair remaining from the uninterrupted array and all show a high level of compensation in the 25-29^{0}C range . Consequently if the *per* gene has an extended *Thr-Gly* array or any Thr-Gly pairs left within the array, temperature compensation between 25-29^{0}C remains intact. Removing the uninterrupted repeat plus some flanking amino acids, as in Δ(*Thr-Gly*) or in deletion *L2* disrupts temperature compensation between 25-29^{0}C

Fig 5.9 a. Constructs with poor overall temperature compensation. (sequences deleted indicated in brackets) **b.** Constructs with compensation at 25-29^{0}C

a.



b.



Interestingly the extended Thr-Gly transgenes [$(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$] and also the *L2* deletion, are well compensated between 18-25^{0}C. In all other transgenes, where the size of the repeat array in relation to the 5' flanking region has been significantly altered, poorer temperature compensation is observed between 18-25^{0}C. From both statistical and experimental evidence, it is known that the length of the repeat

array coevolves with the 5' flanking material (see Chapter 3). Consequently we would predict that removing the 3' flanking sequences, Thr-Gly repeat and the 5' flanking sequences (as in *L2* see Fig 5.8) might have less of an effect in disrupting this coevolution, than in a construct where the size relationships of these regions to each other have been altered. Certainly the fact that the *L2* transformants are perfectly rhythmic (Parkinson, 1997), as opposed to *mps2* transformants (poorly rhythmic chimaeric transgene, with junction between 5' flanking region and repeat, see Chapter 3) argues convincingly that the coevolution might reflect a 'redundant' phenomenon which, once the repeat starts to expand between species, drives the compensation of 5' sequences which then cannot collaborate with heterospecific Thr-Gly lengths. Therefore removing the repeat and 5'sequences brings the system back to a primitive pre-expansion state, as seen in *Musca domestica* for example which has the 5' and 3' flanking regions but only two pairs of Thr-Gly repeats (Nielsen *et.al.*, 1994).

In other words, the coevolution between Thr-Gly and flanking regions may have evolved to facilitate temperature compensation in the lower temperature range, whereas the length or presence of the uninterrupted array itself, determines temperature compensation in this species at the higher temperature range. My results and those of Parkinson (1997) fit this model surprisingly well.

What about the *L3* construct, in which the periods are shorter than 24h? This may reflect the function of the sequence 3' to the Thr-Gly array as a 'brake' on the clock, consequently removing it speeds up the clock. Unfortunately the heterogeneity of the two lines within the *L3* genotype preclude any further discussion. Finally the longer periods of *L4* and *L2* at 18^0C may simply be due to the fact that these two transgenes have the largest deletions. Thus these constructs allow us to separate their effect on temperature compensation and their overall effect on period.

The transgenes containing relatively large deletions give important insights into the behaviour of natural (*Thr-Gly*) variants. It is clear from Fig 5.6 that as the Thr-Gly length of transgenes is decreased [$(Thr\text{-}Gly)_{20,\,17,\,1,\,\Delta}$] so the free-running period increases as temperature rises. These results convincingly support the validity of those based on the natural variants (Sawyer, 1996; Fig 5.1 and 5.2). The $(Thr\text{-}Gly)_{20,\,17}$ transgenes do not reproduce the exact natural $(Thr\text{-}Gly)_{20,\,17}$ variant phenotypes but this can be assumed to be due either to transgenes ectopic location in the genome, or possibly to absence of regulatory elements in the original cloned 13.2 kb *per* $^+$ fragment. These period

alterations, are however in the same direction and appear to be amplifications of effects seen in natural populations (Fig 5.1 and 5.2), with the $(Thr\text{-}Gly)_{20}$ transformants showing better overall temperature compensation than the $(Thr\text{-}Gly)_{17}$ between 18-29^{0}C. Furthermore, the design of the transgenes (Introduction and Chapter 2) means that the associated temperature compensation differences between natural $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ cannot be due to any linkage disequilibrium with the different repeat arrays. The study of the transformants therefore allows us to be confident that the effects shown in natural populations are due to variations in the Thr-Gly region. The evidence from the results of the flanking region deletion constructs (Fig 5.9) suggests that the three different domains of the repeat region (5' flank, repeat and 3' flank) interact to provide an appropriate clock response to changing temperature regimes. The possible molecular mechanism underlying these effects will be discussed in the final chapter.

Chapter 6: How temperature and *per* interact to affect behaviour patterns in *Drosophila*

6.1 Introduction

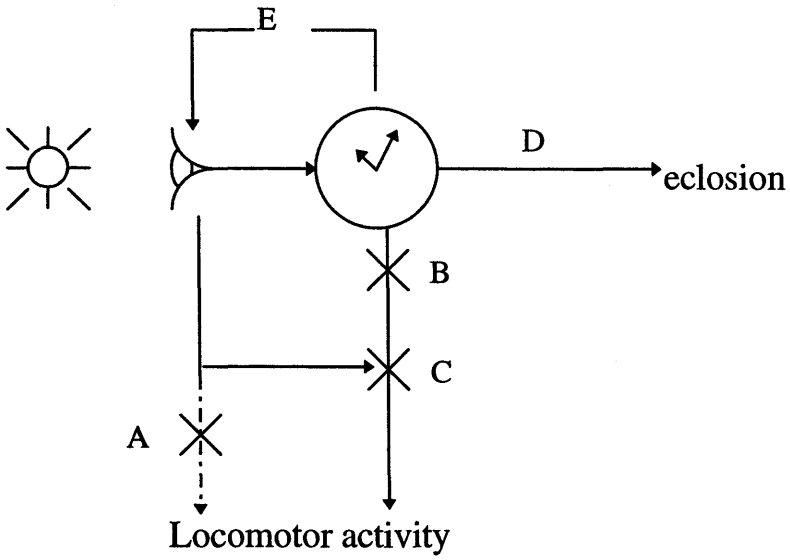
Locomotor activity in many insect species is concentrated around dawn and dusk in field conditions, with a resting period in the middle of the day (Saunders, 1982). Explanations for these locomotor patterns usually have an ecological perspective such as avoidance of midday dessication or connection with feeding behaviour (Saunders, 1982). This 'bimodal' locomotor behaviour has also been extensively observed under artificial light dark cycles for both *D.pseudoobscura* (Engleman and Mack, 1978, Petersen *et al.*, 1988) and *D.melanogaster* (Hamblen *et al.*, 1986; Sawyer, 1996). In *D. melanogaster* this behaviour can be observed in single flies in LD (Newby and Jackson 1991, Emery *et al.*, 1993) and in the average activity profiles of populations of flies in LD (Hamblen-Coyle *et al.*, 1992, Wheeler *et al.*, 1992). Bimodality in these two species is also observed in free-running conditions, and temperature alters the phase relationship of the two activity peaks differentially for the two species (see Chapter 4, Fig 4.2 and Fig 4.3). In these species the phase relationship of their locomotor activity peak levels is determined by *per,* and is therefore related to clock function (see Chapter 4).

For *D.melanogaster* in DD the two peaks appear to combine at 18^0C to present a unimodal early day activity pattern (Newby and Jackson, 1991, and see Fig 4.4). Bimodality becomes apparent as the temperature increases. At 25^0C in DD, two peaks are visible in the middle of the subjective day, while at 29^0C the peaks are completely separate and correspond to the subjective lights on and off positions (see Fig 4.4). Thus in DD both peaks appear to be temporally mobile. For wild type *D.melanogaster* however, all temperatures in LD conditions appear to induce peak activities at the light on and off transitions (Hamblen *et al.*, 1986, Newby and Jackson, 1991, Sawyer, 1996). This shows that the phase relationship between them can be altered by exogenous influences. In LD the per^S and per^L period mutants also have phase alteration in bimodal peak activity (Hamblen-Coyle *et al.*, 1992). In these mutants however it is the 'evening' (or second) peak which moves in relation to the morning peak which is apparently fixed by the lights on/off transition in LD (Hamblen-Coyle *et al.*, 1992). The activity is therefore initiated by a synchronization of the endogenous clock with the environment.

So which elements of the locomotor behaviour that we observe are environmentally flexible in origin, and which are more rigidly genetically programmed by the endogenous clock? The lack of a subjective 'lights on' peak for *D.pseudoobscura* locomotor activity in DD has been used to ascribe a mainly exogenous effect of light to this species dawn peak in LD (Engelman and Mack, 1978, Hamblen-Coyle *et al.*, 1992), and the term 'startle' response has been coined, emphasizing it's environmental origin (see Fig 6.1 point A). Indeed the 'lights on' transition in LD seems to affect *D.pseudoobscura* and *D.melanogaster* very similarly at 18^0C, providing a morning peak even though at this temperature, their free-running profiles are totally dissimilar (see Fig 4.6). This supports the idea of an exogenous effect of light generating these morning activity peak similarities. In addition *per* ^{01} flies also show this lights on response suggesting that a non-functional clock is permissive for this 'startle response' (Hamblen-Coyle *et al.*, 1992). Other research however indicates that the clock is involved in this response (see discussion below).

Mutations at other loci such as *disco* and *ebony*, cause alterations in the bimodal LD activity profile (Newby and Jackson, 1991, Hardin *et al.*, 1992). *disco* mutants have arrhythmic free-running activity, but in LD they can be entrained but show diminished morning peaks ('startle responses'). On the other hand *ebony* mutants show highly variable free running periods of activity but nevertheless show clear morning and evening peaks. Surprisingly, imposing LD conditions in *ebony* flies leads to a deterioration of rhythmicity. In both mutants therefore the endogenous oscillator appears to interfere with the photic input. This suggests that the startle response is not simply a locomotor response to direct photic input bypassing the clock (see point A, Fig 6.1), but rather due to a conflict of photic input with the clock machinery producing a defective output (point B, Fig 6.1), or a conflict between clock output and the photic signals that bypass the clock (point C, Fig 6.1). Interestingly the *ebony* mutants have perfectly normal eclosion rhythms (point D, Fig 6.1) which suggests that the clock is functioning normally. Unfortunately *disco* and *ebony* have visual impairments, and general neuronal disruption of input or output pathways cannot be ruled out as an explanation (Helfrich-Forster, 1998).

Fig 6.1 Speculative scheme connecting environment, clock and behaviour. (A) possible pathway bypassing the clock. (B) Altered clock generated information. (C) Interaction of environment and clock output. (D) Output from clock to eclosion. (E) Feedback from central clock to rhythmic processes in photoreceptor.



However two other more specific clock output mutants without obvious anatomical brain defects have been identified. Mutations in an RNA binding protein, LARK, affect eclosion but not locomotor activity (Newby and Jackson, 1996, Mcneil *et al.*, 1997). The *DCO* mutant codes for a catalytic subunit of cAMP-dependent protein kinase, and like *ebony* it gives the rhythmic eclosion and non-rhythmic locomotor phenotype (Majercak *et al.*, 1997). This cAMP-dependence suggests that the output pathway from the clock to locomotor activity is therefore likely to be available for modulation to affect the activity profile, possibly by photic information and temperature as well as clock function.

What then of other known *bone fide* clock mutants where we can assume, at least initially, no extraneous pleiotropic effects? tim^0 shows complete lack of startle effect (see Fig 7.4). This is the exact opposite of the assumed permissive per^{01} 'startle response', and suggests that the defective tim^0 clock machinery virtually blocks some information between exogenous photic input and locomotor output, and reinforces the observations made for the *ebony* mutants that the clock is affecting the startle response.

We can therefore conclude that the clock can either disrupt locomotor output via mutant clock function or by conflict with environmental input that either may bypass the

clock or act through it (see Fig 6.1). Interaction between the clock and these inputs combine at the locomotor output pathway to generate an activity profile that is appropriate for the time of day and the current environmental conditions. Is it possible for this clock-influenced modulation of activity to be observed? So far we have considered examples of mutants where severe locomotor disruption occurs in either DD (eg. *disco*), LD (eg. *ebony*) or both (eg. *per* ^{L} and *per* ^{s}). The observation of locomotor output in a range of clock genotypes which carry mild clock disturbances may be more informative.

The various (*Thr-Gly*) transformants have a spectrum of temperature inducible effects on clock function (Sawyer *et al.*, 1997; Chapter 5). Would the (*Thr-Gly*) transformants also show temperature dependent differences in an LD cycle? If so, then through dissection of their locomotor profiles, we should be able to investigate the relationship between clock function and environmental input in the production of locomotor activity. What we need then, are instances of profile alteration that are clearly identifiable, related to clock function via the different *per* genotypes, and can be shown to have been modified by the environment. The (*Thr-Gly*) transformants are ideally suited to this investigation as their known clock defects are temperature dependent and temperature alterations may induce novel activity patterns.

Therefore to further investigate the relationship between clock function and locomotor profile, I have analysed the $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_{20}$, $(Thr\text{-}Gly)_{1}$ and $\Delta(Thr\text{-}Gly)$ transformants for activity profile changes in response to temperature variation in DD and LD.

6.2 Methods

The construction of the transgenes used are fully described in Chapter 2, and are the $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_{20}$, $(Thr\text{-}Gly)_{1}$ and $\Delta(Thr\text{-}Gly)$ transformants used in Chapter 5. Manipulation of the flies, entrainment, locomotor activity and analysis was identical to that described in Chapter 4 for DD and LD activity profiles. As before, some of the data has proved to be non-informative when the 48 bins are compressed into 12 windows for statistical analysis. For this reason presentation of the 48 bin data is used and as in Chapter 4, 12 bin fragments of some circadian cycles are also analysed to provide more information.

6.3 Results

The 48 bin locomotor activity profiles of all the transformant lines at 18^0C and 29^0C in DD and LD are presented in Fig 6.2 and Fig 6.3. All genotypes have unimodal patterns of activity at 18^0C and bimodal patterns at 29^0C in DD. As previously shown for *D.melanogaster* (see Fig 4.2) the 18^0C DD unimodality appears to occur in the earlier part of the subjective day for the natural length [$(Thr\text{-}Gly)_{17,20}$] transformants whereas the deletion transformants appear to lose this feature. At 29^0C in DD all genotypes display bimodality where the morning peak occurs at subjective lights on. The evening peak for most lines is 2-3x more active than the morning peak. Although for all genotypes it appears to be close to subjective lights off, it is displaced ~ 2-2.5h into the subjective night phase in both lines of the $(Thr\text{-}Gly)_1$ genotype.

In LD at 18^0C, activity is concentrated within the light phase, with very little activity in the dark. Interestingly the $(Thr\text{-}Gly)_{17}$ genotype appears more bimodal for both lines at 18^0C in LD than the $(Thr\text{-}Gly)_{20}$, the apparent cause being the increased activity at the lights on and off transitions. Similarly $\Delta(Thr\text{-}Gly)$ appears more bimodal than $(Thr\text{-}Gly)_1$ but interline variation in activity profile is evident. At 29^0C in LD, all the genotypes have marked reduction in activity in the middle of the light phase with sharply defined peaks of activity close to the LD transitions. The $(Thr\text{-}Gly)_{17}$ genotype appears conspicuous for the strength of it's peak activities, and is the only genotype with higher evening peak activity in LD than DD for both lines. That peak activities are equivalent or lower in LD than in DD for most of the lines investigated, is relevant to the question of the 'startle' response and will be examined in the Discussion. Perhaps the most remarkable result in LD at 29^0C is the displacement of the $(Thr\text{-}Gly)_1$ evening peak into the dark phase for both lines, which reflects the same activity observed in DD.

Fig 6.4a illustrates the LD results for all the transformants at both temperatures plotted as genotype averages with associated standard errors. The 48 bin data has been changed to the 12 window format for statistical analysis and although this inevitably leads to some smoothing of the 48 bin data, the LD features outlined above are still visible. ANOVA with lines nested within genotypes reveals highly significant Temperature (T) * Window (W) ($p<<0.00001$) and significant Genotype (G) * T * W ($p<0.024$) interactions (see Fig 6.4b).

Fig 6.2 $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ transformants at at 18^0C and 29^0C in LD (Y-axis activity all plotted on the same scale)

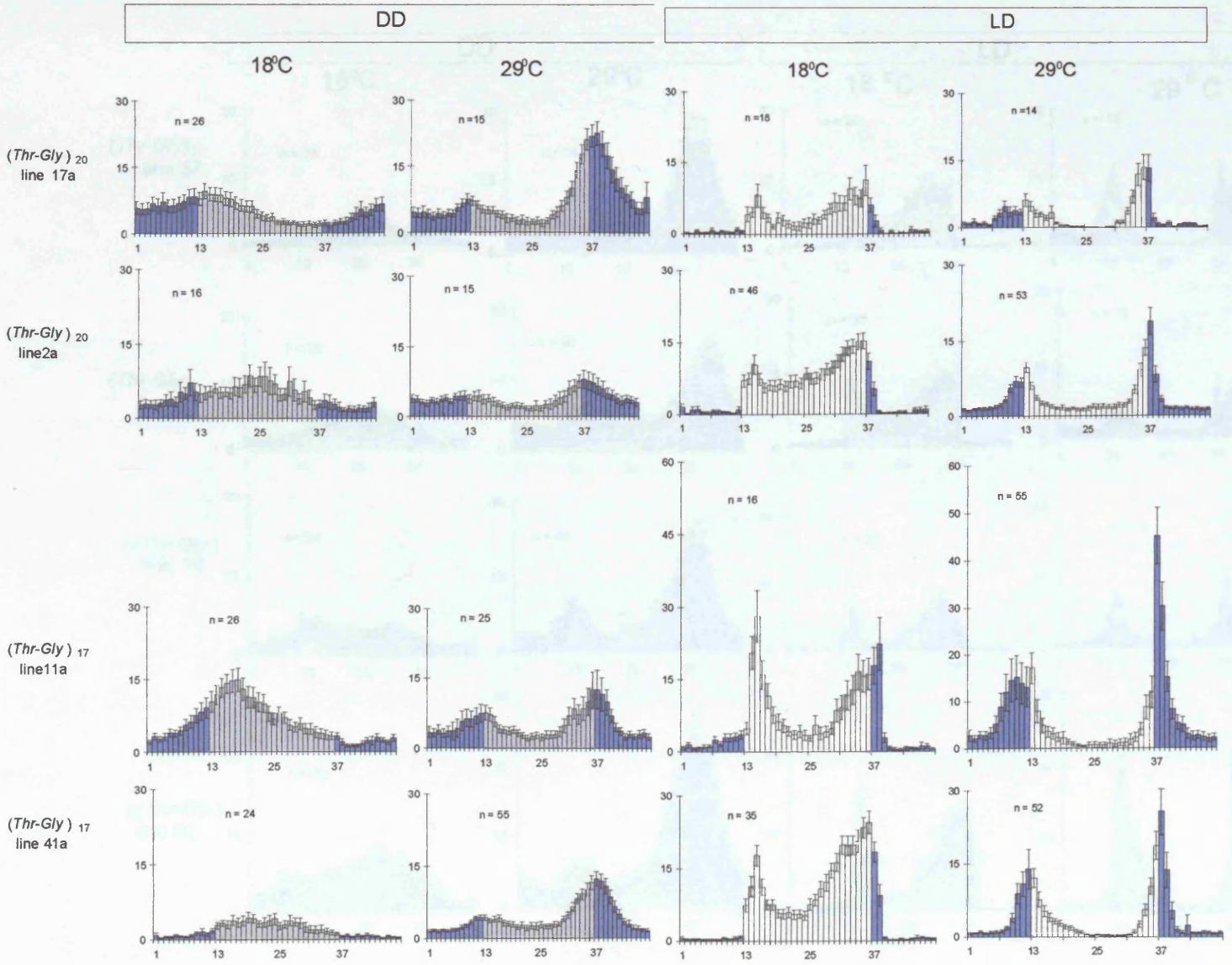


Fig 6.3 (*Thr-Gly*)$_1$ and Δ(*Thr-Gly*) transformants at At 18^{0}C and 29^{0}C in DD and LD (all same scale on y- axis)

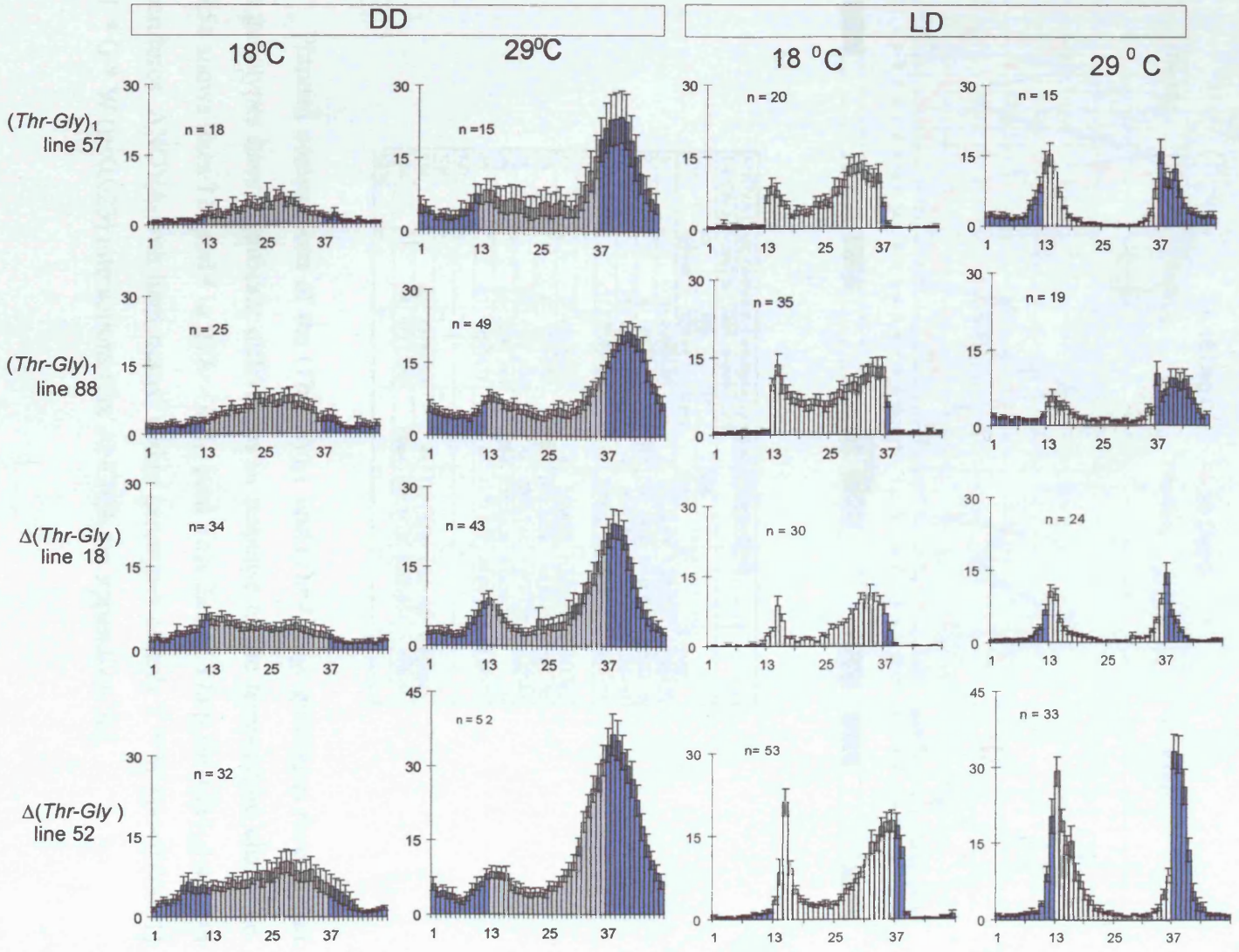
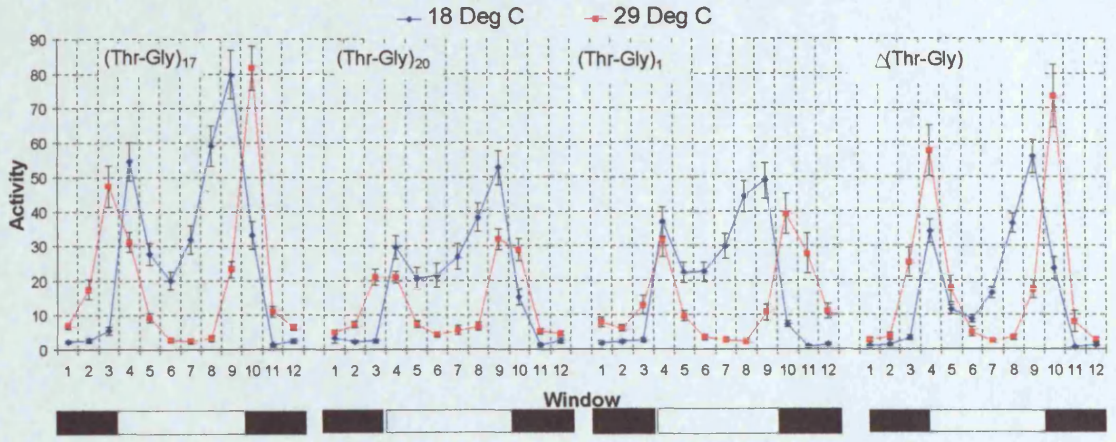


Fig 6.4a All transformants at 18 and 29^{0}C LD (unweighted genotype averages with standard errors) **b** ANOVA with lines nested within genotypes

a.

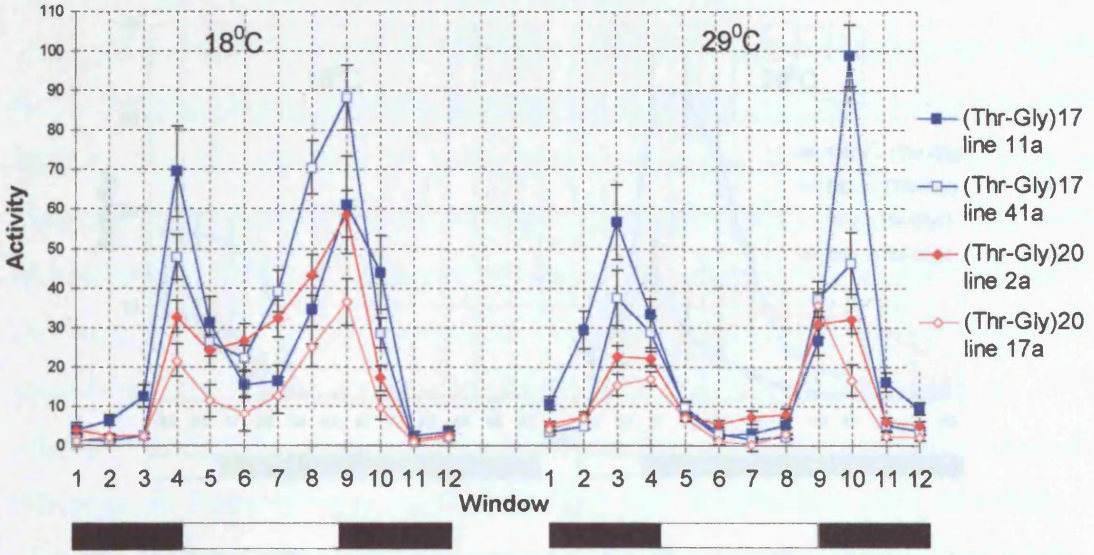


b.

ANOVA Lines nested in Genotype, Lines as random effects						
1-GENO, 2-LINE 3-TEMP, 4-WINDOW						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	3	23594.38	4	18629.54	1.266503	0.398429
2	4	18629.54	5964	538.3428	34.60535	1.33E-28
3	1	12902.04	4	5686.68	2.268817	0.206467
4	11	70201.21	44	3256.341	21.55831	3.26E-14
12	–	–	–	–	–	–
13	3	4642.221	4	5686.68	0.816332	0.548262
23	4	5686.68	5964	538.3428	10.5633	1.58E-08
14	33	4167.384	44	3256.341	1.279775	0.220183
24	44	3256.341	5964	538.3428	6.048824	0
34	11	46896.11	44	1777.368	26.38514	8.2E-16
123	–	–	–	–	–	–
124	–	–	–	–	–	–
134	33	3379.074	44	1777.368	1.901167	**0.023361**
234	44	1777.368	5964	538.3428	3.301554	1.47E-12
1234	–	–	–	–	–	–

Planned comparisons of the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ genotypes reveals that these genotypes have significant differences in response to the temperature challenge. Fig 6.5a shows lines 11a and 41a [$(Thr\text{-}Gly)_{17}$]and lines 2a and 17a [$(Thr\text{-}Gly)_{20}$] at both temperatures. ANOVA with lines nested within genotypes reveals T * W ($p<<0.00001$) and T * G * W ($p<0.023$) interactions (For ANOVA see appendix 6.5a).

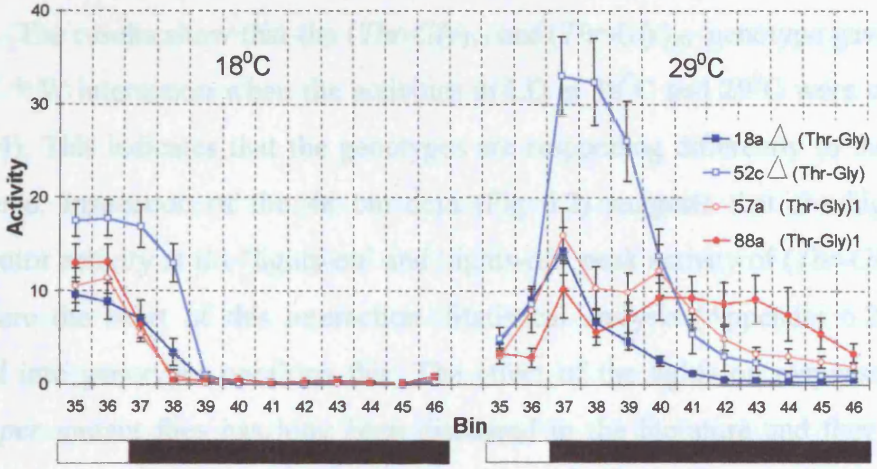
Fig 6.5 Activity of the $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ transformants, LD 12:12



Post hoc Newman-Keuls tests on genotypes (with lines pooled) indicate that for windows 4,8,9 and 10 at 18°C and windows 3 and 10 at 29°C the $(Thr\text{-}Gly)_{17}$ genotype is significantly more active than the $(Thr\text{-}Gly)_{20}$ genotype (at least $p<0.003$ see appendix 6.5b)

Further planned comparisons were then carried out on the $(Thr\text{-}Gly)_1$ and $\Delta(Thr\text{-}Gly)$ deletion transformants in LD conditions in the expectation that the observed displacement of the $(Thr\text{-}Gly)_1$ evening peak at 29°C would reveal additional T * G * W interactions. Surprisingly ANOVA, with lines nested within genotypes on 12 window data for the two genotypes in LD at 18 and 29°C shows no significant T * G * W interaction between the genotypes (see Appendix 6.6a). I can only conclude that the variation within the lines (particularly the $\Delta(Thr\text{-}Gly)$ lines) is masking the apparent intergenotype difference. I therefore performed an ANOVA on the 12 bins (35-46) that seem to show the large variation in LD nightime activity at 29°C only (see Fig 6.6), in the expectation that the finer resolution analysis would give significant results. However again no significant G * W or G * T * W interaction was observed (for ANOVA table see appendix 6.6b)

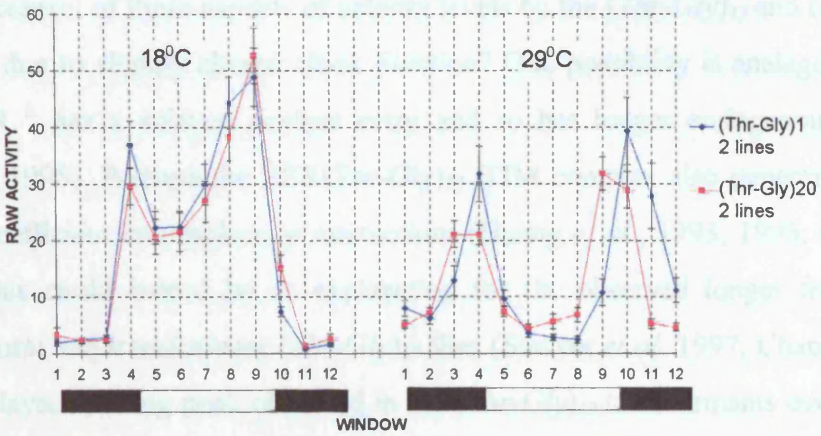
Fig 6.6 Activity profiles in LD 12:12 of the $(Thr\text{-}Gly)_1$ [lines 57a and 88a] and $\Delta(Thr\text{-}Gly)$ [lines 18a and 52c] transformants bins 35-46 only.



To identify differences between the deletion transformants and the 'natural variant' transformants, $(Thr\text{-}Gly)_1$ and $\Delta(Thr\text{-}Gly)$ genotypes were individually compared to the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ genotypes. No significant G * W or G * T * W interactions are observed for $\Delta(Thr\text{-}Gly)$ with either $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ (see Appendix 6.7a).

However Fig 6.7 reveals that if the $(Thr\text{-}Gly)_1$ is compared to $(Thr\text{-}Gly)_{20}$ data then significant G * W ($p<0.024$) and G * T * W ($p<0.04$) interactions are seen. [Lines are pooled within genotypes for clarity of graphical representation, however AVOVA (see Appendix 6.7b) has lines nested within genotype]

Fig 6.7 $(Thr\text{-}Gly)_1$ and $(Thr\text{-}Gly)_{20}$ activity at 18°C and 29°C (lines pooled within genotypes).



In contrast the $(Thr\text{-}Gly)_1$ v $(Thr\text{-}Gly)_{17}$ comparisons are non significant for these interactions (see appendix 6.7c for graph and ANOVA).

6.4 Discussion

The results show that the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ genotype gave a significant G * T * W interaction when the activities in LD at 18^0C and 29^0C were compared (see Fig 6.4). This indicates that the genotypes are responding differently to the temperature challenge. Inspection of the 48 bin data (Fig 6.2) suggests that the higher levels of locomotor activity at the 'lights-on' and 'lights-off' peak activity of $(Thr\text{-}Gly)_{17}$ at 18 and 29^0C are the heart of this interaction. Statistical analysis (Appendix 6.2) of the data pooled into genotypes confirms this. The effect of the lights on response in per^{01} and other *per* mutant flies has long been discussed in the literature and they are generally described as representing a 'startle response' for the morning peak, and being clock mediated for the evening peak (Saunders *et al.*, 1992, Hamblen-Coyle *et al.*, 1992,). My results however indicate that *per* is involved in the modulation of activity levels of both 'morning' and 'evening' peaks.

A recent suggestion relating clock function to activity levels is that "one aspect of strong clock function is to suppress the rapid increases in locomotion that accompany these environmental transitions" (Stanewsky *et al* 1997a, p.695 lines 1-3). If this is correct then the control of activity levels in these acute responses at both temperatures, shown here for the $(Thr\text{-}Gly)_{20}$ (see Fig 6.1) may mean that the $(Thr\text{-}Gly)_{20}$ transformant has a 'stronger' clock function than the $(Thr\text{-}Gly)_{17}$. This idea is of course not without precedent as the natural and transformant $(Thr\text{-}Gly)_{20}$ flies have been shown to have more efficient free-running temperature compensation than the corresponding $(Thr\text{-}Gly)_{17}$ carriers (Sawyer *et al.*,1997, and see Fig 5.7).

Is the control of these aspects of activity levels by the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ transformants due to slightly altered clock function? One possibility is analagous to that by which PER^L has a delayed nuclear entry and so has longer endogenous rhythms (Curtin *et al.*, 1995). Perhaps the PER$(Thr\text{-}Gly)_{17}$/TIM complex also generates a slight delay due to inefficient intermolecular interactions (Huang *et al.*, 1993, 1995; Gekakis *et al.*, 1995). This could indeed be an explanation for the observed longer free-running periods of natural and transformant $(Thr\text{-}Gly)_{17}$ flies (Sawyer *et al.* 1997, Chapter 5), and the slightly delayed evening peak observed in the $(Thr\text{-}Gly)_{17}$ transformants evening peak (Fig 6.2 and Fig 6.5) would support this explanation. In addition it is worth noting that within the $\Delta(Thr\text{-}Gly)$ and $(Thr\text{-}Gly)_1$ genotypes, the line that shows the larger

displacement of activity into the dark phase also has the longest free-running period (Fig 6.3 and Chapter 5 table 1). This means that at an intragenotype level, variation in free-running period may be related to the phase of the 'evening peak', suggesting that some endogenous clock variation does indeed affect at least the phase of peak locomotor activity.

How could this altered endogenous rhythm relate to the postulated (Stanewsky *et al.*, 1997) and observed (Fig 6.2 and Fig 6.5) lack of locomotor suppression associated with these 'weaker' clock components. Perhaps if a specific genotype has an endogenous rhythm more in phase with the exogenous 24h LD cycle than another, then that PER product may be more available for reinforcing its activity because of the synchrony of internal and external factors. This could explain the more distinctive response of the Δ(*Thr-Gly*) genotype to the LD transitions in comparison with the $(Thr\text{-}Gly)_1$, as the free-running patterns for the Δ(*Thr-Gly*) are more closely associated with the LD transitions (Fig 6.3). However visual inspection of the DD data for $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ does not suggest any obvious differences in the timing of their DD activity peaks that could underly the even more exaggerated differences in LD activity levels for these two genotypes. And in any case in the $(Thr\text{-}Gly)_{20}$ (which apparently suppresses peak activities), free-running activity peaks coincide very well with the timing of the LD transitions.

The lack of locomotor profile uniformity between the lines for the $(Thr\text{-}Gly)_1$ and Δ(*Thr-Gly*) genotype is awkward for comparisons, and partly undermines the statistical analysis. However at a qualitative level, $(Thr\text{-}Gly)_1$ clearly has a different evening peak response when compared to Δ(Thr-Gly) [see Fig 6.3 and Fig 6.4]. In both DD and LD $(Thr\text{-}Gly)_1$ has a larger time gap between the morning and evening peaks. Δ(Thr-Gly) in contrast, although slightly delayed has evening peaks more closely related to the subjective (DD) and real (LD) light/dark transitions (Fig 6.3). This 'nightime' peak of $(Thr\text{-}Gly)_1$ in both free run and LD (Fig 6.3) indicates a clock function independent of exogenous influences. In DD the evening activity displacement shows that the clock was not entrained by the LD cycle prior to free-run, and in LD the activity is clearly not just a response to the photic transition to lights off. The similar free-running periods in Δ(*Thr-Gly*) and $(Thr\text{-}Gly)_1$ at 29^0C (see Fig 5.5), but clearly altered evening peak phases, indicates that the relationship between period length and activity pattern as in the per^L analogy, clearly does not hold for all genetic variations in *per*. *per* thus has at least two

apparently separate functions, first the control of endogenous periodicity which at some level is connected to evening peak phase, and second, a separate functional involvement in modulation of locomotor activity pattern.

If we imagine that some effects on activity patterns are not directly related to clock function, then what other possibilities can we consider? An alternative explanation is that *per* is pleiotropic and a *per* variant with an assumed 'weaker' function has less ability to suppress locomotor activity in a dynamic physiological sense, rather than via clock function (see also Chapter 7). This ability could take the form of either direct physiological interaction of PER with other factors at the cellular level, conceivably in the nucleus. *per* is assumed to be involved in transcriptional control of clock controlled genes (*ccg's*)(Van Gelder and Krasnow 1996), and *per* variants could theoretically determine alterations of *ccg* transcripts which may subsequently affect locomotor responses to exogenous influences. However we know that *per* is related to evening peak phase distribution (Hamblen-Coyle *et al.* 1992, Gekakis *et al.*, 1995 and Fig 6.3), and to the species-specific and genotype-specific control of activity patterns (Chapter 4 and Fig 6.5). Yet these activities appear not to be directly related to the free-running locomotor period and by association not to endogenous rhythmicity. In addition the temporal distribution of PER is apparently related to locomotor activity levels in the light phase. High levels of PER occur at the beginning of the light phase at ZT 0-4 and also significant amounts have also accumulated at ZT 10-12 (Edery *et al.*, 1994b; Rutila *et al.*,1996) when most locomotor activity takes place (see Hamblen-Coyle *et al.*, 1992 and Figs 6.2, 4.2 and 4.6). After reduction of PER to very low levels soon after the lights on transition, the build-up of PER starts at around ZT8 in LD (Edery *et al.*, 1994b; Rutila *et al.*, 1996), when activity profiles also show recovery of activity levels from the midday low point (see Hamblen-Coyle *et al.*, 1992 and Figs 6.2, 4.2 and 4.6). If we assume that some clock controlled protein directly initiates and maintains activity levels, then we also have to assume that it is active at phases very similar to the appearance of PER. The most parsimonious explanation for the observed effects is that activity levels are directly affected by PER acting in concert with exogenous influences to control activity levels.

Finally all genotypes show an obvious heat induced 'siesta' effect (Sawyer 1996) where the activity increase after midday is reduced in relation to the activity at the LD transitions. From an ecological perspective the decrease in activity in the middle of the day may have adaptive value. That the $(Thr\text{-}Gly)_1$ effect on activity profile is

distinguishable from the $(Thr\text{-}Gly)_{20}$ profile but not the $(Thr\text{-}Gly)_{17}$ profile yet the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ can be distinguished statistically, indicates (as far as the current paradigm is concerned) that a very small alteration in the Thr-Gly region can produce similar order changes in the organism's ability to respond to environmental stimuli as the drastic reduction in the repetitive region. This may then be an indication of how *D.melanogaster* benefits from the small natural variations in the extended repeat under natural conditions, which are more like laboratory LD conditions than DD.

However before one gets carried away with evolutionary/ecological interpretations it is important to underscore that natural conditions are not faithfully mimicked by laboratory LD conditions. Consequently any differences between these (*Thr-Gly*) genotypes may be considerably amplified in these experiments. Nevertheless these subtle differences between (*Thr-Gly*) variants allow some speculation as to which elements in the LD behavioural repertoire are mediated via the clock and which are simply immediate responses to the environment. These results suggest that PER may also regulate change in the locomotor output directly. This will be persued in the next Chapter.

Chapter 7: Investigating the potential pleiotropic effects of *per*

7.1 Introduction

The questions raised in Chapter 6 and subsequent experimental results provide evidence for a complex interaction of exogenous and endogenous factors in determining the locomotor activity profile. PER was proposed as a direct or indirect factor in the control of this phenotype. To further investigate the question of how the clock or it's components mediate exogenous influences and locomotor activity, another experiment was devised. Instead of providing an LD cycle at different temperatures, light and temperature transitions were dislocated in time to provide a more complex package of exogenous inputs. The basic framework of the LD 12:12 cycle needed to be maintained for two reasons: First I wished to extend the information already gained with Thr-Gly variants (Chapter 6) and draw on the extensive body of work describing LD activity cycles in mutant flies (Hamblen-Coyle *et el.*, 1992). Second, light is thought to be the most important entraining factor (Marrus *et el.*, 1996) and *per* mRNA cycling is locked to lights off (Qiu and Harding, 1996) although this has now been questioned (Liu *et al.*, 1997). Therefore to persue investigation outside the LD 12:12 paradigm was likely to generate results that might be difficult to interpret within the current framework of knowledge, in terms of locomotor responses and the underlying molecular cycles that may be controlling this activity

So how can temperature variation be introduced and yet leave the LD cycle fundamentally intact? If a brief temperature pulse in the middle of the light phase caused an acute startle response similar to the one proposed for the lights on transition, it would be visible and not likely to be confused with the normal LD bimodal pattern where the peak activity is at the LD transitions. Secondly if the temperature pulse also affects the subsequent activity peaks during the LD transitions, then clearly this would not be attributable to 'startle' responses. The use of this approach therefore gives us the opportunity to investigate the impact of acute exogenous influences other than the LD transitions and to examine their immediate and longer term consequences with reference to elements of clock function.

This experiment also has an ecological perspective. In the natural situation, at least in Europe, early and late light phases and the dark phase are generally cool, whereas in the middle of the day temperatures can rise rapidly, particularly during summers in southern

Europe (Hulme *et el.*,1995). Thus in addition to the response to average mean temperature (Sawyer *et al.*,1997 and Chapter 6), the ability to provide an appropriate response to within day temperature variation could also be important to the clock phenotype. It seemed to me that artificial creation of this acute temperature rise within an otherwise cool LD cycle in the laboratory would allow us to investigate the responses needed to cope with the variation in these climate fluctuations.

The question of what 'pulse' temperature and what 'background' temperature to use was influenced by the results in Chapter 5. The data shows that beyond 25^{0}C, *D.melanogaster* temperature compensation improves for natural (*Thr-Gly*) length variants and for the less radical $(Thr\text{-}Gly)_1$ deletion transformants, in that periods become closer to 24h. Δ(*Thr-Gly*) and per^L (Ewer *et al.*, 1990) flies do not show this effect. (see Chapter 5 for a fuller explanation of these phenomena). The possibility therefore that two or more endogenous processes are involved in temperature compensation (one visible between 18-25^{0}C and the other 25-29^{0}C, also discussed in Chapter 5) was likely to confuse the interpretation of any results if a pulse temperature higher than 25^{0}C was used. For the background temperature, 18^{0}C LD was a regime that had already been used extensively by myself and my colleague L.Sawyer to investigate natural and transformant (*Thr-Gly*) variants, as a minimum temperature for our experiments (Chapter 5 and Sawyer, 1996), and so for comparison was ideal. Initially then, the core experiment was to test the flies in an 18^{0}C LD regime where a 25^{0}C 'heat pulse' was introduced for 2h at ZT 5-7, which corresponds to the middle of the light phase.

This experiment was extended to provide its own control conditions of an LD cycle at the background temperature before and after the heat-pulse was introduced. Each of these conditions would be maintained for a number of days to ensure any activity alterations were not transient changes. (For experimental details see Materials and Methods 6.2). Results indicating specific heat pulse responses were observed, and to further investigate these responses the heat pulse experiments were subsequently extended on a limited number of genotypes, by increasing the duration of the heat phase by a further 2 hours (ZT 5-9).

In order to investigate links between temperature and the known clock components, a second experiment was devised where an identical heat pulse would coincide with high levels of PER and TIM. In the middle of the day PER and TIM levels are low while in the middle of the dark phase they are high (Edery *et al.*, 1994b, Rosbash

et al., 1996, Rutila *et al.*, 1996). At ZT 18-20 PER and TIM are also thought to be entering the nucleus as a heterodimer (Hunter-Ensor *et al.*, 1996, Rosbash *et al.*, 1996). Alterations in PER concentration (Edery *et al.*, 1994b), phosphorylation (Rutila *et al.*, 1996) or any other factor that may alter the kinetics of nuclear entry would be predicted to cause phase alteration in the PER/TIM negative feedback (Marrus *et al.*, 1996), and in the regulation of putative downstream clock controlled genes (Van Gelder and Krasnow, 1996). In addition, heat pulses, albeit of 37^0C, at ZT 15 in the night phase, are known to produce large reductions in PER concentration which are replaced over a period of 1-2h (Sidote *et al.*, 1998, A. Rogers pers.comm.) and produce phase shifts in the production and phosphorylation of these clock components on a similar time scale (Sidote *et al.*, 1998). This temporal alteration of PER and TIM abundance is reported to be stable for at least two days and to be associated with phase shifts in locomotor behaviour (Edery *et al.*, 1994a; Sidote *et al.*,1998). A second experiment was therefore undertaken, using the same background and pulse temperature (18 and 25^0C respectively) within an LD cycle as before but with the heat pulse at ZT 15-17 during the rise in cytosolic levels of PER and TIM.

Interestingly with a 37^0C pulse at ZT 21.5, a smaller reduction in PER levels is seen than with the 37^0C pulse at ZT 15. The basis for this phenomenon is either more rapid replacement of PER (Sidote *et al.*, 1998), or possibly heat induced resistance to degradation (A. Rogers pers.comm.). This suggests that some physiological mechanism is more comprehensively ensuring high PER concentrations at this later phase. At ZT 21.5 PER is predicted to be primarily nuclear in *D.melanogaster* and of high molecular weight (Edery *et al.*, 1994b). It may be that at this phase PER is sequestered within the nucleus and protected from degradation. It also may be more than coincidence, that this phase is associated with the anticipatory rise in locomotor activity prior to lights on and prior to the LD transition when high levels of activity are observed (Hamblen-Coyle *et al.*, 1992).

The experiments outlined will therefore provide the opportunity to investigate the effect on locomotor activity of a temperature pulse at times of low and high concentrations of PER and TIM in the wild type and various *per* mutants.

7.2 Materials and methods

Fly stocks

Two lines for each of the major natural *melanogaster* (*Thr-Gly*) variants (*Thr-Gly*)$_{17}$, (*Thr-Gly*)$_{20}$ and (*Thr-Gly*)$_{23}$ homozygous lines (previously isolated from two locations in Northern Italy, San Mattino and Conselve) were reconfirmed by PCR amplification and comparison against sequenced individuals on high concentration agarose gels. [Primers in highly conserved regions on either side of the variable repeat (Sawyer *et al.*,1997) produce DNA fragments that vary in size directly due to (*Thr-Gly*) length variation see Fig 2.6] . The (*Thr-Gly*)$_{20}$, (*Thr-Gly*)$_{17}$ (*Thr-Gly*)$_{1}$, Δ(*Thr-Gly*) and *L2* transgenes are described in Chapter 5 methods. *per* ^{S}, *per* ^{L} and *per* ^{01} and *tim* ^{0} flies were also used.

Activity analysis of combined light/dark and temperature cycles

As for previous activity analyses, male transformants were crossed to *per* ^{01} virgin females to transfer the transgenes to an arrhythmic background, and the male F1 offspring were collected at 1-7 days after eclosion for activity analysis. Each experiment involved flies in LD at 18^{0}C for 5 days (LD condition) with a 25^{0}C heat pulse for either 2 or 4h at ZT5-7 or 2h at ZT15-17 (heat-pulse condition) followed by 5 days in LD at 18^{0}C with no heat pulse (recovery condition). The first day's data for each condition was removed to avoid contamination with transition effects incurred by changes from one condition to another. The pre-heat pulse LD control data were either obtained directly for this experiment [*tim* ^{0} *per* ^{S}, *per* ^{L} and *per* ^{01} and the natural variants (*Thr-Gly*)$_{17}$,(*Thr-Gly*)$_{20}$ and (*Thr-Gly*)$_{23}$] or from previous experiments where the 18^{0}C LD condition had been used [transformant (*Thr-Gly*)$_{20}$, (*Thr-Gly*)$_{17}$ (*Thr-Gly*)$_{1}$, Δ(*Thr-Gly*) and *L2*]. The LD data for the *L2* transformant was provided by Parkinson (1997).

Flies analysed were given no preselection other than for hyper or hypo activity or early death (see Chapter 2). Table 7.1 shows the exact number of flies tested for the LD, ZT 5-7 heatpulse, ZT 5-7 recovery, ZT 15-17 heat pulse and ZT 15-17 recovery conditions.

Table 7.1 Numbers of flies tested at each condition

Genotype	Line	LD	Heat Pulse ZT.5-7	Recovery ZT.5-7	Heat Pulse ZT.5-9	Heat Pulse ZT.15-17	Recovery ZT.15-17
Natural variants							
$(Thr\text{-}Gly)_{17}$	Con78	21	34	34	14	27	
	Smat32	22	21	22		27	
$(Thr\text{-}Gly)_{20}$	Smat27	18	23	24		25	
	Smat35	21	26	27	27	24	
$(Thr\text{-}Gly)_{23}$	Con2	19	24	24	28	28	
	Smat83	14	24	24		24	
Mutants							
per^0		22	24	25	22	26	
per^s		8	23	23	14	29	23
per^L		20	20	20	21	19	20
tim^0		16	30	30		18	
Transformants							
$(Thr\text{-}Gly)_{17}$	11	16	27	27		24	24
	41a	35	27	27		26	26
$(Thr\text{-}Gly)_{20}$	2a	46	20	20		29	18
	17a	17	29	29		21	21
$\Delta(Thr\text{-}Gly)$	18a	30	40	28		23	23
	52c	53	44			27	27
$(Thr\text{-}Gly)_1$	44a		37			28	28
	57a						
	88a	35	32	24		26	26
L.2	19c	23	27	27		34	
	51a		24	24		22	

For all conditions a mean activity for each of 48, 30 min bins across 4 days was calculated by superimposing the data for each of the days, giving an average mean activity per bin for each fly. The genotype means were then calculated per bin from the individual fly means and the 48 bins were subsequently divided into twelve 2h windows, starting from ZT 18, and a mean activity for each 2h window was calculated, exactly as previously described in the methods of Chapter 4. For statistical analysis the maximum number of data samples that could be used to represent one circadian cycle was twelve.

In some cases only one line or replicate was available for one or other of the conditions, and thus the data was pooled into genotypes. Where at least two replicates were tested in all conditions, a more conservative analysis was used where the lines or replicates were nested within genotype (Winer, 1971)

7.3 Results

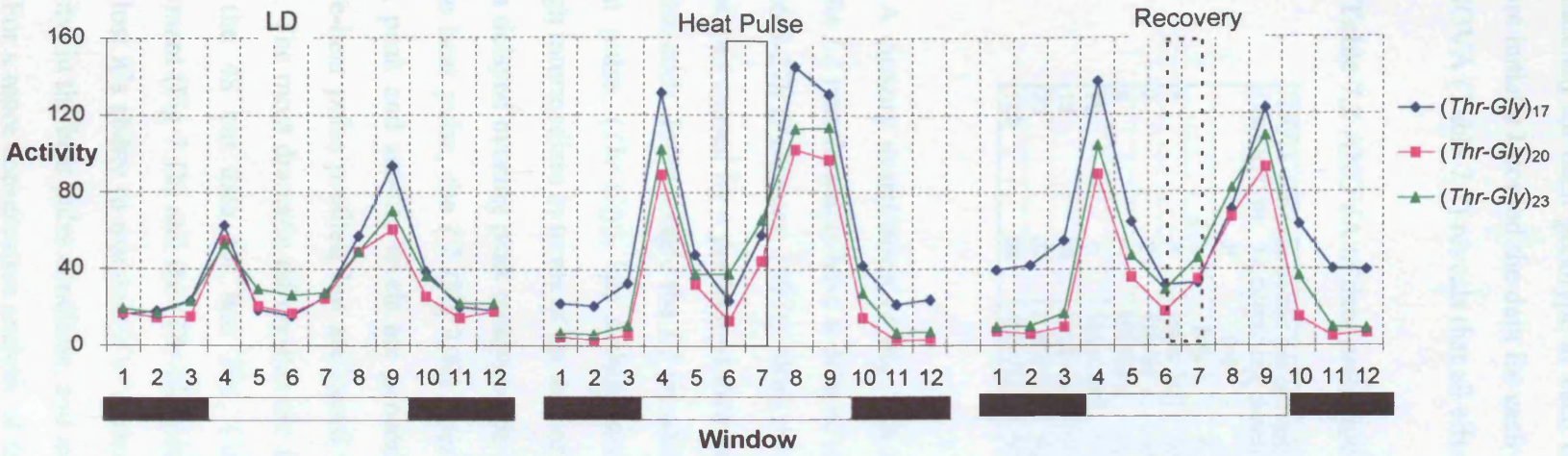
Initially the three most common natural Thr-Gly variants $(Thr\text{-}Gly)_{17,\ 20,\ 23}$ were investigated using two replicate lines for each repeat length. Fig 7 1a shows the effect of the three conditions (LD, heat pulse and recovery) on the behaviour of the natural variants. The nested ANOVA (appendix 7.1) gives highly significant main effects for Windows (W) and Lines (L) [Both $p<<0.00001$] but not genotypes (G) or Conditions (C). The relevant interactions G * C * W and G * W are non-significant, revealing that the different variants react to the conditions in the same way. The C * W interaction is significant ($p<<0.00001$) showing the overall effect of introducing the heat pulse into an LD cycle.

One of the effects of the heat pulse is that the evening peak appears broader in all three genotypes. This is because the flies become active after the heat pulse ends (see raised activity in windows 7 and 8 for heat pulse condition), giving the broader peak which is observed clearly when the 30 min bins are examined [for example see Fig 7.1b $(Thr\text{-}Gly)_{20}$ line SM 35 only]. Unfortunately the smoothing of the data generated by the use of 2h windows (due to limitations of the statistical package) conceals some acute responses to heat and light. I can conclude from this first experiment that attempting to mimic an environmental rise in temperature during the middle of the day does not lead to any detectable differential response (a G * W or a G * C * W interaction) for the (*Thr-Gly*) variants tested using this analysis. It is worth noting however that the $(Thr\text{-}Gly)_{17}$ variants appear to have marginally higher activity peaks at all conditions (see also Chapter 6).

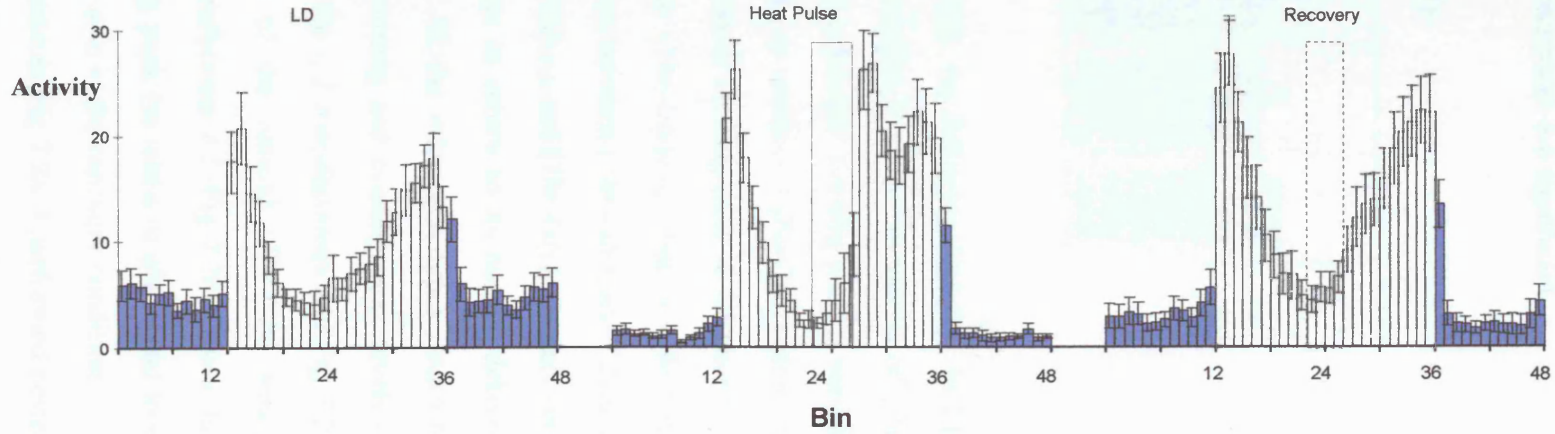
Next I investigated whether a larger difference in (*Thr-Gly*) length might generate behavioural changes in these conditions. I performed the same experiment with the transformants $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_{20}$, $(Thr\text{-}Gly)_{1}$, $\Delta(Thr\text{-}Gly)$, and *L2* I initially examined the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ transformants (2 lines each). ANOVA was performed with lines nested within genotype for the three conditions, and again there were no significant G * C * W interactions (Appendix 7.2a). I therefore pooled these four lines thus creating a transformant data set comprised of the two major naturally occuring (*Thr-Gly*) length variants. I compared the behaviour of these transformants with those carrying one copy of the $(Thr\text{-}Gly)_{1}$, $\Delta(Thr\text{-}Gly)$ and *L2* transgenes.

Fig 7.1a Locomotor activity of natural variants in three LD conditions. **b** $(Thr\text{-}Gly)_{20}$ (line Sm35) 48 bins at three conditions

a.



b.



The experiment was not balanced completely in that I was unable to perform the tests in all three conditions for two lines of each transformant. Approximately 30 flies were examined for each genotype at each condition (for exact numbers see Table 7.1). Therefore initially I pooled the data for each genotype, and Fig 7.2a illustrates the results. The ANOVA (Table 7.2) reveals that all effects and interactions are significant.

Table 7.2 ANOVA of data presented in Fig 7.2a

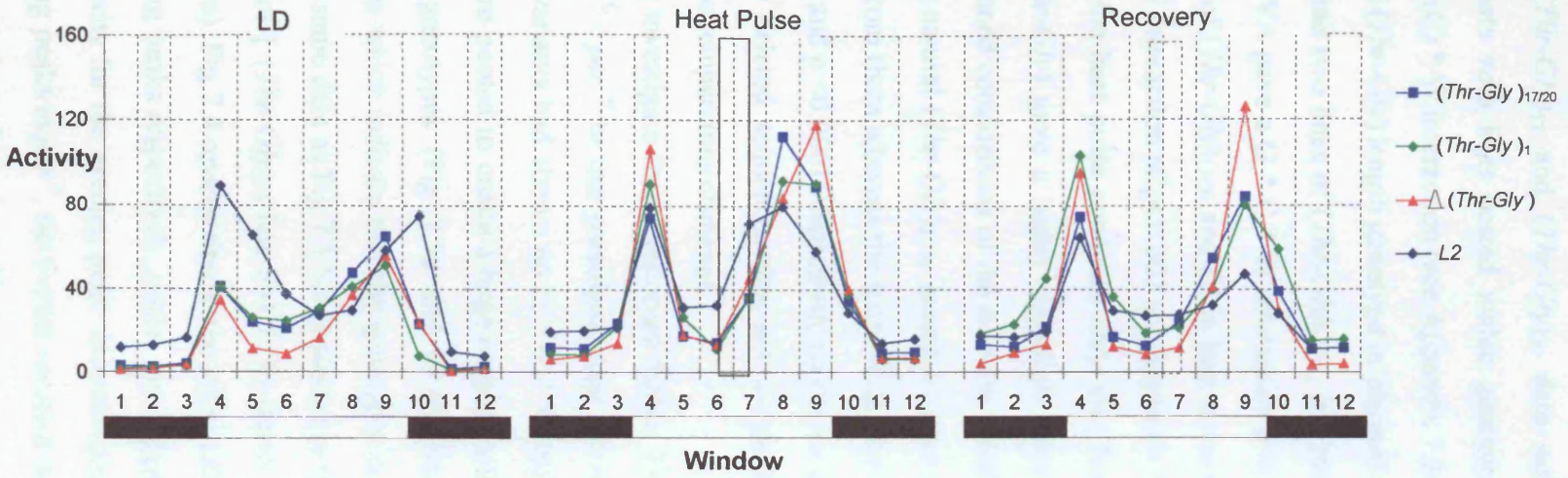
ANOVA with lines pooled into genotypes						
1-GENOTYPE, 2-CONDITION, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	3	10981.62	9072	1029.69	10.66498	**0.0000**
2	2	128767.4	9072	1029.69	125.0545	**0.0000**
3	11	357894.8	9072	1029.69	347.5754	**0.0000**
12	6	18610.26	9072	1029.69	18.07365	**0.0000**
13	33	11294.69	9072	1029.69	10.96902	**0.0000**
23	22	25607.87	9072	1029.69	24.8695	**0.0000**
123	66	4227.626	9072	1029.69	4.105728	**0.0000**

A cursory examination of Fig 7.2a and b reveals the following features: In LD 12:12 the *L2* transformants have a delayed evening peak [the free running period of *L2*at 18^0 is 26.5-27h (Parkinson 1997)]. With the heat pulse a broader evening peak is seen in all genotypes caused by a precocious increase in activity in window 7 (Fig 7.2a) after the heat pulse ends. Interestingly the *L2* transformants delayed evening peak is advanced by the heat pulse. $(Thr\text{-}Gly)_1$ has activity similar to the $(Thr\text{-}Gly)_{17/20}$ but Δ (*Thr-Gly*) [although intermediate in terms of the size of the deletion between $(Thr\text{-}Gly)_1$ and *L2*] now shows a delayed evening peak compared to the $(Thr\text{-}Gly)_{17/20}$ and $(Thr\text{-}Gly)_1$. In recovery after the heat pulse, the *L2* (Fig.7.2b) does not appear to return to its normal delayed evening peak and activity levels are generally lower. All the other genotypes return to their pre-heat pulse profiles, but are raised in both morning and evening peak levels of activity. The most dramatic differences are shown by the *L.2* transformants and Fig 7.2b shows the 48 bin data for line 19c. Comparison of the natural (*Thr-Gly*) length transformant (Fig 7.1b) and the (*Thr-Gly*) deletion transformant *L2* (Fig 7.2b) shows that *L2* has lost it's ability to maintain it's 'normal' evening peak (in terms of phase and level of activity) in the heat pulse condition, and intriguingly also in the recovery condition.

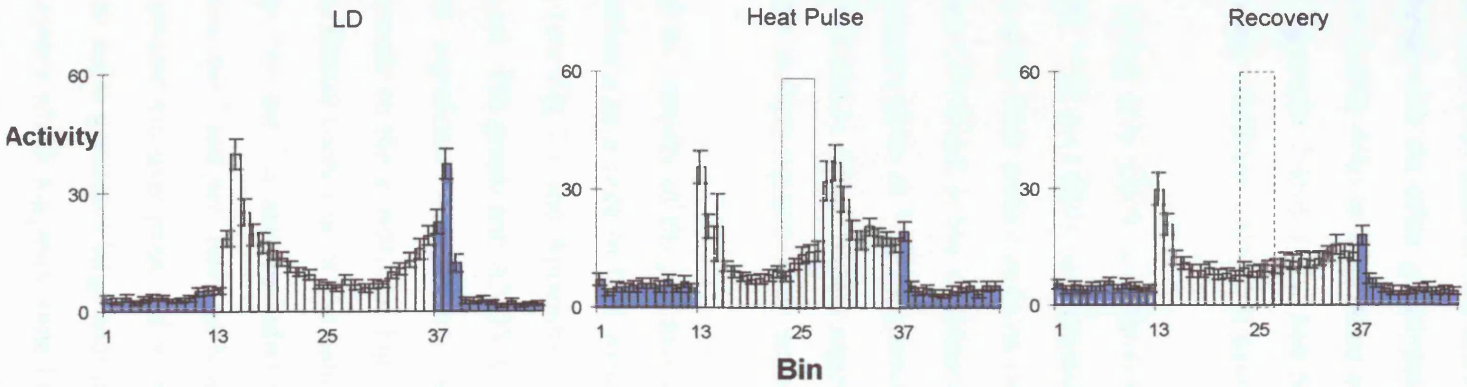
For a more conservative analysis of the data shown in Fig 7.2a, I performed nested ANOVAs on those data for which I had two lines within each genotype. The $(Thr\text{-}Gly)_{17}$ [two lines] and $(Thr\text{-}Gly)_{20}$ [two lines] genotypes were tested seperately in a nested

FIG 7.2 a. Locomotor activity in LD 12:12 of transformants $(Thr\text{-}Gly)_{17/20}$ [pooled], $\Delta(Thr\text{-}Gly)$, $(Thr\text{-}Gly)_1$ and deletion *L2* (all data pooled). **b.** *L2* (line 19c) in 48 bins.

a.



b.



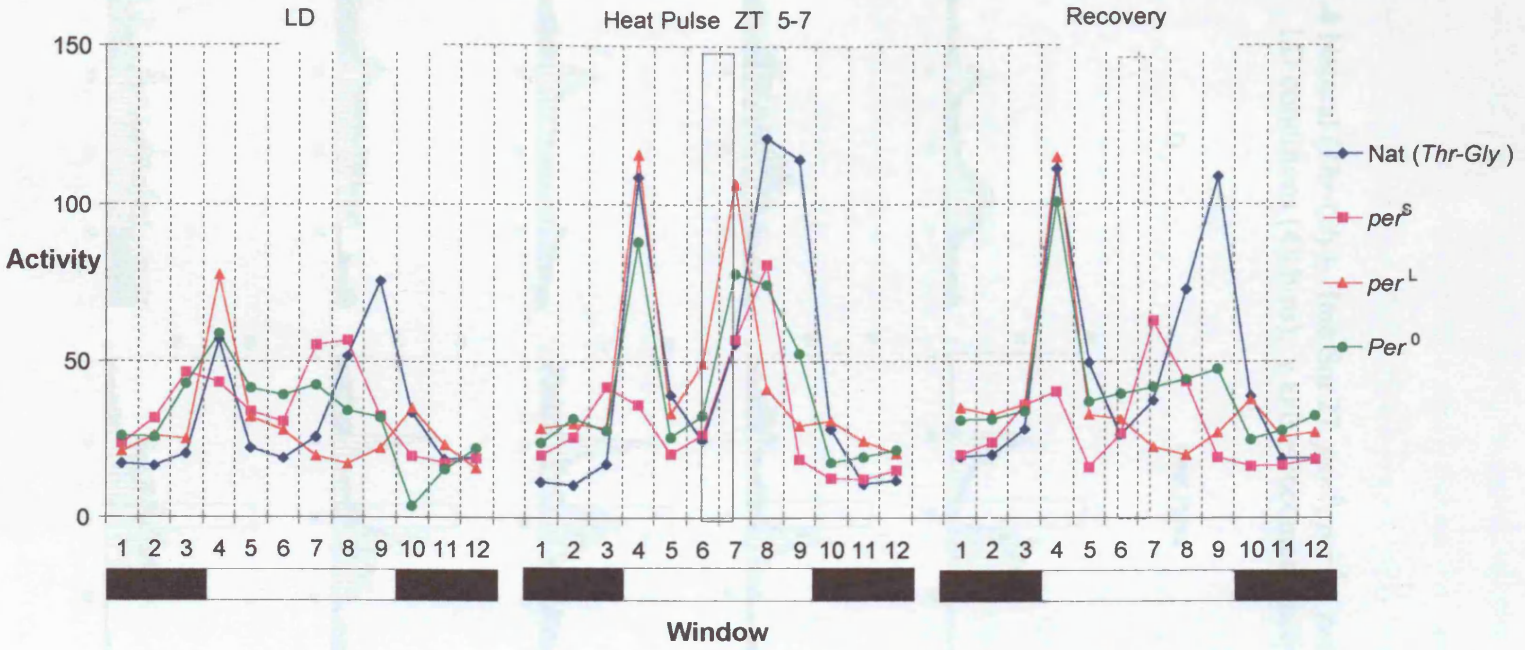
ANOVA (analysis not shown) and no significant Line (L) * Window (W) effects were observed, so the two $(Thr\text{-}Gly)_{17}$ lines and the two $(Thr\text{-}Gly)_{20}$ lines were pooled to create separate $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ data sets [ANOVA on these 'natural length' transformants with lines nested within genotype gave no significant Genotype (G) * Condition (C) * W interaction (see Appendix 7.2a)]. These were then used as two lines of the natural (*Thr-Gly*) length genotype in planned comparisons with the other genotypes. I therefore had two lines of $(Thr\text{-}Gly)_{17/20}$, and two lines of Δ (*Thr-Gly*) in heat pulse and LD [ANOVA gave a G * C * W interaction $p<0.03$ (see Appendix 7.2b)]. I also had two lines each of $(Thr\text{-}Gly)_{17/20}$ and *L2* in heat-pulse and recovery condition [ANOVA gave a G * C * W interaction of $p<0.007$ see Appendix 7.2c].

In the heat pulse condition only, two lines each of *L2*, $(Thr\text{-}Gly)_{17/20}$, $(Thr\text{-}Gly)_1$ and Δ (*Thr-Gly*) gave a highly significant interaction [G * W $p<0.0001$ see appendix 7.2d]. Planned comparisons of the deletion transformants at the heat pulse condition only versus the natural $(Thr\text{-}Gly)_{17/20}$ length variants reveal that $(Thr\text{-}Gly)_1$ is not significantly different from them whereas the Δ (*Thr-Gly*) and L.2 genotypes show G * W interactions of $p<0.01$ and $p<0.0001$ respectively (see appendix 7.2e 1,2 and 3). These results suggest that some deleted sequences (but not a deletion of the perfect repeat) alter activity response to temperature challenge.

To investigate these responses further I examined the activity of the *per* mutants $per^{s}/per^{L}/per^{0}$ in this paradigm using the natural variants as a control [the natural 17/20/23 variants had shown no G * C * W interaction (see Fig 7.1 and Appendix 7.1) and so were pooled to create a large natural variant data set. The graph and ANOVA for all these genotypes (Fig 7.3a and b) indicates highly significant main effects and interactions which indicate that the genotypes react differently to the conditions. Fig 7.4 shows the same data as Fig 7.3 but is plotted in the 48 bin format (only one of the natural variant lines [$(Thr\text{-}Gly)_{20}$ line SM 27]is shown and data for tim^{0} is also included for comparison). Fig 7.4 reveals that in the initial LD condition, per^{S} and per^{L} have early and late evening peaks respectively, while per^{0} shows a moderate morning peak and a bare rise in activity for the evening peak. Interestingly, the heat pulse generates large morning and evening peaks in per^{0}, the former retained during recovery which suggests some kind of entrainment has taken place. The heat pulse appears to delay the per^{S} evening peak, but brings forward that of per^{L}, with a subsequent loss of the later evening peak in per^{L} (similar to deletion *L2*). In recovery, per^{S} and per^{L} return to their normal LD evening

F IG 7.3a Locomotor activity of natural variants $(Thr\text{-}Gly)_{17/20/23}$ (pooled), per^{S}, per^{L}, and per^{0} **b.** ANOVA table

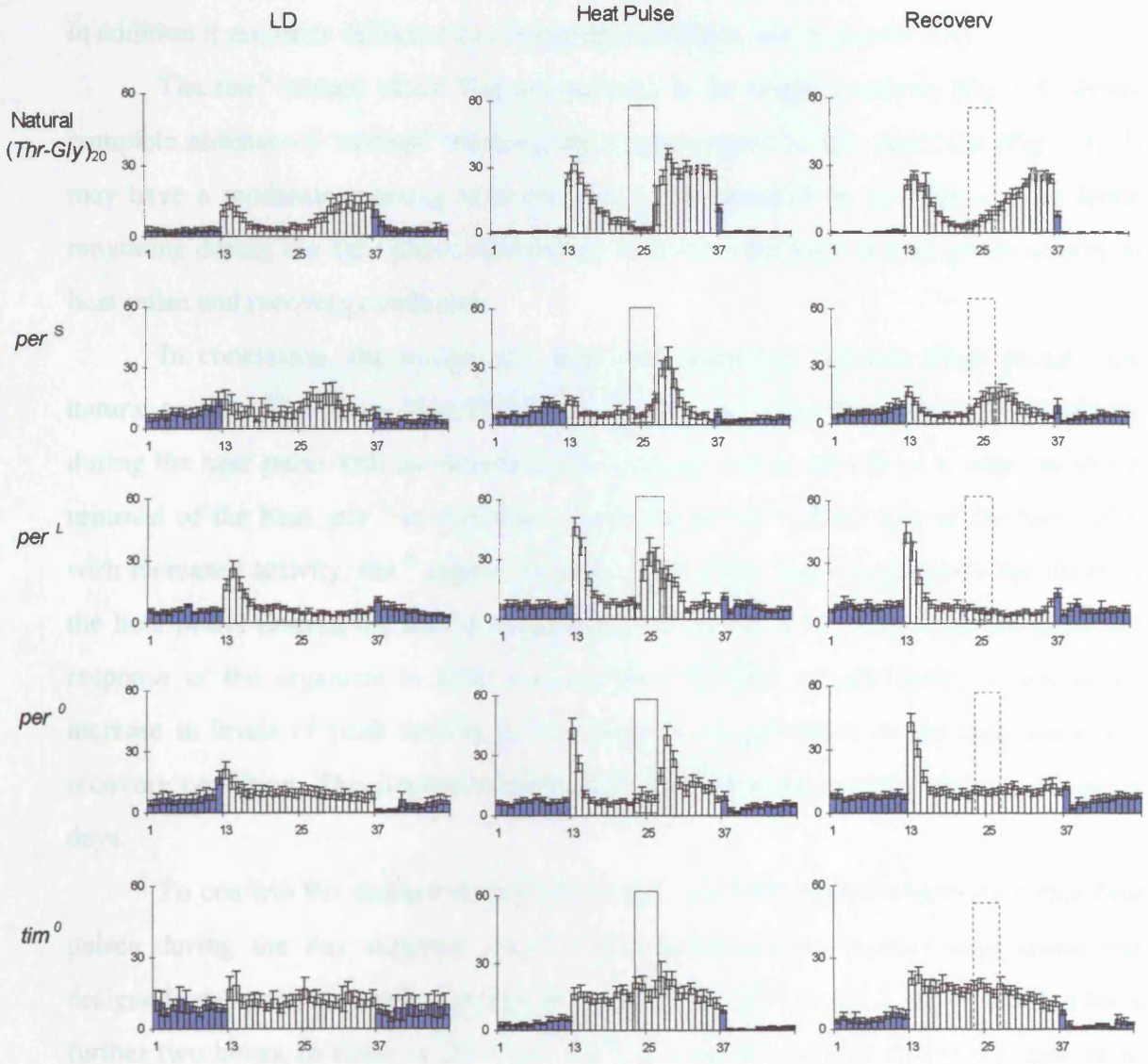
a.



b.

Summary of all Effects						
1-GENO, 2-COND, 3-WINDOW						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	3.00	18911.45	7140.00	1679.91	11.25	0.00000
2	2.00	21710.72	7140.00	1679.91	12.92	0.00000
3	11.00	91675.12	7140.00	1679.91	54.57	0.00000
12	6.00	6209.88	7140.00	1679.91	3.70	0.00114
13	33.00	29554.58	7140.00	1679.91	17.59	0.00000
23	22.00	9442.06	7140.00	1679.91	5.62	0.00000
123	66.00	2403.36	7140.00	1679.91	1.43	0.01286

Fig 7.4 Natural $(Thr\text{-}Gly)_{20}$ line Sm 27, *per* ^{S}, *per* ^{L}, *per* ^{0} and *tim* ^{0} in three LD conditions (48 bins). y axis = locomotor activity events



peaks. Inspection of the data in 30 min bins (see Fig 7.4) clearly shows that *per* ^{L} responds *during* the heat pulse with increased activity, in contrast to all the other genotypes, which have decreased activity during the heat pulse. Interestingly planned comparisons comparing the individual mutants with wild-type (pooled natural (*Thr-Gly*)$_{17/20/23}$ genotypes) reveal each mutant gives a significant G * W interaction (revealing they all have different individual activities in comparison with the natural genotypes, see Appendix 7.4 a, b and c). However only *per*^{L} gives a significant G * C * W interaction (so in addition it responds differently to changing conditions, see Appendix 7.4).

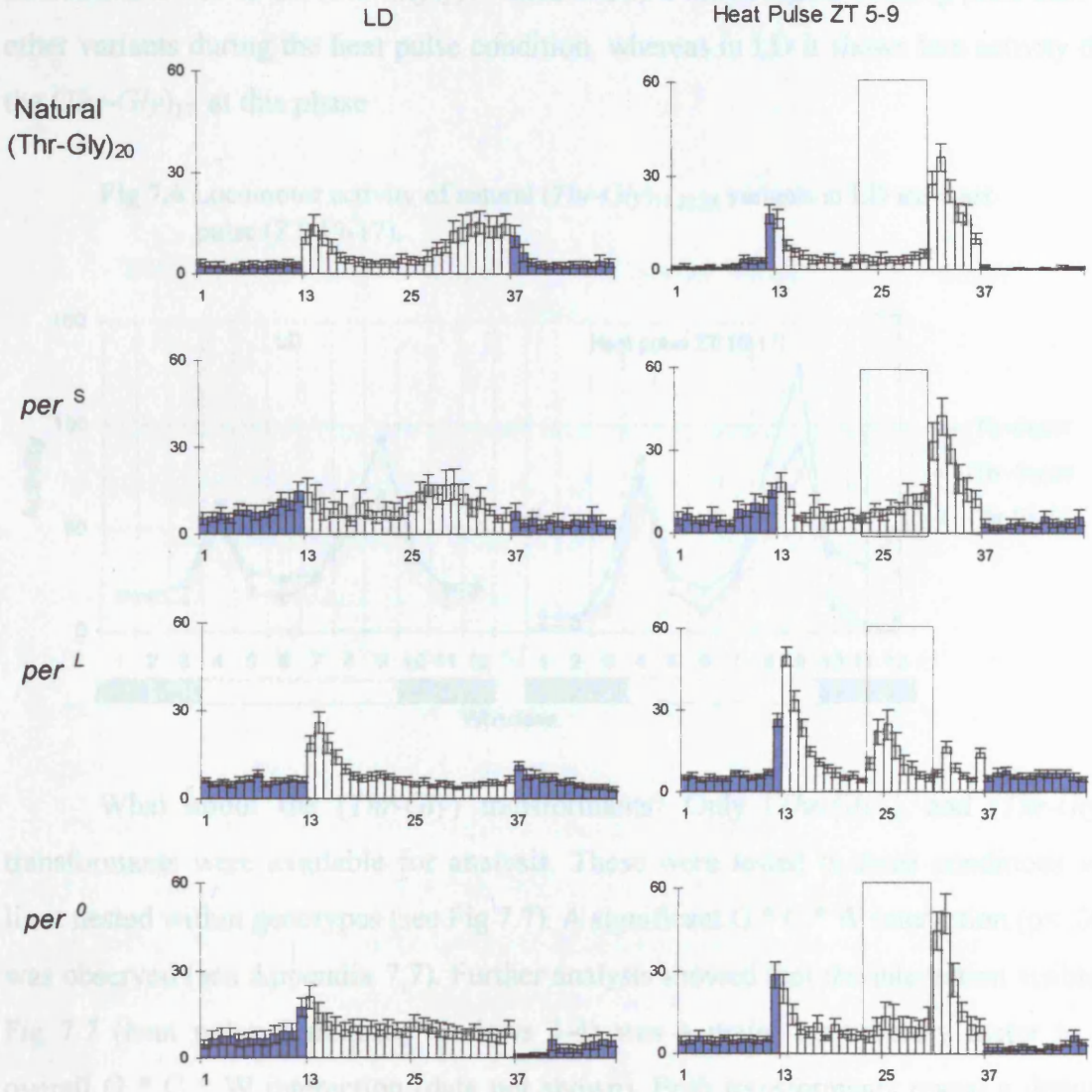
The *tim* ^{0} mutant which was not included in the original analysis (Fig 7.3) shows complete absence of 'normal' morning and evening peaks in any condition (Fig 7.4). It may have a moderate morning response that is then masked by the high activity levels remaining during the light phase. Conversely it shows a decrease in dark phase activity in heat pulse and recovery conditions.

In conclusion, the mutant and wild type genotypes fall into three groups: the natural variant (Thr-Gly)$_{20}$ [line SM27], *per* ^{S} and *per* ^{0} all show suppression of activity during the heat pulse with an immediate and dramatic rise in activity as a response to the removal of the heat. *per* ^{L} in complete contrast responds to the onset of the heat pulse with increased activity. *tim* ^{0} appears to have little, if any, acute response to the onset of the heat pulse. Overall the results indicate that mutations in *per* may influence the acute response of the organism to heat. As with the (*Thr-Gly*) transformants, a generalized increase in levels of peak activity is also seen for all genotypes in the heat pulse and recovery condition. This implies that the heat pulse sensitizes activity patterns for many days.

To confirm the distinctive response of *per* ^{L} and the general observation that heat pulses during the day suppress activity in *D.melanogaster*, another experiment was designed where the heat pulse, instead of returning to 18^{0}C at ZT 7, was extended for a further two hours, to finish at ZT 9 (see Fig 7.5, the pre-pulse LD data is the same as in Fig 7.4 and is re-shown only for ease of comparison). This heat pulse now extends into a phase of the cycle when the wild type fly usually experiences the increase in activity referred to as the 'evening peak'. The experiment conclusively shows that the heat pulse does indeed depress activity for all genotypes other than *per* ^{L} In contrast the transition from high to low temperature induces the acute activity rise in all genotypes other than *per* ^{L}. The fact that *per* ^{L} only maintains it's higher levels of activity during the heat pulse for 60-90 min (a similar time to the post heat-pulse response in the other genotypes),

indicates that the low to high temperature transition was the factor initiating *per* ^{L} 's acute response, rather than high temperature *per se*.

Fig.7.5 Locomotor activity natural (Thr-Gly)$_{20}$, *per* ^{S} *per* ^{L} and *per* ^{0} with 4h heat pulse in LD. Y-axis activity events , X-axis 30 minute bins

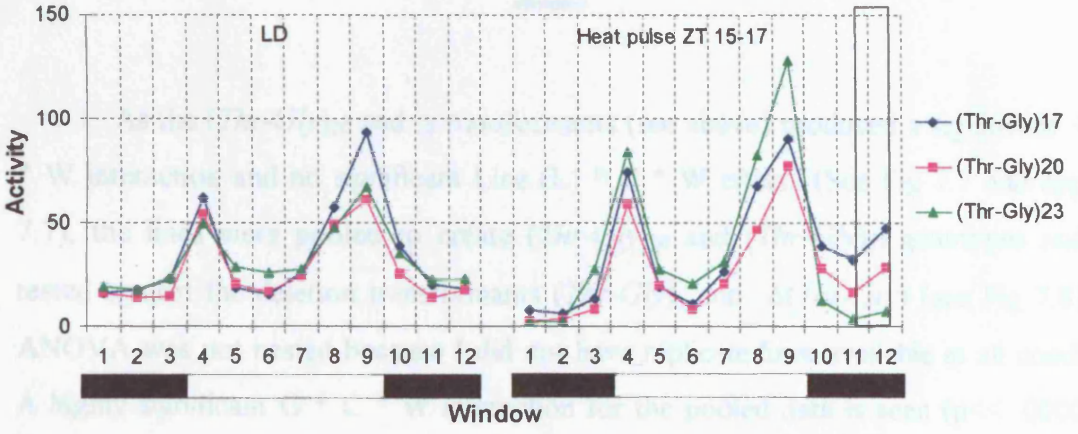


Therefore having a defective *per* product (PER ^{L}), appears in some conditions to have a greater influence on activity than having no *per* product (PER ^{01}).

The experiments above were initially designed to test an ecological hypothesis. I now extended this experiment by giving a heat pulse in the middle of the night phase of the LD cycle at ZT 15-17, when PER levels are rising (Rosbash *et al.*, 1996) and just prior to nuclear entry of the PER/TIM complex. In these experiments, the 2h heat pulse comes at a time which would be unusual under natural conditions. Fig 7.6 shows the results for the natural (*Thr-Gly*) length variants, represented by two lines each of (*Thr-Gly*) $_{17, 20}$ and $_{23}$ in this paradigm (at LD and heat pulse conditions only). The responses of

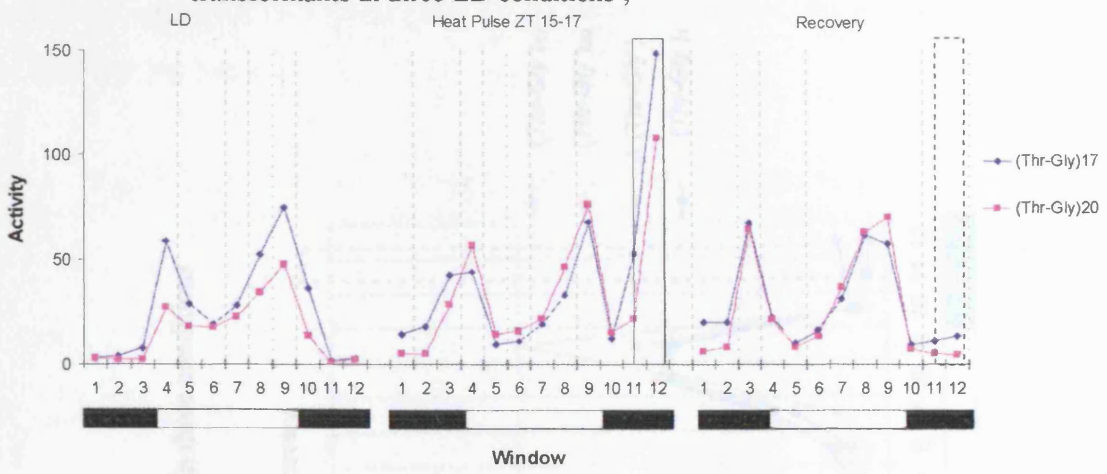
the natural $(Thr\text{-}Gly)_{17, 20}$ and $_{23}$ in a nested ANOVA in the two conditions, LD and heat pulse indicate a marginally significant G * C * W interaction ($p<0.047$ see Appendix 7.6a). A similar analysis using only the $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ data revealed no significant G * C * W interaction ($p<0.94$-see appendix 7.6b). Consequently the interaction is due to the $(Thr\text{-}Gly)_{23}$, which shows a much higher evening peak than the other variants during the heat pulse condition, whereas in LD it shows less activity than the $(Thr\text{-}Gly)_{17}$ at this phase

Fig 7.6 Locomotor activity of natural $(Thr\text{-}Gly)_{17\ 20\ 23}$ variants in LD and heat pulse (ZT 15-17).



What about the (*Thr-Gly*) transformants? Only $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ transformants were available for analysis. These were tested in three conditions with lines nested within genotypes (see Fig 7.7). A significant G * C * W interaction ($p< .001$) was observed (see Appendix 7.7). Further analysis showed that the interaction visible in Fig 7.7 (heat pulse condition, windows 3-4) was a major contributory factor to the overall G * C * W interaction (data not shown). Both transformants reveal a dramatic response to the heat pulse with average activity levels approximately twice as high as the evening lights off activity peak. This large increase in the nighttime activity is at variance with that observed for the natural $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ (see Fig 7.6). This suggests that the *per* transgene may be missing some regulatory material that is normally involved in suppression of this activity.

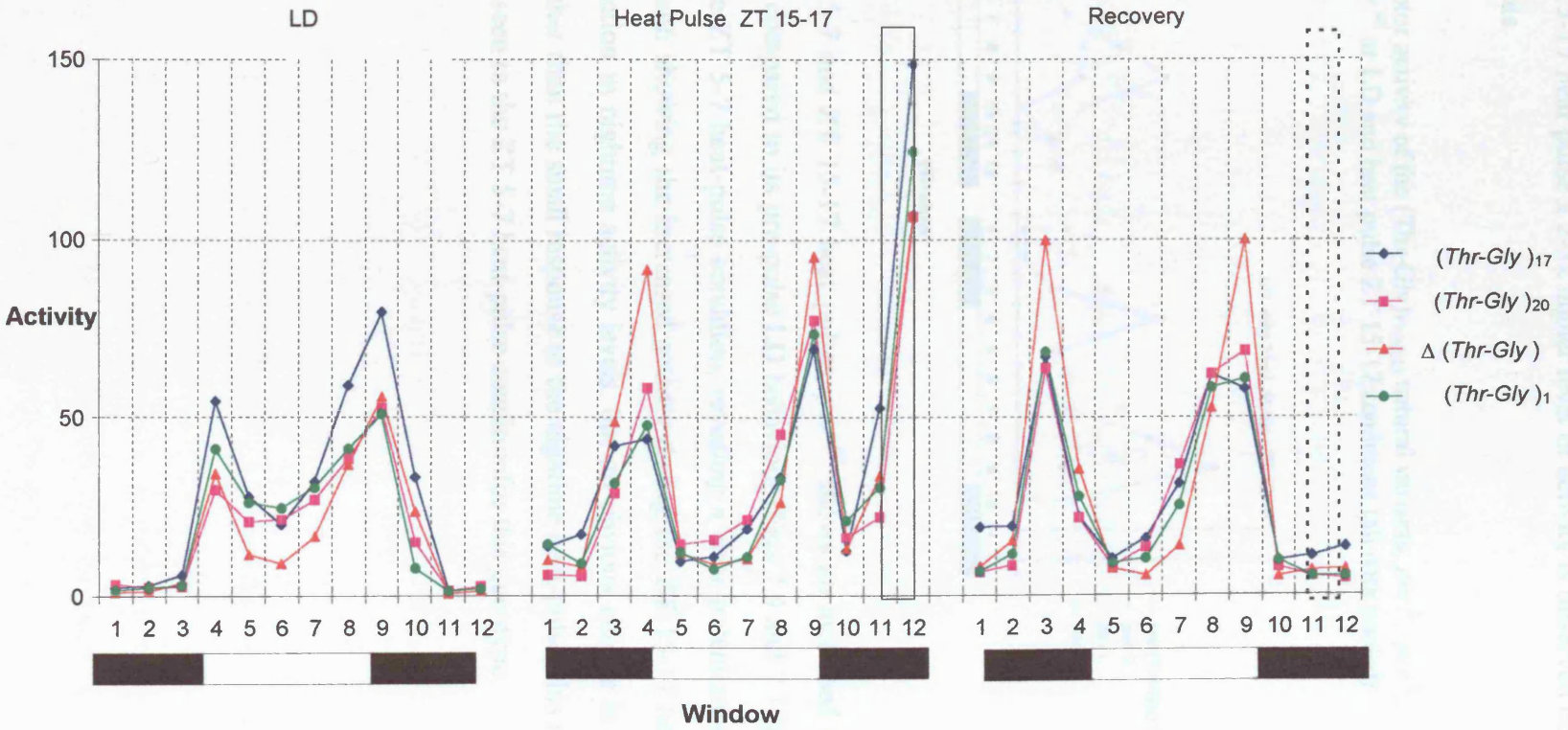
Fig 7.7 Locomotor activity of the $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ transformants in three LD conditions ,



As the $(Thr\text{-}Gly)_{20}$ and $_{17}$ transformants (see above) produced a significant G * C * W interaction and no significant Line (L) * C * W effect (See Fig 7.7 and appendix 7.7), the lines were pooled to create $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{17}$ genotypes and then tested against the deletion transformants $(Thr\text{-}Gly)_1$ and $\Delta(Thr\text{-}Gly)$ (see Fig 7.8). The ANOVA was not nested because I did not have replicate lines available at all conditions. A highly significant G * C * W interaction for the pooled data is seen ($p << .00001$ see Appendix 7.8a). The Δ(Thr-Gly) deletion transformant has higher overall levels of activity in the heat pulse and recovery condition than the other genotypes. This can be seen clearly in Fig 7.8 and further analysis shows Δ(Thr-Gly) to be significantly different to $(Thr\text{-}Gly)_1$ (Appendix 7.8b). Further nested analysis for heat-pulse and recovery conditions confirmed that $(Thr\text{-}Gly)_{20}$, $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_1$ were not significantly different from each other (Appendices 7.8c and d) but Δ(Thr-Gly) was different from the others (Appendix 7.8e).

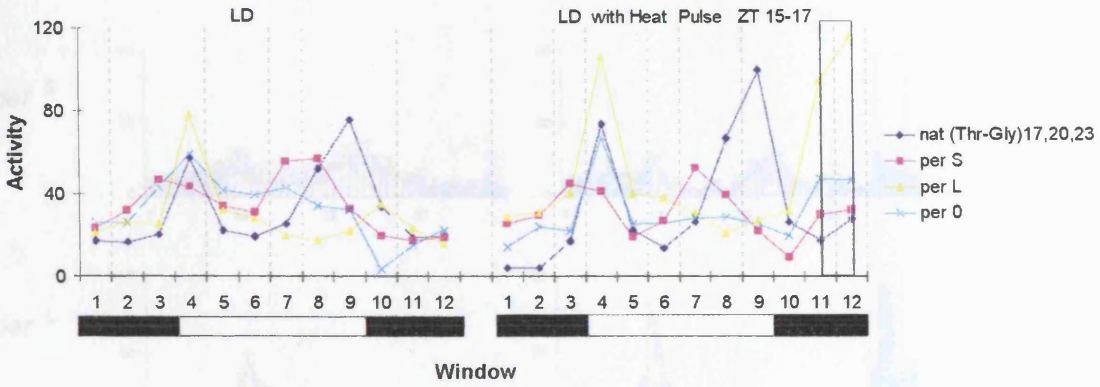
I also examined the activity of $per^{S}/per^{L}/per^{01}$ and the pooled natural $(Thr\text{-}Gly)_{17,20}$ and $_{23}$ variants in LD and heat pulse (ZT 15-17) conditions (see Fig 7.9 for 12 window format and Fig 7.10 for 48 bin data). LD data is the same as used previously and tim^{0} is included in Fig 7.10. ANOVA reveals that all effects were highly significant including a G * C * W interaction ($p<<0.0001$, see Appendix 7.9). The most important difference is that wild type, per^{S}, and per^{01} respond during the heat pulse and decrease their activity immediately after the pulse as does per^{L}. For all the genotypes other than per^{L}, this is completely the reverse of the ZT 5-7 heat pulse effect (see Fig 7.10), in the dark they

Fig 7.8 Locomotor activity of $(Thr\text{-}Gly)_{17/20}$, $(Thr\text{-}Gly)_1$, $\Delta(Thr\text{-}Gly)$ transformants in three conditions lines pooled into genotypes



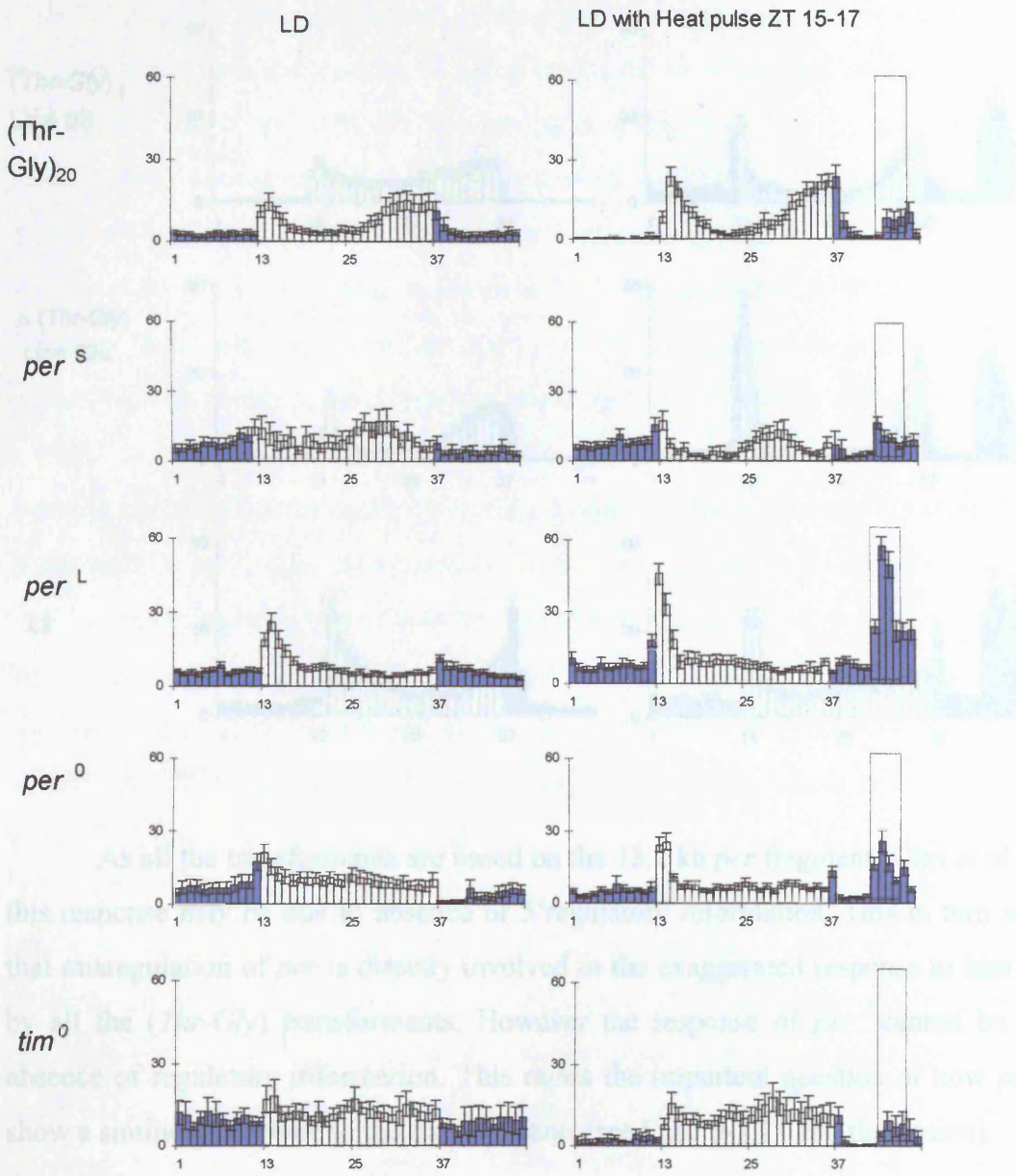
respond to the increase in temperature with increased activity whereas in the daytime they respond to the decrease in temperature with lower activity. The *per* ^{L} response is different in that during the ZT15-17 heat pulse a 2-3x higher level of activity is observed compared to the other *per* variants.

Fig 7.9 Locomotor activity of the $(Thr\text{-}Gly)_{17/20/23}$ natural variants, *per* ^{S}, *per* ^{L}, and *per* ^{01} at LD and heat pulse ZT 15-17 conditions (all data pooled)



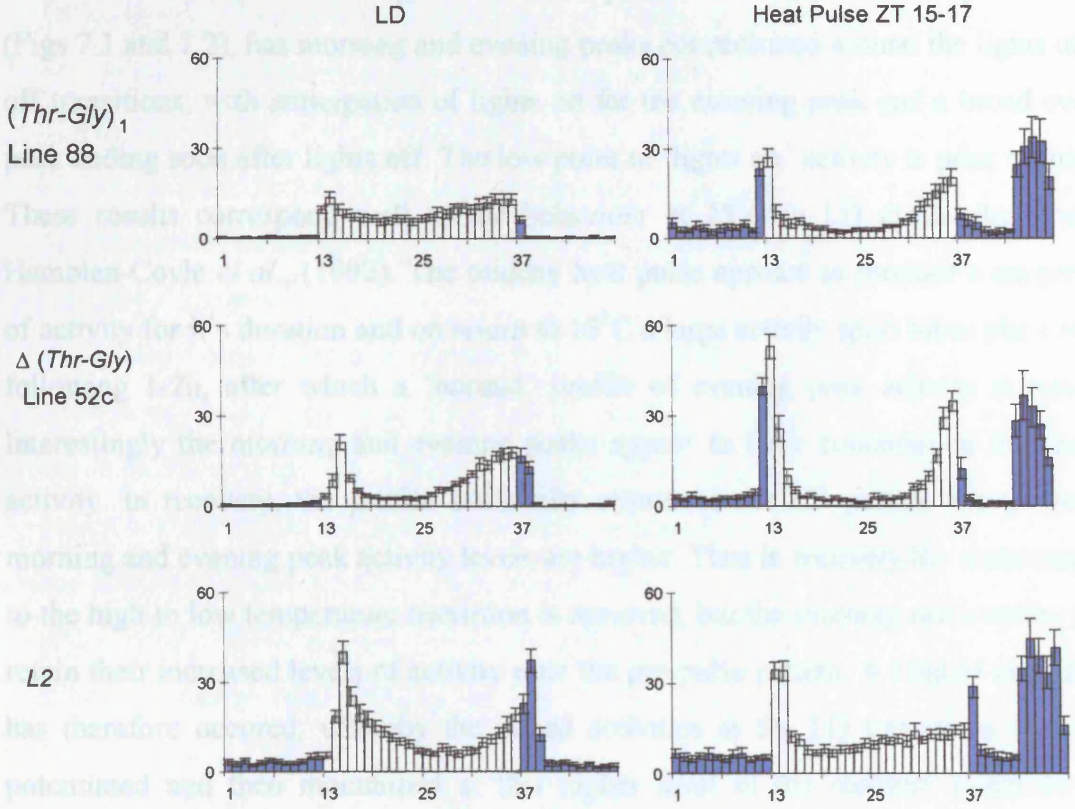
For both ZT 5-7 and ZT 15-17 heat pulses, *per* ^{01} shows an increased morning response to lights on compared to its pre-pulse LD behaviour (Figs 7.4 and 7.10). This is more prominent in the ZT 5-7 heat-pulse condition, revealing a clear potentiation of this response. *tim* ^{0} although showing the increased activity during the ZT 15-17 heat-pulse gives an overall reduction in nightime activity levels, but no obvious change in daytime activity (Fig 7.10). Other than the small response to the nightime heat-pulse, this response is very similar to that seen in the ZT 5-7 heat-pulse conditionfor this genotype.

FIG 7.10 Locomotor activity of $(Thr\text{-}Gly)_{20}$ natural variant [line sm 27], per^S, per^L, per^0 and tim^0 in LD and heat pulse conditions ZT 15-17. **y axis,** activity events; **x axis,** 30 min bins



Finally Fig 7.11 shows the response of the (*Thr-Gly*) deletion transformants to the ZT 15-17 heat pulse. In the pre-pulse LD condition, as the deletion becomes more severe (see deletion *L2*) the bimodality and activity peaks increase, as shown in Chapter 6. In the heat pulse condition, all show large behavioural responses to the temperature rise. However this large rise in activity was also shown by the $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ transformants and so this effect is likely to be peculiar to transformants.

FIG 7.11 Locomotor activity of the (*Thr-Gly*) deletion transformants, $(Thr\text{-}Gly)_1$ Δ(*Thr-Gly*) and *L2* in LD and heat pulse conditions ZT 15-17



As all the transformants are based on the 13.2 kb *per* fragment (Citri *et al.*,1987), this response may be due to absence of 5'regulatory information. This in turn suggests that misregulation of *per* is directly involved in the exaggerated response to heat shown by all the (*Thr-Gly*) transformants. However the response of per^L cannot be due to absence of regulatory information. This raises the important question of how per^L can show a similar phenotype to the transformants (see Fig 7.9-7.11 and discussion).

7.4 Discussion

The natural *D. melanogaster* bimodal pattern of locomotor activity in LD 18^{0}C (Figs 7.1 and 7.2), has morning and evening peaks concentrated around the lights on and off transitions, with anticipation of lights on for the morning peak and a broad evening peak ending soon after lights off. The low point of 'lights on' activity is prior to midday. These results correspond well to the behaviour at 25^{0}C in LD cycles described by Hamblen-Coyle *et al.*, (1992). The midday heat pulse appears to produce a suppression of activity for it's duration and on return to 18^{0}C a large activity spurt takes place for the following 1-2h, after which a 'normal' profile of evening peak activity is resumed. Interestingly the morning and evening peaks appear to have concomitant increases in activity. In recovery, the profile essentially returns to the LD pattern except that the morning and evening peak activity levels are higher. Thus in recovery the acute response to the high to low temperature transition is removed, but the morning and evening peaks retain their increased levels of activity over the pre-pulse pattern. A kind of entrainment has therefore occured, whereby the raised activities at the LD transitions have been potentiated and then maintained at this higher level in the recovery condition (this potentiation effect is reminiscent of the peak enhancement effects noted in Chapter 6, and will be further discussed in Chapter 8).

The extended ZT 5-9 heat pulse confirms that a midday increase in temperature suppresses activity, and in addition shows that the post heat-pulse response is a reaction to the decrease in temperature from 25-18^{0}C, and not a delayed response to temperature rise or elevated temperature *per se*. The temperature changes that cause these responses occur over a period of minutes (I estimate that the incubators will raise or reduce internal temperature by 7^{0}C in ~10 minutes). While this input is not as abrupt as the lights-on transition, it is nevertheless acute in an environmental sense, and the reduction in temperature at the end of the heat pulse, appears to promote a similar response to the lights-on response, in terms of changes in activity levels and the length of time the activity is elevated.

The elevated activity response in *per* ^{01} flies after the end of the heat pulse reveals that it is not *per* related, and the same inference can be made for the lights-on startle response seen in *per* ^{01} in all conditions. In complete contrast, inspection of the *tim* ^{0} profile (Fig 7.4) shows that *tim* ^{0} has only a small and transitory morning peak and does not have the acute post heat-pulse response to the heat-pulse condition. This therefore

suggests that TIM is somehow involved in the generation of the locomotor activity response to acute light and temperature signals. The LD and ZT5-9 heat-pulse experiment further reveals that within the light phase, per^{01} only shows activity changes in response to the 'acute' environmental signals discussed above. Once the response to these acute signals has passed the activity returns to a low level and none of the more complex locomotor patterns seen in the natural variants, are present. This suggests that these more complex activities which can be observed separately from the acute responses to heat and light, are *per* related.

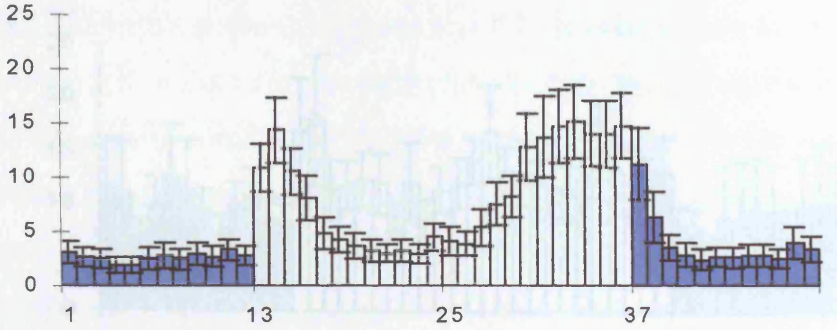
tim ^{0} flies in LD and daytime heat pulse conditions show an overall rise in activity throughout the light phase with little or no morning or evening peaks and no responses to heat transitions (Fig 7.4). The heat pulse does however appear to cause a slight general rise in activity during the light phase and a concomitant reduction during the dark. *tim* ^{0} has low constitutive levels of PER that do not show post translational phosphorylation (Price *et al.*, 1995). So here is a direct indication that the *per* related activity profiles in natural variants (and missing in *tim* ^{0}) are intimately linked to cycling and/or phosphorylated PER.

per^{01} has a response to the morning LD transition and at least at 18^{0}C no obvious evening peak of activity. We know from per^{L} and per^{S}, and the experiments reported in Chapter 4 and Chapter 6, that *per* is involved in the timing and strength of the evening peak. It is also clear from Fig 7.4 that the natural variants have a measured decrease in activity at bins 15-19 (~ZT 2-4) following the lights on peak (this is most obvious in the heat-pulse and recovery conditions). This is in contrast to the abrupt return to basal activity levels after lights on, seen in per^{01} on transition from dark to light.

Could variation in PER concentrations in the light phase be related to activity patterns? At a simplistic level we know that high concentration of PER roughly coincides with high levels of activity for the light phase (Edery *et al.*, 1994b and see Fig 7.12). The representations of PER and TIM in Fig 7.12 are a generalized impression of the protein levels of these two proteins as observed in Western blots from many sources, (for a review see Rosbash *et al.*, 1996). The altering elevation of PER an TIM reflects the known circadian changes in molecular weight resulting from phosphorylation (Edery *et al.*, 1994b; Rutila *et al.*, 1996). Note that when significant amounts of PER are present but TIM is absent, high levels of activity are observed.

Fig 7.12a Locomotor activity profile of natural $(Thr\text{-}Gly)_{20}$ variant at 18^0C in LD (redrawn from Fig.7.4). **b.** Diagramatic representation of PER concentration and phosphorylation. **c.** Diagramatic representation of TIM concentration and phosphorylation. (The PER and TIM representations are not meant to be accurate or to scale)

a.



b.PER

c TIM

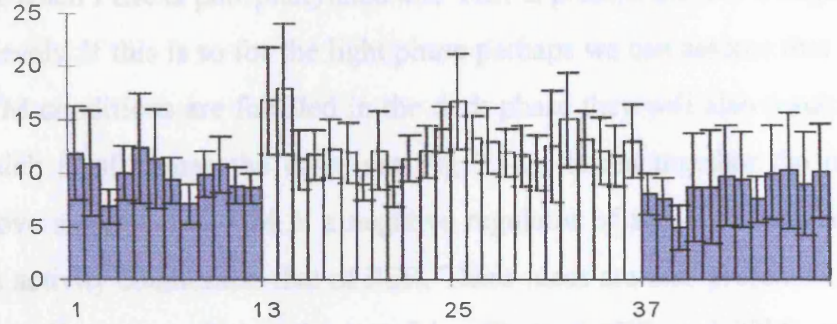
As a working hypothesis, let us assume that PER but not TIM, positively regulates locomotor activity in the light phase. Why then would a daytime heat-pulse suppress locomotor activity? Temperature increases from 18-30^{0}C in the middle of the dark phase when PER concentration is relatively high have been shown to degrade PER levels, by ~40% over one hour, (Sidote *et al.*, 1998). It is quite possible then that the mild heat rise 18-25^{0}C maintained over a period of 4h from the middle of the light phase (the ZT 5-9 experiment) therefore degraded low levels of newly translated PER, and delayed it's accumulation leading to low activity levels. Western analysis of flies subjected to the ZT 5-9 experiment would be required to see whether the 18-25^{0}C heat pulse could prevent the appearance of PER during the light phase.

What of *tim* ^{0}? Like *per*^{0} at 18^{0}C, it has no evening peak but it has a vestigial response to lights on. The one molecular clue that we have is that *tim*^{0} has low levels of constitutive PER and also that the PER product that is visible on a western blot is not phosphorylated cyclically as in wild type flies (Price *et al.*, 1995). This lack of concentration or phosphorylation cycle of PER may therefore cause the absence of a normal locomotor activity profile. The overall slight increase in activity levels and lack

of any locomotor pattern during the light phase seen in Fig 7.13a, therefore fits the hypothesis that PER is a positive regulator of activity in the light phase.

Fig 7.13a. Locomotor activity profile of *tim* ^{0}at 18^{0}C in LD. **b.** diagramatic representation of PER concentration and phosphorylation as seen on a western blot (from Price *et al.*, 1995).

a.



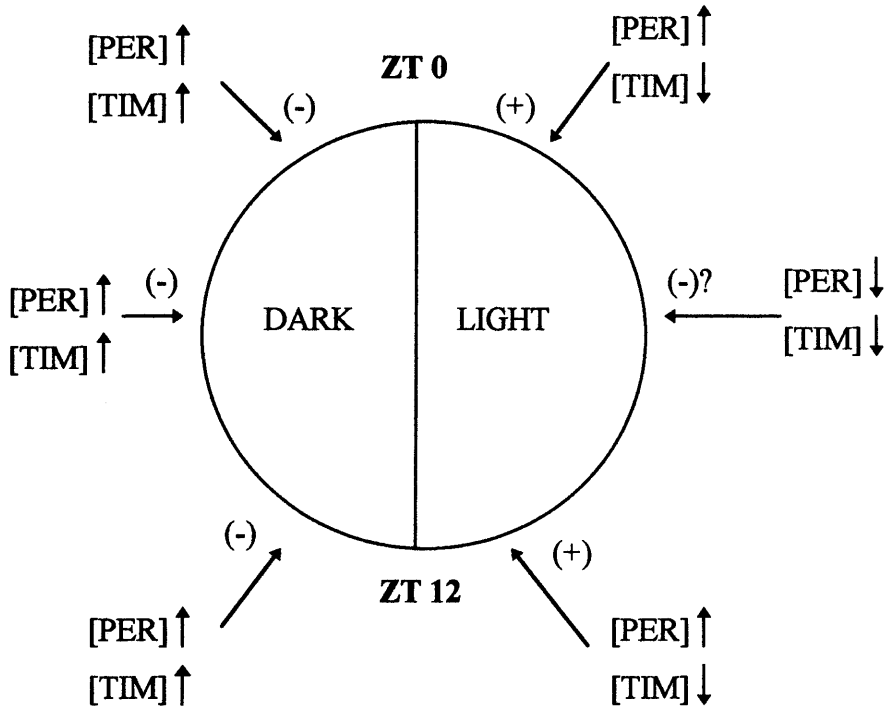
b.

The early evening activity peak of *per* ^{S} (within the light phase, see Fig 7.4) is known to be related to an early peak in PER concentration (Marrus *et al.*, 1996). However the subsequent drop in *per* ^{S} activity in the light phase coincides with the maintained high levels of PER ^{S} (Marrus *et al.*, 1996). This apparent paradox may be explained by the observation that in *per* ^{S} the accumulation of TIM protein is also quite strong at this phase ~ZT 10, some hours earlier than for the natural variants (Marrus *et al.*, 1996). The simultaneous appearance of high levels of PER and TIM may inhibit the otherwise positive regulation of activity levels by PER alone (see night phase Fig 7.12 and 7.14). The *per* ^{S} result (Marrus *et al.*, 1996) is thus a further indication that when PER and TIM are together in the light phase, then inhibition of activity levels occurs. Again the available data fits the hypothesis.

Are there other situations where these conditions occur in the light phase? Compared to PER levels in natural variants PER^{L} has a cycle with intermediate amplitude and poorly defined phosphorylation (Rutila *et al.*,1996). Significant levels of PER are therefore present virtually throughout the LD cycle although a concentration cycle is evident (Rutila *et al.*,1996). According to the working hypothesis therefore, *per*^{L} should have intermediate levels of activity throughout the light phase and this is not the case (see Fig 7.4). *per*^{L} however also has a defective TIM cycle with TIM detectable

throughout the light phase and not phosphorylated in the normal manner (Rutila *et al.*, 1996). It is possible therefore that PER in these conditions cannot act in it's normal putative regulatory role in terms of locomotor behaviour. As in the per^{S} case outlined above (Marrus *et al.*, 1996), per^{L} finds itself in conditions which are in some respects more like those experienced by the natural variant in the early dark phase. This, and the concentrations of TIM and PER in the later phase of the per^{S} cycle does however indicate that when PER is phosphorylated and TIM is present there is a negative impact on activity levels. If this is so for the light phase perhaps we can assume that when these PER and TIM conditions are fulfilled in the dark phase they will also result in reduced activity, which is of course the case (see Fig 7.12). Taken together the observations outlined above suggest that TIM is a negative regulator of activity in the light or dark and that it's activity counteracts that of PER. These ideas are also presented in Fig 7.14, a speculative diagram outlining the possible effect of PER and TIM on locomotor activity in the light and dark phases.

Fig 7.14 Speculative correlation between PER and TIM concentration and negative (-) or positive (+) effects on activity levels at different clock phases.



What of the increase in activity noted for the per^{L} genotype during the daytime heat pulse which is not observed for any of the other genotypes (see Fig 7.4)? We know that heat can degrade both PER and TIM (Sidote *et al.*, 1998) and PER^{L} is more readily detectable than TIM at this phase (Rutila *et al.*, 1996). Also the PER^{L}/TIM interaction is known to be weak (Gekakis *et al.*, 1995). It is plausible therefore that a temperature rise may result in the release of free PER^{L} that is available to upregulate locomotor activity.

What of the relative concentrations of PER and TIM? Fig 7.12 suggests that for PER to actively increase locomotor activity then it needs to be at a much higher relative concentration than TIM, and that whenever the concentrations of the two molecules are approximately equal then lower activity levels occur. Does this make any sense? We know that PER and TIM function in their nuclear entry role as a dimer. This would suggest that whenever the two molecules exist in the cell, some dimer formation is likely to occur. If the concentration of PER far exceeds that of TIM then it is likely that all the available TIM will be in the heterodimeric form. Perhaps when in this form TIM is unavailable for other pleiotropic functions involving locomotor activity, such as the proposed negative regulation of activity. If this hypothetical sequence of events is accurate, then the breakdown of TIM by light may allow for the maintainance of daytime activity levels via relatively higher levels of PER.

Puting aside the problem that TIM is at very low levels during the light phase, when the post heat pulse acute response occurs, if we assume that the TIM protein in some way enables the acute responses in per^{01}, we must ask what prevents the action of PER in tim^{0}. PER requires TIM to attain higher phosphorylated states, absolutely in the case of the tim^{0} (Price *et al.*, 1995) and partially in the case per^{L} where a poorly phosphorylated PER^{L} can be partly compensated by a possibly hyperphosphorylated TIM^{SL} (tim^{SL} is a *tim* mutant that restores temperature compensation to per^{L} -Rutila *et al.*, 1996). In both per^{01} and tim^{0} there is defective regulation of activity levels. So perhaps the hypothesis need to be amended slightly such that PER will only function as a positive regulator of activity if it is hyperphosphorylated or if TIM is being produced with it. Fig 7.12 shows that these conditions are probably met.

If we return to the tim^{0} and per^{01} heat pulse experiment, we can see that for per^{01} which has LD driven (though low amplitude, non-phosphorylated) TIM cycling (Zeng *et al.*, 1996) there are acute responses to both the LD transition and heat pulse challenges, whereas tim^{0} has little or no response to either challenge. This suggests that per^{01} flies

have a factor for these responses that the *tim* ^{0} flies do not possess. The most obvious candidate is the TIM protein. This suggests that TIM in addition to being a negative regulator of activity, it is also a facilitator of the acute response, and that it can do this without the presence of PER (see Fig 7.4 and 7.14).

The initial simple hypothesis (Fig 7.14) has therefore been changed such that phosphorylated PER or PER in the presence of low levels of TIM activates locomotor activity. We can complete this by suggesting that TIM is a suppressor of activity when at a sufficiently high relative concentration to PER. This situation only usually occurs in the night phase and it is the action of light which degrades TIM leading to the release of PER to maintain the post acute response to lights on. As PER levels are maintained in the first few hours of daytime this explains the extended morning peak in activity in wild type flies. For the natural variants this explains low levels of activity in the night and fits well with PER's hypothesized role in activity patterns during the day. For per^{S} relatively high levels of TIM and PER within the light phase explains how the activity levels of the evening peak return to basal levels while still in the light phase.

The small increases seen in locomotor activity in natural variants and the mutants (except per^{L}) in response to the onset of heat at ZT 15-17 (Fig 7.9b) cannot be explained by this hypothesis as similar results are seen in per^{0} (with high TIM levels) and tim^{0} (low PER levels). However the major activity change displayed by per^{L} during the heat phase suggests that PER and/or TIM are involved perhaps the heat pulse releases large amounts of free PER being released from the weak PER^{L}/TIM dimer.

The similar large effects produced by the nightime heat pulse noted for all the (*Thr-Gly*) transformants (natural variant lengths and deletions) must be due to the ectopic location of the transgenes within the genome, as the natural variants do not show this phenomenon (see Fig 7.9). Could this be due to alteration in transcriptional repression of these *per* genes via the negative feedback model (Rosbash *et al.*, 1996)? The similar response in per^{L} gives a tantalizing view of how this may occur. PER is thought to act negatively on it's own transcription (Rosbash *et al.*, 1996), and we know that PER^{L} has poor interactions with TIM and is retarded entering the nucleus (Rutila *et al.*, 1996). Therefore maybe PER^{L} is unable to inhibit it's own transcription efficiently at this time. The coincidence of a similar per^{L} nightime heat response in activity to the (Thr-Gly) transgenes may also be due to inefficient negative regulation but this time due to the absence of cis acting regulatory factors. Inefficient negative regulation via

inefficient interactions at the *per* promoter may allow more transcription of *per*, maintaining higher PER concentrations and hence higher levels of activity. To speculate further, in the transformants an increase in temperature at a time when normally the *per* transcription rates are going down, may result in poorer negative regulation of *per* (because of missing cis acting factors) but efficient negative regulation of *tim* (as endogenous *tim* has its normal cis acting regions). This would result in a change in the balance between PER and TIM levels favouring PER which may according to the hypothesis account for the activity rises seen during the heat pulse.

The observed rapid replacement of PER after heat pulsing (Sidote *et al.*, 1988 and A Rogers pers.comm) has been noted to be more apparent at ZT21.5 (2.5 hours prior to lights on) than at ZT 17 before nuclear entry (Sidote *et al.*, 1988). This may thererefore ensure the morning activity response is developed into the normal pattern of activity by maintaining sufficient PER. Excess PER may also be available to absorb any remaining TIM molecules after the LD transition, thus the ensuring removal of inhibitory function the remaining TIM may have on activity.

What of the daytime post heat pulse acute response? As this is present in per^0 then no involvement of PER can be invoked. Involvement of TIM is likely as this effect is missing in tim^0 flies but present in all the other genotypes, where low levels of TIM are present. Could it be then that the very low levels of non phosphorylated TIM present in the middle of the day and not some third factor is behind this phenomenon? If we suppose that some third factor were responsible for the midday acute response then this factor may also be responsible for the lights on acute response. As both responses are not present in tim^0 flies then this postulated third factor may be a TIM partner that is not functional in the absence of TIM in the manner of PER. This is perhaps an indication of how TIM may be part of a general response to acute environmental challenge. With the information available it is not possible to comment further other than to note that the level and duration of the morning and post heat pulse responses are intriguingly similar.

Other explanations for PER's apparent involvement in activity will be suggested in the next and final chapter.

Chapter 8 conclusions

I have investigated the enigmatic (*Thr-Gly*) repeat region of the *Drosophila per* gene in a number of ways. The region and it's flanking environment are clearly important for temperature compensation. A very diverged (*Thr-Gly*) repeat region from another species competently substitutes for the corresponding *D.melanogaster* region as long as the foreign flanking sequences are carried over in the chimaeric *per* constructs (Peixoto *et al.*, 1998; Piccin, 1998). The selective pressure on the organism to maintain a reliable temperature compensated clock may have led to the coevolutionary changes seen in the (*Thr-Gly*) flanking regions of other Drosophila species as they struggle to maintain PER conformation despite huge variations in the repeat length and structure (Peixoto et al., 1993; Nielsen et al., 1994).

The existence of two stages of temperature compensation, with one phase apparent at 18-25^{0}C and another at 25-29^{0}C (Chapter 5) indicates that at least two separate processes are involved in temperature compensation. The differences in temperature profiles for the (*Thr-Gly*) deletions (Chapter 5) also indicate that it may be possible to assign roles to functional subdomains within the repeat region, reinforcing the idea that temperature compensation is a multistage process. Speculatively these processes may be occuring at any point in the feedback cycle; for example PER/TIM affinity (Gekakis *et al.*, 1995), PER/TIM phosphorylation (Rutila *et al.*, 1996) nuclear entry (Curtin *et al.*, 1995) or PER degradation (Dembinska *et al.*, 1997).

The modulation of responses to external stimuli such as 'lights on and off' are obviously very important in the provision of a coherent response of the animal to its external environment. Chapter 6 (Fig 6.5) shows that *per* is intimately involved in this modulation. Surprisingly the results also suggest that small alterations in the (*Thr-Gly*) repeat can produce disproportionate effects in comparison to larger deletions of the same region. It seems very unlikely that the increased activity at the LD transitions at higher temperature observed for the $(Thr\text{-}Gly)_{17}$ transgene is due to the free-running period deviations in the $(Thr\text{-}Gly)_{17}$ clock (Fig 5.7), especially in an LD regime. consequently *per* may have pleiotropic functions that can be divorced from basic clock activity.

These same LD locomotor effects (Fig 6.5) due to small variations in the perfect repeat of *D.melanogaster* may also help us understand how these species may benefit

from small evolutionary changes in the repeat, and hence illuminate the European cline in *D.melanogaster*'s (*Thr-Gly*) length variation. An initial objective of the work presented in the previous chapters was to try to explain the north-south cline in natural (*Thr-Gly*) length variants where the $(Thr\text{-}Gly)_{20\ allele}$ predominate in the north and the $(Thr\text{-}Gly)_{17}$ in the south. Because of the results outlined above, a variety of plausible explanations are now possible.

In DD $(Thr\text{-}Gly)_{20}$ is a molecular component that leads to a 'better' endogenous clock (Chapter 5). ['better' here meaning a clock that keeps time rigorously and is not affected by temperature variation (Chapter 5)]. In northern Europe, where the annual temperature range is greater than the south (Hulme *et al.*, 1995), the more robust temperature compensation properties of the $(Thr\text{-}Gly)_{20}$ allele may be at a premium for survival. Conversely in the warmer south, the less thermally rigorous endogenous rythm of the $(Thr\text{-}Gly)_{17}$ variant may aid survival by allowing temperature rise to lengthen the period of the clock (Sawyer *et al.*, 1997) so that it resonates with the environmental cycle. Furthermore in DD high temperature $(Thr\text{-}Gly)_{17}$ peak locomotor activity is displaced into the dark phase which involves moving activity to a colder part of the circadian cycle in the wild (Fig 6.2). Thus a balancing selection scenario can be envisaged, whereby the $(Thr\text{-}Gly)_{20}$ are particularly adapted to the colder, more thermally variable clines and the $(Thr\text{-}Gly)_{17}$ to warmer environments. In fact, both in Europe (Peixoto *et al.* 1993) and Australia (L.A.Sawyer and C.P. Kyriacou, ms in prep.), the $(Thr\text{-}Gly)_{17}$ allele generally predominates over the $(Thr\text{-}Gly)_{20}$, and only starts to fall in frequency at the more extreme, cooler regions within these continents.

The differences in temperature compensation associated with the different (*Thr-Gly*) lengths, are also fully consistent with the coevolutionary dynamics that are proposed to explain the disruption of temperature compensation with interspecific chimeric *per* genes (see Chapter 3). Thus, within *D. melanogaster*, because the differences in repeat length are not compensated by changes in flanking haplotypes, small but detectable differences in temperature compensation are observed.

The LD (Chapter 6) and combined LD/temperature cycle (Chapter 7) data allow us to speculate futher about the maintenance of the cline. The $(Thr\text{-}Gly)_{20}$ variant responds less than the $(Thr\text{-}Gly)_{17}$ to stimuli presented by the LD transitions and I speculated that this might be due to a stronger repression of the response by $(Thr\text{-}Gly)_{20}$ (Fig 6.5). This may be another factor in survival in colder regimes where the LD transitions do not necessarily correspond to an immediate rise in ambient temperature.

Conversely where higher temperatures prevail the greater response of $(Thr\text{-}Gly)_{17}$ to the transitions may mean that bimodality is emphasized and avoidance of midday heat by quiescent behaviour is more likely.

The repetitive (Thr-Gly) tract therefore provides an important component of the fly's circadian temperature compensation system, it influences immediate locomotor responses to acute environmental changes and partially controls locomotor profile throughout a combined teperature/light cycle (Chapter 5 and 7). It is perhaps not surprising therefore that in various environments throughout Europe some (*Thr-Gly*) variants are more advantaged than others. The identification of a coevolved 'cassette' (which includes the repeat) enables the PER protein to function properly in a variety of responses to environmental challenge (Chapter 3 and 4), and the deterioration in these functions when disrupted, indicates the integral role of *per* in these responses.

The chimaeric transgenes have an added dimension, revealing species-specific activity that can perhaps be divorced from basic clock function *per se*. The *mps3* transgene for example, which displays *D.melanogaster* type responses for free-running period (Fig 3.9) and DD locomotor profile (Figs 4.4, 4.5), indicates that this molecule is allowing the clock to function competently and virtually indistinguishably from the *D.melanogaster* clock even though the C terminal half of *per* is from *D.pseudoobscura* The subtle but identifiable species-specific effect of *mps3* in response to an LD cycle indicates that this molecule also has novel activity in the *D.melanogaster* background. This justifies my earlier view with regard to *D.melanogaster* (*Thr-Gly*) variants, where I suspect that processes other than basic clock function may be important for the maintenance of the latitudinal cline.

What other evidence is there for pleiotropic effects of PER? Comparison of LD and combined LD/temperature cycles ('midday heat pulse' ZT 5-7, Chapter 7) shows that altered activity patterns at the light/dark boundaries is caused by a temporally remote temperature challenge. The midday heat pulse resulted in a greater activity response at later LD transitions, in comparison to the transition response in normal LD cycles (Fig 7.4). This effect is also maintained in a 'without heat pulse' recovery phase. An explanation may be that a factor is induced by the heat pulse and is active at the following lights-on phase, or alternatively it is induced and can remain in an activated state for a period of some days. Another possibility is that a second oscillator has been entrained by the heat pulse and this oscillator reinforces the activity output of the light entrained oscillator at subsequent LD transitions (see Fig 8.1). The increase in morning peak

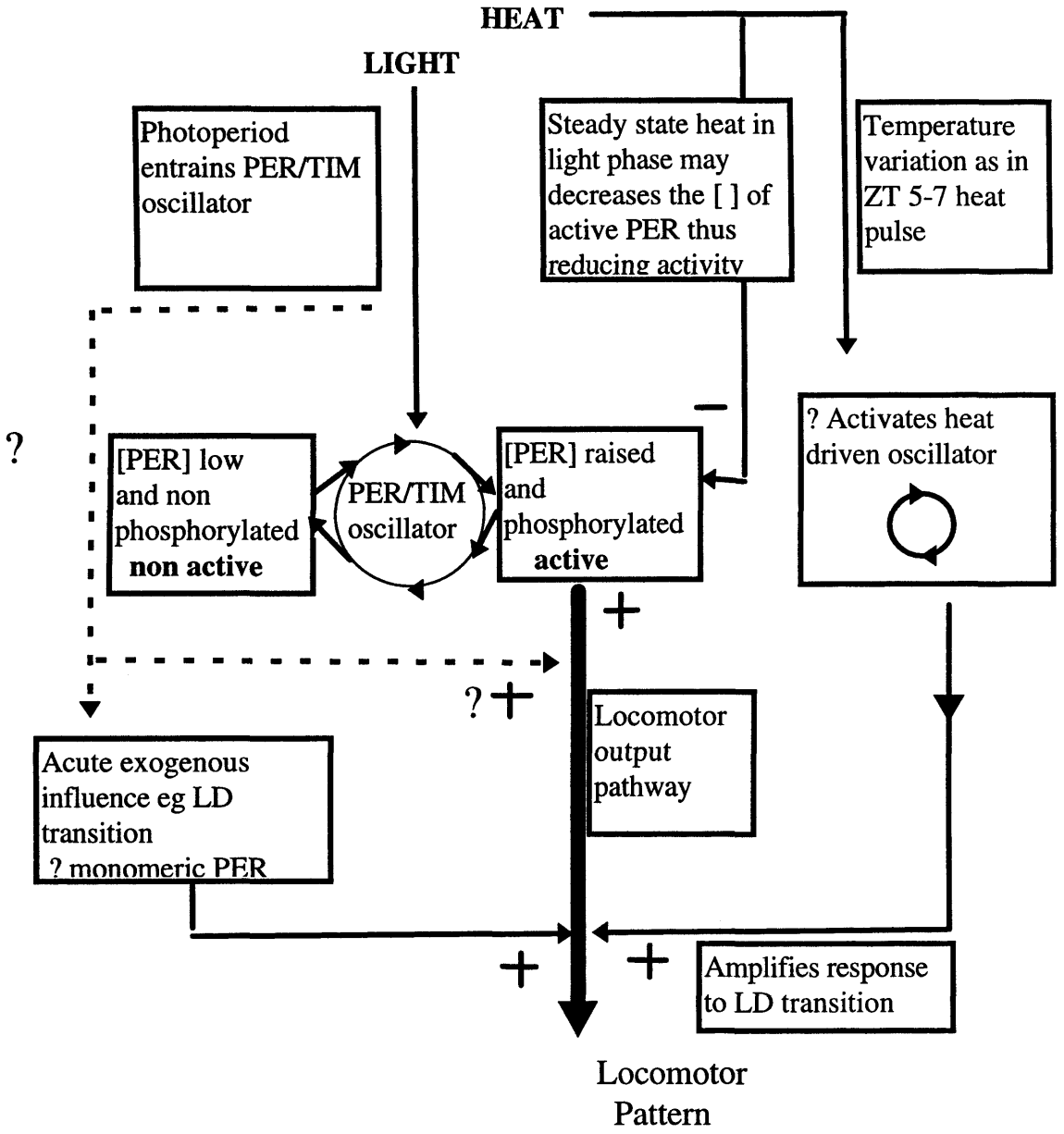
activity in per^0 after midday heat treatment (Fig 7.4) provides some tentative evidence for this view and is reminiscent of Pittendrigh's idea of A and B oscillators (Pittendrigh, 1981)

Increasing temperature at ZT 5-7 also decreases activity at this time and strongly suggests that the decrease in ambient temperature from a midday high may initiate the evening peak of bimodal activity. In Chapter 7, I suggest that the negligible levels of locomotor activity, also seen during the ZT 5-9 heat pulses (Fig 7.5), may be due to increased degradation of the already low levels of PER that are produced at this phase. This was based on the observation that per^L, which has substantial levels of PER at this phase (Rutila *et al.*, 1996), shows a different response during the heat pulse thereby implicating PER in this transient locomotor phenotype.

The observations above illustrate a recurrent theme running through the data presented in Chapters 3-7. The free-running *Drosophila* bimodal activity pattern at higher temperatures (Fig 4.2) seems to loosely fit a natural temperature light cycle. It is however extremely maleable and whatever distortions are presented in terms of variations in light and temperature (within the boundaries present in nature), the fly seems to be able to accomodate them and provide itself with an ecologically sound activity profile. Does the clock merely provide the overall timing structure within which other physiological components direct the locomotor patterns, or do the clock molecules affect activity patterns in a different manner from which they direct the feedback loop?

In an attempt to answer this final question I have suggested that PER may take a more active and direct role in the generation of locomotor output profile. The main evidence for this comes first from the species-specific activity of the chimaeric transgenes, second, the differential response of the (*Thr-Gly*) variant transgenes to the LD transitions, third, the effect of heat on activity levels in the midday phase which is plausibly related to the degradation of free PER and finally the coincidence reported in the literature of active PER being at high concentration at times of maximum fly activity, at least in the light phase. These ideas are summarized in Fig 8.1.

Fig 8.1 Summary of hypothetical effects of PER, heat and light on locomotor activity patterns



APPENDIX

Appendices for Chapter 3

Appendix 3.4 Newman-Keuls *post hoc* comparisons for the free running periods of the mel transformants at 18, 25 and 29^{0}C

			{1}	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}	{12}
			24.61415	25.03591	24.80111	24.06516	24.34474	24.17615	24.21067	25.02231	24.92219	23.50683	25.51800	25.42452
mel 2a	18	{1}		0.46271	0.46341	0.19785	0.29070	0.31433	0.25324	0.37814	0.44833	**0.00022**	**0.00721**	**0.01855**
mel 2a	25	{2}	0.46271		0.79371	**0.00446**	0.07310	**0.01708**	**0.02070**	0.95746	0.89617	**0.00001**	0.14134	0.12750
mel 2a	29	{3}	0.46341	0.79371		0.04507	0.17296	0.10208	0.09445	0.66084	0.63490	**0.00003**	0.05564	0.10359
mel 17a	18	{4}	0.19785	**0.00446**	0.04507		0.69171	0.66335	0.83576	**0.00433**	0.01369	**0.02855**	**0.00002**	**0.00002**
mel 17a	25	{5}	0.29070	0.07310	0.17296	0.69171		0.78602	0.59903	0.06050	0.10638	**0.00900**	**0.00014**	**0.00047**
mel 17a	29	{6}	0.31433	**0.01708**	0.10208	0.66335	0.78602		0.89234	**0.01580**	0.04021	**0.02356**	**0.00002**	**0.00004**
mel 34a	18	{7}	0.25324	**0.02070**	0.09445	0.83576	0.59903	0.89234		**0.01825**	0.04200	**0.02948**	**0.00002**	**0.00008**
mel 34a	25	{8}	0.37814	0.95746	0.66084	**0.00433**	0.06050	**0.01580**	**0.01825**		0.69457	**0.00001**	0.20969	0.25543
mel 34a	29	{9}	0.44833	0.89617	0.63490	**0.01369**	0.10638	**0.04021**	0.04200	0.69457		**0.00003**	0.13335	0.19936
mel116a	18	{10}	**0.00022**	**0.00001**	**0.00003**	**0.02855**	**0.00900**	**0.02356**	**0.02948**	**0.00001**	**0.00003**		**0.00002**	**0.00001**
mel116a	25	{11}	**0.00721**	0.14134	0.05564	**0.00002**	**0.00014**	**0.00002**	**0.00002**	0.20969	0.13335	**0.00002**		0.71391
mel116a	29	{12}	**0.01855**	0.12750	0.10359	**0.00002**	**0.00047**	**0.00004**	**0.00008**	0.25543	0.19936	**0.00001**	0.71391	

Appendix 3.7 Newman-Keuls *post hoc* comparisons for the free running periods of the *mps3* transformants at 18, 25 and 29^{0}C

			{1}	{2}	{3}	{4}	{5}	{6}
			23.32800	24.36969	25.01083	24.14400	25.70691	25.62231
67A	18	{1}		**0.00003**	**0.00001**	**0.00024**	**0.00002**	**0.00002**
67A	25	{2}	**0.00003**		**0.00388**	0.30940	**0.00001**	**0.00002**
67A	29	{3}	**0.00001**	**0.00388**		**0.00029**	**0.00490**	0.00590
65C	18	{4}	**0.00024**	0.30940	**0.00029**		**0.00002**	**0.00001**
65C	25	{5}	**0.00002**	**0.00001**	**0.00490**	**0.00002**		0.70320
65C	29	{6}	**0.00002**	**0.00002**	**0.00590**	**0.00001**	0.70320	

Appendix 3.8 Newman-Keuls *post hoc* comparisons for the free running periods of the *mps4* transformants at 18, 25 and 29^{0}C

			{1}	{2}	{3}	{4}	{5}	{6}
			24.27467	24.81500	24.83855	23.92500	25.14706	24.71762
6F	18	{1}		**0.03033**	**0.04082**	0.10107	**0.00042**	**0.03779**
6F	25	{2}	**0.03033**		0.91208	**0.00018**	0.26442	0.64791
6F	29	{3}	**0.04082**	0.91208		**0.00019**	0.14797	0.83762
16B	18	{4}	0.10107	**0.00018**	**0.00019**		**0.00002**	**0.00060**
16B	25	{5}	**0.00042**	0.26442	0.14797	**0.00002**		0.18278
16B	29	{6}	**0.03779**	0.64791	0.83762	**0.00060**	0.18278	

Appendix 3.9a Newman-Keuls *post hoc* comparisons for the free running periods of the pooled lines of the *mel, mps1, mps3 and mps4* transformants at 18, 25 and 29^{0}C

			{1}	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}	{12}
			24.09301	25.04318	24.85299	28.26474	28.60086	29.20572	23.73600	25.12865	25.32880	24.09419	24.96757	24.80795
mel	18	{1}		**0.04407**	0.09464	**0.00001**	**0.00001**	**0.00001**	0.27687	**0.02689**	**0.00416**	0.99712	0.05947	0.07504
mel	25	{2}	**0.04407**		0.83119	**0.00001**	**0.00002**	**0.00002**	**0.00135**	0.79464	0.65929	**0.03150**	0.81787	0.89059
mel	29	{3}	0.09464	0.83119		**0.00002**	**0.00003**	**0.00003**	**0.00604**	0.83556	0.59563	0.05426	0.72710	0.89091
mps1	18	{4}	**0.00001**	**0.00001**	**0.00002**		0.30594	**0.01158**	**0.00001**	**0.00002**	**0.00001**	**0.00003**	**0.00002**	**0.00003**
mps1	25	{5}	**0.00001**	**0.00002**	**0.00003**	0.30594		0.06543	**0.00001**	**0.00001**	**0.00002**	**0.00001**	**0.00002**	**0.00003**
mps1	29	{6}	**0.00001**	**0.00002**	**0.00003**	**0.01158**	0.06543		**0.00002**	**0.00002**	**0.00001**	**0.00001**	**0.00003**	**0.00001**
mps3	18	{7}	0.27687	**0.00135**	**0.00604**	**0.00001**	**0.00001**	**0.00002**		**0.00060**	**0.00005**	0.51964	**0.00243**	**0.00603**
mps3	25	{8}	**0.02689**	0.79464	0.83556	**0.00002**	**0.00001**	**0.00002**	**0.00060**		0.54211	**0.02024**	0.87578	0.86579
mps3	29	{9}	**0.00416**	0.65929	0.59563	**0.00001**	**0.00002**	**0.00001**	**0.00005**	0.54211		**0.00322**	0.68937	0.60757
mps4	18	{10}	0.99712	**0.03150**	0.05426	**0.00003**	**0.00001**	**0.00001**	0.51964	**0.02024**	**0.00322**		**0.03908**	**0.02971**
mps4	25	{11}	0.05947	0.81787	0.72710	**0.00002**	**0.00002**	**0.00003**	**0.00243**	0.87578	0.68937	**0.03908**		0.87788
mps4	29	{12}	0.07504	0.89059	0.89091	**0.00003**	**0.00003**	**0.00001**	**0.00603**	0.86579	0.60757	**0.02971**	0.87788	

Appendix 3.9b Two way ANOVA for the free-running periods of the *mps3*, and *mel* (pooled lines) at 18, 25 and 29^{0}C

Source	MS	d.f.	F	P
Genotype (G)	.37800	1	.36667	.545096
Temperature(T)	49.46194	2	47.97971	.000000
G*T	3.99075	2	3.87116	.021449
Error	1.03089	510		

Appendix 3.9c Two way ANOVA for the free running periods of the *mps4*, and *mel* (pooled lines) at 18, 25 and 29^{0}C

Source	MS	d.f.	F	P
Genotype (G)	.15041	1	.15955	.689731
Temperature(T)	25.64375	2	27.20216	.000000
G*T	.03964	2	.04204	.958831
Error	.942710	532		

Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 4475–4480, April 1998
Genetics

Molecular coevolution within a *Drosophila* clock gene

(*per*/coevolution/Thr-Gly/repeat/circadian)

ALEXANDRE A. PEIXOTO*†‡, J. MICHAEL HENNESSY*†§, IAN TOWNSON*, GAITI HASAN¶‖, MICHAEL ROSBASH¶, RODOLFO COSTA§, AND CHARALAMBOS P. KYRIACOU*,**

*Department of Genetics, University of Leicester, Leicester, LE1 7RH, United Kingdom; §Dipartimento di Biologia, Universita di Padova, via Trieste 75, Padova 35121, Italy; and ¶Howard Hughes Medical Institute, National Science Foundation Science and Technology Center for Biological Timing, Department of Biology, Brandeis University, Waltham, MA 02254

Communicated by Howard A. Nash, National Institute of Mental Health, Bethesda, MD, January 16, 1998 (received for review November 4, 1997)

ABSTRACT **The *period* (*per*) gene in *Drosophila melanogaster* provides an integral component of biological rhythmicity and encodes a protein that includes a repetitive threonine-glycine (Thr-Gly) tract. Similar repeats are found in the *frq* and *wc2* clock genes of *Neurospora crassa* and in the mammalian *per* homologues, but their circadian functions are unknown. In Drosophilids, the length of the Thr-Gly repeat varies widely between species, and sequence comparisons have suggested that the repeat length coevolves with the immediately flanking amino acids. A functional test of the coevolution hypothesis was performed by generating several hybrid *per* transgenes between *Drosophila pseudoobscura* and *D. melanogaster*, whose repetitive regions differ in length by about 150 amino acids. The positions of the chimeric junctions were slightly altered in each transgene. Transformants carrying *per* constructs in which the repeat of one species was juxtaposed next to the flanking region of the other were almost arrhythmic or showed a striking temperature sensitivity of the circadian period. In contrast, transgenes in which the repeat and flanking regions were conspecific gave wild-type levels of circadian rescue. These results support the coevolutionary interpretation of the interspecific sequence changes in this region of the PER molecule and reveal a functional dimension to this process related to the clock's temperature compensation.**

Genes that are essential circadian clock components have been identified at the molecular level in *Drosophila* (*per* and *tim*), *Neurospora* (*frq, wc-1,* and *wc-2*), and the mouse (*Clock*) (reviewed in ref. 1). *per, frq,* and *wc-2* encode regions with repetitive sequences, which include runs of Thr-Gly or Ser-Gly dipeptides (2–4). Putative mammalian homologues of the fly *per* gene also encode similar repeats (5–8). The function of these repeats is unknown, but structural studies of $(\text{Thr-Gly})_n$ dipeptides reveal that they generate β-turn conformations that are extremely flexible and dynamic when challenged with different temperatures and polarities (9). In addition, the Thr-Gly repeats in *Drosophila melanogaster* are polymorphic in length and are distributed as a significant latitudinal cline in Europe, with high levels of the shorter length variants distributed predominantly in the south (10). This suggests that natural selection might be predisposing the length alleles to different types of environments. Statistical analyses of the Thr-Gly and surrounding DNA sequences with a variety of models designed to reveal the signature of natural selection have suggested that in both *D. melanogaster* (11) and *Drosophila simulans* (12) the frequencies observed of different natural Thr-Gly haplotypes are not consistent with the expectations of neutrality or drift. In addition, interspecific sequence analyses of the repetitive regions of over a dozen *Drosophila* species have suggested that the immediate flanking regions of the repeat coevolve with repeat length (13, 14). The evolutionary dynamics being played out in this region of the PER molecule suggest that the repeat "domain" may play a functional role.

To perform an experimental test for the proposed coevolution, we have generated a number of hybrid *per* genes between two species that have very different repeat lengths. We juxtaposed the repeat of one species next to the flanking region of the other, thereby breaking up the suspected coevolutionary interaction between the two regions. It would be predicted that this might have some effect on circadian behavior. A potential problem is that selection can act on fitness increments that are as small as the inverse of the effective population size (15), about 10^6 in *D. melanogaster* (16). Therefore, any disruption of the proposed coevolution between the two regions may be phenotypically undetectable given the limitations of laboratory experiments. Nevertheless, we selected two *per* genes, one from *Drosophila pseudoobscura* (17) and the other from *D. melanogaster* (2); the former has the longest repetitive region of all *per* genes so far identified. It consists of about 10 copies of a degenerate Thr-Gly motif to which is added a pentapeptide cassette, which has been derived by slippage from the dipeptide Thr-Gly repeat (14). There are 30–35 copies of this pentapeptide in different *D. pseudoobscura* strains (18), giving the repetitive region a length in excess of 200 amino acids (13, 14, 17, 18). In comparison, the *D. melanogaster* repetitive region is composed of about 20 pairs of Thr-Gly, although different strains also have different repeat copy number (18). Fig. 1 illustrates the Thr-Gly repeat and its proposed coevolving 5′ flanking sequences (13, 14), which are labeled as block P and consist of amino acids 665–695 in the *D. melanogaster* sequence (2). Block H (amino acids 639–664) cannot be aligned between *Drosophila* species (13, 14, 17) but may nevertheless represent additional sequences that could be involved in the putative interaction between the repeat and the immediately adjacent region. The 40-amino acid region immediately upstream of block H (amino acids 598–638) does not appear to coevolve with interspecific repeat length (14).

Four chimeric *per* transgenes were created (*mps2–5*; see Fig. 1) by using sequences from *D. melanogaster* and *D. pseudoobscura*. All hybrid *per* genes except *mps4* encode the 5′ half of *D. melanogaster* and the 3′ half of *D. pseudoobscura per* (Fig. 1). The position of the chimeric junction between the coding regions of the two species was manipulated (Fig. 1). The repeat and its immediate flanking regions H and P were made

Abbreviations: LD, light/dark; ZT, Zeitgeber time.

†A.A.P. and J.M.H. contributed equally to this work.
‡Present address: Department of Biology, Brandeis University, Waltham, MA 02254.
‖Present address: National Centre for Biological Science, TIFR CTR, Bangalore 560012, India.
**To whom reprint requests should be addressed. e-mail: cpk@leicester.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

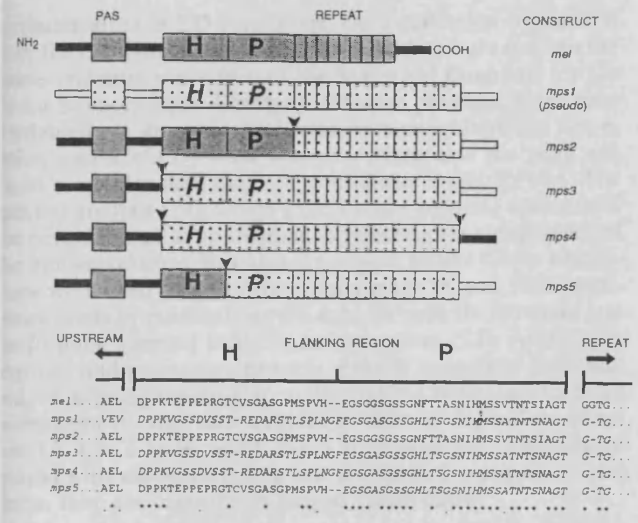


FIG. 1. (*Top*) *per* transgenes of *D. melanogaster* N- and C-terminal sequences (filled bars) *D. pseudoobscura* (open bars). The approximate positions of the N-terminal PAS domain (39) are shown as filled or speckled boxes for the two species. *D. melanogaster* sequences in blocks H (amino acids 639–664, ref. 2 for numbering) and P (amino acids 665–695) are in gray, and corresponding sequences in *D. pseudoobscura* are speckled and italicized. Thr-Gly repeats of *D. melanogaster* (amino acids 696–737) and *D. pseudoobscura* are shown as narrow repetitive units, 9 for *D. melanogaster*, representing the 20 pairs of Thr-Gly dipeptides, and 5 for *D. pseudoobscura*, representing the 10 imperfect dipeptide repeats in this species (13, 14). The 5 broader repeat units in *D. pseudoobscura* represent the 30 copies of the pentapeptide encoding sequence (see refs. 13 and 14). Arrows show positions of chimeric junctions. (*Bottom*) Coevolving amino acid blocks (marked H and P, see text). A few amino acids N-terminal to block H and the first few residues of the Thr-Gly repeat (C-terminal to block P) are shown. The length of this repeat is 66 amino acids in *D. melanogaster* and 209 amino acids in *D. pseudoobscura* (13). Asterisks indicate identical amino acids, and dots indicate conservative substitutions. *D. melanogaster* amino acids are shown in roman type, and *D. pseudoobscura* sequences are shown in italic type.

conspecific for *D. pseudoobscura* (chimeric constructs *mps3 and mps4*), or the repeat was heterospecific compared with *D. melanogaster* flanking regions H and P (construct *mps2*), or the repeat and region P were conspecific for *D. pseudoobscura* but heterospecific with regard to *D. melanogaster* region H and all upstream *per* sequences (*mps5*). Our results reveal striking phenotypic differences between the transformants, which supports the idea of an intragenic coevolution between the repeat and the flanking regions.

MATERIALS AND METHODS

Generation of Chimeric *per* Genes. The *mel* transgene consists of the 13.2-kb *per* transcription unit from *D. melanogaster* inserted into the cp20.1 transformation vector marked with *rosy*$^+$ and has been described previously (2). This fragment was also inserted into the pW8 transformation vector marked with *white*$^+$ (19), and two further independently transformed *mel* lines were generated. *mps1* has also been described previously (20). It contains the complete coding sequence of the *D. pseudoobscura per* gene fused to the 5′ upstream region of *D. melanogaster* at a position close to the 3′ end of the large first intron. These sequences are also carried within the cp20.1 transformation vector. To generate the novel *mps2–5* chimeric transgenes, PCR strategies were used. The *D. melanogaster* primers are represented with nucleotide numbering as in ref. 2 and *D. pseudoobscura* primers with numbers from the amino acid sequence from ref. 17. Transgenes *mps3–5* were inserted into pW8 and were generated in Leicester, whereas *mps2* was inserted into cp20.1 and made at Brandeis.

mps2. A 3′ junction primer, 5′-CGTGACCGTACCAGTGCCAGTGCCGGCAATGCTCG-3′, with the *D. pseudoobscura* sequence (amino acids 659–664) and *D. melanogaster* sequence (amino acids 5090–5107) was used with an upstream *D. melanogaster* 5′ primer, 5′-AAAGAGCTCGATCCGCCCAAAACG-3′ (*Sst*I site underlined). Included were three extra As that were digested before ligation. The *D. pseudoobscura* 3′ primer, 5′-AAATCTAGAGTTATCGGCTCG-3′, also had three extra As, which were digested before ligation. The chimeric *Sst*I-*Xba*I fragment obtained (21) was sequenced for errors and then ligated to the *Xba*I-*Eco*RI 3′ *D. pseudoobscura* fragment. This fragment was then inserted in front of the relevant *Sst*I site of the cp20.1 vector carrying the 5′ *D. melanogaster Bam*HI-*Sst*I fragment to give the final *mps2* clone. The *mps2* construct was confirmed with diagnostic restriction analysis and sequencing of the chimeric junction.

mps3. A 5′ *D. pseudoobscura* primer, 5′-CACCCGTGGAGCTCGACCCG-3′ (amino acids 619–638, *Sst*I site underlined, mismatch bold), was used with a 3′ *D. pseudoobscura* primer 5′-TTCTCCATCTCGTCGTTGTG-3′ (amino acids 878–884) to generate a 0.6-kb fragment, which was sequenced for errors and cut with *Sst*I and *Xba*I. This was ligated to a *D. pseudoobscura* 2.4-kb *Xba*I-*Eco*RI fragment reconstituting the 3′ *D. pseudoobscura* sequences. A *Bam*HI-*Sst*I *D. melanogaster per* fragment representing the 5′ part of the construct was ligated to the *D. pseudoobscura Sst*I-*Eco*RI fragment generating the *mps3* transgene in pW8. Diagnostic restriction analysis and DNA sequencing of the chimeric junction confirmed the correct construction of the chimeric gene.

mps4. The *mps3* 0.6-kb *D. pseudoobscura* fragment above was coamplified with a *D. melanogaster* fragment with a 5′ primer, 5′-AAGCACAACGACGAGATGGA-3′ (amino acids 5333–5352), and a 3′ primer, 5′-GCTACGCCTGTTCCGGATCC-3′ (amino acids 5627–5646, *Bam*HI site underlined). The *D. pseudoobscura* 3′ primer and the *D. melanogaster* 5′ primer are complementary. Further amplification of the initial products in the presence of the external 5′ and 3′ primers generates a chimeric fragment, which was restricted and used to replace the corresponding *D. melanogaster Sst*I-*Bam*HI fragment in pW8.

mps5. A *D. pseudoobscura* fragment was generated with a 5′ primer, 5′-GAGGGCAGTGGCGCCAGTGG-3′ (amino acids 628–634), and 3′ primer, 5′-TTCTCCATCTCGTCGTTGTG-3′ (amino acids 878–884). A *D. melanogaster* product was amplified with 5′ primer, 5′-AACTATAACGAGAACCTGCT-3′ (4874–4893), and 3′ primer, 5′-*CCACTGGCGCCACTGCCCTC*GTGGACGGGACT-3′ (italics, complementary to amino acids 628–634, 5002–5013). The two fragments, the 5′ *D. melanogaster* and 3′ *D. pseudoobscura* primers, amplify a chimeric fragment, which was sequenced for errors and digested with *Sst*I and *Xba*I to give a 0.6-kb fragment, which replaces the corresponding fragment in *mps3*.

P-Element Transformation and Behavioral Analysis. The four hybrid genes and two further *per* constructs carrying the cloned parental *D. melanogaster* (*mel*) and *D. pseudoobscura* (*mps1*) coding sequences were transformed into the appropriate hosts by using standard methods (22). *per*^{01} males carrying an autosomal copy of the transgene were examined with respect to their circadian locomotor activity under free running conditions (DD, constant darkness) at 18, 25, and 29°C. The activity of each fly was analyzed by using autocorrelation (23) and a high resolution spectral analysis (24, 25). Flies were monitored for 7 days in constant darkness (DD) after a previous entrainment period in a light/dark 12:12 cycle (LD 12:12). Data were collected in 30-min bins in an automated event recorder similar to ones described previously (26). Transformants were reared at 25°C in an LD 12:12 cycle, and monitoring was carried out at 18, 25, or 29°C after 2 days of

acclimatization in LD conditions. Data collection began 18 h after the last light to dark transition. The data were fed into the autocorrelation procedure of the Statistical Programs for the Social Sciences (SPSS, Chicago) statistical package. Significant rhythmicity in an autocorrelogram was one where the correlation coefficient (r) itself showed cycling, and the peak was equal or greater than the 95% confidence limits ($2/\sqrt{n}$). The spectral analysis (24), which gave a more accurate estimate of the period as well as confirming or rejecting the significance of the autocorrelation, was also employed. Monte-Carlo simulations were used to generate approximate 95 and 99% confidence limits by randomizing the data for each fly 100 times and performing spectral analyses on these data (27). Autocorrelograms and spectrograms were given a numerical code and judged as being rhythmic or arrhythmic on the basis of a blind assessment of their correlogram by three independent assessors (A.A.P., J.M.H., and C.P.K.). If a record gave a significant period with autocorrelation, but not spectral analysis, or vice versa, then the record was judged "arrhythmic."

RESULTS AND DISCUSSION

All four lines of the *D. melanogaster per*$^{+}$ transgene *mel* rescue rhythms in a high percentage of *per*^{01} hosts with periods in the circadian 24–25-h range at all three temperatures (Table 1). The *D. pseudoobscura per* coding region (*mps1*) is not as effective in rescuing *per*^{01} behavior, yet 50–75% of the transformants show significant rhythmicity at 18 and 25°C (see Fig. 2 and Table 1) with periods of *ca.* 28 h. These periods are stable, even at the highest temperature, when only 15% of the transformants are rhythmic. Previous results with the same *D. pseudoobscura* (*mps1*) construct (20) found fewer rhythmic individuals at 25°C, but this was because of the use of a less sensitive statistical measure of rhythmicity (20, 23, 24). The *mps2* construct (three lines) generated a small percentage of weakly rhythmic flies, barely above the level found for *per*^{01} mutants (Table 1; Figs. 2 and 3). Thus in *mps2* transformants, where the repeat of *D. pseudoobscura* lies directly adjacent to the heterospecific 5′ flanking material of *D. melanogaster* (Fig. 1), *per* gene function is severely disrupted. In contrast, the *mps3* gene (two lines) produced excellent rescue with 75–95% of transformants showing statistically significant rhythms and with mean periods in the 23–25-h range at all temperatures (Figs. 2 and 3). In *mps3* the repeat of *D. pseudoobscura* lies next to its conspecific 5′ flanking sequences (blocks H and P in Fig. 1).

The dramatic phenotypic differences observed between the *mps2* and *mps3* transformants strongly support the idea of an intragenic coevolution between the length of the repeat and the immediate flanking upstream ≈60 amino acids in blocks H and P (13, 14). In addition, the *mps5* transgene generated a phenotype intermediate between *mps2* and *mps3*, in that both lines showed robust rhythms, but the periods were extremely temperature-dependent, remarkably so in line 25 (Table 1; Fig. 2). The phenotype displayed by these *mps5* transgenic flies is reminiscent of the classic *per*^{L1} mutation, which also gives a temperature-sensitive lengthening of period (26, 28). However, even *mps5* transformant line 23, which gives the least dramatic phenotype of the two (Table 1), shows a larger change in period than that reported in *per*^{L1} between 17 and

Table 1. Free running circadian locomotor activity periods of various *per* transformants at different temperatures based on spectral analysis (24)

	18°C			25°C			29°C		
Line	*N*	*n*	Period, h	*N*	*n*	Period, h	*N*	*n*	Period, h
mel									
2a	61	41	24.6 ± 0.1	25	23	25.0 ± 0.1	53	45	24.8 ± 0.1
17a	59	31	24.1 ± 0.2	24	20	24.3 ± 0.7	46	26	24.2 ± 0.2
34a	51	30	24.2 ± 0.2	47	39	25.0 ± 0.1	32	32	24.9 ± 0.2
116a	50	41	23.5 ± 0.1	33	30	25.5 ± 0.1	33	31	25.4 ± 0.2
Pooled	221	143	24.1 ± 0.1	129	112	25.0 ± 0.1	164	134	24.9 ± 0.1
mps1									
I20	26	17	29.4 ± 0.8	31	18	28.3 ± 1.2	61	13	29.0 ± 1.2
I26	25	21	27.4 ± 0.6	37	17	28.9 ± 0.8	82	8	29.5 ± 1.0
Pooled	51	38	28.3 ± 0.5	68	35	28.6 ± 0.6	143	21	29.2 ± 0.8
mps2									
6	31	4	22.7 ± 4.1	61	5	33.0 ± 1.6	65	5	29.8 ± 1.6
9	28	6	24.6 ± 2.1	29	8	28.6 ± 2.0	39	7	28.6 ± 2.7
22	28	11	27.5 ± 2.2	68	8	28.1 ± 2.7	34	5	28.8 ± 2.4
Pooled	87	21	25.8 ± 1.5	158	21	29.5 ± 1.3	138	17	29.0 ± 1.3
mps3									
65c	19	15	24.1 ± 0.1	44	42	25.7 ± 0.1	15	13	25.6 ± 0.2
67a	22	15	23.3 ± 0.2	32	32	24.4 ± 0.1	12	12	25.0 ± 0.2
Pooled	41	30	23.7 ± 0.1	76	74	25.1 ± 0.1	27	25	25.3 ± 0.2
mps4									
6f	17	15	24.3 ± 0.1	23	30	24.8 ± 0.1	66	62	24.8 ± 0.1
16b	27	16	23.9 ± 0.1	20	17	25.1 ± 0.2	24	21	24.7 ± 0.1
Pooled	44	31	24.1 ± 0.1	43	37	25.0 ± 0.1	90	83	24.8 ± 0.1
mps5									
23	25	20	24.0 ± 0.1	68	65	27.2 ± 0.1	51	48	28.3 ± 0.2
25	67	47	24.9 ± 0.3	73	48	30.9 ± 0.7	97	74	38.9 ± 1.0
Pooled	92	67	24.6 ± 0.2	141	113	28.8 ± 0.4	148	122	34.7 ± 0.8
per^{01}	31	5	22.6 ± 1.1	32	3	19.3 ± 1.9	57	4	21.5 ± 2.1

Results based on autocorrelation-derived periods are almost identical (data not shown, but see Fig. 2). Two-way ANOVA omitting the largely arrhythmic *mps2* transformants revealed significant effects of temperature (T), genotype (G), and T × G interaction ($P < 0.001$ in all cases). The Newman–Keuls *a posteriori* procedure revealed significant changes in period over temperature for both lines of the *mps5* transformants ($P < 0.002$ for both lines). *N*, number of males examined; *n*, number of significant rhythmic males. Results are mean ± SEM.

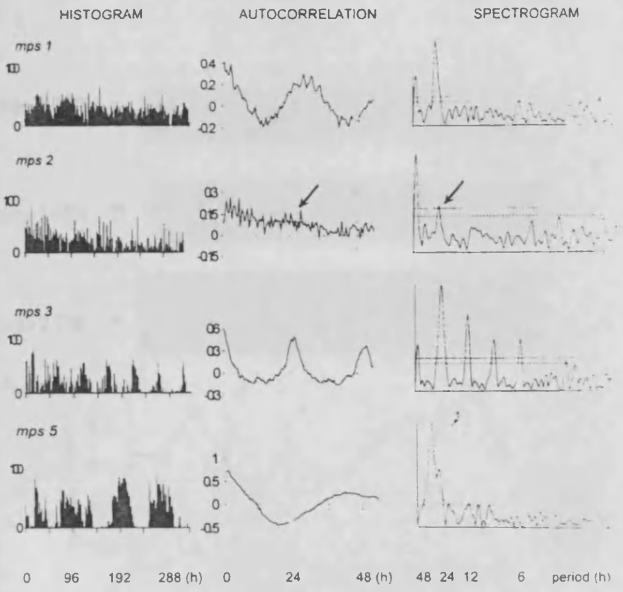


FIG. 2. Free running locomotor activity histograms, autocorrelation, and spectral analysis for four male per^{01} transformants monitored at 29°C carrying the *mps1, mps2, mps3,* and *mps5* transgenes. Histograms give raw activity events in 30-min bins. The correlograms show r and 95% confidence limits either side of $r = 0$. Spectrograms show 95% (dashed, *Lower*) and 99% (dotted, *Upper*) confidence limits (27). From top to bottom are shown: *mps1,* 28-h cycle; *mps2,* one of few males rhythmic by our criteria (see *Materials and Methods*), and barely so, period *ca.* 26 h (arrow); *mps3,* typical of *mel* and *mps4* transformants (not shown), period *ca.* 24.5 h; *mps5,* >45-h period.

25°C (28). These *mps5* results confirm that the unalignable amino acids in block H must also play a role in the coevolutionary interaction between the flanking region and the repeat, and suggest that these residues are under positive natural selection rather than drift. The *mps4* transgene produced essentially wild-type rhythms in terms of strength of rescue, period length, and period stability over temperature, revealing that a *D. pseudoobscura* repeat plus its flanking sequences represents a "coevolved" module, which can function perfectly well within an otherwise *D. melanogaster* PER protein.

Within the *per* genes of Diptera, the Thr-Gly encoding hexamer ACNGGN has mutated to give both perfect and imperfect repetitive motifs of various lengths (13, 14, 17, 18). *D. pseudoobscura* represents the species with the longest repetitive region that has been identified to date and contains both the Thr-Gly dipeptide and a related pentapeptide (13, 14, 17, 18). Secondary structure predictions of the repetitive regions in a large number of *Drosophila* species suggest that,

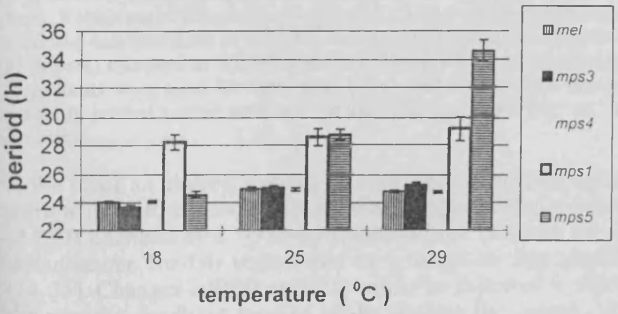


FIG. 3. Pooled mean (±SEM) free running locomotor activity periods for the different transformants at 18, 25, and 29°C. Data for the individual lines are in Table 1. *mps5* gives a dramatic increase in period at higher temperatures. *mps2* results are not shown as most transformants were arrhythmic, and those that were rhythmic showed a wide range of periodicities (see Table 1).

irrespective of sequence composition, the amino acids in these regions generate a series of flexible turns (13, 18). Conformational analysis of poly(Thr-Gly) peptides, the major component of the *D. melanogaster* repeat, reveals that Thr-Gly repeats generate type II or type III β-turns (9). Consequently, it would appear that it is the length of the repeat that coevolves with the flanking region rather than the species-specific amino acid composition of the arrays.

Overall, these results suggest that the coevolutionary interaction between the repeat and its flanking sequences may serve a function related to the clock's cardinal property of temperature compensation (29). The improved rescue of *mps3* over *mps1* shows that the *D. melanogaster* N-terminal half of PER is also required for robust behavioral rhythms in addition to a coevolved Thr-Gly region. This may reflect the fact that the PAS dimerization domain of PER (1) encoded in *mps3* is conspecific with the Timeless (TIM) partner molecule of the host (30). Furthermore, it is important to note that in spite of the *mps1* transgene's relatively poor rescue, its temperature compensation is maintained, even at the highest temperature. Therefore it does not follow that a poorly rescuing *per* gene will inevitably also suffer from defects in temperature compensation.

The sequences flanking the repeat in blocks H and P (Fig. 1) contain 27 amino acid substitutions between the two parental species, most of which are distributed in block H (Fig. 1). Comparing this flanking region to sequences in the databases did not reveal any significant similarities to motifs that might illuminate the function of this region. To learn more about the *mps* products, Western blots were performed with head extracts from *mps2* and *mps3* transformants over the circadian cycle in an LD 12:12 regime at 25°C, with a polyclonal anti-PER antibody (Fig. 4). Serial dilutions of total protein were used to quantify the amount of mps2PER signal in relation to wild-type and mps3PER (data not shown). Whereas *mps3* and wild-type PER produced similar levels of signal [compare mps3PER at Zeitgeber time (ZT 22) with wild-type at ZT 0 in lane c], mps2PER detected was 5–7 times less abundant (Fig. 4*A*, see top panel) and similar to the levels of wild-type PER found in tim^0 mutants (31). The *mps3* transformants show significant differences in PER levels over the circadian cycle, with all three individual mps3PER blots showing a peak in intensity just before dawn at ZT 22 (Fig. 4). This temporal profile is similar to that observed with wild-type PER (e.g., ref. 32, and data not shown). Wild-type *D. melanogaster* PER also shows a change in mobility and a "broadening" of the PER band over the circadian cycle because of phosphorylation (32). This was present in our wild-type control blots (data not shown) and can also be appreciated by inspecting the broader wild-type control bands that can be seen in lanes c of Fig. 4, which represents PER mobility at ZT 0. However, the changes in mobility because of phosphorylation appeared to be considerably reduced with mps3PER (Fig. 4), even though these transformants gave robust behavioral rhythms.

In sharp contrast, the mps2PER product appears to be expressed both in the light and dark phases, with little evidence for any robust circadian cycling or for any mobility shifts indicative of posttranslational modifications (Fig. 4). The apparent failure of mps2PER to cycle suggests either altered stability of the chimeric protein or less effective synthesis due perhaps to low activity of mps2PER, which in turn could lead to a defective cycle, low mRNA levels, and low synthesis of mps2PER. Because the levels of mps2PER are so low, we cannot state definitively that cycling was not present. On the other hand, *mps2* transformant head extracts, blotted with an anti-TIM antibody (29), revealed a low amplitude (3× peak-to-trough) TIM cycling (Fig. 4), similar to that of per^{01} mutants under LD 12:12 conditions (33).

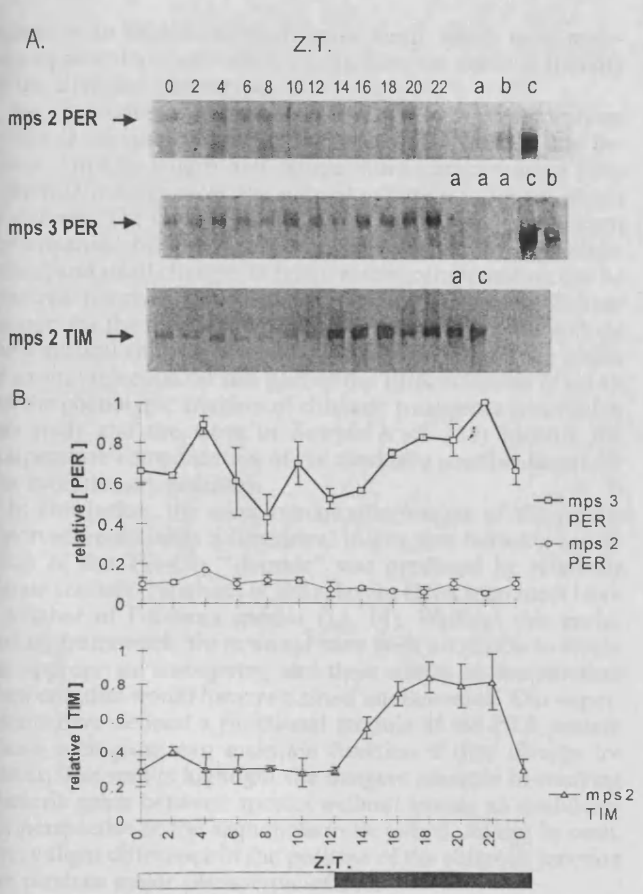


FIG. 4. (*A*) Western blots of *mps2* and *mps3* transformant heads collected at different Zeitgeber times (ZT 0, lights on; ZT 12, lights off) in LD 12:12 at 25°C. Lanes on the right show per^{01} (a), tim^{0} (b), and per^{+} Canton-S (c) controls at ZT 0, except for the TIM blot where the controls were taken at ZT 18. From top to bottom are shown: *mps2* blotted with anti-PER; *mps3* blotted with anti-PER; *mps2* blotted with anti-TIM. The predicted primary M_r of the mps2–5PER product is 136,000, heavier than either *D. melanogaster* (128,000) or *D. pseudoobscura* (132,000) proteins. However, *D. melanogaster* PER runs at about 180,000 (32), and the chimeric MPS proteins run a little higher. (*B*) Densitometry analysis for three separate anti-PER and anti-TIM blots. The highest intensity PER or TIM band within each blot was given a value of unity. Means ± SEM are shown. Control experiments in which serial amounts of total protein were loaded for mps3PER, wild-type PER, and mps2PER at ZT 0, followed by Western blotting with anti-PER, revealed an approximate 5- to 7-fold reduction in mps2PER intensity for the same amount of protein (data not shown). No characteristic wild-type PER cycling (32) was ever observed in *mps2* blots, although all three blots showed the highest protein levels at ZT 4. Western blotting was performed as in ref. 32 with minor modifications. Rabbit anti-PER antibody (gift of J. Hall and R. Stanewsky) was used at a concentration of 1:10,000, and rat anti-TIM antibody (gift of M. Myers) was used at 1:1,000. On each Western blot, equal amounts of proteins were used for each time point and for controls, but the amounts loaded varied between 50 and 200 μg, depending on the genotype.

We favor an altered stability of mps2PER, based on recent work with PER–reporter fusions, which suggests that a region of PER encoded by a ≈700-bp fragment that contains the *D. melanogaster* Thr-Gly region may be a target for degradation (34, 35). Changes in PER stability would be expected to affect the negative feedback loop of *per* by altering the period (36). Perhaps then, the proximal cause for the dramatic lack of phenotypic temperature compensation in *mps5* flies lies in a subtle temperature-sensitive change in PER degradation. The arrhythmicity of *mps2* transformants might therefore reflect a more serious instability in the mps2PER molecule. A dysfunction in the mRNA, although possible, is unlikely given the statistical analysis of DNA and protein sequences that gave rise to these experiments (13, 14). These revealed a significant correlation between amino acid changes in the flanking regions, but not between synonymous nucleotide positions, and repeat length differences between pairs of *Drosophila* species. As the substitution rate in the synonymous position largely reflects the molecular clock, then the correlation between flanking region amino acid changes and repeat length differences was not simply because of the evolutionary time elapsed because any two species shared a common ancestor (13, 14). If RNA structure was involved in this coevolutionary process, we might expect that any two species with different repeat lengths would also have more changes in their flanking RNA sequences and so generate a positive correlation between the silent nucleotide position and repeat length.

This coevolution could thus represent the signature of natural selection as it attempts to maintain the appropriate circadian degradation kinetics for PER in the presence of a relatively high mutation rate in the repetitive region (11). It is not clear whether mutations in the flanking region may have imposed selection pressure on the length of the repeat or vice versa. However, because mutations in the repeat and flanking region are tightly linked, they conform to models where compensatory neutral mutations can become fixed even when the individual mutations are deleterious (37). We can envisage a scenario where a small change in the length of the repeat, for example by a gain or loss of a single repeat unit, may be slightly deleterious under certain conditions (see ref. 38). A mutation in the flanking region may restore the status quo, even though this compensating mutation may be slightly deleterious in the absence of the initial change. After many rounds of mutation, heterospecific combinations of flanking regions and repeat arrays could become incompatible with function, as in the case of the two species studied here. This process would be enhanced in a region associated with a high mutation rate such as the Thr-Gly repeat (11) and if the domain determined an adaptive character, which circadian temperature compensation is likely to represent.

Like the *mps5* transformant, the per^{L1} mutant also shows significant increases in the period with higher temperatures (26, 28). The mutation maps to the PAS region of the PER protein, which is involved in protein-protein interactions (39, 40) and which may mediate the negative feedback mechanism believed to be at the heart of *per* function (41). PER^{L1}/TIM protein–protein interactions are temperature-sensitive in a yeast assay (41), so perhaps the *mps2* and *mps5* transgenes produce defective PER–TIM dimerization (but see ref. 42). In addition, it may be more than coincidence that the frq^{3} and frq^{7} mutations in *Neurospora crassa*, which disrupt the temperature compensation of the circadian period, map to the immediate flanking regions of the fungal Thr-Gly/Ser-Gly motif (43). Perhaps these repetitive regions within the two clock genes serve a common function related to temperature compensation. In *Neurospora* this could be in addition to the different forms of FRQ product that are differentially translated at high and low temperatures (44).

Coevolution within molecules is rarely studied, although a notable example concerns the *Drosophila Adh* gene, where compensatory mutations appear to maintain long range interactions between the 5′ and 3′ regions of the transcript (45). The interspecific compensatory changes in the flanking regions that track the length changes in the Thr-Gly repeat seem to be driven, at least in part, by selection aimed at maintaining the temperature compensation of the circadian oscillator in the face of the instability provided by the rapidly evolving repetitive region. However, this might form part of a bigger overall picture, which has the Thr-Gly domain as a key feature of the degradation kinetics of the PER protein. Thus the Thr-Gly and associated regions may not provide part of the mechanism for the temperature compensation of the clock *per se* but rather

represent an evolutionarily dynamic motif, which must maintain a particular conformation to mediate the thermal stability of the circadian phenotype.

Recent studies of the natural Thr-Gly length polymorphism within *D. melanogaster* have also revealed a relationship between Thr-Gly length and temperature compensation (38). Within *D. melanogaster*, the natural variation in repeat length of different Thr-Gly alleles (from 14 to 24 Thr-Gly pairs) is not accompanied by any significant flanking amino acid alterations, and small changes in temperature compensation can be detected between the different variants generating further support for the coevolution hypothesis (38). Thus not only do the statistical analyses of Thr-Gly sequences suggest the action of natural selection on this part of the PER molecule (11–14), but the phenotypic analyses of chimeric transgenes reported in this study and the work of Sawyer *et al.* (38) identify the temperature compensation of the clock as a possible target for this evolutionary pressure.

In conclusion, the most remarkable feature of the results reported here is that a functional interaction between subregions of the Thr-Gly "domain" was predicted by relatively simple statistical analyses of the relevant DNA sequences from a number of Dipteran species (13, 14). Without this evolutionary framework, there would have been no reason to create the appropriate transgenes, and their effects on temperature compensation would have remained undiscovered. Our experiments have defined a functional module of the PER protein whose subregions can maintain function if they diverge together. The results highlight the dangers inherent in studying chimeric genes between species without having an evolutionary perspective of the sequences to be joined. As can be seen, a very slight difference in the position of the chimeric junction can produce major phenotypic effects.

We thank the Science and Engineering Research Council, Biotechnology and Biological Sciences Research Council, and Human Frontiers Science Program for grants (to C.P.K.) and the Commission of the European Communities for a grant (to C.P.K. and R.C.). We acknowledge a Ministerio dell'Universitá e della Ricerca Scientifica e Technologia-British Council award, a Biotechnology and Biological Sciences Research Council studentship (to I.T.), and the Brazilian Government for a Conselho Nacional de Pesquisas scholarship (to A.A.P.). M.R. and G.H. were supported by National Institutes of Health Grant GM33205.

1. Rosato, E., Piccin, A. & Kyriacou, C. P. (1997) *BioEssays* **19,** 1075–1082.
2. Citri, Y., Colot, H. V., Jacquier, A. C., Yu, Q., Hall, J. C., Baltimore, D. & Rosbash, M. (1987) *Nature (London)* **326,** 42–47.
3. McClung, C. R., Fox, B. A. & Dunlap, J. C. (1989) *Nature (London)* **339,** 558–562.
4. Linden, H. & Macino, G. (1997) *EMBO J.* **16,** 98–109.
5. Sun, Z. S., Ulbrecht, U., Zhuchenko, O., Bailey, J., Eichele, G. & Lee, C. C. (1997) *Cell* **90,** 1003–1011.
6. Tei, H., Okamura, H., Shigeyoshi, Y., Fukuhara, C., Ozawa, R., Hirose, M. & Sakaki, Y. (1997) *Nature (London)* **389,** 512–516.
7. Albecht, U., Sun, Z. S., Eichele, G. & Lee, C. C. (1997) *Cell* **91,** 1055–1064.
8. Shearman, L. O., Zylka, M. J., Weaver, D. R., Kolakowski, L. F., Jr. & Reppert, S. M. (1997) *Neuron* **19,** 1261–1269.
9. Castiglione-Morelli, M. A., Guantieri, V., Villani, V., Kyriacou, C. P., Costa, R. & Tamburro, A. M. (1995) *Proc. R. Soc. London Ser. B* **260,** 155–163.
10. Costa, R., Peixoto, A. A., Barbujani, G. & Kyriacou, C. P. (1992) *Proc. R. Soc. London Ser. B* **258,** 43–49.
11. Rosato, E., Peixoto, A. A., Costa, R. & Kyriacou, C. P. (1997) *Genet. Res.* **69,** 89–99.
12. Rosato, E., Peixoto, A. A., Barbujani, G., Costa, R. & Kyriacou, C. P. (1994) *Genetics* **138,** 693–707.
13. Peixoto, A. A., Campesan, S., Costa, R. & Kyriacou, C. P. (1993) *Mol. Biol. Evol.* **10,** 127–139.
14. Nielsen, J., Peixoto, A. A., Piccin, A., Costa, R., Kyriacou, C. P. & Chalmers, D. (1994) *Mol. Biol. Evol.* **11,** 839–853.
15. Kimura, M. (1962) *Genetics* **47,** 713–719.
16. Aquadro, C. F. (1992) *Trends Genet.* **8,** 355–362.
17. Colot, H., Hall, J. C. & Rosbash, M. (1988) *EMBO J.* **7,** 3929–3937.
18. Costa, R., Peixoto, A. A., Thackeray, J. R., Dalgleish, R. & Kyriacou, C. P. (1991) *J. Mol. Evol.* **32,** 238–246.
19. Klemenz, R., Weber, U. & Gehring, W. J. (1987) *Nucleic Acids Res.* **15,** 3947–3959.
20. Petersen, G., Hall, J. C. & Rosbash, M. (1988) *EMBO J.* **7,** 3939–3947.
21. Yon, J. & Fried, M. (1989) *Nucleic Acids Res.* **17,** 4895.
22. Rubin, G. M. & Spradling, A. C. (1982) *Science* **218,** 348–352.
23. Dowse, H. B. & Ringo, J. M. (1989) *J. Theor. Biol.* **139,** 487–515.
24. Roberts, D. H., Lehar, J. & Dreher, J. W. (1987) *Astron. J.* **93,** 968–989.
25. Kyriacou, C. P. & Hall, J. C. (1989) *Anim. Behav.* **37,** 850–859.
26. Ewer, J., Hamblen-Coyle, M., Rosbash, M. & Hall, J. C. (1990) *J. Neurogenet.* **7,** 31–73.
27. Konopka, R. J., Kyriacou, C. P. & Hall, J. C. (1996) *J. Neurogenet.* **11,** 117–140.
28. Konopka, R. J., Pittendrigh, C. & Orr, D. (1989) *J. Neurogenet.* **6,** 1–10.
29. Pittendrigh, C. S. (1954) *Proc. Natl. Acad. Sci. USA* **40,** 1018–1029.
30. Myers, M. P., Wager-Smith, K., Rothenfluh-Hilfiker, A. & Young, M. W. (1996) *Science* **271,** 1736–1740.
31. Price, J. L., Dembinska, M. A., Young, M. W. & Rosbash, M. (1995) *EMBO J.* **14,** 4044–4049.
32. Edery, I., Zweibel, L. J., Dembinska, M. E. & Rosbash, M. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2260–2264.
33. Zeng, H., Qian, Z., Myers, M. P. & Rosbash, M. (1996) *Nature (London)* **380,** 129–135.
34. Dembinska, M. E., Stanewsky, R., Hall, J. C. & Rosbash, M. (1997) *J. Biol. Rhythms* **12,** 157–172.
35. Stanewsky, R., Frisch, B., Brandes, C., Hamblen-Coyle, M. J., Rosbash, M. & Hall, J. C. (1996) *J. Neurosci.* **17,** 676–696.
36. Hardin, P. E., Hall, J. C. & Rosbash, M. (1990) *Nature (London)* **343,** 536–540.
37. Kimura, M. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 5969–5971.
38. Sawyer, L., Hennessy, J. M., Peixoto, A. A., Rosato, E., Parkinson, H. E., Costa, R. & Kyriacou, C. P. (1997) *Science* **278,** 2117–2120.
39. Huang, Z. J., Edery, I. & Rosbash, M. (1993) *Nature (London)* **364,** 259–262.
40. Huang, Z. J., Curtin, K. D. & Rosbash, M. (1995) *Science* **267,** 1169–1172.
41. Gekakis, N., Saez, L., Delahaye-Brown, A. M., Myers, M. P., Sehgal, A., Young, M. W. & Weitz, C. J. (1995) *Science* **270,** 811–815.
42. Rutila, J. E., Zeng, H., Le, M., Curtin, K. D., Hall, J. C. & Rosbash, M. (1996) *Neuron* **17,** 921–929.
43. Aronson, B. D., Johnson, K. A. & Dunlap, J. C. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 7683–7687.
44. Liu, Y., Garceau, N. Y., Loros, J. J. & Dunlap, J. C. (1997) *Cell* **89,** 477–486.
45. Parsch, J., Tanda, S. & Stephan, W. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 928–933.

Appendices for Chapter 4

Appendix 4.1 Recovery and free-running periods of *D.pseudoobscura* (Ay, PA), *D.melanogaster* (Br,17m), *mps1, mps2 and mps3* in DD. Total numbers tested = N, rhythmic individuals = n, average period of n = t and standard error of the mean = se. [see Fig 1 Ch.3 for data on individual transformant lines]

Temperature	12 °C				15 °C				18 °C				21 °C				25 °C				29 °C			
Recovery	N	n	t	se	N	n	t	se	N	n	t	se	N	n	t	se	N	n	t	se	N	n	t	se
pseud (Ay)	35	30	23.8	0.13	27	19	23.8	0.17	40	38	24.0	0.15	19	17	24.1	0.16	15	15	24.7	0.32	43	27	26.3	0.71
pseud (PA)	12	11	23.5	0.18	11	11	23.5	0.15	11	11	23.7	0.40	5	4	24.3	0.70	14	13	24.9	0.23	54	31	26.9	0.30
mel (Br)	20	7	22.3	0.86	29	16	22.5	0.54	42	32	23.2	0.13	15	14	23.4	0.19	32	29	24.3	0.33	36	31	23.2	0.22
mel (17m)									40	36	22.8	0.06					15	14	22.9	0.09	32	28	22.6	0.30
mps1 (2lines)									51	38	28.3	0.53					65	35	28.6	0.59	143	21	29.2	0.84
mps3 (2lines)									41	30	23.7	0.14					76	74	25.1	0.11	27	25	25.3	0.16
mps4 (2lines)									44	31	24.1	0.09					43	37	25.0	0.10	90	83	24.8	0.08

Appendix 4.3 Planned comparison of peak bin values of *D.pseudoobscura* strains and *D.melanogaster* lines pooled into genotypes

Planned comparison *pseud* (Ay + Pa) v *mel* (Br + 17m)						
1-GENO						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	1039.109	67	55.59193	18.69173	5.22E-05

Appendix 4.4a Planned comparison of peak bin values of *D.pseudoobscura*, *D.melanogaster* and the transformants mps1, mps3 and 13.2 with strains and lines nested under genotypes

Planned comparison of all genotypes						
Lines as random effects and nested within genotype						
1-GENO, 2-LINE						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	4	447.1301	5	55.00224	8.129307	0.020534
2	5	55.00224	195	70.18255	0.783702	0.562538
12	--	--	--	--	--	--

4.4b Planned comparison of peak bin values of *D.pseudoobscura* and the transformant *mps1*, with strains and lines nested under genotype.

Planned comparison of *pseud* (Ay + Pa) v *mps1* (l20 + l26)						
Lines as random effects and nested within genotype						
1-GENO, 2-LINE						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	159.0885	2	25.49832	6.239174	0.129795
2	2	25.49832	54	80.26421	0.31768	0.729187
12	--	--	--	--	--	--

4.4c Planned comparison of peak bin values of *D.melanogaster* and the transformants *13.2* and *mps3*, with strains and lines nested under genotypes.

Planned comparison of :						
mel (Br +17m) v *13.2* (2a +17a) v *mps3* (65c +67a)						
Lines as random effects and nested within genotype						
1-GENO, 2-LINE						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	2	7.966892	3	74.67151	0.106693	0.902065
2	3	74.67151	141	66.32149	1.125902	0.34079
12	--	--	--	--	--	--

4.4d Planned comparison of peak bin values of the transformants *mps1* and *13.2*, with strains and lines nested under genotypes.

Planned comparison of *mps1* (I20+I26) v *13.2* (2a+17a)						
Lines as random effects and nested within genotype						
1-GENO, 2-LINE						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	478.3405	2	13.75415	34.77791	0.02757
2	2	13.75415	94	78.0436	0.176237	0.838696
12	--	--	--	--	--	--

Appendix 4.8a Comparison of *D.pseudoobscura* strains Ay and Pa over bins 29 - 37

Two way ANOVA of *D.pseudoobscura* lines (Ay v Pa) bins 29 to 37						
1-LINE, 2-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	3764.487	504	431.6617	8.720922	0.003293
2	8	3411.971	504	431.6617	7.904272	4.83E-10
12	8	406.0103	504	431.6617	0.940575	0.482392

4.8b *mps1* Comparison of *mps1* lines I20 and I26 over bins 29 - 37

Comparison of *mps1* lines I20 and I26 over bins 29 - 36

Two way ANOVA of mps1 lines (I20 and I26) bins 29 to 36						
1-LINE, 2-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	1804.085	336	60.36358	29.88698	8.95E-08
2	7	66.45059	336	60.36358	1.100839	0.362095
12	7	54.63901	336	60.36358	0.905165	0.502503

4.8c Comparison of *D.melanogaster* strains Ret 4 and Ret 9 over bins 29 - 37

Two way ANOVA of *D.melanogaster* lines (Ret 4 v Ret 9) bins 29 to 37						
1-LINE, 2-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	846.6457	387	43.32056	19.54374	1.28E-05
2	8	802.9879	387	43.32056	18.53595	1.52E-23
12	8	17.77432	387	43.32056	0.410298	0.91459

4.8d Comparison of *13.2* transformant lines *2a,11a 17a and 41a* over bins 29 - 37

Two way ANOVA of *13.2* lines (2a v 11a v 17a v 41a) bins 29 to 37						
1-LINE, 2-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	3	5917.547	990	114.8588	51.52018	0
2	8	619.4595	990	114.8588	5.393225	1.16E-06
12	24	68.94851	990	114.8588	0.600289	0.935531

4.8e Comparison of *mps3* transformant lines 65c *and* 67a over bins 29 - 37

Two way ANOVA of *mps3* lines (65c v 67a) bins 29 to 37						
1-LINE, 2-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	4651.317	270	158.563	29.3342	1.35E-07
2	8	336.5725	270	158.563	2.122643	0.033967
12	8	75.79992	270	158.563	0.478043	0.871304

4.8f Comparison of *mps4* transformant lines 6f *and* 16b over bins 29 - 37

Two way ANOVA of *mps4* lines (6f v 16b) bins 29 to 37						
1-LINE, 2-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	3.678033	234	90.3963	0.040688	0.840317
2	8	761.8007	234	90.3963	8.427344	4.69E-10
12	8	22.63372	234	90.3963	0.250383	0.980392

Appendix 4.9a Pairwise comparison of *D.pseudoobscura* and *mps1* transformants over bins 29-36 with strains and lines nested under genotype.

ANOVA *pseud v mps1* lines nested within genotypes (bins 29 to 36)						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	72888.49	2	2445.496	29.8052	0.031952
2	2	2445.496	784	282.838	8.646278	0.000193
3	7	812.4897	14	255.9645	3.174228	0.03132
12	--	--	--	--	--	--
13	7	513.3448	14	255.9645	2.005532	0.126843
23	14	255.9645	784	282.838	0.904986	0.553101
123	--	--	--	--	--	--

4.9b Pairwise comparison of *D.melanogaster* and *13.2* transformants over bins 29-36 with strains and lines nested under genotype.

ANOVA *melanogaster v 13.2* lines nested within genotypes (bins 29 to 36)						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	1747.453	2	4481.25	0.389948	0.596067
2	2	4481.25	1240	99.81434	44.89586	1.5E-19
3	7	1111.108	14	41.08772	27.04235	4.25E-07
12	--	--	--	--	--	--
13	7	49.91157	14	41.08772	1.214756	0.357011
23	14	41.08772	1240	99.81434	0.411641	0.971794
123	--	--	--	--	--	--

Appendix 4.10a Pairwise comparison of *13.2* transformants and *D.melanogaster* over bins 29-36 with lines nested under genotype

ANOVA *13.2 v mel* lines nested within genotypes (bins 29 to 36)						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	1747.453	2	4481.25	0.389948	0.596067
2	2	4481.25	1240	99.81434	44.89586	1.5E-19
3	7	1111.108	14	41.08772	27.04235	4.25E-07
12	--	--	--	--	--	--
13	7	49.91157	14	41.08772	1.214756	0.357011
23	14	41.08772	1240	99.81434	0.411641	0.971794
123	--	--	--	--	--	--

4.10b Pairwise comparison of *mps3* transformants and *D.melanogaster* over bins 29-36 with lines nested under genotype.

ANOVA *mps3 v mel* lines nested within genotypes (bins 29 to 36)						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	2358.676	2	2776.101	0.849636	0.453963
2	2	2776.101	584	98.40958	28.20966	2.02E-12
3	7	241.5641	14	39.43343	6.125871	0.002024
12	--	--	--	--	--	--
13	7	289.5074	14	39.43343	7.341676	0.000832
23	14	39.43343	584	98.40958	0.400707	0.97476
123	--	--	--	--	--	--

4.10c Pairwise comparison of *mps3* and *13.2* transformants over bins 29-36 with lines nested under genotype.

ANOVA *mps3 v 13.2* lines nested within genotypes (bins 29 to 36)						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	383.3319	2	6357.958	0.060292	0.828934
2	2	6357.958	1136	130.7393	48.6308	5.44E-21
3	7	332.6934	14	68.61273	4.848858	0.00588
12	--	--	--	--	--	--
13	7	227.8991	14	68.61273	3.321528	0.026636
23	14	68.61273	1136	130.7393	0.524806	0.919845
123	--	--	--	--	--	--

4.10d Pairwise comparison of *mps3* and *mps4* transformants over bins 29-36 with lines nested under genotype.

ANOVA *mps3 v mps4* lines nested within genotypes (bins 29 to 36)						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	988.7635	2	2330.595	0.424254	0.581665
2	2	2330.595	448	138.0832	16.8782	8.57E-08
3	7	466.1472	14	45.65292	10.21068	0.000145
12	--	--	--	--	--	--
13	7	256.5646	14	45.65292	5.619894	0.003032
23	14	45.65292	448	138.0832	0.330619	0.989936
123	--	--	--	--	--	--

4.10e Pairwise comparison of *mps4* transformants and *D.melanogaster* over bins 29-36 with lines nested under genotype

ANOVA *mps4 v mel* lines nested within genotypes (bins 29 to 36)						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	171.3819	2	453.8884	0.377586	0.601489
2	2	453.8884	552	67.23039	6.751239	0.001268
3	7	926.8586	14	18.12791	51.12881	6.67E-09
12	--	--	--	--	--	--
13	7	146.6788	14	18.12791	8.091322	0.000505
23	14	18.12791	552	67.23039	0.269639	0.99654
123	--	--	--	--	--	--

4.10f Pairwise comparison of *mps4* and *13.2* transformants over bins 29-36 with lines nested under genotype.

ANOVA *mps4 v 13.2* lines nested within genotypes (bins 29 to 36)						
lines as random effects and nested within genotypes						
1-GENO, 2-LINE, 3-BIN						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	401.1438	2	4035.745	0.099398	0.782409
2	2	4035.745	1104	116.0868	34.76488	2.28E-15
3	7	1150.163	14	47.30722	24.31264	8.35E-07
12	--	--	--	--	--	--
13	7	114.3556	14	47.30722	2.417298	0.075746
23	14	47.30722	1104	116.0868	0.407516	0.973032
123	--	--	--	--	--	--

Appendices for Chapter 5

Appendix 5.3 Newman - Keuls *post hoc* comparisons between the period lengths of Δ(Thr-Gly) lines 52C and 18A at 18, 25 and 29^{0}C

Newman-Keuls test; PERIOD (LINE 18A AND 52C Δ(Thr-Gly))							
Probabilities for Post Hoc Tests							
INTERACTION: 1 x 2							
		{1}	{2}	{3}	{4}	{5}	{6}
		25.02296	25.67000	26.15700	24.58156	26.30778	27.38073
18A	18^{0}C {1}		0.0532	**0.0020**	0.1872	**0.0007**	**0.0000**
18A	25^{0}C {2}	0.0532		0.1456	**0.0033**	0.1370	**0.0000**
18A	29^{0}C {3}	**0.0020**	0.1456		**0.0000**	0.6523	**0.0008**
52C	18^{0}C {4}	0.1872	**0.0033**	**0.0000**		**0.0000**	**0.0000**
52C	25^{0}C {5}	**0.0007**	0.1370	0.6523	**0.0000**		**0.0013**
52C	29^{0}C {6}	**0.0000**	**0.0000**	**0.0008**	**0.0000**	**0.0013**	

Appendix 5.4 Newman - Keuls *post hoc* comparisons between the period lengths of $(Thr\text{-}Gly)_1$ lines 44A , 57A and 88A at 18, 25 and 29^{0}C

Newman-Keuls test; PERIOD (Line 57 88 and 44 (*Thr-Gly*)$_1$)										
Probabilities for Post Hoc Tests										
INTERACTION: 1 x 2										
		{1}	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}
		24.33526	26.83455	26.24937	24.92769	26.78268	26.58286	24.76273	26.70296	26.04972
57	18^{0}C {1}		**0.0000**	**0.0000**	**0.0148**	**0.0000**	**0.0000**	**0.0444**	**0.0000**	**0.0000**
57	25^{0}C {2}	**0.0000**		**0.0469**	**0.0000**	0.8074	0.6374	**0.0000**	0.8099	**0.0031**
57	29^{0}C {3}	**0.0000**	0.0469		**0.0000**	0.0587	0.1169	**0.0000**	**0.0833**	0.3478
88	18^{0}C {4}	**0.0148**	**0.0000**	**0.0000**		**0.0000**	**0.0000**	0.4380	**0.0000**	**0.0000**
88	25^{0}C {5}	**0.0000**	0.8074	0.0587	**0.0000**		0.6152	**0.0000**	0.7078	**0.0052**
88	29^{0}C {6}	**0.0000**	0.6374	0.1169	**0.0000**	0.6152		**0.0000**	0.5723	**0.0326**
44A	18^{0}C {7}	0.0444	**0.0000**	**0.0000**	**0.4380**	**0.0000**	**0.0000**		**0.0000**	**0.0000**
44A	25^{0}C {8}	**0.0000**	0.8099	0.0833	**0.0000**	0.7078	0.5723	**0.0000**		**0.0114**
44A	29^{0}C {9}	**0.0000**	**0.0031**	0.3478	**0.0000**	**0.0052**	**0.0326**	**0.0000**	**0.0114**	

Appendix 5.5 Newman - Keuls *post hoc* comparisons between the period lengths of $(Thr\text{-}Gly)_{17}$ lines 11A and 41A at 18, 25 and 29^{0}C

Newman-Keuls test; PERIOD (line 11A and 41A (*Thr-Gly*)$_{17}$)							
Probabilities for Post Hoc Tests							
INTERACTION: 1 x 2							
		{1}	{2}	{3}	{4}	{5}	{6}
		23.43600	23.77385	24.26583	24.30123	25.22633	25.47518
11A	18^{0}C {1}		0.1476	**0.0011**	**0.0012**	**0.0000**	**0.0000**
11A	25^{0}C {2}	0.1476		**0.0350**	0.0615	**0.0000**	**0.0000**
11A	29^{0}C {3}	**0.0011**	**0.0350**		0.8794	**0.0001**	**0.0000**
41A	18^{0}C {4}	**0.0012**	0.0615	0.8794		**0.0001**	**0.0000**
41A	25^{0}C {5}	**0.0000**	**0.0000**	**0.0001**	**0.0001**		0.2862
41A	29^{0}C {6}	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.2862	

Appendix 5.6 Newman - Keuls *post hoc* comparisons between the period lengths of all (Thr-Gly) variants (pooled to create genotype means) at 18, 25 and 29^{0}C

Newman-Keuls test; PERIOD (All GENOTYPES)												
Probabilities for Post Hoc Tests												
INTERACTION: 1 x 2												
	{1}	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}	{12}
	24.78356	26.01440	26.79800	24.71359	26.76266	26.34290	24.00287	24.92177	25.20393	24.09301	25.04318	24.85299
Δ(Thr-Gly) 18°C {1}		**0.0000**	**0.0000**	0.6713	**0.0000**	**0.0000**	**0.0000**	0.6792	0.0800	**0.0001**	0.3932	0.6737
Δ(Thr-Gly) 25°C {2}	**0.0000**		**0.0000**	**0.0000**	**0.0000**	**0.0463**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Δ(Thr-Gly) 29°C {3}	**0.0000**	**0.0000**		**0.0000**	0.8303	**0.0159**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
$(Thr\text{-}Gly)_1$ 18°C {4}	0.6713	**0.0000**	**0.0000**		**0.0000**	**0.0000**	**0.0001**	0.5868	**0.0349**	**0.0002**	0.2664	0.6747
$(Thr\text{-}Gly)_1$ 25°C {5}	**0.0000**	**0.0000**	0.8303	**0.0000**		**0.0109**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
$(Thr\text{-}Gly)_1$ 29°C {6}	**0.0000**	**0.0463**	**0.0159**	**0.0000**	**0.0109**		**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
$(Thr\text{-}Gly)_{17}$ 18°C {7}	**0.0000**	**0.0000**	**0.0000**	**0.0001**	**0.0000**	**0.0000**		**0.0000**	**0.0000**	0.5846	**0.0000**	**0.0000**
$(Thr\text{-}Gly)_{17}$ 25°C {8}	0.6792	**0.0000**	**0.0000**	0.5868	**0.0000**	**0.0000**	**0.0000**		0.2009	**0.0000**	0.4615	0.6765
$(Thr\text{-}Gly)_{17}$ 29°C {9}	0.0800	**0.0000**	**0.0000**	**0.0349**	**0.0000**	**0.0000**	**0.0000**	0.2009		**0.0000**	0.3296	0.1440
$(Thr\text{-}Gly)_{20}$ 18°C {10}	**0.0001**	**0.0000**	**0.0000**	**0.0002**	**0.0000**	**0.0000**	0.5846	**0.0000**	**0.0000**		**0.0000**	**0.0000**
$(Thr\text{-}Gly)_{20}$ 25°C {11}	0.3932	**0.0000**	**0.0000**	0.2664	**0.0000**	**0.0000**	**0.0000**	0.4615	0.3296	**0.0000**		0.4811
$(Thr\text{-}Gly)_{20}$ 29°C {12}	0.6737	**0.0000**	**0.0000**	0.6747	**0.0000**	**0.0000**	**0.0000**	0.6765	0.1440	**0.0000**	0.4811	

Appendix 5.7 Newman - Keuls *post hoc* comparisons between the period lengths of $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ (lines pooled into genotypes) at 18, 25 and 29^{0}C

Newman-Keuls test; PERIOD ($(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$)							
Probabilities for Post Hoc Tests							
INTERACTION: 1 x 2							
		{1}	{2}	{3}	{4}	{5}	{6}
		24.00287	24.92177	25.20393	24.09301	25.04318	24.85299
$(Thr\text{-}Gly)_{17}$	18°C {1}		**0.0000**	**0.0000**	0.5448	**0.0000**	**0.0000**
$(Thr\text{-}Gly)_{17}$	25°C {2}	**0.0000**		0.1399	**0.0000**	0.4147	0.6440
$(Thr\text{-}Gly)_{17}$	29°C {3}	**0.0000**	0.1399		**0.0000**	0.2802	0.0854
$(Thr\text{-}Gly)_{20}$	18°C {4}	0.5448	**0.0000**	**0.0000**		**0.0000**	**0.0000**
$(Thr\text{-}Gly)_{20}$	25°C {5}	**0.0000**	0.4147	0.2802	**0.0000**		0.4077
$(Thr\text{-}Gly)_{20}$	29°C {6}	**0.0000**	0.6440	0.0854	**0.0000**	0.4077	

Reprint Series
9 December 1997, Volume 278, pp. 2117–2120

SCIENCE

Natural Variation in a *Drosophila* Clock Gene and Temperature Compensation

Lesley A. Sawyer,* J. Michael Hennessy,*
Alexandre A. Peixoto,† Ezio Rosato, Helen Parkinson,
Rodolfo Costa, and Charalambos P. Kyriacou‡

Natural Variation in a *Drosophila* Clock Gene and Temperature Compensation

Lesley A. Sawyer,* J. Michael Hennessy,* Alexandre A. Peixoto,† Ezio Rosato, Helen Parkinson, Rodolfo Costa, Charalambos P. Kyriacou‡

The threonine-glycine (Thr-Gly) encoding repeat within the clock gene *period* of *Drosophila melanogaster* is polymorphic in length. The two major variants *(Thr-Gly)17* and *(Thr-Gly)20* are distributed as a highly significant latitudinal cline in Europe and North Africa. Thr-Gly length variation from both wild-caught and transgenic individuals is related to the flies' ability to maintain a circadian period at different temperatures. This phenomenon provides a selective explanation for the geographical distribution of Thr-Gly lengths and gives a rare glimpse of the interplay between molecular polymorphism, behavior, population biology, and natural selection.

The clock gene *period* (*per*) in *Drosophila melanogaster* is an essential component of circadian rhythmicity, and its product is involved in a negative autoregulatory feedback loop with the Timeless protein [reviewed in (*1*)]. The *per* gene has a repetitive region, which encodes alternating pairs of predominantly threonine-glycine, but also serine-glycine dipeptide pairs (*2*). This repetitive region is conserved in the mammalian *per* homolog, suggesting that it may play an important functional role in circadian phenotypes (*3*). However, the only role assigned for the Thr-Gly region is to convey the species-specific characteristics of the ultradian male courtship song cycle (*4*).

Within natural populations of *D. melanogaster* and *D. simulans*, the Thr-Gly repeat is polymorphic in length (*5*). In *D. melanogaster*, Thr-Gly alleles that encode 14, 17, 20, and 23 dipeptide pairs [termed *(Thr-Gly)14*, *(Thr-Gly)17*, and so on] make up about 99% of European variants (*6*). The *(Thr-Gly)17* and *(Thr-Gly)20* alleles are distributed as a highly significant latitudinal cline, with high occurrences of the former observed in the southern Mediterranean and the latter predominating in northern Europe (*6*). In both *D. melanogaster* and *D. simulans*, analyses of intraspecific Thr-Gly haplotypes aimed at testing neutral models suggest that the repetitive regions are under selection (*7*, *8*). Furthermore, several studies revealed that Thr-Gly repeat length coevolves with the immediate flanking amino acids (*9*, *10*). If selection is shaping variation in the repetitive region, then the Thr-Gly cline in Europe implicates temperature as a possible selective agent.

Therefore, we studied the temperature responses of natural *Thr-Gly* length variants, which have the sequences shown in Fig. 1 (*11*). For simplicity, the *(Thr-Gly)17c* allele, which has the downstream (Thr-Gly)2 deletion, is referred to as *(Thr-Gly)15*. The Ser-to-Phe replacement is the only amino acid polymorphism that has been encountered in the immediate flanking regions surrounding the repeat in European (*11*) and other populations (*12*).

Free-running circadian locomotor activity rhythms of males from 37 different attached-X lines, whose *per*-carrying X chromosomes originated from eight European and North African localities, were examined at 18° and 29°C (Table 1) (*13*). A further attached-X line whose original male carried the *(Thr-Gly)23* allele from the American Canton-S laboratory strain was also added. This Thr-Gly haplotype is also found in European populations (*11*). The results based on spectral analysis (*14*) are presented in Table 1 and Fig. 2, A and B. Similar results were obtained with autocorrelation (Table 1) (*15*), but are not presented in detail. Two-way analysis of variance (ANOVA), performed with the 38 lines and temperature as the variables, gave significant Line and Temperature × Line interactions ($P < 0.01$), and further ANOVA of the data pooled into genotypes also gave significant Genotype and Temperature × Genotype interaction (both $P < 0.001$). Planned comparisons revealed that six of the 38 lines showed significant period differences between the two temperatures, whereas nine were significantly different when periods were determined with autocorrelation (Ta-

L. A. Sawyer, J. M. Hennessy, A. A. Peixoto, E. Rosato, H. Parkinson, C. P. Kyriacou, Department of Genetics, University of Leicester, LE1 7RH, UK.
R. Costa, Dipartimento di Biologia, Universitá di Padova, Via Trieste 75, 35121 Padova, Italy.

*These authors contributed equally to this work.
†Present address: Departamento de Bioquimica e Biologia Molecular Fundacao Oswaldo Cruz, Avenue Brasil, 4365-Manquinhos, CEP 21045-900, Rio de Janeiro, Brazil.
‡To whom correspondence should be addressed. E-mail: cpk@leicester.ac.uk

1). These significant differences tended to fall in the comparisons involving the shorter (Thr-Gly) variants (Table 1). However, these results should be treated with some caution, because the critical P value of 0.05 was not adjusted for the 38 planned comparisons.

(Thr-Gly)20 variants showed the most efficient circadian temperature compensation, with no overall significant difference between the periods at the two temperatures (Fig. 2A; $P = 0.47$). However, the mean periods at both temperatures were somewhat shorter than 24 hours. The *(Thr-Gly)17* lines, on the other hand, produced a period closer to 24 hours at the warmer temperature, but the period shortened significantly at the colder temperature ($P = 0.029$). Nevertheless, the *(Thr-Gly)17* and *(Thr-Gly)20* variants, which make up 70% of natural alleles (6), are better compensated than the others. In addition, the periods of the *(Thr-Gly)14* and *(Thr-Gly)17* lines became longer as temperature increased (the direction of the arrow in Fig. 2A shows the change in period at warmer temperatures), whereas the converse was seen with *(Thr-Gly)23* (Fig. 2A). The *(Thr-Gly)15-21-24* variants, which are structurally "out of phase" with the (Thr-Gly)3 interval of the *(Thr-Gly)14-17-20-23* allelic series, appear less predictable in this respect as can be seen from Fig. 2B, which illustrates temperature differences in period as computed from the pooled means for each genotype. The "in-phase" *(Thr-Gly)14-17-20-23* series of variants, which differ by units of (Thr-Gly)3, fall close to the regression line ($r = -0.98$, $P < 0.02$; for these four variants only), whereas the out-of-phase variants fall further from the illustrated regression line (overall $r = -0.57$, not significant; for all seven variants). However, because of the unavoidably small sample sizes for the rare variants, we also weighted the correlation, using the period differences shown for each strain (Table 1). A significant correlation was obtained ($r = -0.328$, $P = 0.044$, $n = 38$). Removing the five lines with out-of-phase 15-21-24 lengths again strengthened the correlation ($r = -0.365$, $P = 0.037$, $n = 33$), but not significantly. Thus, it appears that an approximately linear relationship exists between Thr-Gly length and temperature compensation; this is particularly evident with the *(Thr-Gly)14-17-20-23* series, which make up the vast majority of natural variants (6). Structural studies of Thr-Gly peptides show that a (Thr-Gly)3 peptide represents a conformational monomer, generating a β turn (16). Perhaps then, the relationship between Thr-Gly length and temperature compensation has a structural component related to the dynamic properties of the (Thr-Gly)3 motif (16).

The phenotypic effects associated with these very small changes in natural Thr-Gly length are marginal. To test their validity, we generated *per* transgenes in which internal deletions of the repetitive tract from a cloned *(Thr-Gly)20 per* gene were made. *(Thr-Gly)17* and *(Thr-Gly)1* transgenes were constructed, and a *Δ(Thr-Gly)* transgene was included (Fig. 1) (17, 18). The free-running circadian locomotor rhythms were examined in two to four independent-

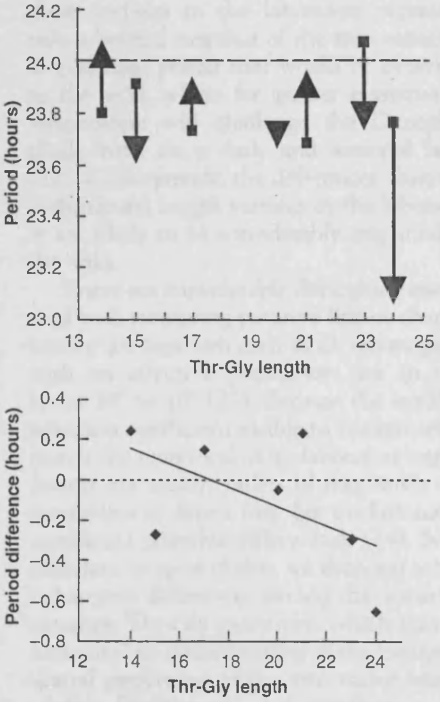


Fig. 2. (**Top**) Mean-free running periods at 18° (box) and 29°C (arrowhead) of males carrying different Thr-Gly length alleles. The means represent the pooled period averaged across the number of individuals within each Thr-Gly length (Table 1). The arrow reflects the change in direction of the period at 29°C. The *(Thr-Gly)20* variants show almost identical periods at the two temperatures. (**Bottom**) The same data as (A) is used, but is plotted as the mean period obtained at 18°C subtracted from that obtained at 29°C. The regression line is plotted.

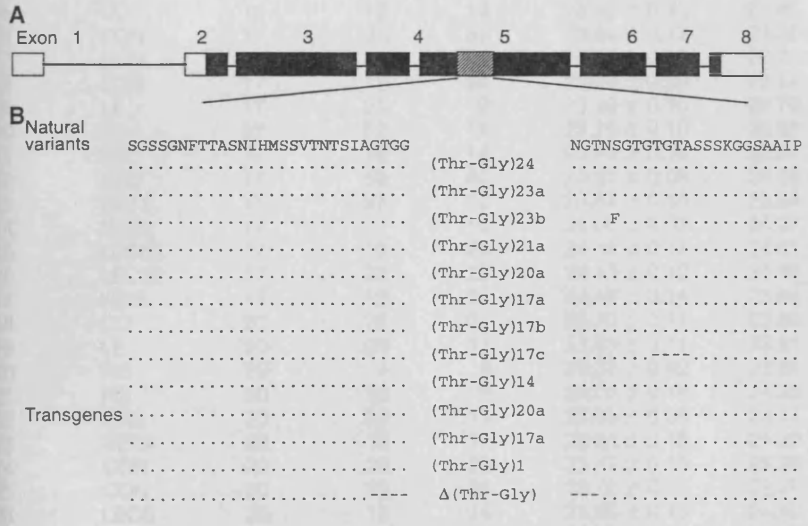


Fig. 1. (**A**) Intron and exon structure of *per* showing the position of the Thr-Gly repeat. Filled boxes represent translated exons, and the hatched box represents the Thr-Gly region. (**B**) Amino acid sequences of the Thr-Gly region from the natural strains and from the Thr-Gly transgenes (27). The uninterrupted Thr-Gly repeat length is given, and a, b, or c identify different isolength DNA sequences (5, 11). Dots denote identical amino acids; dashes show deletions. All European *(Thr-Gly)23b* variants show a fixed substitution (Ser to Phe) in the 3′ flanking region. The *(Thr-Gly)17c* variant has a downstream deletion of two Thr-Gly pairs and is therefore referred to as *(Thr-Gly)15* in the text. Fly populations were sampled from Europe and North Africa (5, 6), and isofemale lines were established immediately. One male from each isofemale line was crossed to attached-*X* females, generating a stable line in which the males carry the original paternal *X* chromosome. The length of the Thr-Gly encoding minisatellite within *per* was examined in the males of each attached-*X* line by PCR, by heteroduplex formation, and by subsequent DNA sequencing.

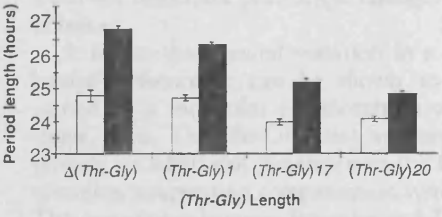


Fig. 3. Mean (and SEM) for free-running periods of per^{01} transformants at 18° (open bars) and 29°C (filled bars), which carry a single copy of a *Thr-Gly* transgene. The spectrally derived data from the different independently transformed lines within each *Thr-Gly* genotype have been pooled (19).

ly transformed lines for each Thr-Gly transgene, on a per^{01} background (Fig. 3). Two-way ANOVA on spectrally derived data gave significant Line, Temperature, and Line×Temperature interaction effects (all $P < 0.001$). A posteriori tests revealed that all lines for the Δ*(Thr-Gly)*, *(Thr-Gly)1*, and *(Thr-Gly)17* transgenes gave significantly longer periods at 29°C ($P \ll 0.01$ for each case). In contrast, only two of the four *(Thr-Gly)20* lines gave significant lengthening of the period at 29°C, and the temperature differences were smaller than for the other genotypes (*19*). These results convincingly support those based on the natural variants, even to the extent that the *(Thr-Gly)20* transformants show overall better temperature compensation than do the *(Thr-Gly)17* variants. Furthermore, the design of the transgenes (*17*) means that the associated temperature compensation differences cannot be due to any linkage disequilibrium with the different repeat arrays (*7*), but are caused by changes in the number of Thr-Gly pairs, with similar implications for the natural Thr-Gly variants.

A free-running circadian period of 24 hours may be optimal in *Drosophila*, reducing the physiological "cost" of a daily resetting of the circadian clock (*20*). Thus, at warmer temperatures, the *(Thr-Gly)17* variant has a period very close to 24 hours (Fig. 2A) and may enjoy an advantage, whereas at colder temperatures, its period shortens significantly. The more robust temperature compensation of the *(Thr-Gly)20* allele might therefore be at a premium in colder, more thermally variable environments, such as in northern Europe (*21*). Consequently, a balancing selection scenario can be envisaged, whereby the *(Thr-Gly)17* circadian periods are particularly adapted to warmer environments and the *(Thr-Gly)20* to colder climates. In fact, in Europe (*6*) and Australia (*22*), the *(Thr-Gly)17* allele generally predominates over the *(Thr-Gly)20* and only starts to fall in frequency at the more extreme, cooler regions within these continents. The behavioral differences we see in these variants in the laboratory represent only a limited snapshot of the true variation in circadian period that would be observed in the wild, where far greater extremes of temperature will challenge the *Drosophila* clock, both on a daily and seasonal basis (*21*). Consequently, the differences observed with natural length variants in the laboratory are likely to be considerably amplified in the wild.

Table 1. Spectrally-derived (*14*) free-running periods in constant darkness (DD) of males from attached-*X* lines carrying various Thr-Gly length alleles at different temperatures. The origins of the different lines are Cognac (CO), France; Conselve (CON), Pietrastornina (PI), and Lecce (LEC), Italy; Leiden (LE), Netherlands; Casablanca (CAS), Morocco; Rethimnon (RET), Greece; Canton-S (Cant.s), United States; and North Wooton (NW), United Kingdom. Significant differences for a priori comparisons based on ANOVA are highlighted (*$P < 0.05$,**$P < 0.01$; see text). These spectrally derived results are more conservative than those obtained from autocorrelation (*28*), in which significant period differences at the two temperatures were obtained in nine of the comparisons, six of which involved *(Thr-Gly)14* and *(Thr-Gly)17* variants.

Line number	Origin	(Thr-Gly) length	*n*		Period (in hours) ± SEM	
			18°C	29°C	18°C	29°C
1	CON	14	33	29	23.97 ± 0.15	24.12 ± 0.12
2	CON	14	28	28	23.43 ± 0.16	24.00 ± 0.14**
3	CO	14	29	28	23.92 ± 0.13	24.16 ± 0.13
4	LEC	14	13	16	23.85 ± 0.17	23.69 ± 0.17
5	CO	15	12	19	23.92 ± 0.20	23.65 ± 0.16
6	CON	17	31	32	23.94 ± 0.13	24.55 ± 0.15**
7	CON	17	15	29	23.31 ± 0.19	23.71 ± 0.15
8	CON	17	19	24	23.66 ± 0.29	23.91 ± 0.20
9	LE	17	21	9	23.49 ± 0.15	23.78 ± 0.36
10	PI1	17	12	14	23.75 ± 0.10	23.91 ± 0.25
11	PI9	17	15	14	23.99 ± 0.16	23.81 ± 0.22
12	CAS	17	59	40	23.21 ± 0.08	23.09 ± 0.17
13	RET9	17	27	12	23.84 ± 0.19	23.94 ± 0.15
14	RET2	17	7	10	24.20 ± 0.32	24.37 ± 0.20
15	LEC60	17	19	20	24.18 ± 0.17	24.31 ± 0.17
16	LEC12	17	28	22	24.13 ± 0.12	24.14 ± 0.17
17	NW6	17	13	20	24.48 ± 0.14	23.84 ± 0.14*
18	CO	20	25	29	23.70 ± 0.11	23.39 ± 0.18
19	LE	20	20	12	23.81 ± 0.11	23.91 ± 0.30
20	PI5	20	4	8	23.37 ± 0.42	23.61 ± 0.22
21	PI8	20	10	8	24.01 ± 0.18	24.28 ± 0.16
22	CAS	20	38	13	23.69 ± 0.08	23.12 ± 0.15*
23	RET4	20	16	11	23.81 ± 0.18	24.20 ± 0.29
24	CON	20	25	30	23.47 ± 0.13	23.28 ± 0.14
25	CON	20	22	26	23.72 ± 0.17	23.79 ± 0.14
26	LEC6	20	15	14	23.90 ± 0.13	24.40 ± 0.19
27	LEC3	20	12	14	23.88 ± 0.14	24.27 ± 0.12
28	NW1	20	10	18	23.73 ± 0.23	23.20 ± 0.17
29	CO	21	16	17	23.47 ± 0.21	23.85 ± 0.20
30	CON	21	26	25	23.81 ± 0.16	23.93 ± 0.13
31	CON	21	27	22	23.71 ± 0.10	23.96 ± 0.16
32	CO	23	19	11	24.38 ± 0.19	24.78 ± 0.15
33	CON	23	30	24	23.99 ± 0.15	23.75 ± 0.14
34	CON	23	21	23	24.03 ± 0.14	24.03 ± 0.14
35	Cant.S	23	16	19	24.00 ± 0.19	23.19 ± 0.21**
36	LEC79	23	12	13	23.95 ± 0.17	23.49 ± 0.19
37	LEC30	23	8	10	24.18 ± 0.17	23.70 ± 0.34
38	NW5	24	15	16	23.77 ± 0.15	23.11 ± 0.14*

There are considerable difficulties associated with measuring putative fitness characters for an organism such as *D. melanogaster* with an effective population size (n_e) of about 10^5 to 10^6 (*23*). Because the smallest selection coefficient visible to natural selection is the reciprocal of n_e, laboratory experiments are usually orders of magnitude too insensitive to detect tiny, but evolutionarily significant adaptive differentials (*24*). Nevertheless, in spite of this, we detected subtle behavioral differences among the naturally occuring Thr-Gly genotypes, which may illuminate our understanding of the European spatial patterning of the two major length alleles. Furthermore, our conclusions are buttressed by studies of linkage disequilibria involving the Thr-Gly repeat, which have revealed patterns of associations that are consistent with the major *Thr-Gly* alleles as being under selection (*7*). Finally, the differences in temperature compensation associated with the different Thr-Gly lengths are consistent with the coevolutionary dynamics that have been shown to act in this region (*9*, *10*). Within *D. melanogaster*, the differences in repeat length are not compensated by changes in flanking haplotypes, so small but detectable phenotypic changes are observed.

It is rare that natural variation in a behavioral phenotype can be shown to be caused by a molecular polymorphism at a single locus. The Thr-Gly array appears to provide an additional dimension to the fly's circadian temperature compensation system. This association between behavior and Thr-Gly polymorphism may "fine-tune" the circadian clock to different thermal environments and leads us to propose a simple selective explanation for the clinal pattern of Thr-Gly length variation seen in Europe (*6*).

REFERENCES AND NOTES

E. Rosato, A. Piccin, C. P. Kyriacou, *Bioessays*, in press.

C. P. Kyriacou, L. A. Sawyer, A. Piccin, M. E. Couchman, D. Chalmers, *Semin. Cell Dev. Biol.* **7**, 803 (1996).

Z. H. Sun *et al.*, *Cell* **90**, 1003 (1997); H. Tei *et al.*, *Nature* **389**, 512 (1997).

D. A. Wheeler *et al.*, *Science* **251**, 1082 (1991).

R. Costa, A. A. Peixoto, J. R. Thackeray, R. Dalgleish, C. P. Kyriacou, *J. Mol. Evol.* **32**, 238 (1991).

R. Costa, A. A. Peixoto, G. Barbujani, C. P. Kyriacou, *Proc. R. Soc. London Ser. B.* **258**, 43 (1992).

E. Rosato, A. A. Peixoto, R. Costa, C. P. Kyriacou, *Genet. Res.* **69**, 89 (1997).

E. Rosato, A. A. Peixoto, G. Barbujani, R. Costa, C. P. Kyriacou, *Genetics* **138**, 693 (1994).

A. A. Peixoto, S. Campesan, R. Costa, C. P. Kyriacou, *Mol. Biol. Evol.* **10**, 127 (1993); J. Nielsen *et al.*, *ibid.* **11**, 839 (1994).

A. A. Peixoto *et al.*, in preparation.

E. Rosato, A. Gallippi, A. A. Peixoto, C. P. Kyriacou, R. J. Costa, *Mol. Evol.* **42**, 392 (1996).

R. M. Kliman and J. Hey, *Genetics* **133**, 375 (1993).

Flies were reared in 12-hour cycles of light and darkness (LD 12:12) at 25°C with lights on at 9 a.m. and off at 9 p.m. Males 2 to 7 days old were placed in the locomotor activity tubes at either 18°C or 29°C and exposed to two additional LD cycles, after which data collection in constant darkness (DD) was initiated. Locomotor activity data was collected in 30-min "time windows" for 7 days. Autocorrelation and spectral analysis were carried out on each male's activity record (*14, 15*).

R. J. Konopka, C. P. Kyriacou, J. C. Hall, *J. Neurogenet.* **11**, 117 (1997).

H. B. Dowse and J. M. Ringo, *J. Theor. Biol.* **139**, 487 (1989).

M. A. Castiglione-Morelli *et al.*, *Proc. R. Soc. London Ser. B.* **260**, 155 (1995).

Two *(Thr-Gly)20* and two Δ*(Thr-Gly)* transformant lines in which the *per* transgenes were inserted within the cp20.1 vector marked with the *rosy*$^+$ gene were used (*18*). In addition, the *(Thr-Gly)20 per* transgene was also ligated into the pW8 vector marked with *white*$^+$ to further produce two transformant lines. The (*Thr-Gly*)1 construct was made by amplifying a 716–base pair (bp) fragment from the (*Thr-Gly*)20 clone and incorporating a deletion encoding 19 Thr-Gly pairs. This was done with 5′ primer (A), 5′-AACTATAACGAGAACCTGCT-3′ (4874 to 4893); with 3′ primer (B), 5′-ATTGCCGGTACCAC-CAGTGCCGGCAATGCT-3′ (5094 to 5113); with 5′ primer (C), 5′-GCACTGGTGGTACCGGCAATG-GAACAAATTCCGG-3′ (5231 to 5250); and with 3′ primer (D), 5′-GCTACGCCTGTTCCGGATCC-3′ (5627 to 5646), using hybrid polymerase chain reaction (PCR) methods (*25*). Numbers in brackets denote the nucleotide positions corresponding to the *per* sequence described (*26*). Primer B and C create a Kpn I site (GGTACC) using the two underlined mismatch bases. Primer A ends 37 bp upstream of an Sst I site, and primer D incorporates a Bam HI site (underlined); these sites were used to substitute a 607-bp Thr-Gly deleted fragment into a 7-kb Xba fragment, which encodes the 3′ half of the *per* gene. This Xba fragment was then ligated to a 5′ Bam HI–Xba fragment in several cloning steps, thereby reconstituting the 13.2-kb *per* transcription unit within the pW8 transformation vector. The *(Thr-Gly)17* construct was generated by amplifying a 364-bp Thr-Gly fragment using a natural *(Thr-Gly)17a* variant as template and 5′ primer (A), in conjunction with 3′ primer 5′-CATTGCCGGTACCAGTGCCT-3′ (5199 to 5215 and 5233 to 5236), which carried two mismatch bases (underlined) forming a Kpn I site. The amplified product was digested with Sst I (cuts at position 4930) and Kpn I, and was used to replace the Sst I–Kpn I fragment in the (*Thr-Gly*)1 construct.

Q. Yu *et al.*, *Nature* **326**, 765 (1987).

The mean free-running spectrally derived periods of the two Δ(*Thr-Gly*) transformant lines were 25.02 and 24.58 hours at 18°C, and 26.16 and 27.38 hours at 29°C, respectively; those for the three (*Thr-Gly)1* lines were 24.76, 24.34, and 24.93 hours at 18°C, and 26.05, 26.25, and 26.58 hours at 29°C, respectively; for the two (*Thr-Gly)17* lines, periods were 23.4 and 24.3 hours at 18°C, and 24.27 and 25.48 hours at 29°C, respectively; periods for the four (*Thr-Gly)20* lines were 24.61, 24.07, 24.21, and 23.51 hours at 18°C, and 24.80, 24.18, 24.92, and 25.43 hours at 29°C, respectively. Periods derived from autocorrelation were almost identical.

C. S. Pittendrigh, *Proc. Natl. Acad. Sci. U.S.A.* **40**, 1018 (1954); ——— and D. H. Minis, *ibid.* **69**, 1537 (1972); C. S. Pittendrigh, *Annu. Rev. Physiol.* **55**, 17 (1993).

M. Hulme *et al.*, *Int. J. Climatol.* **15**, 1333 (1995).

L. A. Sawyer and C. P. Kyriacou, unpublished results.

C. F. Aquadro, *Trends Genet.* **8**, 355 (1992).

M. Kimura, *Genetics* **47**, 713 (1962).

J. Yon and M. Fried, *Nucleic Acids Res.* **17**, 4895 (1989).

Y. Citri *et al.*, *Nature* **326**, 42 (1987).

Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

L. A. Sawyer *et al.*, unpublished results.

We thank the U.K. Biotechnology and Biological Sciences Research Council (BBSRC), the U.K. National Environment Research Council (NERC), and the Human Frontier Science Program for grants to C.P.K. and the European Commission for a grant to C.P.K and R.C. We acknowledge a NERC studentship to L.S., a BBSRC studentship to H.P., a Brazilian CNPq scholarship to A.A.P., and a Ministero Universitá Ricerca Scientifica Tecnologica–British Council award for Anglo-Italian cooperation to C.P.K. and R.C.

22 July 1997; accepted 5 November 1997

APPENDIX 6.5a ANOVA $(Thr\text{-}Gly)_{17}$ [two lines] $(Thr\text{-}Gly)_{20}$ [two lines] at 18 and 29^{0}C (lines nested within genotypes)

$(Thr\text{-}Gly)_{17}$ v $(Thr\text{-}Gly)_{20}$ 2 lines each						
1-TEMP, 2-GENO, 3-LINE, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	18564.45	2	8079.25	2.297794	0.268806
2	1	59035.54	2	8770.734	6.730969	0.121974
3	2	8770.734	3324	585.045	14.99156	3.3E-07
4	11	43197.43	22	3676.419	11.74986	8.06E-07
12	1	388.4675	2	8079.25	0.048082	0.846779
13	2	8079.25	3324	585.045	13.80962	1.06E-06
23	--	--	--	--	--	--
14	11	22745.75	22	1350.036	16.84826	3.01E-08
24	11	5862.511	22	3676.419	1.594625	0.169186
34	22	3676.419	3324	585.045	6.283994	2.13E-18
123	--	--	--	--	--	--
124	11	3629.33	22	1350.036	2.688322	0.023249
134	22	1350.036	3324	585.045	2.307576	0.000496
234	--	--	--	--	--	--
1234	--	--	--	--	--	--

APPENDIX 6.5b 1. Newman-Keuls tests on $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ pooled genotypes at 18^{0}C. 2.Newman-Keuls tests on $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ pooled genotypes at 29^{0}C.

1.

Newman-Keuls test; DATA ((Thr-Gly) 17 and 20 lines pooled into genotypes 18 °C)													
Probabilities for Post Hoc Tests													
INTERACTION: 1 x 2													
		{1}	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}	{12}
		2.298086	2.784314	5.610178	54.60901	27.75922	20.01482	31.94550	59.15943	79.73810	33.24627	1.431139	2.486228
17tg	1 {1}		1.0000	0.9918	0.0000	0.0000	0.0042	0.0000	0.0000	0.0000	0.0000	0.9809	0.9991
17tg	2 {2}	1.0000		0.8148	0.0000	0.0000	0.0019	0.0000	0.0000	0.0000	0.0000	0.9999	0.9977
17tg	3 {3}	0.9918	0.8148		0.0000	0.0001	0.0054	0.0000	0.0000	0.0000	0.0000	0.9929	0.9620
17tg	4 {4}	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.3262	0.0000	0.0000	0.0000	0.0000
17tg	5 {5}	0.0000	0.0000	0.0001	0.0000		0.4520	0.6382	0.0000	0.0000	0.6370	0.0000	0.0000
17tg	6 {6}	0.0042	0.0019	0.0054	0.0000	0.4520		0.1338	0.0000	0.0000	0.0820	0.0030	0.0030
17tg	7 {7}	0.0000	0.0000	0.0000	0.0000	0.6382	0.1338		0.0000	0.0000	0.7790	0.0000	0.0000
17tg	8 {8}	0.0000	0.0000	0.0000	0.3262	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000
17tg	9 {9}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000
17tg	10 {10}	0.0000	0.0000	0.0000	0.0000	0.6370	0.0820	0.7790	0.0000	0.0000		0.0000	0.0000
17tg	11 {11}	0.9809	0.9999	0.9929	0.0000	0.0000	0.0030	0.0000	0.0000	0.0000	0.0000		0.9994
17tg	12 {12}	0.9991	0.9977	0.9620	0.0000	0.0000	0.0030	0.0000	0.0000	0.0000	0.0000	0.9994	
20tg	1 {13}	0.9999	0.9192	0.6113	0.0000	0.0000	0.0017	0.0000	0.0000	0.0000	0.0000	0.9999	0.9984
20tg	2 {14}	0.9994	1.0000	0.9966	0.0000	0.0000	0.0051	0.0000	0.0000	0.0000	0.0000	0.8522	1.0000
20tg	3 {15}	0.9784	0.9998	0.9834	0.0000	0.0000	0.0037	0.0000	0.0000	0.0000	0.0000	0.9965	0.9892
20tg	4 {16}	0.0000	0.0000	0.0000	0.0000	0.6805	0.2962	0.6230	0.0000	0.0000	0.7201	0.0000	0.0000
20tg	5 {17}	0.0030	0.0016	0.0065	0.0000	0.4171	0.8909	0.1435	0.0000	0.0000	0.0938	0.0020	0.0023
20tg	6 {18}	0.0017	0.0010	0.0054	0.0000	0.3701	0.9432	0.1620	0.0000	0.0000	0.1156	0.0010	0.0013
20tg	7 {19}	0.0000	0.0000	0.0001	0.0000	0.8642	0.4375	0.7052	0.0000	0.0000	0.6566	0.0000	0.0000
20tg	8 {20}	0.0000	0.0000	0.0000	0.0013	0.1487	0.0025	0.3490	0.0000	0.0000	0.2696	0.0000	0.0000
20tg	9 {21}	0.0000	0.0000	0.0000	0.6796	0.0000	0.0000	0.0001	0.3436	0.0000	0.0001	0.0000	0.0000
20tg	10 {22}	0.1047	0.0394	0.0406	0.0000	0.0691	0.2889	0.0068	0.0000	0.0000	0.0029	0.0924	0.0710
20tg	11 {23}	0.9962	1.0000	0.9953	0.0000	0.0000	0.0031	0.0000	0.0000	0.0000	0.0000	0.9729	0.9998
20tg	12 {24}	0.9999	0.9612	0.9127	0.0000	0.0000	0.0023	0.0000	0.0000	0.0000	0.0000	0.9999	0.9875

APPENDIX 6.5b 2.

Newman-Keuls test; DATA ((Thr-Gly) 17 and 20 lines pooled into genotypes 29 °C)												
Probabilities for Post Hoc Tests												
INTERACTION: 1 x 2												
		{1}	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}
		7.117411	17.55553	47.42763	31.17371	9.290442	2.735623	2.500893	3.280391	23.24759	81.65278	11.06609
17tg	1 {1}		0.0903	0.0000	0.0000	0.9473	0.9845	0.9863	0.9888	0.0016	0.0000	0.8579
17tg	2 {2}	0.0903		0.0000	0.0079	0.0936	0.0156	0.0141	0.0224	0.4781	0.0000	0.1021
17tg	3 {3}	0.0000	0.0000		0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
17tg	4 {4}	0.0000	0.0079	0.0001		0.0000	0.0000	0.0000	0.0000	0.1130	0.0000	0.0000
17tg	5 {5}	0.9473	0.0936	0.0000	0.0000		0.9136	0.9129	0.9372	0.0059	0.0000	0.6547
17tg	6 {6}	0.9845	0.0156	0.0000	0.0000	0.9136		0.9529	0.8909	0.0001	0.0000	0.7027
17tg	7 {7}	0.9863	0.0141	0.0000	0.0000	0.9129	0.9529		0.9790	0.0001	0.0000	0.6955
17tg	8 {8}	0.9888	0.0224	0.0000	0.0000	0.9372	0.8909	0.9790		0.0001	0.0000	0.7584
17tg	9 {9}	0.0016	0.4781	0.0000	0.1130	0.0059	0.0001	0.0001	0.0001		0.0000	0.0183
17tg	10 {10}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000		0.0000
17tg	11 {11}	0.8579	0.1021	0.0000	0.0000	0.6547	0.7027	0.6955	0.7584	0.0183	0.0000	
17tg	12 {12}	0.9889	0.1021	0.0000	0.0000	0.9831	0.9797	0.9841	0.9826	0.0013	0.0000	0.9167
20tg	1 {13}	0.9958	0.0633	0.0000	0.0000	0.9799	0.9761	0.9869	0.9685	0.0004	0.0000	0.8906
20tg	2 {14}	0.9835	0.0688	0.0000	0.0000	0.8583	0.9894	0.9903	0.9930	0.0014	0.0000	0.7643
20tg	3 {15}	0.0104	0.6493	0.0000	0.0533	0.0250	0.0005	0.0004	0.0008	0.5836	0.0000	0.0568
20tg	4 {16}	0.0086	0.3864	0.0000	0.0771	0.0169	0.0005	0.0004	0.0008	0.8374	0.0000	0.0332
20tg	5 {17}	0.9977	0.0506	0.0000	0.0000	0.6293	0.9913	0.9916	0.9946	0.0013	0.0000	0.6212
20tg	6 {18}	0.9978	0.0523	0.0000	0.0000	0.9808	0.9011	0.9605	0.7661	0.0003	0.0000	0.8847
20tg	7 {19}	0.9839	0.0690	0.0000	0.0000	0.9713	0.9899	0.9931	0.9907	0.0006	0.0000	0.8770
20tg	8 {20}	0.9419	0.0976	0.0000	0.0000	0.9719	0.9830	0.9858	0.9867	0.0015	0.0000	0.8943
20tg	9 {21}	0.0000	0.0058	0.0001	0.8617	0.0000	0.0000	0.0000	0.0000	0.1314	0.0000	0.0000
20tg	10 {22}	0.0000	0.0352	0.0000	0.5630	0.0000	0.0000	0.0000	0.0000	0.1561	0.0000	0.0001
20tg	11 {23}	0.9906	0.0619	0.0000	0.0000	0.9731	0.9879	0.9926	0.9870	0.0004	0.0000	0.8752
20tg	12 {24}	0.9956	0.0501	0.0000	0.0000	0.9747	0.9665	0.9849	0.9422	0.0003	0.0000	0.8686

Appendix 6.6a ANOVA of the $(Thr\text{-}Gly)_1$ [lines 57a and 88a] and $\Delta(Thr\text{-}Gly)$ [lines 18a and 52c] transformants, 12, 2h window data, LD 12:12. (lines nested within genotypes).

(Thr -Gly)$_1$ and Δ(Thr-Gly) at 18 and 29°C						
1-TEMP, 2-GENO, 3-LINE, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	700.748	2	3294.109	0.212728	0.689938
2	1	9.996881	2	28488.36	0.000351	0.986755
3	2	28488.36	2640	463.861	61.41571	8.5E-27
4	11	30565.51	22	3918.447	7.800415	2.51E-05
12	1	5331.637	2	3294.109	1.618537	0.331202
13	2	3294.109	2640	463.861	7.1015	0.00084
23	--	--	--	--	--	--
14	11	24448.37	22	2222.874	10.99854	1.44E-06
24	11	3380.386	22	3918.447	0.862685	0.586075
34	22	3918.447	2640	463.861	8.447458	7.84E-27
123	--	--	--	--	--	--
124	11	1885.367	22	2222.874	0.848166	0.598266
134	22	2222.874	2640	463.861	4.792113	1.41E-12
234	--	--	--	--	--	--
1234	--	--	--	--	--	--

Appendix 6.6b ANOVA of the $(Thr\text{-}Gly)_1$ [lines 57a and 88a] and $\Delta(Thr\text{-}Gly)$ [lines 18a and 52c] transformants bins 35-46 only, LD 12:12 (lines nested within genotypes).

Summary of all Effects; delta (Thr-Gly) v (Thr-Gly)$_1$, Bins 35-46						
1-GENO, 2-LINE, 3-TEMP, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	498.222	2	5221	0.09543	0.7866
2	2	5221	2640	72.641	71.8741	0
3	1	8377.73	2	837.983	9.9975	0.08715
4	11	3396.85	22	591.791	5.73995	0.00025
12	--	--	--	--	--	--
13	1	91.5174	2	837.983	0.10921	0.77245
23	2	837.983	2640	72.641	11.5359	1E-05
14	11	904.13	22	591.791	1.52779	0.19118
24	22	591.791	2640	72.641	8.14679	1.2E-25
34	11	1826.25	22	356.497	5.12277	0.00056
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	11	213.987	22	356.497	0.60025	0.80866
234	22	356.497	2640	72.641	4.90766	5.2E-13
1234	--	--	--	--	--	--

Appendix 6.7a **1.** ANOVA of the $(Thr\text{-}Gly)_{17}$ [lines 11a and 41a] and $\Delta(Thr\text{-}Gly)$ [lines 18a and 52c] transformants, 12, 2h window data, LD 12:12 (lines nested within genotypes). **2.** ANOVA of the $(Thr\text{-}Gly)_{20}$ [lines 2a and 17a] and $\Delta(Thr\text{-}Gly)$ [lines 18a and 52c] transformants, 12window data, LD 12:12(lines nested within genotypes).

1.

Summary of all Effects; delta (Thr-Gly) v (Thr-Gly)$_{17}$						
1-GENO, 2-LINE, 3-TEMP, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	39282.2	2	30868.4	1.27257	0.37641
2	2	30868.4	3432	669.984	46.0733	1.8E-20
3	1	3404.94	2	9290.44	0.3665	0.60646
4	11	72439.1	22	5963.77	12.1465	6E-07
12	--	--	--	--	--	--
13	1	12265.2	2	9290.44	1.3202	0.36942
23	2	9290.44	3432	669.984	13.8667	1E-06
14	11	1404.29	22	5963.77	0.23547	0.99169
24	22	5963.77	3432	669.984	8.90137	6.5E-29
34	11	40172.6	22	2710.04	14.8236	9.9E-08
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	11	5274.08	22	2710.04	1.94612	0.08854
234	22	2710.04	3432	669.984	4.04494	7.2E-10
1234	--	--	--	--	--	--

2.

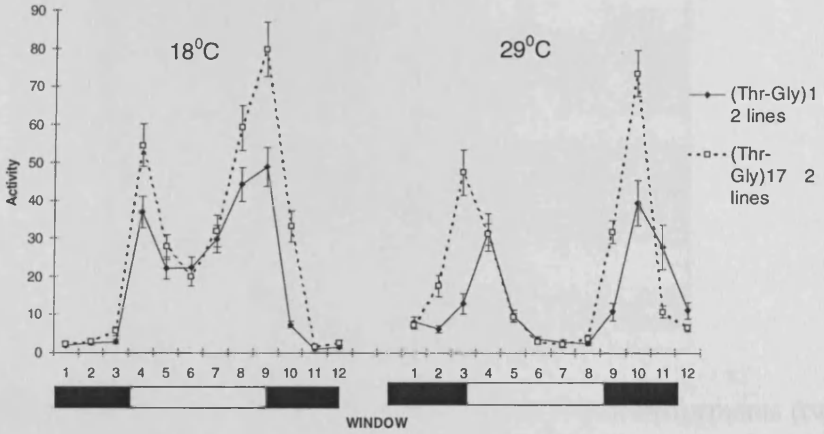
Summary of all Effects; delta (Thr-Gly) v (Thr-Gly)20						
1-GENO, 2-LINE, 3-TEMP, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	3665.04	2	34813.9	0.10528	0.77638
2	2	34813.9	3132	424.583	81.9954	0
3	1	1177.51	2	3311.21	0.35561	0.61146
4	11	34499.2	22	3868.43	8.91813	8.5E-06
12	--	--	--	--	--	--
13	1	6877.87	2	3311.21	2.07715	0.28623
23	2	3311.21	3132	424.583	7.79872	0.00042
14	11	4807.97	22	3868.43	1.24287	0.31814
24	22	3868.43	3132	424.583	9.11112	1.1E-29
34	11	15074.7	22	2038.51	7.39497	3.8E-05
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	11	2998.86	22	2038.51	1.47111	0.21194
234	22	2038.51	3132	424.583	4.80119	1.2E-12
1234	--	--	--	--	--	--

Appendix 6.7b. ANOVA of the $(Thr\text{-}Gly)_{20}$ [lines 2a and 17a] and $(Thr\text{-}Gly)_1$ [lines 57a and 88a] transformants, 12, 2h window data, LD 12:12(lines nested within genotypes).

(Thr-Gly)1 v (Thr-Gly)20						
1-TEMP, 2-GEN, 3-LIN, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	9844.473	2	2082.917	4.726291	0.161752
2	1	3281.194	2	6390.732	0.51343	0.548032
3	2	6390.732	2532	359.9099	17.75648	2.2E-08
4	11	17369.95	22	548.9074	31.6446	6.32E-11
12	1	35.76308	2	2082.917	0.01717	0.907741
13	2	2082.917	2532	359.9099	5.78733	0.003107
23	--	--	--	--	--	--
14	11	14234.98	22	844.6915	16.85229	3E-08
24	11	1461.916	22	548.9074	2.663321	0.024289
34	22	548.9074	2532	359.9099	1.525125	0.055641
123	--	--	--	--	--	--
124	11	1945.523	22	844.6915	2.303236	0.046113
134	22	844.6915	2532	359.9099	2.346953	0.000388
234	--	--	--	--	--	--
1234	--	--	--	--	--	--

Appendix 6.7c **1.** $(Thr\text{-}Gly)_{20}$ [lines 2a and 17a] and $(Thr\text{-}Gly)_1$ [lines 57a and 88a] transformants, 12, 2h window data, LD 12:12 (lines pooled within genotypes). **2.** ANOVA of the $(Thr\text{-}Gly)_{20}$ [lines 2a and 17a] and $(Thr\text{-}Gly)_1$ [lines 57a and 88a] transformants, 12, 2h window data, LD 12:12(lines nested within genotypes).

1.



2.

(Thr-Gly)1 v (Thr-Gly)17						
1-TEMP, 2-GEN, 3-LIN, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	15417.02	2	8062.152	1.912271	0.300866
2	1	29567.39	2	2445.222	12.0919	0.073677
3	2	2445.222	2832	649.5367	3.764564	0.023294
4	11	36320.28	22	3726.437	9.746653	4.04E-06
12	1	636.3056	2	8062.152	0.078925	0.805156
13	2	8062.152	2832	649.5367	12.41216	4.29E-06
23	--	--	--	--	--	--
14	11	31661.39	22	1534.405	20.63431	4.36E-09
24	11	6633.026	22	3726.437	1.779992	0.120278
34	22	3726.437	2832	649.5367	5.737069	3.37E-16
123	--	--	--	--	--	--
124	11	2198.127	22	1534.405	1.43256	0.227238
134	22	1534.405	2832	649.5367	2.362307	0.000346
234	--	--	--	--	--	--
1234	--	--	--	--	--	--

Appendices for Chapter 7

Appendix 7.1 ANOVA on natural variants $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{23}$ (two lines each) at three conditions. (lines nested within genotypes)

Lines as random effects, nested within genotype						
1-GENOTYPE, 2-LINE, 3-CONDITION, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	2	131790.1	3	71963.16	1.831355	0.302138
2	3	71963.16	4848	1717.189	41.90752	9.94E-27
3	2	86743.09	6	43799.86	1.980443	0.218555
4	11	387910.5	33	4428.636	87.59141	3.76E-21
12	--	--	--	--	--	--
13	4	14071.61	6	43799.86	0.321271	0.854307
23	6	43799.86	4848	1717.189	25.50671	5.41E-30
14	22	4374.982	33	4428.636	0.987885	0.501965
24	33	4428.636	4848	1717.189	2.579003	2.03E-06
34	22	27456	66	2120.064	12.95055	2.89E-16
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	44	2266.175	66	2120.064	1.068918	0.397456
234	66	2120.064	4848	1717.189	1.234613	0.096665
1234	--	--	--	--	--	--

Appendix 7.2a ANOVA on $(Thr\text{-}Gly)_{17}$ and $(Thr\text{-}Gly)_{20}$ transformants (two lines each) at three conditions (lines nested within genotype)

$(Thr\text{-}Gly)_{17}$ v $(Thr\text{-}Gly)_{20}$ lines nested within genotype						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	36279.31	2	8053.035	4.505048	0.167806
2	2	8053.035	3696	975.6281	8.254206	0.000265
3	2	72382.86	4	8285.098	8.736512	0.0347
4	11	203497.9	22	4301.556	47.30799	1.02E-12
12	--	--	--	--	--	--
13	2	3421.137	4	8285.098	0.412927	0.687024
23	4	8285.098	3696	975.6281	8.492065	8.1E-07
14	11	1232.5	22	4301.556	0.286524	0.981995
24	22	4301.556	3696	975.6281	4.409012	3.19E-11
34	22	13434.52	44	1065.185	12.61238	1.58E-12
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	22	1601.211	44	1065.185	1.503224	0.123445
234	44	1065.185	3696	975.6281	1.091794	0.313927
1234	--	--	--	--	--	--

7.2b ANOVA on $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_{20}$ and $\Delta(Thr\text{-}Gly)$ transformants (two lines each) in LD and Heat pulse ZT5-7 conditions only (lines nested within genotype)

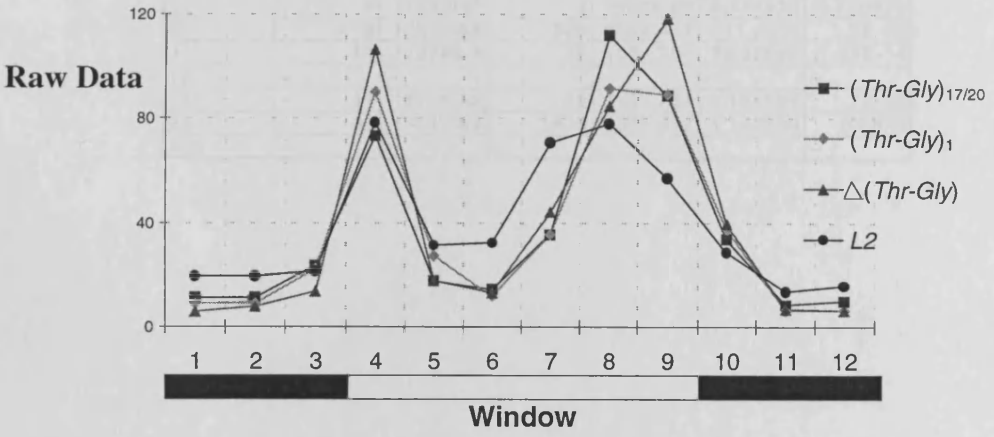
(Thr-Gly)17,20 transformants V Δ(Thr-Gly)						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	7858.975	2	15095.38	0.520621	0.545528
2	2	15095.38	4512	1029.182	14.66735	4.47E-07
3	1	396780.1	2	13203.44	30.05127	0.031703
4	11	286400.8	22	2552.477	112.205	1.1E-16
12	--	--	--	--	--	--
13	1	24653.13	2	13203.44	1.867175	0.305143
23	2	13203.44	4512	1029.182	12.82906	2.78E-06
14	11	5596.044	22	2552.477	2.192397	0.056382
24	22	2552.477	4512	1029.182	2.480102	0.000147
34	11	38357.61	22	3183.775	12.04784	6.46E-07
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	11	8224.741	22	3183.775	2.58333	0.027957
234	22	3183.775	4512	1029.182	3.093499	1.52E-06
1234	--	--	--	--	--	--

7.2c ANOVA on $(Thr\text{-}Gly)_{17}$, $(Thr\text{-}Gly)_{20}$ and *L2* deletion transformants (two lines each) Heat pulse ZT5-7 and recovery conditions only (lines nested within genotype)

(Thr-Gly)17,20 transformants V 3.1						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	675.7242	2	30550.1	0.022119	0.895414
2	2	30550.1	3600	1118.788	27.30642	1.7E-12
3	1	58272.18	2	52.99607	1099.557	0.000908
4	11	150213.5	22	1801.567	83.37937	2.63E-15
12	--	--	--	--	--	--
13	1	7317.573	2	52.99607	138.0776	0.007165
23	2	52.99607	3600	1118.788	0.047369	0.953736
14	11	18490.79	22	1801.567	10.26373	2.61E-06
24	22	1801.567	3600	1118.788	1.610285	0.035595
34	11	17096.8	22	507.5553	33.6846	3.36E-11
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	11	1728.44	22	507.5553	3.405423	0.006957
234	22	507.5553	3600	1118.788	0.453665	0.986355
1234	--	--	--	--	--	--

7.2d 1. Transformants $(Thr\text{-}Gly)_{17/20}$, $(Thr\text{-}Gly)_1$, $\Delta(Thr\text{-}Gly)$, and *L2* (two lines each) in heat pulse ZT 5-7 condition only. **2.** ANOVA of data in 7.2d1 (lines nested within genotypes)

1.



2.

lines as random factors, nested within genotypes						
1-GENO, 2-LINE, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	3	1479.155	4	10306.29	0.14352	0.92868
2	4	10306.29	3588	1385.248	7.440032	5.78E-06
3	11	308840.1	44	2789.714	110.7067	2.62E-28
12	--	--	--	--	--	--
13	33	9882.381	44	2789.714	3.542436	5.56E-05
23	44	2789.714	3588	1385.248	2.013873	9.06E-05
123	--	--	--	--	--	--

7.2e **1.** ANOVA on $(Thr\text{-}Gly)_{17/20}$ and $(Thr\text{-}Gly)_1$ **2.** ANOVA on $(Thr\text{-}Gly)_{17/20}$ and $\Delta(Thr\text{-}Gly)$ **3.** ANOVA on $(Thr\text{-}Gly)_{17/20}$ and *L2* deletion $(Thr\text{-}Gly)_1$ [All analyses are at heat pulse ZT 5-7 condition only, all genotypes have two lines and ANOVA have lines nested within genotypes].

1.

1TG V TRANS						
1-GENO, 2-LINE, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	53.54335	2	5005.191	0.010698	0.927059
2	2	5005.191	2016	1294.619	3.866151	0.021094
3	11	195082.7	22	1471.201	132.601	1.82E-17
12	--	--	--	--	--	--
13	11	3121.152	22	1471.201	2.121499	0.064168
23	22	1471.201	2016	1294.619	1.136397	0.298512
123	--	--	--	--	--	--

2.

DELTA V TRANS						
1-GENO, 2-LINE, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	2350.565	2	6876.711	0.341815	0.617951
2	2	6876.711	2196	1579.633	4.35336	0.012975
3	11	260408	22	3392.143	76.76798	6.32E-15
12	--	--	--	--	--	--
13	11	12192.99	22	3392.143	3.594481	0.005142
23	22	3392.143	2196	1579.633	2.147425	0.001488
123	--	--	--	--	--	--

3.

3.1 V TRANS						
1-GENO, 2-LINE, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	1772.989	2	15425.63	0.114938	0.766878
2	2	15425.63	1800	1352.472	11.4055	1.2E-05
3	11	113766.3	22	1551.704	73.31699	1.03E-14
12	--	--	--	--	--	--
13	11	12710.83	22	1551.704	8.191529	1.7E-05
23	22	1551.704	1800	1352.472	1.147309	0.2874
123	--	--	--	--	--	--

Appendix 7.5 1. ANOVA on *per*^{0} *v* natural variants. **2.** ANOVA on *per*^{L} *v* natural variants. **3.** ANOVA on *per*^{S} *v* natural variants. (all analyses at three conditions and all data pooled)

1.

. Per0 v Natural Variants						
1-GENO, 2-COND, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	6765.769	5844	1853.547	3.650174	0.056112
2	2	42957.36	5844	1853.547	23.17576	9.43E-11
3	11	130995.1	5844	1853.547	70.67268	0
12	2	2337.135	5844	1853.547	1.260899	0.283476
13	11	29291.54	5844	1853.547	15.80297	0
23	22	11839.81	5844	1853.547	6.387652	5.29E-19
123	22	1944.424	5844	1853.547	1.049029	0.397642

2.

PerL v Natural Variants						
1-GENO, 2-COND, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	12140.33	5712	1824.306	6.654768	0.009914
2	2	53355.13	5712	1824.306	29.24681	2.31E-13
3	11	111459.4	5712	1824.306	61.0969	0
12	2	3494.663	5712	1824.306	1.915612	0.147346
13	11	43581.85	5712	1824.306	23.88955	0
23	22	11731.61	5712	1824.306	6.430723	3.61E-19
123	22	3799.715	5712	1824.306	2.082828	0.002149

3.

PERS v Natural variants						
1-GENO, 2-COND, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	46931.62	5640	1852.729	25.33108	4.98E-07
2	2	3969.649	5640	1852.729	2.142596	0.117445
3	11	63423.39	5640	1852.729	34.23243	0
12	2	14256.53	5640	1852.729	7.694881	0.00046
13	11	32079.47	5640	1852.729	17.31471	0
23	22	5129.257	5640	1852.729	2.768488	1.77E-05
123	22	2063.296	5640	1852.729	1.113652	0.322048

Appendix 7.6 a. ANOVA on natural variants $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{23}$ at LD and heat pulse ZT 15-17 only. **b.** ANOVA on natural variants $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ at LD and heat pulse ZT 15-17 only. [All genotypes have two lines and analyses have lines nested within genotypes)

a.

Natural (Thr-Gly) $_{17, 20}$ and $_{23}$						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	2	20850.78	3	42230.29	0.49374	0.652581
2	3	42230.29	3096	1175.05	35.93916	7.96E-23
3	1	2439.616	3	17658.31	0.138157	0.734821
4	11	151464.1	33	5887.466	25.72653	4.63E-13
12	--	--	--	--	--	--
13	2	1295.062	3	17658.31	0.07334	0.9309
23	3	17658.31	3096	1175.05	15.02772	1.04E-09
14	22	3397.995	33	5887.466	0.577157	0.909909
24	33	5887.466	3096	1175.05	5.010397	6.31E-19
34	11	9209.376	33	2702.467	3.407767	0.003076
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	22	5119.453	33	2702.467	1.894363	0.047106
234	33	2702.467	3096	1175.05	2.299874	3.71E-05
1234	--	--	--	--	--	--

b.

Naturals (Thr-Gly)17 and 20						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	36139.41	2	10052.48	3.595073	0.198412
2	2	10052.48	2124	1026.234	9.795507	5.83E-05
3	1	362.5183	2	17948.89	0.020197	0.900012
4	11	94611.82	22	3703.126	25.54918	5.38E-10
12	--	--	--	--	--	--
13	1	1307.376	2	17948.89	0.072839	0.812544
23	2	17948.89	2124	1026.234	17.49005	2.92E-08
14	11	2129.037	22	3703.126	0.57493	0.828583
24	22	3703.126	2124	1026.234	3.608461	3.07E-08
34	11	3856.059	22	2614.344	1.474962	0.21046
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	11	1037.6	22	2614.344	0.396887	0.942612
234	22	2614.344	2124	1026.234	2.547512	9.83E-05
1234	--	--	--	--	--	--

Appendix 7.7 ANOVA on $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ transformants (two lines each) at all three conditions (lines nested within genotypes).

Summary of all Effects; design:						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	1	29060.79	2	8143.165	3.568734	0.199468
2	2	8143.165	3492	927.2312	8.782238	0.000157
3	2	74264.81	4	1597.015	46.50226	0.0017
4	11	93854.22	22	5318.907	17.6454	1.95E-08
12	--	--	--	--	--	--
13	2	4496.768	4	1597.015	2.815733	0.172479
23	4	1597.015	3492	927.2312	1.722348	0.142115
14	11	1909.075	22	5318.907	0.358922	0.959138
24	22	5318.907	3492	927.2312	5.736333	2.84E-16
34	22	52846.03	44	1191.798	44.34145	3.96E-23
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	22	4653.913	44	1191.798	3.904952	5.95E-05
234	44	1191.798	3492	927.2312	1.28533	0.098903
1234	--	--	--	--	--	--

Appendix 7.8 a Anova on $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ and $(Thr\text{-}Gly)_{23}$ natural variants, $(Thr\text{-}Gly)_1$ and $\Delta(Thr\text{-}Gly)$. **b** Anova on $(Thr\text{-}Gly)_1$ and $\Delta(Thr\text{-}Gly)$ [All at three conditions with all data pooled]

a.

Summary of all Effects; design: 1-GENO, 2-COND, 3-WINDOW						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	3	11054.57	7404	878.3028	12.58629	3.33E-08
2	2	174945.3	7404	878.3028	199.1857	0
3	11	232690.2	7404	878.3028	264.9316	0
12	6	5057.563	7404	878.3028	5.758336	5.44E-06
13	33	5348.709	7404	878.3028	6.089823	1.47E-25
23	22	110330.4	7404	878.3028	125.6177	0
123	66	3735.733	7404	878.3028	4.253354	1.31E-27

b.

Delta TG v 1 TG ZT 15-17 1-GENO, 2-COND, 3-WINDOW						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	3381.328	3840	800.0886	4.226192	0.039872
2	2	107572.5	3840	800.0886	134.4507	0
3	11	133318.8	3840	800.0886	166.63	0
12	2	5763.658	3840	800.0886	7.203775	0.000754
13	11	7858.126	3840	800.0886	9.82157	8.45E-18
23	22	56798.27	3840	800.0886	70.98997	0
123	22	3392.989	3840	800.0886	4.240767	1.32E-10

Appendix 7.8 c. Anova on $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ transformants. **d.** Anova on $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ transformants and $(Thr\text{-}Gly)_1$. **e.** Anova on $(Thr\text{-}Gly)_{17}$ $(Thr\text{-}Gly)_{20}$ transformants and $\Delta(Thr\text{-}Gly)$ [c,d and e heat-pulse 15-17 and recovery only, all genotypes have two lines nested within genotype for analyses]

c.

17TG V 20TG (trans), HS and Rec 1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df Effect	MS Effect	df Error	MS Error	F	p-level
1	1	7410.397	2	3135.07	2.363711	0.264015
2	2	3135.07	2172	1143.617	2.741363	0.064706
3	1	54689.51	2	1360.407	40.20084	0.023984
4	11	92237.9	22	4275.096	21.57563	2.83E-09
12	--	--	--	--	--	--
13	1	705.2513	2	1360.407	0.518412	0.546294
23	2	1360.407	2172	1143.617	1.189565	0.304552
14	11	5655.906	22	4275.096	1.322989	0.276423
24	22	4275.096	2172	1143.617	3.738223	1.07E-08
34	11	69073.7	22	1278.556	54.02478	2.57E-13
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	11	2257.126	22	1278.556	1.765371	0.123567
234	22	1278.556	2172	1143.617	1.117993	0.318075
1234	--	--	--	--	--	--

7.8d

17TG V 20 TG (trans) V 1TG, HS and Rec						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	2	6135.518	3	2108.323	2.910141	0.198362
2	3	2108.323	3420	1070.406	1.969648	0.11634
3	1	85282.74	3	920.2031	92.67817	0.002379
4	11	143182.6	33	3373.372	42.44496	2.85E-16
12	--	--	--	--	--	--
13	2	357.6457	3	920.2031	0.38866	0.707793
23	3	920.2031	3420	1070.406	0.859677	0.461283
14	22	3307.417	33	3373.372	0.980449	0.509628
24	33	3373.372	3420	1070.406	3.151488	4.13E-09
34	11	108837.3	33	992.8024	109.6263	1.07E-22
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	22	1373.779	33	992.8024	1.383738	0.194888
234	33	992.8024	3420	1070.406	0.927501	0.586684
1234	--	--	--	--	--	--

7.8e

17TG V 20TG (Trans), V Delta TG, HS and Rec						
1-GENO, 2-LINE, 3-COND, 4-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	2	5173.479	3	2343.952	2.207161	0.25738
2	3	2343.952	3324	1115.721	2.100841	0.098
3	1	79297.42	3	1194.467	66.38728	0.003866
4	11	171565.9	33	3348.219	51.24094	1.62E-17
12	--	--	--	--	--	--
13	2	434.0664	3	1194.467	0.363398	0.722235
23	3	1194.467	3324	1115.721	1.070579	0.360271
14	22	10664.17	33	3348.219	3.185029	0.001313
24	33	3348.219	3324	1115.721	3.000947	2.27E-08
34	11	102742.6	33	1261.611	81.43763	1.19E-20
123	--	--	--	--	--	--
124	--	--	--	--	--	--
134	22	2582.609	33	1261.611	2.047071	0.030443
234	33	1261.611	3324	1115.721	1.130759	0.278618
1234	--	--	--	--	--	--

APPENDIX 7.9 Anova on natural variants $(Thr\text{-}Gly)_{17/20/23}$, per^{S}, per^{L} and per^{0} ^{in} LD and heat pulse ZT 15-17. (data pooled)

Summary of all Effects; design:						
1-GENO, 2-COND, 3-WINDOW						
	df	MS	df	MS		
	Effect	Effect	Error	Error	F	p-level
1	3	7278.36	4632	1235.679	5.890172	0.000522
2	1	12702.99	4632	1235.679	10.28017	0.001354
3	11	25923.42	4632	1235.679	20.9791	0
12	3	15831.85	4632	1235.679	12.81227	2.46E-08
13	33	19294.85	4632	1235.679	15.61478	0
23	11	8566.847	4632	1235.679	6.932909	9.88E-12
123	33	4136.867	4632	1235.679	3.34785	3.85E-10

Bibliography

Albrecht, U., Sun, Z.S., Eichele, G. and Lee C.C. (1997) A differential response of two putative mammalian circadian regulators, *mper1* and *mper2*, to light. Cell 91 : 1055-1064

Allada, R., White, N.E., So, W.V., Hall, J.C. and Rosbash, M. (1998). A mutant *Drosophila* homolog of mammalian *clock* disrupts circadian rhythms and transcription of *period* and *timeless*. Cell 93 : 791-804

Antoch, M.P., Song, E.-J., Chang, A.M., Vitaterna, M.H., Zhao, Y., Wilsbacher, L.D., Sangoram, A.M., King, D.P., Pinto, L.H., and Takahashi, J.S. (1997) Functional identification of the mouse circadian *Clock* gene by transgenic BAC rescue. Cell 89 : 655-667

Aronson B.D., Johnson K.A., Loros J.J. and Dunlap J.C. (1994a) Negative feedback defining a circadian clock: autoregulation of the clock gene *frequency*. Science 263 : 1578-1584

Aronson B.D., Johnson K.A. and Dunlap J.C. (1994b) Circadian clock locus *frequency* protein encoded by a single open reading frame defines period length and temperature compensation . Proc. Natl. Acad. Sci. USA 91 : 7693-7687

Arpaia, G., Loros, J.J., Dunlap, J.C., Morelli, G. and Macino, G. (1993) The interplay of light and the circadian clock- independent dual regulation of clock controlled gene *ccg-2* (*eas*). Plant Physiology 102 : 1299-1305.

Ashburner, M. (1988) *Drosophila* a laboratory manual. Cold Spring Harbor Press, USA

Bargiello T.A., Saez L., Bayles M.K., Gasik G., Young M.W. and Spray D.C. (1987) The *Drosophila* clock gene *per* affects intercellular junctional communication. Nature 328 : 686-691.

Baylies M.K., Bargiello T.A., Jackson F.R and Joung M.W. (1987) Changes in abundance or structure of the *per* gene product can alter periodicity of the *Drosophila* clock. Nature 326 : 390-392.

Castiglione-Morelli, M.A., Guantieri, V., Villani, V., Kyriacou, C.P., Costa, R. and Tamburro, A.M. (1995) Conformational study of the Thr-Gly repeat in the *Drosophila* clock protein PERIOD. Proc. R. Soc. London 260 : 155-163

Chou, P.Y. and Fasman, U.D. (1974) Prediction of protein conformation. Biochemistry 13 : 211-215

Citri Y., Colot H.V., Jacquier A.C., Yu Q., Hall J.C., Baltimore, D. and Rosbash M. (1987) A family of unusually spliced biologically active transcripts encoded by a *Drosophila* clock gene. Nature 326 : 42-47

Colot, H.V., Hall, J.H. and Rosbash M. (1988) Interspecific comparison of the *period* gene of *Drosophila* reveals large blocks of non-conserved coding DNA. EMBO J. 7 : 3929-3937.

Costa R.A., Peixoto, A.A., Thackeray, J.R., Dalgleish, R. and Kyriacou, C.P. (1991) Length polymorphism in the threonine-glycine encoding repeat region of the *period* gene in *Drosophila*. J. Mol. Evol. 32 : 238-246

Costa R.A., Peixoto, A.A., Barbujani, G. and Kyriacou, C.P. (1992) A latitudinal cline in a *Drosophila* clock gene. Proc. R. Soc. London 250 : 43-49

Coyne, J.A. (1992) Genetics and Speciation. Nature 355 : 511-515

Coyne, J.A. and Orr, H.A. (1989) Patterns of speciation in Drosophila. Evolution 43: 262-281

Crosthwaite, S.K., Dunlap, J.C. and Loros, J.J. (1997) *Neurospora wc-1* and *wc-2*: transcription, photoresponses and the origins of circadian rhythmicity. Science 276 : 763-769

Curtin, K.D., Huang, Z.J. and Rosbash, M. (1995) Temporally regulated nuclear entry of the *Drosophila period* protein contributes to the circadian clock. Neuron 14 : 365-372

Darlington, T.K., Wager-Smith, K., Ceriani, M.F., Staknis, D., Gekakis, N., Steeves, T.D.L., Weitz, C.J., Takahashi, J.S. and Kay, S.A. (1998) Closing the circadian loop: CLOCK-induced transcription of its own inhibitors *per* and *tim*. Science 280 : 1599-1603

Dembinska, M.E., Stanewsky, R., Hall, J.C. and Rosbash, M. (1997) Circadian cycling of a PERIOD-beta-galactosidase fusion protein in Drosophila: Evidence for cyclical degradation. J. Biol. Rhythms 12 : 157-172

Diggle, P.J. (1990) Time Series: A biostatistical introduction. Oxford Science Publications. Clarendon Press, Oxford, UK

Dowse, H.B., Hall, J.C. and Ringo, J.M. (1987) Circadian and ultradian rhythms in *period* mutants of *Drosophila melanogaster*. Behav. Genetics. 17 : 19-35

Dowse, H.B. and Ringo, J.M. (1989) The search for hidden periodicities in biological time series revisited. J. Theor. Biol. 139 : 487-515

Dunlap, J.C. (1996). Genetic and molecular analysis of circadian rhythms. Annu. Rev. Genetics 30 : 579-601

Dunlap, J. (1998). An end in the beginning. Science 280 : 1548-1549

Edery, I., Rutila, J.E. and Rosbash, M.(1994a) Phase shifting of the circadian clock by induction of the *Drosophila period* protein. Science 263 :237-240

Edery, I., Zwiebel, L.J., Dembinska, M.E. and Rosbash, M. (1994b) Temporal phosphorylation of the *Drosophila* period protein. Proc. Natl. Acad. Sci. USA 91 : 2260-2264

Emery, P.T.J., Morgan, E. and Birley A.J. (1994) An investigation of natural genetic variation in the circadian system of *Drosophila melanogaster* : rhythm characteristics and methods of quantification. Chronobiol. International 11 (2) : 72-84

Engleman, W., and Mack, J.(1978). Different oscillators control the circadian rythm of eclosion and activity in *Drosophila*. J. Comp. Physiol. 127 : 229-237

Erlich, H.A., (1989) PCR Technology: Principles and applications for DNA amplification.Macmillan Publishers LTD, England.

Ewer, J., Hamblen-Coyle, M., Rosbash, M. and Hall, J.C. (1990) Requirement for *period* gene expression in the adult and not during development for locomotor activity rhythms of imaginal *Drosophila melanogaster*. J. Neurogenet. 7 : 31-73.

Ewer, J., Frisch, B., Hamblen-Coyle, M.J., Rosbash, M. and Hall, J.C. (1992) Expression of the *period* clock gene within different cell types in the brain of *Drosophila* adults and mosaic analysis of these cells' influence on circadian behavioral rhythms. J. Neurosci. 12 : 3321-3349.

Feldman, J.F. and Hoyle, M.N. (1973). Isolation of circadian clock mutants of *Neurospora crassa*. Genetics 75 : 605-613.

Frisch, B., Hardin, P.E., Hamblen-Coyle, M.J., Rosbash, M. and Hall, J.C. (1994) A promoterless *period* gene mediates behavioral rhythmicity and cyclical *per* expression in a restricted subset of the *Drosophila* nervous system. Neuron 12 : 555-570

Gaily, D. and Hall, J.C. (1989) Behaviour and cytogenetics of *fruitless* in *Drosophila melanogaster* : different courtship defects caused by separate , closely linked lesions. Genetics 121 : 773-476

Garceau, N.Y., Liu, Y., Loros, J.J. and Dunlap, J.C, (1997) Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein *frequency*. Cell 89 : 469-476

Gardner, G.F. and Feldman, J.F.(1981) Temperature compensation of cicadian periodicity in clock mutants of *Neurospora crassa*. Plant Physiol. 68 : 1244-1248

Gekakis, N., Saez, L., Delahaye-Brown, A.M., Myers, M.P., Sehgal, A., Young, M.W. and Weitz, C.J. (1995) Isolation of *timeless* by *per* protein interaction, defective interaction between *timeless* protein and long-period mutant per^{L} . Science, 270 : 811-815

Gekakis, N., Staknis, D., Nguyen, H.B., Davis F.C., Wilsbacher, L.D., King, D.P., Takahashi, J.S. and Weitz, C.J. (1988) Science 280 : 1564-1569

Giebultowicz, J.M. and Hege, D.M. (1997) A circadian clock in Malpighian tubules. Nature 386 : 664

Gloor, G. and Engels, W. (1990) Single-fly DNA preps for PCR. *Drosophila* Information Newsletter 1

Goldbeter, A. (1995). Amodel for circadian oscillations in the Drosophila period protein (PER). Proc. Roy. Soc. London 261 : 319-324

Gray, I. and Jeffreys, A.J. (1991) Evolution transience of hypervariable minisatellites in man and primates. Proc R. Soc. London. Biol. 243 : 241-253

Grobbelar, N., Hung, T-C., Lin, H.Y. and Chow, T.J. (1986) Dinitrogen-fixing endogenous rhythm in *Synechococcus* RF-1. FEMS Microbiol. Lett. 37 173-177

Gussow, D. and Clackson, T. (1989) Direct clone characterization from plaques and colonies by the polymerase chain reaction. Nucleic Acids research 17 : N° 10

Hamblen, M., Zehring, W.A., Kyriacou, C.P., Reddy, P., Yu, Q., Wheeler, D.A., Zwiebel, L.J., Konopka, R.J., Rosbash, M. and Hall, J.C. (1986) Germ-line transformation involving DNA from the *period* locus in *Drosophila melanogaster*: overlapping genomic fragments that restore circadian and ultradian rhythmicity to per^0 and per^- mutants. J. Neurogenet. 3 : 249-291

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

Hancock, J.M. and Dover, G.A. (1990). Compensatory slippage in the evolution of ribosomal RNA genes. Nucl. Acids Res. 18 : 5949-5953

Hao, H., Allen, D.L. and Hardin, P.E. (1997) A circadian enhancer mediates PER-dependent mRNA cycling in *Drosophila* melanogaster. Mol. and Cell. Biol. 17(7) : 3687-3693

Hardin, P.E., Hall, J.C. and Rosbash, M. (1990) Feedback of the *Drosophila period* gene product on circadian cycling of its messenger RNA levels. Nature 343 : 536-540.

Hardin, P.E., Hall, J.C. and Rosbash, M. (1992a) Circadian oscillations in *period* gene mRNA levels are transcriptionally regulated. Proc. Natl. Acad. Sci. USA 89 : 11711-11715.

Hardin, P.E., Hall, J.C. and Rosbash, M. (1992b) Behavioral and molecular analyses suggest that circadian output is disrupted by *disconnected* mutants in *D.melanogaster*. EMBO J. 11 : 1-6

Helfrich-Forster, C. Robust circadian rythmicity of *Drosophila melanogaster* requires the presence of lateral neurons: a brain-behavioural study of *disconnected* mutants. J. Comparative Physiol. 1998 182 : 435-453

Hennessy, J.M. (1991). The construct of a chimaeric period gene to investigate the genetic basis of species-specific behaviour in *Drosophila*. MSc Thesis, University of Leicester

Hennig W. (1981) Insect phylogeny. Pitman Press, Bath, England

HATZOPOLOU, A. (1994). Characterisation of transgenic *D.melanogaster* lines by *in situ* hybridization. Laurea Thesis, University of Padova

Huang, Z.J., Curtin, K.D. and Rosbash, M. (1995) PER protein interactions and temperature compensation of a circadian clock in *Drosophila*. Science 267 : 1169-1172

Huang, Z.J., Edery, I. and Rosbash, M. (1993) PAS is a dimerization domain common to *Drosophila* Period and several transcription factors. Nature 364 : 259-262

Hulme, M., Conway, D., Jones, P..D., Jiang T., Barrow, E.M. and Turney, C. (1995) Construction of a 1961-90 European climatology for climate change modelling and impact applications. International Journal of Climatology 15 : 1333-1363.

Hunter-Ensor, M., Ousley, A. and Sehgal, A. (1996) Regulation of the *Drosophila* protein *timeless* suggests a mechanism for resetting the circadian clock by light. Cell 84 : 677-685

Ishiura, M.,Kutsuna, S., Aoki, S., Iwasaki, H., Andersson C., Tanabe, A., Golden, S., Johnson, C., Kondo, T., (1998) Expression of a gene cluster *kaiABC* as a circadian feedback proces in Cyanobacteria. Science 281:1519-1523

Jeffreys, A.J., Wilson, V., Neumann, R. and Keyte, J. (1988) Amplification of human minisatellites by polymerase chain reaction: towards DNA fingerprinting of single cells. N.A.R. 16 : 10953-10971

Kimura, M. (1991). Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. Proc. Natl. Acad. Sci. USA 88 : 5969-5971.

King, D.P., Vitaterna, M.H., Chang, A.M. Dove, W.F., Pinto, L.H., Turek, F.W., and Takahashi, J.S. (1997a) Positional cloning of the mouse circadian Clock gene. Cell 89 : 641-953

King, D.P., Vitaterna, M.H., Chang, A.M., Dove, W.F., Pinto, L.H., Turek, F.W. and Takahashi, J.S. (1997b) The mouse *Clock* mutation behaves as an antimorph and maps within the W^{19H} deletion distal of *Kit*. Genetics 146 : 1049-1060

Klemenz, R., Weber, U. and Gehrig, W. J. (1987) The white gene as a marker in a new P-element vector for gene transfer in *Drosophila*. N.A.R. 10 : 39474-3959

Kliman, R.M. and Hey, J. (1993) DNA sequence variation at the *period* locus within and among species of the *Drosophila melanogaster* complex. Genetics 133 : 375-387

Kloss, B., Price, J.L., Saez, L., Blau., J., Rothenfluh, A., Wesley, C.S. and Young, M.W. (1998) The *Drosophila* clock gene *double-time* encodes a protein closely related to human casein kinase Iε. Cell 94 : 97-107

Kondo, T., Strayer, C.A., Kulkarni, R.D., Taylor, W., Ishiura, M., Golden, S.S. and Johnson, C.H. (1993) Circadian rhythms in prokaryotes: Luciferase as a reporter of circadian gene expression in cyanobacteria. Proc. Natl. Acad. Sci. USA. 90 : 5672-5676.

Kondo, T. and Ishiura, M. (1994) Circadian rhythms of cyanobacteria: monitoring the the biological clocks of individual colonies by bioluminescence. J. Bact. 176 : 1881-1885

Konopka, R.J. and Benzer, S. (1971) Clock mutants of *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA 68 : 2112-2115

Konopka, R.J., Pittendrigh, C.S. and Orr, D. (1989) Reciprocal behaviour associated with altered homeostasis and photosensitivity of *Drosophila* clock mutants. J. Neurogenet. 6 : 1-10

Kyriacou, C.P. and Burnet, B. (1979) Genetic analysis of the phototaxis near the upper limit of the visual spectrum of *Drosophila melanogaster.* Behav. Genetics 9 : 123-128

Kyriacou, C.P. and Hall, J.C. (1980) Circadian rhythms mutations in *Drosophila melanogaster* affect short term fluctuations in the male's courtship song. Proc. Natl. Acad. Sci. USA 77 : 6929-6933.

Kyriacou, C.P. and Hall, J.C. (1986) Interspecific genetic control of courtship song production and reception in *Drosophila.* Science 232 : 494-497

Kyriacou, C.P. and Hall, J.C. (1989) Spectral analysis of *Drosophila* courtship song rhythms. Anim. Behav. 37 : 850-859

Lee, C., Parikh, T., Itsukaichi, T., Bae, K. and Edery, I. (1996) Resetting the *Drosophila* clock by photic regulation of PER and PER-TIM complex. Science 271 : 1740-1744

Liu X., Lorenz L., Yu Q., Hall J.C. and Rosbash M (1988) Spatial and temporal expression of the *period* gene in *Drosophila melanogaster.* Genes Dev. 2 : 228-238.

Liu, X., Zwiebel, L.J., Hinton, D., Benzer, S., Hall, J.C. and Rosbash, M. (1992) The *period* gene encodes a predominantly nuclear protein in the adult *Drosophila.* J. Neurosci. 12 : 2735-2744.

Liu, Y., Merrrow, M., Loros, J.J. and Dunlap, J.C. (1997) How temperature changes reset a circadian oscillator. Science 281 : 825-829

Majercak, J., Kalderon, D. and Edery, I. (1997) Drosophila melanogaster deficient in Protein Kinase A manifests behaviour-specific arrythmia but normal clock function. Mol. and Cell. Biol. 1997 : 5195-5922

Marrus, S.B., Zeng, H. And Rosbash, M. (1996) Effect of constant light and circadian entrainment of *per*^{S} flies: evidence for light mediated delay of the negative feedback loop in *Drosophila.* EMBO J. 15 : 6877-6886

McClung, C.R., Fox, B.A. and Dunlap, J.C. (1989) *frequency*, a clock gene in *Neurospora* shares a sequence element with the *Drosophila* clock gene *period.* Nature 339 : 558-562

McNeil, G.P., Zhang, X., Genova G. and Jackson F.R. (1998) A molecular rhythm mediating circadian clock output in *Drosophila.* Neuron 20 : 297- 303

Merrow, M.W., Garceau, N.Y. and Dunlap, J.C. (1997) Dissection of a circadian oscillation into discrete domains. Proc. Natl. Acad. Sci. USA 94 : 3877-3882

Myers, M.P., Wagner-Smith, K., Wesley, C.S., Young, M. and Sehgal, A. (1995) Positional cloning and sequence analysis of the *Drosophila* clock gene, *timeless.* Science 270 : 805-808.

Myers, M.P., Wager-Smith, K., Rothenfluh-Hilfiker, A. and Young, M.W. (1996) Light induced degradation of TIMELESS and entrainment of the *Drosophila* circadian clock. Science 271 : 1763-1740

Newby, L.M. and Jackson, F.R. (1991). *Drosophila ebony* mutants have altered circadian activity rhythms but normal eclosion rhythms. J. Neurogenetics 7 : 85-101

Newby, L.M. and Jackson, F.R. (1996). Regulation of a specific circadian clock output pathway by *lark*, a putative RNA-Binding Protein with repressor activity. J. Neurobiol. 31 117-128

Nielsen, J., Peixoto, A.A., Piccin, A., Costa, R., Kyriacou, C.P. and Chalmers, D. (1994) Big Flies, Small Repeats - the Thr-Gly Region Of the Period Gene In Diptera. Mol. Biol. Evol. 11 : 839-853

Parkinson, H.D. (1997) Mutagenisis of the clock gene *period* in *Drosophila melanogaster* PhD Thesis, University of Leicester.

Parsch, J., Tanda, S., and Stephan, W. (1997) Site directed mutations reveal long range compensatory interactions in the Adh gene of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 94, 928- 933

Peixoto, A.A (1993) Molecular evolution of a repetitive region within a clock gene in Drosophila. PhD Thesis, University of Leicester.

Peixoto, A.A., Costa, R., Wheeler, D.A., Hall, J.C. and Kyriacou, C.P. (1992) Evolution of the Threonine-Glycine repeat region of the *period* gene in the *melanogaster* species subgroup of *Drosophila*.J. Mol. Evol. 35 : 411-419

Peixoto, A.A., Campesan, S., Costa, R. and Kyriacou, C.P. (1993) Molecular evolution of a repetitive region within the *per* gene of *Drosophila*. Molecular Biology and Evolution 10 : 127-139.

Peixoto, A.A., Hennessy, J.M., Townson, I., Hasan, G., Rosbash, M., Costa, R. and Kyriacou, C.P. (1998) Molecular coevolution within a *Drosophila* clock gene. Proc. Natl. Acad. Sci. USA 95 : 4475-4480

Petersen, G., Hall, J.C. and Rosbash, M. (1988) The *period* gene of *Drosophila* carries species-specific behavioural instructions. EMBO Journal. 7 : 3939-3947

Piccin, A. (1998). The clock gene *period* in the housefly, *Musca Domestica* : a molecular analysis. PhD thesis , University of Leicester.

Pittendrigh, C.S (1954) On the temperature independence in the clock system controlling emergence time in *Drosophila*. Proc. Natl. Acad. Sci. USA 40 : 1018-1029

Pittendrigh, C.S (19581). Circadian organization and the photoperiodic phenomena. Biological Clocks in reproductive cycles Ed B. K. Follet.Published: John Wright, Bristol.

Plautz, J.D., Kaneko, M., Hall, J.C. and Kay, S.A. (1997) Independant photoreceptive circadian clocks throughout *Drosophila*. Science 278 : 1632-1635

Price, J.L., Dembinska M.E., Young M.W. and Rosbash M. (1995) Suppression of period protein abundance and circadian cycling by the *Drosophila* clock mutation *timeless*. EMBO J. 14 : 4044-4049

Price, J.L., Blau, J., Rothenfluh, A., Abodeely, M., Kloss, B. and Young M.W. (1998) *double-time* is a novel *Drosophila* clock gene that regulates PERIOD protein accumulation. Cell 94 : 83-95

Qiu, J. and Hardin, P.E. (1996) *per* mRNA cycling is locked to lights off under photoperiodic conditions that support circadian feedback loop function. Mol. Cell. Biol. 16 : 4182-4188

Quinn, W.G. and Greenspan, R.J. (1984). Learning and courtship in *Drosophila* : two stories with mutants. Annu. Rev. Neurosci. 21 : 67-93

Ralph, M., (1996) Circadian rhythms-mammalian aspects. Cell and Developmental Biology. 7 : 821-830

Roberts, D.B. (1986) Drosophila a practical approach. Eds Rchwood D. and Hames B.D. IRL Press, Oxford.

Roberts, D. H., Lehar, J., and Dreher, J. W. (1987) Time series with CLEAN. I Derivation of a spectrum. J. Astronomy 93 : 968-989

Robertson, H.M., Preston, C.R., Phillips, R.W., Johnson-Chlitz,D.M. Benz,W.K. and Engels, W.R. (1988) A stable source of P-element transposase in *Drosophila melanogaster*. Genetics 118 : 461-470

Rogers, S., Wells, R. and Rechsteiner, M. (1986) Amino acid sequences common to rapidly degraded proteins. Science 234 : 364-368

Rosato, E., Peixoto, A.A., Barbujani, G., Costa, R. and Kyriacou, C.P. (1994) Molecular polymorphism in the *period* gene of *Drosophila simulans*. Genetics 138 : 693-707.

Rosato, E., Peixoto, A.A., Gallippi, A., Kyriacou, C.P. and Costa, R. (1996) Mutational mechanisms, phylogeny, and evolution of a repetitive region within a clock gene of *Drosophila melanogaster*. J. Mol. Evol. 42 : 392-408.

Rosato, E., Peixoto, A.A., Costa, R. and Kyriacou, C.P. (1997) Linkage disequilibrium , mutational analysis and natural selection in the repetitive region of the clock gene, *period*, in *Drosophila melanogaster*. Genet. Res., Camb. 69 : 89-99

Rosbash, M., Allada, R., Dembinska, M., Guo, W.Q., Le, M., Marrus, S., Qian, Z., Rutila, J., Yaglom, J. and Zeng, H. (1996) A *Drosophila* circadian clock. Cold Spring Harb. Sym. Quant. Biol. 61 : 265-278

Rousset, F., Pelandakis, M. and Solignac, M. (1991) Evolution of compensatory substitutions through G.U intermediate state in Drosophila rRNA. Proc. Natl. Acad. Sci. USA. Vol 88, 10032-10036

Russo, C.A.M., Takezaki, N. and Nei, M. (1995) Mol. Biol. Evol. 12 : 391-404 Molecular phylogeny and divergence times of Drosophilid species.

Rutila, J.E., Zeng, H., Le, M., Curtin, K.D., Hall, J.C. and Rosbash, M. (1996) The tim^{SL} mutation of the *Drosophila* rhythm gene *timeless* manifests allele specific intercations with *period* gene mutants. Neuron 17 :921-929

Rutila , J.E., Suri, V., Le, M., So W.V., Rosbash, M. and Hall, J.C. (1998) Cycle is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of *Drosophila period* and *timeless*. Cell 93 : 805-814

Saez, L. and Young, M.W. (1988) *In situ* localisation of the *per* clock protein during development of *D. melanogaster* Mol. Cell. Biol. 8 : 5378-5385

Saez, L. and Young M.W. (1996) Regulation of nuclear entry of the *Drosophila* clock proteins PERIOD and TIMELESS. Neuron 17 : 911-920

Sambrook, Maniatis and Frisch (1989) Molecular Cloning, a laboratory manual. Cold Spring Harbour Laboratory Press, USA

Saunders, D.S. (1982). Insect Clocks. 2nd ed. Pergamon Press, Oxford

Saunders, D.S., Gillanders, S.W. and Lewis, R.D. (1994) Light pulse phase response curves for the locomotor activity rhythm in *period* mutants of *D. melanogaster*. J Insect Physiology 40 : 957-968.

Sawyer, L.A. (1996) Naural variation within a clock gene of *Drosophila melanogaster* : a phenotypic and molecular analysis. PhD Thesis, University of Leicester

Sawyer, L.A., Hennessy. J.M., Peixoto, A.A., Rosato, E., Parkinson, H.E. Costa, R. and Kyriacou, C.P. (1997). Natural variation in a *Drosophila* clock gene and temperature compensation. Science 278 : 2117-2120

Sehgal, A., Price, J.L., Man, B. and Young, M.W. (1994) Loss of circadian behavioural rhythms and *per* RNA oscillations in the *Drosophila* mutant *timeless*. Science 263 : 1603-1606.

Sehgal, A., Rothenfluh-Hilfiker, A., Hunter -Ensor, M., Chen, Y.F., Myers, M.P. and Young, M.W. (1995) Rhythmic expression of *timeless* - a basis for promoting circadian cycles in *period* gene autoregulation. Science 270 : 808-810

Serghini, M.A., Ritzenthaler, C. and Pinck, L. (1989) A rapid and efficient 'miniprep' for isolation of plasmid DNA. N.A.R. 17 : 3604

Shearman, L., Zylka, M., Weaver, D., Kolakowski, L., Reppert, S., (1977) Two *period* homologs: Circadian and photic regulation in the suprachiasmatic nuclei. Neuron 19, 1261-1269

Sidote, D., Majercak, J., Parikh, V., Edery, I. (1998) Differential effects of light and heat on the *Drosophila* circadian clock proteins PER and TIM. Mol. Cell. Biol. Vol 18, N° 4 : 2004-2013

Siwicki, K.K., Eastman, C., Petersen, G, Rosbash, M. and Hall, J.C. (1988) Antibodies to the *period* gene product of *Drosophila* reveal diverse tissue distribution and rhythmic changes in the visual system. Neuron 1 : 141-150.

Siwicki, K.K., Schwartz, W.J. and Hall, J.C. (1992) An antibody to the *Drosophila* period protein labels antigens in the suprachiasmatic nucleus of the rat. J Neurogen 8 : 33-42

Stanewsky, R., Frisch, B., Brandes, C., Hamblen-Coyle, M. J., Rosbash, M. and Hall, J.C. (1997a) Temporal and spatial expression patterns of transgenes containing increasing amounts of the *Drosophila* clock gene *period* and a *lacZ* reporter : mapping elements of the PER protein involved in circadian cycling. J. Neurosci. 17(2) : 676-696

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

So, W.V. and Rosbash, M. (1997) Post translational regulation contributes to Drosophila clock gene mRNA cycling. EMBO J. 23 : 7146-7155

Spradling, A.C. (1986) P-element mediated transformation . Drosophila: Apractical approach (D.B. Roberts ed.). IRL press, Oxford. pp 175-197

Suri, V., Qian, Z., Hall, J.C. and Rosbash, M. (1998) Evidence that the TIM light response is relevant to light-induced phase shifts in *Drosophila melanogaster.* Neuron 21: 225-235

Takumi, T., K.,Miyake, S.,Sakakida, Y.,Takashima, N.,Matsubara, C.,Maebayashi, Y.,Okumura, K.,Takekida, S.,Yamamoto, S.,Yagita, K.,Yan, L., Young, MW., Okamura, H (1998). A light-independent oscillatory gene *mPer3* in mouse SCN and OVLT. EMBO JOURNAL 17, 4753-4759

Tei, H., Okamura, H., Shigeyoshi, Y., Fukuhara, C., Ozawa, R., Hirose, M., and Sakaki, Y., (1997) Circadian oscillation of a mammalian homologue of the *Drosophila* period gene. Nature 389 : 512-516

Tomlinson, A. and Ready, D. (1986) Sevenless : a cell specific homeotic mutationof the *Drosophila* eye. Science 231 : 400-402

Van Gelder, R.N. and Krasnow, M.A. (1996) A novel circadianly expressed *D. melanogaster* gene dependent on the *period* gene for its rhythmic expression. EMBO J. 15 : 1625-1631

Van Gelder, R.N., Bae, H., Palazzolo, M. and Krasnow, M.A. (1995) Extent and character of circadian gene expression in *D. melanogaster.* Identification of 20 oscillating mRNAs in the fly head. Current Biol. 5 : 1424-1436

Vosshall, L. B., Price, J.L., Sehgal, A., Saez, L. and Young, M.W. (1994) Block in nuclear localization of *period* protein by a second clock mutation, *timeless.* Science 263 : 1606-1609

Vosshall, L. and Young, M.W. (1995) Circadian rhthms in *Drosophila* can be driven by *period* expression in a restricted group of central brain cells. Neuron 12 : 555-570

Wang R.L. and Hey J. (1996) Gene flow and natural selection in the origin of Drosophila pseudoobscura and close relatives. Genetics 147 : 1091- 1106

Wheeler, D.A., Kyriacou, C.P., Greenacre, M.L., Yu, Q., Rutila, J.E., Rosbash, M. and Hall, J.C. (1991) Molecular transfer of a species-specific behaviour from *Drosophila simulans* to *Drosophila melanogaster.* Science 251 : 1082-1085.

Wheeler, D.A., Hamblen-Coyle, M.J., Dushay, M.S. and Hall, J.C. (1993) Behaviour in light-dark cycles of *Drosophila* mutants that are arrhythmic , blind, or both. J. biol. Rhythms 8 : 67-94

Winer, B.J. (1971) Statistical principles in experimental design. Tokyo, McGraw-Hill Kogakusha

Yang , Z., Emerson, M., Su, H.S. and Sehgal, A. (1998) Response of the *timeless* protein to light correlates with behavioural entrainment and suggests a nonvisual pathway for circadian photoreception. Neuron 21 : 215-223

Yon and Fried (1989) Precise gene fusion by PCR. Nucl. Acids Res. Vol 17 N° 12 4895-

Young M.W., Wager-Smith, K., Vosshall, L., Saez, L., and Myers, M.P. (1996) Moelcular biology of a light sensitive circadian pacemaker in *Drosophila*. Cold Spring Harbor Symposium on Quantitative Biology. Volume LXI

Yu, Q., Jacquier, A.C., Citri, Y., Hamblen, M., Hall, J.C. and Rosbash, M. (1987a). Molecular mapping of point mutations in the *period* gene that stop or speed up biological clocks in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 84 : 784-788.

Yu, Q., Colot, H.V., Kyriacou, C.P., Hall, J.C. and Rosbash, M. (1987b) Behaviour modification by *in vitro* mutagenesis of a variable region within the *period* gene of *Drosophila*. Nature 326 : 765-769

Zeng, H., Hardin, P.E. and Rosbash, M. (1994) Constitutive overexpression of the *Drosophila period* protein inhibits *period* mRNA cycling. EMBO Journal 13 : 3590-3598.

Zeng, H., Quian, Z., Myers, M.P. and Rosbash, M. (1996) A light entrainment mechanism for the *Drosophila* circadian clock. Nature 380 : 129-135

Zerr, D.M., Hall, J.C., Rosbash, M. and Siwiki, K.K. (1990) Circadian fluctuations of period protein immunoreactivity in the CNS and the visual system of *Drosophila*. J. Neuroscience 10 : 2749-2762.