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 $(Thr\text{-}Gly)_n$ 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.

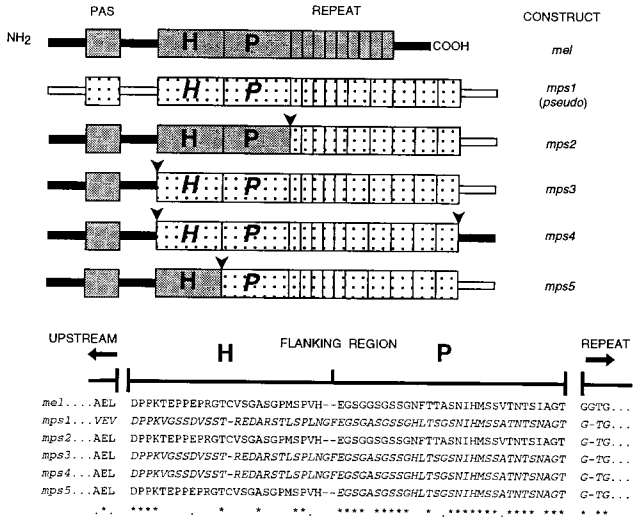

FIG. 1. (*Top*) *per* transgenes of *D. melanogaster* N- and C-terminal sequences (filled bars) *D. pseudoobscura* (open bars). The approximate positions of the N-terminal PAS domain (39) are shown as filled or speckled boxes for the two species. *D. melanogaster* sequences in blocks H (amino acids 639–664, 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.

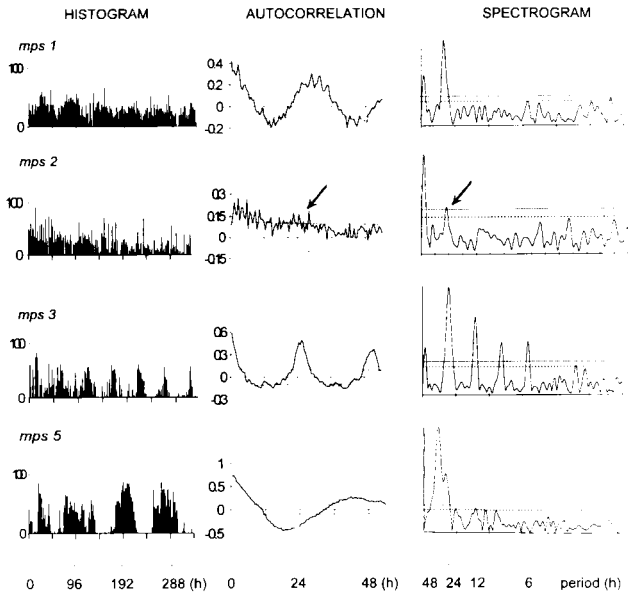

FIG. 2. Free running locomotor activity histograms, autocorrelation, and spectral analysis for four male per^{01} transformants monitored at 29°C carrying the *mps1, mps2, mps3,* and *mps5* transgenes. Histograms give raw activity events in 30-min bins. The correlograms show r and 95% confidence limits 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,

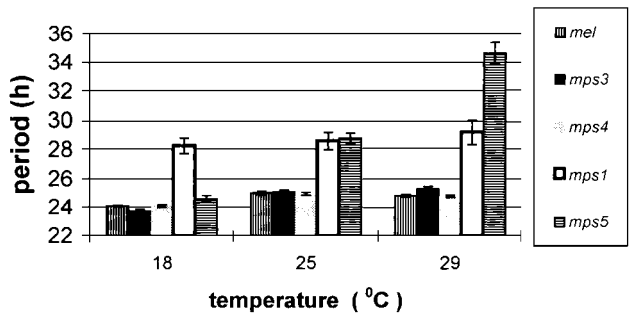

FIG. 3. Pooled mean (±SEM) free running locomotor activity periods for the different transformants at 18, 25, and 29°C. Data for the individual lines are in Table 1. *mps5* gives a dramatic increase in period at higher temperatures. *mps2* results are not shown as most transformants were arrhythmic,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

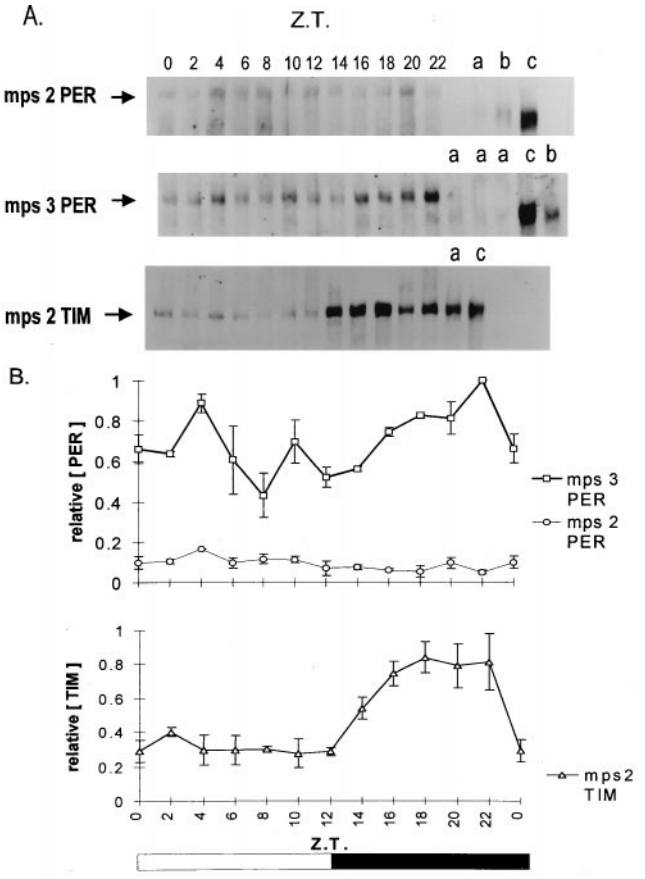

FIG. 4. (*A*) Western blots of *mps2* and *mps3* transformant heads collected at different Zeitgeber times (ZT 0, lights on; ZT 12, lights off) in LD 12:12 at 25°C. Lanes on the right show per^{01} (a), tim^{0} (b), and per^{+} Canton-S (c) controls at ZT 0, except for the TIM blot where the controls 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.