Intergeneric complementation of a circadian rhythmicity defect: phylogenetic conservation of structure and function of the clock gene *frequency*

Martha W. Merrow and Jay C. Dunlap¹

Department of Biochemistry, Dartmouth Medical School, Hanover, NH 03755-3844, USA

¹Corresponding author

Communicated by G. Dover

The *Neurospora crassa frequency* locus encodes a 989 amino acid protein that is a central component, a state variable, of the circadian biological clock. We have determined the sequence of all or part of this protein and surrounding regulatory regions from additional fungi representing three genera and report that there is distinct, preferential conservation of the *frequency* open reading frame (ORF) as compared with non-coding sequences. Within the coding region, many of the domain hallmarks of the *N. crassa* protein are highly conserved, especially an internal region bearing the causative mutations in *frq¹* and *frq⁷*, the most extreme alleles in the *frequency* allelic series. Despite considerable diversity among the strains analyzed in terms of morphology, growth, circadian clock output and *frq* sequence, the ORF from the most distantly related fungus included in this study (*Sordaria fimicola*) rescues rhythmicity in a *N. crassa frequency* null strain. Both sequence conservation, and the ability of *frequency* from a genus displaying one developmental program to complement circadian defects in a separate genus with a distinct, clock-regulated developmental program, are consistent with a central role of the *frequency* gene product in a general circadian oscillator capable of controlling diverse outputs in a variety of systems.

Key words: circadian/conidia/*frequency*/*Neurospora*/*Sordaria*

Introduction

Circadian biological clocks are observed in both prokaryotic and eukaryotic species, having been characterized in organisms as diverse as cyanobacteria, plants, mollusks, fungi, algae, flies and mammals (Edmunds, 1988; Dunlap, 1993; Hall and Rosbash, 1993). A circadian clock is manifested by a rhythmic change in some physiological, biochemical or behavioral parameter which recurs every 24 h under typical environmental conditions, and approximately once during that period under constant conditions. The mechanism by which time is kept has not yet been characterized, although its complexity is inferred by the existence of multiple distinct circadian rhythms within a unicellular organism (Roenneberg and Morse, 1993). Because clocks function at an intracellular level, simple organisms with modest differentiative potential represent attractive model systems. The filamentous fungus *Neurospora crassa* is ideal for this purpose. During the

asexual vegetative growth cycle *Neurospora* reproduces as an undifferentiated mycelium, periodically giving rise to aerial hyphae which bear masses of orange-colored asexual spores called conidia. Under alternating light/dark cycles, or in constant darkness following entrainment with light, the developmental shift which results in conidium production is under the control of the circadian clock (Pittendrigh *et al.*, 1959). Thus, in the *band* strain of *N. crassa* intense conidiation occurs once per 21.5 h in constant darkness (Sargent *et al.*, 1966), and this rhythm persists without re-entrainment by environmental cues.

The genetic underpinning of the oscillator has been demonstrated by the identification of mutant strains with period lengths which deviate from wild-type. Fourteen clock mutants have been isolated from *N. crassa*, the majority of which (eight) map to the *frequency* (*frq*) locus (Dunlap, 1993). *frq* was cloned by identifying the minimal DNA fragment capable of rescuing rhythmicity in the recessive *frq⁹* null mutant strain (McClung *et al.*, 1989). At least two scarce and divergent transcripts of ~1500 and 4500 nt arise from this region. The small transcript is unlikely to encode a protein (Johnson, 1993), while the large transcript contains a continuous open reading frame (ORF) encoding a 989 amino acid protein. All of the *frq* mutations have been identified as sequence alterations contained within this extended ORF (Aronson *et al.*, 1994a).

One approach to determining how a gene functions is analysis of gene homologs from different species. As has been amply borne out in the study of cell cycle genes, molecules critical for the execution of universal cellular processes often share functional and, therefore, molecular conservation (Sikorski *et al.*, 1990). Thus, homologous genes cloned from related species have served as an analytical tool, permitting determination of probable functional domains based on preferential phylogenetic conservation. We have taken this approach to analyze the clock gene *frq*.

Neurospora crassa belongs to the family *Sordariaceae*, which includes the genera *Neurospora*, *Gelasinospora* and *Sordaria*, among others. Members of these genera were targeted for analysis of their *frq* loci by DNA sequence determination. An estimate of the phylogenetic relatedness among these genera was obtained by sequence analysis of PCR products, and a *frq* homolog was then cloned from the most distantly related species, *Sordaria fimicola*. Molecular analysis of this clone demonstrates that the ORF is preferentially conserved relative to non-coding sequences, and further, that some regions within the ORF are rigorously maintained between the species. Transformation of a recessive *N. crassa frq* null strain with a construct bearing the *S. fimicola frq* ORF in the context of *N. crassa* non-coding sequences results in restoration of a circadian clock phenotype with properties similar to the *N. crassa* circadian rhythm. This complementation is the first example of transfer

of a clock phenotype across genera, and it occurs despite the fact that *S.fimicola* does not exhibit the developmental program producing macroconidia, the most salient biological manifestation of the circadian clock in *N.crassa*. These data confirm the importance of the protein encoded by the long ORF of *frq* as a key element of oscillator function, and will serve to highlight the importance of discrete parts of the FRQ protein.

Results

Molecular analysis of *frq* homologs

The *N.crassa frq* locus was cloned by a chromosome walk between the *oligomycin* and *formate* loci on linkage group VIIR (McClung *et al.*, 1989). Full phenotypic rescue of the recessive mutant *frq*⁹ strain was achieved with an 8.7 kb *Cl*I fragment containing the *frq* locus (Figure 1). The entire *Cl*I fragment was used to probe Southern blots of genomic DNA from a variety of fungi, and within the *Sordariaceae* family (*N.intermedia*, *N.sitophila*, *N.discreta*, *Gelasinospora calospora*, *G.cerealis*, *G.reticulospora* and *S.fimicola*) cross-hybridization with heterologous *N.crassa frq* was consistently and readily visualized (data not shown). The weakest hybridization among this group was between *S.fimicola* and *N.crassa*, suggesting that *S.fimicola* was the most distantly related to *N.crassa*. This inference was borne out by preliminary PCR and sequence data (see below). Taken together, these data suggested that the *S.fimicola frq* locus would contain the greatest divergence within the *Sordariaceae* genera which were evaluated. In order to clone *S.fimicola frq*, the locus was characterized by restriction enzyme digests and a size selected genomic DNA pool was created (see Materials and methods). A 6.3 kb genomic DNA fragment which hybridized to *N.crassa frq* was obtained (Figure 1B). This clone (pSOR3.1) was sequenced in its entirety, and was found to encode regions corresponding to the *N.crassa frq* large ORF, and to 5' and 3' flanking sequences.

Sequence comparisons revealed that the region containing the extended ORF within the large transcript of *N.crassa frq* is the most highly conserved part of the *frq* locus. Specifically, throughout this area the DNA sequence of *S.fimicola frq* is 86% identical to that of *N.crassa frq* (Figure 2). There are nucleotide insertions or deletions at four regions within the *Sordaria* sequence, all of which occur in multiples of 3 bp. As with *N.crassa frq*, the ORF contains no introns, and in this case it encodes 997 amino acids. Both *frq* sequences appear to utilize the same translational start site. The *S.fimicola frq* DNA contains two additional ATG codons upstream of the putative initiating methionine, neither of which is in-frame with the extended ORF, nor in a good consensus context for translation initiation. In contrast, the methionine codons commencing the long ORF of both *frq* genes are bounded by identical nucleotides (GTG-AACATGG), which fulfill a minimal Kozak consensus sequence [GC(A/G)C(A)AC(A)CATGG; Kozak, 1987], a feature common to many eukaryotic translation start sites.

The deduced protein sequence of the *S.fimicola frq* homolog (Figure 2) also shows a high degree of general conservation: 85% overall identity and 92% similarity based on conservative substitutions (Bestfit). However, interestingly, there are distinct domains possessing

Fig. 1. The *frq* loci of *N.crassa* and *S.fimicola*. (A) The 8.6 kb *Cl*I (*C*) *N.crassa* genomic DNA clone which rescues the arrhythmic *frq*⁹ phenotype. At least two transcripts arise from the locus, and they are divergently transcribed, as indicated by the arrows below the diagram. The longer transcript encodes a 989 amino acid ORF. The expanded region above the diagram represents the ORF, approximate positions of the *frq* mutations are indicated by their allele number and an asterisk. The nuclear localization signal (NLS), the hyperacidic region (HA) and the strongest PEST sequences contained within are marked. (B) The insert of the genomic DNA clone of the *frq* homolog from *S.fimicola*, pSOR3.1. The conserved *Bgl*III (*Bg*) site falls within DNA corresponding to the smaller transcript of *N.crassa*. The conserved *Sph*I (*Sp*) restriction site near the beginning of both *N.crassa* and *S.fimicola* ORFs is shown. The *Xba*I (*X*) site outside the 3' end of the DNA encoding the extended ORF facilitated cloning. (C) Plasmids used in phenotype rescue experiments. pKAJ120 contains the 8.7 kb *N.crassa Cl*I fragment cloned into pDE3ΔBH, a vector which permits gene targeting to the *his-3* locus. The *N.crassa* ORF is indicated. pMM203 is identical to pKAJ120 except that DNA from the *S.fimicola frq* clone corresponding to the ORF replaces the *N.crassa* sequence in this region. Additional restriction endonuclease sites are indicated: A, *Age*I; Sn, *Sna*BI; M, *Mlu*I; B, *Bss*HII.

preferential conservation or divergence, such that locally the degree of identity ranges from 76 to 96%, when calculated in lengths of 50 amino acids (Figure 3). Similar calculations for the DNA sequence reveal less overall fluctuation (80–92% identity as calculated in 150 nt windows; Bestfit, gap weight = 5.0) relative to the variability noted for the translation product. Third base substitutions within codons account for 61% of all DNA sequence variability between the *S.fimicola* and *N.crassa frq* ORFs. Therefore, the amino acid sequence conservation within the ORF probably reflects biological selection of gene function. Two internal regions that are exceptionally highly conserved correspond to amino acids 158–230 (containing two conservative substitutions) and 425–502 (one conservative change). Although there are no known or theoretical functions assigned to either region, the latter overlaps the sites of the two most extreme mutant alleles of *frq*, *frq*¹ (period length or τ = 16 h) and *frq*⁷ (τ = 29 h at 25°C). All of the mutant codons in the *N.crassa frq* alleles are maintained as wild-type *N.crassa* FRQ sequence in the homologous *S.fimicola frq* gene product, and the predicted nuclear localization signal is also intact. Several PEST regions, which correlate with rapid protein turnover (Rechsteiner, 1990), were identified using PESTFIND (Rogers *et al.*, 1986). There are four PEST regions in *N.crassa*, which are intact in *S.fimicola frq*, and one additional PEST sequence is located within the *Sordaria* homolog (amino acids 650–697, *S.fimicola* sequence, PEST

[illegible]

Fig. 2. The DNA sequence of the extended ORF from pSOR3.1 and its 997 amino acid translation product are compared with the *N.crassa* *frq* sequence. The top two lines are the *S.fimicola* protein and DNA sequences, respectively, and numbering corresponds to those sequences. The DNA and protein sequences of *N.crassa* are noted in the next two rows only where differences between the two species occur. The highest scoring PEST regions which are conserved within *N.crassa* and *S.fimicola* FRQ are underlined (PEST scores are 18.9 and 10.0, in order of location), and the hyperacidic region found within the most C-terminal PEST sequence is italicized in bold print. The putative NLS is boxed and shaded, and the locations of the *frq* mutations are boxed with their corresponding numbers. The Genbank accession number for the pSOR3.1 sequence is L14467.

Fig. 3. The degree of divergence between *S.fimicola* and *N.crassa* *frq* is variable throughout the extended ORF. The percent identity (▲) and similarity (X) of the translation products were determined in non-overlapping blocks of 50 amino acids according to parameters of the Bestfit program (GCG package). The locations of the hallmarks within the *frq* ORF as listed in Figures 1 and 2 are noted on the x-axis.

Fig. 4. The *S.fimicola* *frq* locus contains numerous insertions and deletions in its DNA sequence relative to *N.crassa*. The DNA sequences were compared with Bestfit (gap weight = 3.0; Genetics Computer Group, 1991) and the net increase or decrease in base pairs of *Sordaria* versus *Neurospora* was calculated for blocks of 50 nt. Insertions are represented as bars above the zero line and deletions are bars extending below it. The abscissa represents the entire length of the *S.fimicola* clone, and the relative locations of homologous *N.crassa* *frq* DNA are indicated.

score = 3.1). As in *N.crassa* FRQ, the PEST sequence nearest to the C-terminus contains a hyperacidic stretch. The net charge of this domain is virtually the same in both species (*N.crassa* = -15; *S.fimicola* = -14), despite numerous amino acid substitutions. Thus, the regions in *N.crassa* FRQ which have been linked to potential functional activities by sequence analysis are maintained in *S.fimicola* FRQ.

The *S.fimicola* *frq* genomic clone contains DNA which flanks the ORF, including DNA homologous to the smaller *frq* transcript, the intergenic sequence between the two transcripts, an unusually long (~1500 bp) 5' untranslated region (UTR) preceding the ORF of the large *frq* transcript, and DNA 3' to the large ORF (See Figure 1). These *Sordaria* flanking sequences contain numerous insertions and deletions relative to the *N.crassa* DNA sequence (Figure 4; McClung *et al.*, 1989). However, the gaps are interspersed with regions of clear identity, indicating homology. In contrast to the ORF, insertions and deletions in the non-coding regions do not conform to any specific pattern in terms of size or multiples of three. Specifically, the *S.fimicola* sequence 5' to the ORF is 81% identical between *N.crassa* and *S.fimicola*, with a net increase of 166 bp in the *Sordaria*

A

<i>N. crassa</i>	71	NSSGE-----	SEPTNWFNQ	SNRNPAAAFH	DESHIMEVDP	PFYQKETDSS	114
<i>N. discreta</i>	>						
<i>N. intermedia</i>	>						
<i>G. cerealis</i>							
<i>S. fimicola</i>	D....	SNDTG	Q.D.KS..D.		V...N	N.....	
<i>N. crassa</i>	115	NEESRYPPGR	NPVHPGGVQ	LPGRFPVAAH	SSAADDYRSV	IDDLTVENKR	164
<i>N. discreta</i>							
<i>N. intermedia</i>							
<i>G. cerealis</i>							
<i>S. fimicola</i>		D....					
<i>N. crassa</i>	165	LKEELKRYKQ	FGSDVMRKEK	LFEIKVHGLP	RRKKRELEAT	LRDFAASLGD	214
<i>N. discreta</i>							
<i>N. intermedia</i>							
<i>G. cerealis</i>							
<i>S. fimicola</i>							
<i>N. crassa</i>	215/216	SS					
<i>N. intermedia</i>		..95/92					
<i>G. cerealis</i>		..94/95					
<i>S. fimicola</i>		..84/86					
<i>N. crassa</i>	311	TDKEKKLVV	RRLEQLFTGK	ISGRNMQRNQ	SMPSMDAPLA	PEG	353
<i>N. discreta</i>	>						
<i>S. fimicola</i>							
<i>N. crassa</i>	440	LDPRVQIPSEN					451
<i>N. discreta</i>	>						
<i>S. fimicola</i>							
<i>N. crassa</i>	564	RKKRKTQQAS	SEIGRFGPSR	SPSDTFHYKP	MFVHRNSSSI	ETS	606
<i>N. discreta</i>	>						
<i>S. fimicola</i>							
<i>N. crassa</i>	631	FSGSGTTQQR	RKRRYDGAIV	YYTGAPFCTD	LSGDPGDMSP	TAQMTAGR	678
<i>N. discreta</i>	>						
<i>S. fimicola</i>							

B

Fig. 5. Predicted amino acid sequences from portions of fungal *frq* homologs. (A) The predicted amino acid sequence for homologs of *frq* from *N.discreta*, *N.intermedia*, *G.cerealis* and the relevant portions of *N.crassa* and *S.fimicola* are aligned. A dot (.) indicates identity, a dash (-) symbolizes a deleted residue and the numerical quotient at the end of the homologous sequence represents the percent identity relative to *N.crassa* at the amino acid and DNA levels, respectively (aa/DNA). An arrow (>) denotes the start of a DNA sequence. (B) A schematic diagram of the *N.crassa* ORF illustrates the relative location of the putative translation products aligned in (A).

sequence over a span of 2.7 kb of *N.crassa* DNA (Bestfit, gap weight = 3.0). The region 3' of the ORF is 80% identical, with a net decrease of 59 bp in *Sordaria* over a span of 0.5 kb of *N.crassa* sequence (Bestfit, gap weight = 3.0). Thus, while the *S.fimicola* DNA which flanks the ORF is clearly homologous to the *N.crassa* *frq* locus, it is neither conserved as stringently as the ORF, as indicated by the pattern of insertions and deletions, nor is it conserved for potential translation products. It may be deduced that the remainder of the smaller *frq* transcript, and regions 3' to it, are poorly conserved relative to the large *frq* ORF, since only the single 6.3 kb band corresponding to pSOR3.1 was visualized on Southern blot hybridizations of *S.fimicola* genomic DNA probed with the entire *N.crassa* *frq* locus (data not shown).

frq homologs derived from other *Sordariaceae* family members were also analyzed, as noted above, using PCR (Figure 5). As with the *Sordaria* and *N.crassa* *frq* sequence,

Fig. 6. Predicted phylogenetic relationships based on analysis of FRQ homolog sequences shown in Figure 5A. A distance matrix was determined for these data with the PROTDIST (Dayhoff PAM) program (PHYLP3.5c, Felsenstein, 1993), and these calculations were used to infer a phylogenetic tree by the Fitch–Margoliash method (Fitch and Margoliash, 1967; FITCH in PHYLP3.5c, Felsenstein, 1993). Branch lengths are proportional to the distance determination, as indicated by the scale bar. The average percent standard deviation for this tree is 5.5. The numbers located at the nodes represent the results of bootstrap analysis of 100 replicates of the data set. That is, 86% of the time *N.intermedia* and *N.crassa* are located at the same node; 99% of the time *N.discreta* is grouped with the other two *Neurospora* species. *S.fimicola* is identified as the most distant relative 100% of the time.

continuous ORFs were identified from these DNA fragments. The predicted translation products show that *frq* homologs, in these strains, are more highly conserved relative to *N.crassa* than is the sequence from *S.fimicola*. Nevertheless, together these data demonstrate identical trends relative to *Sordaria*; that is, more or less highly conserved regions can be identified, and these correspond to similar regions identified through comparison of *N.crassa* and *S.fimicola frq*. For example, amino acids 125–145 (Figure 5, *N.crassa* sequence), which correspond to a divergent region between *S.fimicola* and *N.crassa*, show only seven residues conserved among all strains, whereas this is immediately followed by a highly conserved region (38 of 40 residues conserved among all strains).

Because these data are among the first resulting from a systematic comparison of nucleotide and protein coding sequences from a broad spectrum of the *Sordariaceae*, the sequences of the putative translation products were used to construct phylogenetic trees for these genera of *Sordariaceae*. Parsimony [PROTPARS (PHYLP3.5c) and PAUP], distance (Neighbor joining and Fitch–Margoliash in PHYLP3.5c) and UPGMA analyses (Genetics Computer Group, 1991; Swofford, 1991; Felsenstein, 1993) of the combined amino acid sequence data produced trees with the same fundamental branching topology as the tree shown in Figure 6. Fitch–Margoliash analysis of 100 bootstrapped data sets yielded a consensus tree with the identical structure. This tree is also consistent with endonuclease restriction site analysis which indicates that *N.crassa* and *N.intermedia* are closely related with respect to other heterothallic species in the same genus (Taylor and Natvig, 1988). In summary, these data confirm evolutionary relationships which were previously based primarily upon morphological or physiological criteria.

Functional analysis of the *S.fimicola frq* homolog

In order to put the sequence differences detected among the *Sordariaceae* species in a meaningful biological context we

Table I. Strains used in this study

Species	Strain	Genotype	Source
<i>N.crassa</i>	30-7	<i>band (bd) A</i>	lab strain (Sargent and Briggs, 1967)
<i>N.crassa</i>	93-4	<i>his-3; bd; frq¹⁰ A</i>	Aronson <i>et al.</i> (1994a); <i>his-3</i> crossed from Fungal Genetics Stock Center (FGSC) # 6103
<i>N.crassa</i>	93-4/pDE3ΔBH-1	<i>his-3 (his-3⁺)^{hom} pDE3ΔBH; bd; frq¹⁰ A</i>	this work
<i>N.crassa</i>	93-4/pKAJ120-6	<i>his-3 (his-3⁺ frq⁺)^{hom} pKAJ120; bd; frq¹⁰ A</i>	lab strain
<i>N.crassa</i>	93-4/pMM203-5	<i>his-3 (his-3⁺ frq⁺)^{hom} pMM203; bd; frq¹⁰ A</i>	this work
<i>N.crassa</i>	87-12	<i>his-3; bd A</i>	lab strain
<i>N.crassa</i>	87-12/pDE3ΔBH-1	<i>his-3 (his-3⁺)^{hom} pDE3ΔBH; bd A</i>	this work
<i>N.crassa</i>	87-12/pKAJ120-7	<i>his-3 (his-3⁺ frq⁺)^{hom} pKAJ120; bd A</i>	lab strain
<i>N.crassa</i>	87-12/pMM203-2	<i>his-3 (his-3⁺ frq⁺)^{hom} pMM203; bd A</i>	this work
<i>N.intermedia</i>	435 (FGSC #)	wild-type	FGSC
<i>N.intermedia</i>	1766	wild-type	FGSC
<i>N.discreta</i>	3228	wild-type	FGSC
<i>N.discreta</i>	3269	wild-type	FGSC
<i>N.sitophila</i>	2216	wild-type	FGSC
<i>G.cerealis</i>	959	wild-type	FGSC
<i>G.calospora</i>	958	wild-type	FGSC
<i>G.reticulospora</i>	960	wild-type	FGSC
<i>S.fimicola</i>	2918	wild-type	J.Taylor, UCA, Berkeley; FGSC

Fig. 7. Molecular analysis of *N. crassa frq* null mutants rescued by transformation with *S. fomicola/N. crassa* hybrid constructs. (A) Southern blot analysis of *his-3* targeted transformants. Genomic DNA was digested with *Sna*BI, electrophoresed, blotted and probed with the *Cla*I fragment of *N. crassa frq* (see B). The molecular weight of the hybridizing bands in kb, based on analysis of standards, is shown on the left. Strain designations are as follows (also see Table I): 93-4 is the null *frq* mutant (*frq*¹⁰) crossed into the *his-3* background; 93-4/pDE3ΔBH has been rescued for histidine prototrophy by homologous recombination of a portion of the wild-type *his-3* gene at the *his-3* locus; 93-4/pKAJ120-6 has been rescued for histidine prototrophy with a plasmid containing *N. crassa frq* in addition to the *his-3* gene; and 93-4/pMM203 has been rescued with a plasmid containing a chimeric construct consisting of the *N. crassa frq* locus, with the *S. fomicola frq* extended ORF replacing the homologous *N. crassa* region (see Figure 1). (B) The *frq* locus on linkage group VIIR in *frq*⁺ is shown as the genomic *Cla*I fragment. The null mutant *frq*¹⁰ strain was constructed by insertion of a hygromycin-resistance cassette into the region which includes the promoter and almost all of the large *frq* transcript (Aronson *et al.*, 1994c). Maps of the *his-3* locus with either *N. crassa frq* (pKAJ120) or the chimeric *frq* (pMM203) transformed by homologous recombination are shown schematically on linkage group I. Restriction enzymes listed include: Sn, *Sna*BI; Cl, *Cla*I; Bg, *Bgl*II.

carried out studies designed to assess the ability of the *S. fomicola* gene to functionally complement *N. crassa frq* mutants. In the construct pMM203 the *S. fomicola* ORF was fused in-frame to the first 11 amino acids of the *N. crassa* FRQ ORF. The protein coding region of this chimeric construct is thus under the regulation of the *N. crassa* promoter and 5' untranslated sequences (Figure 1).

The fusion construct was made so as to permit targeted insertion of the uninterrupted fusion gene into the *his-3* locus

of *N. crassa*. The parent vector (pKAJ120, containing wild-type *N. crassa frq*) or the chimeric construct, pMM203, were transformed into the recessive null *frq* mutants, *his-3, frq*¹⁰ (strain 93-4, see Table I) or *his-3, frq*⁹ (data not shown). To confirm correct insertion, genomic DNA was prepared for Southern analysis from transformants which rescued histidine prototrophy (Figure 7). The endogenous genomic *frq*¹⁰ bands are observed at 1.2 and 3.8 kb, and appropriate insertion at the *his-3* locus was verified for *N. crassa frq* (pKAJ120) by observation of a 14 kb band following a *Sna*BI digest (lane 3, 93-4/pKAJ120-6). The *Sordaria frq* ORF contains a *Sna*BI restriction site not present in *N. crassa*, resulting in bands of 7.9 and 6.2 kb in pMM203 transformed strains (lane 4, 93-4/pMM203-5). Of 32 transformants in *frq*⁹ and *frq*¹⁰ strains, 17 showed the correct insertion of *frq* at the *his-3* locus. Appropriately targeted transformants of pMM203 and the control pKAJ120 were analyzed for clock function.

As is typical for *his-3* targeted transformation of *frq*⁺ into *frq* null mutants, not all transformants were rescued for conidiation and banding. Specifically, 60% of pKAJ120 transformants (Johnson, 1993) and 50% of the pMM203 transformants displayed circadian regulation of conidiation. Data from two separate experiments is combined and presented in Table II. The wild-type strain, *band A*, was determined to have a period length of 21.6 h. When the null mutant *frq*¹⁰ is rescued for histidine prototrophy with vector alone (Figure 8, 93-4/pDE3ΔBH) scarce conidiation and no evidence of circadian banding are observed. When *N. crassa frq* is transformed into *frq*¹⁰ the transformant (93-4/pKAJ120-6) displays a normal period length of 21.3 h. The observed conidial banding pattern is less discrete than that seen in wild-type strains, a trait which is characteristic of the recessive *frq* mutant strains following transformation and rescue. Transformation of the null mutant with the chimeric pMM203 also rescues conidiation and banding (for example, 93-4/pMM203-5, Table II and Figure 8), albeit with a decrease in abundance of conidia, somewhat greater variability in period length, and a significantly (>99% confidence level) shorter period length (19.9 h, *z* value = -5.7; Devore, 1987).

In order to see whether an additional copy of the *S. fomicola frq* gene could interfere with normal clock function in *Neurospora*, the pMM203 chimeric construct was transformed into strains with a *frq*⁺ background (87-12 series, see Table II). While a second copy of endogenous *frq* has no effect on the wild-type period length, the *S. fomicola frq* gene results in a shortened (>99% confidence level) period (22.4 versus 23.3 h, *z* value = 6.8; Devore, 1987).

Discussion

Circadian biological clocks are a widely characterized phenomenon found throughout nature. However, the identity of a bona fide component of the oscillator has only recently been established (Aronson *et al.*, 1994b). Little is known regarding the mechanistic process that both generates and maintains circadian rhythmicity. Although *frq* encodes an integral component of the clock, as demonstrated by the effect of single amino acid substitutions on period length (Dunlap, 1993; Aronson *et al.*, 1994a), it has thus far been

Table II. Period length of transformant strains

Strain ^a	τ^b	SD	Range	n^c	Phased ^d
bdA	21.6	0.4	21.3–22.2	68	1.4
93-4/pKAJ120-6	21.3	1.0	19.4–23.2	55	22.9
93-4/pMM203-5	19.3	1.61	6.3–22.1	38	22.3
93-4/pMM203(all) ^e	19.9	1.9	16.3–24.1	84	23.7
87-12/pDE3ΔBH	23.2	0.5	22.5–23.9	43	23.0
87-12/pKAJ120-7	23.3	0.4	22.5–24.0	58	1.4
87-12/pMM203-2	22.4	0.9	21.1–23.2	55	0.4

^aSee Table I for strain information. 93-4 derived strains are transformants into the *frq*¹⁰ background and 87-12 derived strains are transformants into the *frq*⁺ background.

^b τ = average period length between conidial bands expressed in hours. See Materials and methods for a description of calculations.

^c n is the total number of period lengths on which calculations were based.

^dThe phase represents the time at which the first band occurs following entrainment.

^eAll transformants tested in two separate experiments (a total of five separate strains) were pooled for these calculations.

impossible to assign a discrete biochemical function to this gene or its product. Thus, analysis of its domain structure, in conjunction with the identification of what portions of the protein are phylogenetically conserved, may indicate functional components. To this end, we have cloned and/or sequenced part or all of the *frq* DNA encoding the ORF from four species taken from three genera of the *Sordariaceae*, and additional non-coding sequences from one of those species. Sequence comparisons indicate that the overall gene structure including the lack of introns is conserved. All of the previously identified molecular hallmarks of FRQ including a putative NLS, PEST sequence and hyperacidic domain are conserved, as were all of the mutation sites in the existing *frq* alleles. Finally, we have shown that the ORF from the most distantly related *frq* homolog among the *Sordariaceae*, that from *S.fimicola*, is capable of rescuing a circadian conidial banding rhythm in a *N.crassa frq* null strain.

When the entire length of the homologous clone is examined the differential conservation and integrity of the *S.fimicola frq* DNA sequence (Figure 4 and Results) supports the central importance of the protein encoded by the ORF within the large *frq* transcript of *N.crassa*. Identification of the location of all of the allelic *frq* mutations within this ORF (Aronson *et al.*, 1994a, and see Figure 1 and 2) further bolsters this hypothesis. Not only is the ORF the most highly conserved segment, but the area surrounding the *frq*¹ and *frq*⁷ mutations is the most highly conserved region within the ORF, highlighting its importance in clock function.

The several motifs which have been identified within the ORF by their sequence are conserved in the homologous genes, despite significant DNA and amino acid sequence variability. The putative nuclear localization signal suggests that *frq* may be sequestered in the nuclear compartment, and experiments are in progress in this lab to identify FRQ in nuclear fractions. The only other clock gene cloned, *period* (*per*) from *Drosophila melanogaster*, has been shown to accumulate as a protein product in the nucleus (Liu *et al.*, 1992). One theoretical mode of activity of such a nuclear protein which is involved in regulating a cascade of events at the intracellular level is that of transcriptional activator. Transcriptional activators are a diverse group of proteins, many of which have a hyperacidic region (Ptashne, 1988) similar to that identified and conserved in FRQ. It is interesting to note that this is one of the lesser conserved

Fig. 8. The *S.fimicola frq* ORF restores conidial banding in the *N.crassa* null *frq* mutant, *frq*¹⁰. *his-3* targeted transformants were analyzed on race tubes (see Material and methods) for circadian conidial banding period lengths. band A is the wild-type strain; 93-4/pDE3ΔBH-1 is *frq*¹⁰ rescued only for histidine prototrophy; 93-4/pKAJ120-6 is transformed with *N.crassa frq*; 93-4/pMM203-5 is transformed with the chimeric construct which contains the *S.fimicola frq* ORF.

regions of the ORF amino acid sequence by simple identity (78%), however, the similarity quotient is very high (93%), reflecting the numerous conservative changes (see Figures 2 and 3, amino acids 868–894, *S.fimicola* sequence). Conversely, we have previously noted the presence of numerous potential phosphorylation sites within *frq* (McClung *et al.*, 1989), several of which are not conserved within these sequences. Also, similarities in FRQ corresponding to the TG repeat region of the *D.melanogaster per* gene are not strictly conserved, although the amino acid substitutions in this region (such as TG to TV or TA) reflect conservative substitutions (see Figure 2, amino acids 389–436 *S.fimicola* sequence). Thus, the *S.fimicola frq* homolog is sufficiently diverged from *N.crassa frq* to illustrate contrasting phylogenetic conservation, information which will continue to be utilized as biochemical characterization of FRQ proceeds.

To our knowledge this is the first gene cloned from *S.fimicola*, and the second from any *Sordaria* species. The *ura5* gene was cloned from *S.macrospora* (Le Chevanton, 1989) and analysis of the putative ORF revealed similarities to *N.crassa* codon usage in highly expressed genes [that is, a bias against purines, especially A, in the third position (Orbach *et al.*, 1986; Chary *et al.*, 1990)]. *frq* mRNA, however, rather than being abundant, can be expressed weakly under conditions where the clock is robust and active, and FRQ protein, similarly, appears to be scarce under all

culture conditions (N.Garceau and J.C.Dunlap, unpublished data). Consistent with this, *N.crassa* and *S.fimicola* *frq* ORFs utilize pyrimidines (57 and 53%, respectively) and purines (43 and 47%, respectively) at approximately equal frequencies in the third position. Thus, the codon usage in the *S.fimicola* ORF is typical of a weakly expressed gene (Bennetzen *et al.*, 1982).

The most important finding of these experiments is the functional complementation of an *N.crassa* *frq* null strain by the *S.fimicola* *frq* homolog. The rescuing properties of *S.fimicola* *frq* are interesting, especially when considering the life cycle of *S.fimicola*, a fungus which does not engage in conidium production. This coprophilic fungus is homothallic, and utilizes ascospores for reproduction (Ingold, 1971). However, cultures at the appropriate developmental stage (~12–18 days) release ascospores in a manner which appears to be regulated by an endogenous circadian rhythm (Austen, 1968). Ascospore release in *Neurospora* species also appears to be directed by the circadian clock [S.Brody, unpublished data cited in Lakin-Thomas *et al.* (1990) and R.Brambl, personal communication]. Conidiation in *N.crassa* occurs at circadian time 0 (CT 0, subjective dawn), while ascospore release in both *Neurospora* and *S.fimicola* appears to peak near CT 6 (subjective noon). Similarities between the rhythmic ascospore release of *S.fimicola* and the conidial banding pattern of *N.crassa* are (i) they are blue light inducible and are not affected by red light (Ingold and Dring, 1957) and (ii) they are diurnal. Although precise calculations are difficult, it appears that the period length of ascospore release by *S.fimicola* in constant darkness is ~24 h (Austen, 1968).

The qualitative appearance of conidial banding in null strains transformed with *S.fimicola* *frq* was generally less distinct than that seen in strains rescued with *N.crassa* *frq*; nevertheless, the complementation is impressive since a single amino acid substitution within *frq* can change the period of the *N.crassa* clock from 21.5 to 16 or 29 h (Dunlap, 1993). Here, the chimeric *frq* ORF is altered at 15% of its amino acids and still regulates circadian banding in the null *frq* mutant, with a mean period length of 19.9 h, versus 21.6 h for transformants containing a single copy of the *N.crassa* *frq* allele. Thus, the heterologous *frq* gene must be capable of integration with components of the *Neurospora* circadian clock pathway. However, the greater variation in period lengths (Figure 8 and Table II) may indicate deficiencies in the ability to do so efficiently. The observation that period lengths are similarly shortened in *frq*¹⁰ and *frq*⁺ strains which are transformed with the chimeric construct, versus the *N.crassa* *frq* construct, suggests that the heterologous gene is co-dominant relative to endogenous *frq*, and that *S.fimicola* *frq* is behaving genetically as a short period length *frq* allele.

Molecular analysis of phylogenetic conservation of a circadian clock gene was previously applied to the *D.melanogaster* gene *per*, where there are also specifically conserved domains within the coding region (Colot *et al.*, 1988; Thackeray and Kyriacou, 1990; Peixoto *et al.*, 1993). The long period mutation *per*^{LI} occurs in a region conserved among drosophilids and possibly involved in protein–protein interactions (Huang *et al.*, 1993). A short period mutation, *per*^S, also occurs within a highly conserved region in which mutations typically result in short period lengths (Baylies *et al.*, 1992; Rutila *et al.*, 1992). In

experiments similar to those reported here, the ability of *per* homologs to rescue a *D.melanogaster* null mutant (*per*⁰¹) was assayed using homologs from *D.simulans* (Hall *et al.*, 1992) and *D.pseudoobscura* (Petersen *et al.*, 1988). The well conserved (compared with *D.melanogaster*) *D.simulans* gene efficiently restored circadian locomotor activity in the *D.melanogaster* arrhythmic mutant. However, despite the existence of a well characterized circadian clock in *D.pseudoobscura* ($\tau = 23.3$ h), this less well conserved *per* homolog rescued rhythmicity with a long period (~29 h) when one or two copies of the gene were integrated into the genome of *per*⁰¹ flies (Petersen *et al.*, 1988). The period length of these transformants thus resembled neither *D.melanogaster* nor *D.pseudoobscura*. In these experiments, endogenous (*D.melanogaster*) 5' regulatory regions were included to drive expression of the transgene, and despite the transfer of some species-specific behavioral characteristics regulated by *per*, the penetrance of the chimeric gene in rescuing circadian rhythmicity was low (9 and 24% for one or two gene doses, respectively); this suggests difficulty in integrating with the rest of the *D.melanogaster* clock apparatus. In contrast to the *S.fimicola* *frq* homolog, this *per* homolog is interspersed with blocks of non-conserved DNA (Colot *et al.*, 1988), a factor that may have contributed to inefficient circadian phenotypic rescue (Petersen *et al.*, 1988; Wheeler *et al.*, 1991).

Finally, the choice of *S.fimicola* as a species from which to clone a homolog was based on practical considerations resulting from preliminary data acquisition. This was necessary in part because there has been little data published regarding the phylogeny of the filamentous fungi. Our phylogenetic studies, which include *Neurospora*, *Gelasinospora* and *Sordaria* species, commenced as Southern blot hybridizations, which were then supplemented with DNA sequence analysis, making quantitative determinations possible. Restriction fragment length polymorphisms (data not shown) and inspection of DNA sequence and phylogenetic trees confirm, on a molecular level, the expectation that the *Neurospora* species are more closely related to each other than to *Gelasinospora*, and that *Sordaria* is further diverged. These data are presented as a basis for expanding the phylogenetic relationships among the *Sordariaceae*. Further, they suggest that *frq* is a clock component which may be utilized by a biologically diverse collection of organisms.

The intergeneric complementation of the clock phenotype reported here demonstrates that *frq* is not a developmental gene strictly involved in asexual development in *Neurospora*. Rather, *frq* must be a central, control gene that is essential to the operation of a circadian clock that can be used to initiate diverse developmental pathways. Taken together with experiments which demonstrate that FRQ determines the phase of the *Neurospora* circadian clock (Aronson *et al.*, 1994b) we conclude that this protein product is a component of the oscillator which constitutes a circadian biological clock.

Materials and methods

Strains and media

The strains used in this study and their sources are listed in Table I. The *band* gene facilitates visualization of the circadian conidiation phenotype of wild-type *N.crassa*, but does not alter the basic properties of the clock

itself. All *N. crassa* strains used in these studies contain the *band* mutation. In addition, *frq*⁺ and *frq* mutant strains were crossed with *N. crassa his-3* (FGSC #462) to facilitate targeted integration using the pDE3 vector (Ebbole, 1990). The null strain (*frq*⁰) was constructed in this lab (Aronson, 1994a; and general method described in Aronson *et al.*, 1994c) by deletion of the two central *Bgl*III fragments of a genomic DNA *frq* clone (see Figure 7). A hygromycin resistance cassette (Staben *et al.*, 1989) was inserted and used as an aid in selecting for homologous recombination events. This disrupted *frq* allele behaves as a normal genetic marker in crosses. *frq*⁰ was crossed with a *his-3* strain (FGSC #6103) to create 93-4 for use as a recipient for plasmids such as those described below.

Media supporting vegetative growth included minimal media: 1×Vogel's salts, 1.5% sucrose, 1.5% agar (Davis and deSerres, 1970), with or without 0.5 mg/ml l-histidine (Sigma Chemical). Race tubes contained 1×Vogel's, 0.1% glucose, 0.17% arginine, 1.5% agar, with the pH adjusted to 6.0 (West, 1975). Spheroplasts were prepared for transformation of plasmid DNA according to Vollmer and Yanofsky (1986), and histidine prototrophs were selected.

Neurospora mycelial tissue was propagated for DNA minipreps by inoculating conidia in 5 ml Horowitz Complete medium (Davis and deSerres, 1970). Cultures were incubated for 2 days at 30°C with shaking. DNA was prepared from *Sordaria* by germinating ascospores in 1×Vogel's supplemented with 2.4% yeast extract and 1.2% tryptone. Cultures were incubated as for DNA minipreps.

DNA preparation

Genomic DNA purified on a cesium chloride gradient was generously supplied by D. Natvig (University of New Mexico) for the strains *N. intermedia* (FGSC #435), *N. discreta* (#3228) and *G. cerealis* (#959). Genomic DNA from *S. fimicola* was purified using the CTAB method (Zolan and Pukkila, 1986; Taylor and Natvig, 1987). PCR amplification of genomic DNA was utilized to generate fragments for sequencing (Ausubel *et al.*, 1989). Specific conditions were: 1.5 μM Mg²⁺, 40 pM of each primer, an annealing temperature of 45°C, for 35 cycles. Primers used in PCR were complementary to sequences found within *N. crassa frq*, and included the following: BDA1 5'-ACTGGGAGTGACCGAAT-3'; BDA2 5'-AACAAGCGCTTAAGGA-3'; BDA3 5'-CCCAAAGGTAGAA-AAC-3'; BDA4 5'-GGACTGGAAGCGGAGAT-3'; BDA5 5'-GGTCG-GAAAATCCGCTG-3'; BDA17 5'-CTCCTCTGCGATGTGCGATTC-3'; MM1 5'-GCTGCAGCACATGTTCAAC-3'; (Molecular Genetics Center at Dartmouth Medical School). PCR products were purified either by GeneClean (Bio 101) or by spinning agarose gel slices over siliconized glass wool and concentrating the eluate by ethanol precipitation. Such fragments were not cloned; sequence was determined for the population of fragments, thus avoiding the requirement for sequencing multiple samples. Additionally, >50% of the DNA sequence generated from PCR products was verified by determination from independent PCR experiments. Genomic DNA for Southern analysis was obtained from minipreps, which were performed according to Yarden *et al.* (1992; see Total RNA method), modified by an incubation with 25 μg/ml RNase A. Single-stranded DNA was prepared for sequencing by standard methods (Ausubel *et al.*, 1989).

DNA sequencing

DNA sequencing by the dideoxy chain termination method using SequenaseII was carried out according to manufacturer's instructions (US Biochemical). Where double-stranded DNA was utilized as a template, alkali treatment was used for denaturation, except in the case of PCR products. PCR products were heat denatured and snap frozen (Ausubel *et al.*, 1989). Primers used for sequencing were all specific for *N. crassa frq* sequence, except in the case of the *S. fimicola frq* homolog, pSOR3.1, where a number of *Sordaria*-specific primers were synthesized (Molecular Genetics Center at Dartmouth Medical School and DNA Express). DNA sequence was recorded with the aid of the computer program DNA Parrot (T & T Research). DNA sequence of pSOR3.1 and PCR products corresponding to *N. intermedia*, *N. discreta* and *G. cerealis frq* have been submitted to Genbank (pSOR3.1, L14467; *N. intermedia*, L23950, L23951; *N. discreta*, L23952, L23953, L23954, L23955, L23956; *G. cerealis*, L23948, L23949).

Southern blot analysis, cloning and subcloning

Restriction enzymes were used according to manufacturer's instructions (New England Biolabs). Southern blot analysis was performed according to standard methods (Maniatis *et al.*, 1982). DNA probes were prepared by random oligonucleotide priming (Feinberg and Vogelstein, 1983).

To facilitate cloning the *S. fimicola frq* homolog, a genomic DNA library was constructed by digestion of *S. fimicola* DNA with *Bgl*III, *Xba*I and *Ssp*I followed by separation on a 1% agarose gel (SeaKem). DNA fragments included in the 6–9 kb range were isolated and ligated to pBluescriptII

SK[−] (Stratagene) which had been digested with *Bam*HI and *Xba*I. Blue/white colony selection was utilized and colony lifts were probed with a 1.0 kb *Eco*RI fragment from the *frq* ORF (purified from pCRM134).

Subcloning was achieved using standard methods (Sambrook *et al.*, 1989). The vector for *his-3* targeting, pDE3ΔBH, was created from pDE3 by excising the *Bam*HI to *Hind*III fragment containing the *lacZ* reporter gene (K.A. Johnson and J.C. Dunlap, unpublished data). pKAJ122 contains *N. crassa frq* genomic DNA from the *Mlu*I (near the 5' end of the small *frq* transcript) to *Clal* (3' of the stop codon for the large *frq* ORF) restriction enzyme sites (positions of relevant restriction sites are noted in Figure 1) cloned into pDE3ΔBH. The large *frq* ORF from pSOR3.1 was exchanged into pKAJ122 utilizing the *Sph*I and *Age*I sites, to produce pMM103. This intermediate plasmid consists of the *his-3* targeting vector, pDE3DBH, *N. crassa frq* sequence from ~180 bp upstream of the 5'UTR of the *frq* transcript through the *Sph*I site (33 bp downstream of the ATG), *S. fimicola frq* DNA through the ORF to the *Age*I site (~100 bp outside the stop codon), followed by *N. crassa frq* flanking DNA to the *Clal* site (0.5 kb). The ORF from this construct was then cloned into a full-length *frq* construct pKAJ120 utilizing the *Sph*I and *Bss*HII sites, to produce the chimeric construct for *his-3* targeting, pMM203 (Figure 1).

Computer analysis

DNA sequence was analyzed with a variety of programs included in the University of Wisconsin Genetics Computing Group package including Bestfit, Pileup, Plotdata and Publish (Genetics Computer Group, 1991). Gap weight scores for Bestfit analysis reflect relative stringency in assignment of gaps in the homologous sequences. For DNA sequence cited in the text a gap weight of 5.0 was routinely used, except where insertions, deletions or poor homology required a lower value. The Pileup program constructs trees based on the UPGMA algorithm (Sneath and Sokal, 1973). Other phylogeny programs used include PHYLIP3.5c (Felsenstein, 1993) and PAUP (Swofford, 1991). Additional programs used include Genetics Construction Kit (TextCo), DNA Inspector (TextCo), DNA Strider (C. Marck) and MacPattern (R. Fuchs, EMBL Data Library).

Functional (race tube) analysis of the *Neurospora circadian biological clock*

Race tubes were inoculated with conidia and incubated in constant light at room temperature (~25°C) for 24 h. The tubes were placed in constant darkness at 25°C and the growth fronts were marked at that time and every 24 h thereafter. Conidial bands were noted and the period length was calculated with the aid of a computer program as previously described (McClung *et al.*, 1989) and a digitizer board (Kurta Corp.). Period lengths are expressed as the average number of hours which elapse between conidial bands. Period lengths and standard deviations were calculated based on averaged data for individual race tubes, which are pooled for each strain, or where indicated, for groups of transformants. Statistical analyses (z tests) were performed as described in Devore (1987).

Acknowledgements

We thank D. Natvig and J. Taylor for materials and helpful advice. D. Natvig for his gift of genomic DNAs from *N. intermedia*, *N. sitophila*, *N. discreta*, *G. cerealis*, *G. calospora* and *G. reticulospora* and J. Taylor for the *S. fimicola* strain; J. Felsenstein and W. Fitch for their suggestions on phylogenetic analysis; C. P. Kyriacou and members of our laboratory for critical reading of the manuscript, and the latter for constructs and transformants of control strains. This work was supported by NIH grant GM34985 to J. C. Dunlap and by core grant support to the Norris Cotton Center at Dartmouth Medical School.

References

- Aronson, B.D., Johnson, K.A. and Dunlap, J.C. (1994a) *Proc. Natl Acad. Sci. USA*, in press.
- Aronson, B.D., Johnson, K.A., Loros, J.J. and Dunlap, J.C. (1994b) *Science*, **263**, 1578–1584.
- Aronson, B.D., Lindgren, K.M., Dunlap, J.C. and Loros, J.J. (1994c) *Mol. Gen. Genet.*, **263**, 1578–1584.
- Austen, B. (1968) *Ann. Bot.*, **32**, 261–278.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1989) *Current Protocols in Molecular Biology*. John Wiley and Sons, New York.
- Baylies, M.K., Voshall, L.B., Sehgal, A. and Young, M.W. (1992) *Neuron*, **9**, 575–581.
- Bennetzen, J.L. and Hall, B.D. (1982) *J. Biol. Chem.*, **257**, 3026–3031.

- Chary,P., Hallewell,R.A. and Natvig,D.O. (1990) *J. Biol. Chem.*, **265**, 18961–18967.
- Colot,H.V., Hall,J.C. and Rosbash,M. (1988) *EMBO J.*, **7**, 3929–3937.
- Davis,R.L. and DeSerres,D. (1970) *Methods Enzymol.*, **27A**, 79–143.
- Devore,J.L. (1987) *Probability and Statistics for Engineering and the Sciences*. 2nd edn. Belmont, CA.
- Dunlap,J.C. (1993) *Annu. Rev. Physiol.*, **55**, 683–728.
- Ebbola,D. (1990) *Fung. Genet. News.*, **37**, 15–16.
- Edmunds,L.E., Jr (1988) *Cellular and Molecular Bases of Biological Clocks*. Springer-Verlag, New York.
- Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **137**, 266–269.
- Felsenstein,J. (1993) *PHYLIP* (Phylogeny Inference Package). Version 3.5c. Department of Genetics, University of Washington, Seattle.
- Fitch,W.M. and Margoliash,E. (1967) *Science*, **155**, 279–284.
- Genetics Computer Group (1991) *Program Manual for the GCG Package*. GCG, 575 Science Drive, Madison, WI 53711, USA.
- Hall,J.C. and Rosbash,M. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 5382–5383.
- Hall,J.C., Hamblen-Coyle,M.C., Moroz,L., Rutila,J.E., Yu,Q., Rosbash,M. and Wheeler,D.A. (1992) *Drosophila Inf. Serv.*, **71**, 204–207.
- Huang,Z.J., Edery,I. and Rosbash,M. (1993) *Nature*, **364**, 259–262.
- Ingold,C.T. (1971) *Fungal Spores: Their Liberation and Dispersal*. Clarendon Press, Oxford.
- Ingold,C.T. and Dring,V.J. (1957) *Ann. Botany*, **21**, 465–477.
- Johnson,K.A. (1993) Ph.D. Thesis, Dartmouth Medical School.
- Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8147.
- Lakin-Thomas,P., Cote,G. and Brody,S. (1990) *CRC Crit. Rev. Microbiol.*, **17**, 365–416.
- Le Chevanton,L. (1989) *Gene*, **77**, 39–49.
- Liu,X., Zweibel,L.J., Hinton,D., Benzer,S., Hall,J.C. and Rosbash,M.R. (1992) *J. Neurosci.*, **12**, 2735–2744.
- Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- McClung,C.R., Fox,B.A. and Dunlap,J.C. (1989) *Nature*, **339**, 558–562.
- Orbach,M.J., Porro,E.A. and Yanofsky,C. (1986) *Mol. Cell. Biol.*, **6**, 2452–2461.
- Peixoto,A.A., Campesan,S., Costa,R. and Kyriacou,C.P. (1993) *Mol. Biol. Evol.*, **10**, 127–139.
- Petersen,G., Hall,J.C. and Rosbash,M. (1988) *EMBO J.*, **7**, 3939–3947.
- Pittendrigh,C.S., Bruce,V.G., Rosenzweig,N.S. and Rubin,M.L. (1959) *Nature*, **184**, 169–170.
- Ptashne,M. (1988) *Nature*, **335**, 683–689.
- Rechsteiner,M. (1990) *Semin. Cell Biol.*, **1**, 433–440.
- Roenneberg,T. and Morse,D. (1993) *Nature*, **362**, 362–364.
- Rogers,S., Wells,R. and Rechsteiner,M. (1986) *Science*, **234**, 364–368.
- Rutilla,J.E., Edery,I., Hall,J.C. and Rosbash,M. (1992) *J. Neurogenet.*, **8**, 101–113.
- Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sargent,M.L. and Briggs,W.R. (1967) *Plant Physiol.*, **42**, 1504–1510.
- Sargent,M.L., Briggs,W.R. and Woodward,D.O. (1966) *Plant Physiol.*, **41**, 1343–1349.
- Sikorski,R.S., Boguski,M.S., Goebel,M. and Hieter,P. (1990) *Cell*, **60**, 307–317.
- Sneath,P.H.A. and Sokal,R.R. (1973) *Numerical Taxonomy*. W.H. Freeman and Co., San Francisco, CA.
- Staben,C., Jensen,B., Singer,M., Pollock,J., Schechtman,M., Kinsey,J. and Selker,E.U. (1989) *Fungal Genet. Newslett.*, **36**, 79–81.
- Swofford,D.L. (1991) *PAUP: Phylogenetic Analysis Using Parsimony*, Version 3.0s. Illinois Natural History Survey, Champaign, IL.
- Taylor,J.W. and Natvig,D.O. (1987) *Isolation of Fungal DNA*. Southeastern Publishing Co., Athens, GA.
- Taylor,J.W. and Natvig,D.O. (1988) *Myc. Res.*, **93**, 257–272.
- Thackeray,J.R. and Kyriacou,C.P. (1990) *J. Mol. Evol.*, **31**, 389–401.
- Vollmer,S.J. and Yanofsky,C. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 4869–4873.
- West,D.J. (1975) *J. Bacteriol.*, **123**, 387–389.
- Wheeler,D.A., Kyriacou,C.P., Greenacre,M.L., Yu,Q., Rutila,J.E., Rosbash,M. and Hall,J.C. (1991) *Science*, **251**, 1082–1085.
- Yarden,O., Plamann,M., Ebbola,D.J. and Yanofsky,C. (1992) *EMBO J.*, **11**, 2159–2166.
- Zolan,M.E. and Pukkila,P.J. (1986) *Mol. Cell. Biol.*, **6**, 195–200.

Received on December 22, 1993; revised on February 24, 1994