The EMBO Journal vol.11 no.3 pp.1145-1153, 1992

Demonstration of retrotransposition of the Tf1 element in fission yeast

Henry L.Levin and Jef D.Boeke[1]

Department of Molecular Biology and Genetics, The Johns Hopkins School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205, USA

[1]Corresponding author

Communicated by D.Nathans

Tf1, a retrotransposon from fission yeast, has LTRs and coding sequences resembling the protease, reverse transcriptase and integrase domains of retroviral *pol* genes. A unique aspect of Tf1 is that it contains a single open reading frame whereas other retroviruses and retrotransposons usually possess two or more open reading frames. To determine whether Tf1 can transpose, we overproduced Tf1 transcripts encoded by a plasmid copy of the element marked with a *neo* gene. Approximately 0.1-4.0% of the cell population acquired chromosomally inherited resistance to G418. DNA blot analysis demonstrated that such strains had acquired both Tf1 and *neo* specific sequences within a restriction fragment of the same size; the size of this restriction fragment varied between different isolates. Structural analysis of the cloned DNA flanking the Tf1-*neo* element of two transposition candidates with the same regions in the parent strain showed that the ability to grow on G418 was due to transposition of Tf1-*neo* and not other types of recombination events.

Key words: fbp promoter/retrotransposition/retrotransposon/*Schizosaccharomyces pombe*/Tf1

Introduction

Retrotransposons have been isolated from a wide variety of eukaryotic hosts including human, mouse, *Drosophila, Arabidopsis, Neurospora* and *Saccharomyces cerevisae*. For only a few of these is there direct evidence that a given transposon copy is transpositionally active. A major limitation in studying retrotransposition is the lack of assay systems; this has until now been overcome in only three host systems, *S.cerevisiae* (Boeke *et al.*, 1985, 1988; Hansen and Sandmeyer, 1990; Curcio and Garfinkel, 1991), *Drosophila* (Jensen and Heidmann, 1991; Pelisson *et al.*, 1991) and mouse (Heidmann *et al.*, 1988; Heidmann and Heidmann, 1991). In metazoans like *Drosophila* and mouse, only relatively small numbers of transposition events can be conveniently studied, whereas transposition in microorganisms like fungi can be studied with much higher sensitivity simply because much larger populations can be manipulated.

We recently reported the isolation of retrotransposons Tf1 and Tf2 from *Schizosaccharomyces pombe*. As is true of most retrotransposons, Tf1 was cloned on the basis of its repetitive nature and there was little evidence that the cloned copies were functional. Tf1-107 has 385 base long terminal repeats (LTRs) and predicted protein sequences that resemble the protease (PR), reverse transcriptase (RT) and integrase (IN) domains of retroviruses. Comparison of the amino acid sequences of Tf1 and other RTs revealed that Tf1 is clearly a member of the *gypsy* family, which contains several *Drosophila* elements as well as Ty3 from *S.cerevisiae* (Levin *et al.*, 1990). In general, *gypsy* family elements exhibit some level of target site specificity. Examination of sequences flanking *Drosophila gypsy* elements such as *gypsy* itself, *297* and *17.6* revealed a consensus target site sequence (reviewed by Sandmeyer *et al.*, 1990). Ty3 has a unique non-sequence specific form of target site selection in that all transposition events occur at pol III promoters (Chalker and Sandmeyer, 1990; Sandmeyer *et al.*, 1990). Tf1, however, has no obvious insertion sequence preference or pattern of integration at pol III promoters.

The chromosomal locations and total numbers of Tf1 differ greatly in various isolates of *S.pombe*. Tf2 has been observed in the genomes of all wild-type and laboratory strains examined. Although there are no endogenous copies of Tf1 in the derivatives of the common laboratory strains 975 or 972, 20 to 40 copies exist in other wild-type isolates such as NCYC 132 (Levin *et al.*, 1990).

The nucleotide sequence of Tf1-107 revealed a unique property of Tf1: it contains a single open reading frame (ORF) of 1340 amino acids (Levin *et al.*, 1990). Other LTR-containing retrotransposons and retroviruses typically contain two or more ORFs whose products are expressed at different levels. The first ORF, *gag*, produces abundant virion structural components, whereas the second ORF, *pol*, produces smaller amounts of enzymatic components like RT and IN. One of the few retrotransposons with a single ORF, *copia*, uses splicing to produce different levels of *gag* and *gag-pol* products (Miller *et al.*, 1989; Brierley and Flavell, 1990; Yoshioka *et al.*, 1990). Recently, we have obtained direct evidence that Tf1, in contrast to all other retrotransposons studied, produces a single primary translation product (D.C.Weaver, H.L.Levin and J.D.Boeke, in preparation). A second unusual feature of Tf1 is that there is no obvious candidate for a primer, such as a tRNA, identifiable by a binding site just downstream of the 5′ LTR, where complementarity to the CCA 3′ terminus of tRNAs is usually present. Because of the unusual structural features of Tf1-107, the possibility exists that this element was not a functional copy, as is often the case with multicopy eukaryotic transposons (Boeke *et al.*, 1988; Hansen and Sandmeyer, 1990; Voytas *et al.*, 1990).

In order to determine whether Tf1-107 was an active or an inactive transposon copy, we sought evidence for its transposition. Initially, strains related to 975 (lacking Tf1) were transformed with high copy number plasmids that contained Tf1. Following a growth period and plasmid segregation, genomic DNA blotting of random colonies was carried out to identify transposition events. Because no transposition events were observed with this low sensitivity

technique (data not shown), we developed an assay that can be used to screen large numbers of candidate strains efficiently.

In previous studies, high level promoter fusions to the *S.cerevisiae* retrotransposon Ty1 increased the frequency of transposition by as much as 50- to 200-fold (Boeke *et al.*, 1985). To assay for transposition activity of Tf1, we fused the *S.pombe fbp* promoter to Tf1-107, such that Tf1 transcripts of native structure were overproduced. This plasmid fusion was marked with the *neo* gene just downstream of the Tf1 ORF. Potential transposition events were detected by testing colonies for chromosomally encoded resistance to G418 conferred by *neo*. After strains were exposed to a defined growth regimen, ~0.1–4% of the cells acquired such resistance to G418. Analysis of genomic DNA demonstrated that each candidate transposition strain had gained a Tf1–*neo* within a different fragment of genomic DNA. The Tf1–*neo* genomic sequences representing two candidate transposition events were cloned and the corresponding sequences flanking the insertion were determined. Sequence analysis of the target sites in the parent strains demonstrated that these candidates represent true transposition events and not other forms of recombination.

Results

Mapping the 5′ end of Tf1 transcripts

To determine the exact site of Tf1 transcriptional initiation, we compared the results of primer extension reactions with those of nuclease S1 protection assays. Figure 1 (lane 2) shows the products of a primer extension reaction with oligo JB54.3 (ATCACAAGAGTTCAGTTA) that hybridized to Tf1 transcripts just downstream of the 5′ LTR. The exact base of the 5′ end of the Tf1 transcript was determined by adding the extension product to sequence reactions that were produced using ^{32}P-labeled JB54.3 on a Tf1 DNA template. Because JB54.3 hybridized to Tf1 as well as Tf2 transcripts (Levin *et al.*, 1990), the single primer extension band was produced for RNA that contained both Tf1 and Tf2 transcripts, suggesting that the two related families of elements initiate transcription at the same site. The initiation site of Tf1 transcription was confirmed by primer extension with a Tf1 specific oligo (see below). The position of the start site was corroborated by the results of S1 nuclease reactions as shown in Figure 1, lanes 3 and 4.

A plasmid that overexpresses a neo-marked copy of Tf1

In order to observe Tf1 transposition, we increased the level of Tf1-107 transcription by fusing the element to the *S.pombe* fructose bis-phosphatase (*fbp*) promoter on a high copy vector (Figure 2). The resultant plasmid, pHL234, contained the *URA3, ars* and stabilization fragments from pFL20, which is maintained in *S.pombe* at 74–80 copies/cell (Heyer *et al.*, 1986). The level of Tf1 transcript produced from the *fbp* promoter was compared by RNA blot analysis to the amount produced by the Tf1 promoter. The RNA blot shown in Figure 3a contains RNA from YHL351, a transformant that carried Tf1 with its own LTR promoter (lane 2) and RNA from YHL448, a strain that contained the *fbp* promoter fusion (lane 1). The difference in signal of the two lanes indicated that the *fbp* promoter produced substantially more transcript than the Tf1 promoter.

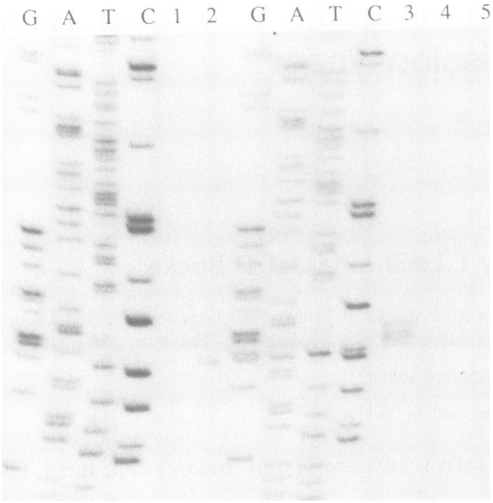

Fig. 1. RNA isolated from YHL334 was used in primer extension assays with an oligo that hybridizes next to the 5′ LTR of Tf1-107. The first set of sequence reactions was produced using ^{32}P kinased JB54.3 and double-stranded template. RNA was either pretreated with DNase and RNase (lane 1) or DNase alone (lane 2) before addition to the primer extension reaction. The second set of sequence reactions were identical to the first except that the product from lane 2 was added to each lane in the set. The superimposition of these samples demonstrated that Tf1 transcripts begin at base 189 of Tf1. S1 nuclease protection assays were carried out with YHL334 RNA pretreated with DNase (lane 3), RNase (lane 4) or DNase plus RNase (lane 5).

One reason for selecting the *fbp* promoter for use in the Tf1 fusion plasmids is that transcription of *fbp* can be repressed in medium containing 8% glucose and induced as much as 50-fold when 3% glycerol is the carbon source (Hoffman and Winston, 1989). Because growth of these transformants bearing these plasmids on SC-ura medium containing 8% glucose was erratic, we used medium containing 3% glucose. To determine whether Tf1 transcription is induced when cells were grown in glycerol, equal amounts of total RNA were extracted from glucose and glycerol grown cells, electrophoretically separated on a formaldehyde–agarose gel, transferred to nitrocellulose and reacted with a Tf1 specific probe. Figure 3b indicates that the same quantity of Tf1 RNA is produced by plasmid pHL234 whether cells were grown in media containing glycerol or glucose. For the experiments described below, unless otherwise indicated, the medium contained 3% glucose, which resulted in a faster growth rate.

To provide a means for the detection of transposition events, the plasmid pHL234 was further modified by the addition of the bacterial *neo* gene from Tn*903* to the region of Tf1-107 just upstream of the 3′ LTR. The structure of the *neo* containing plasmid, pHL297, is shown in Figure 4A. Cells that contained pHL297 grew on YE-G418 (which contains 500 μg/ml G418), whereas the untransformed strain failed to grow on as little as 150 μg/ml G418. We have previously shown that when this *neo* fragment from Tn**903** was inserted in single copy in the *Bgl*II site just 3′ to the *ura4* coding region in *S.pombe* (Grimm *et al.*, 1988), high levels of G418 resistance were observed (data not shown).

As indicated by Figure 4A, the position of the *fbp* promoter with respect to the transposon sequence in pHL297 was placed so that the transcript from the *fbp* promoter

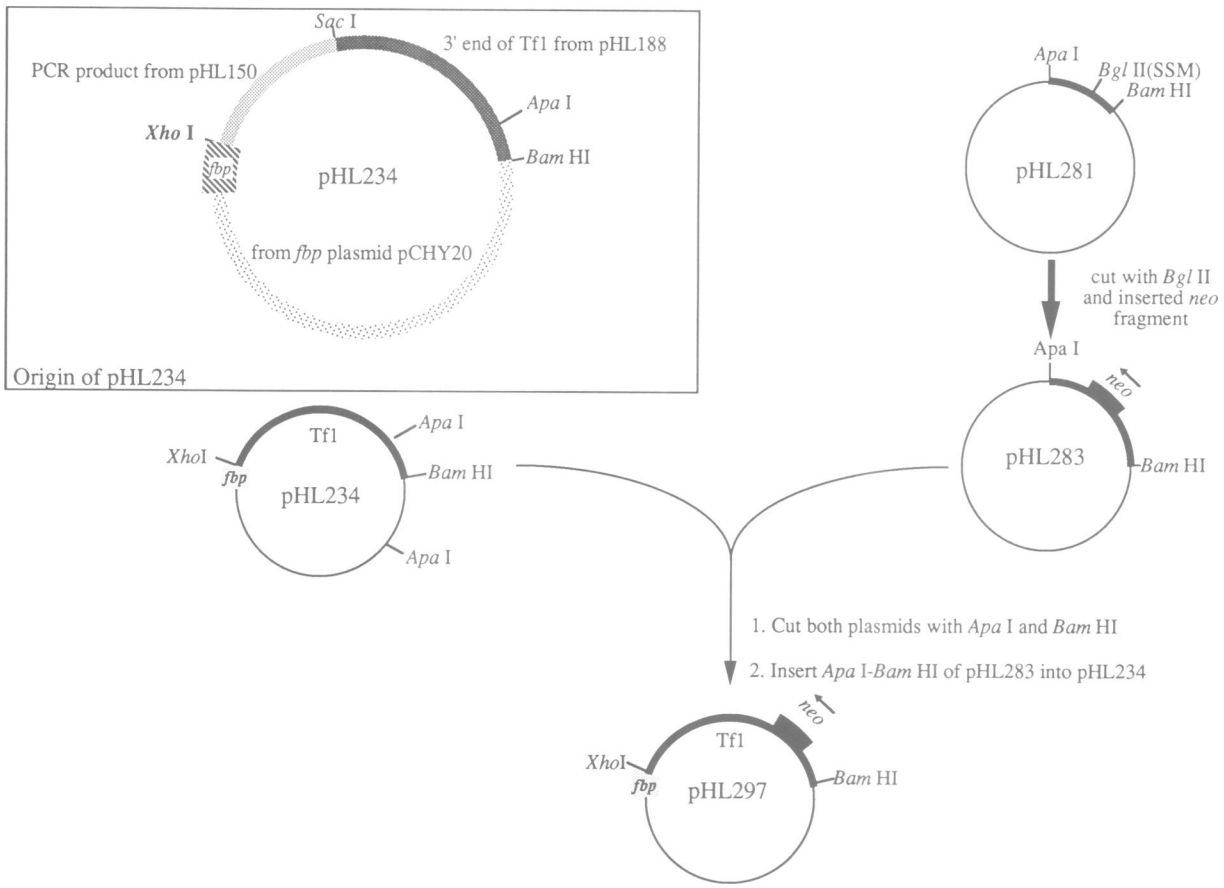

Fig. 2. Construction of the *fbp*−Tf1−*neo* plasmid pHL297. SSM in pHL281 indicates the position of the *Bgl*II site created by site specific mutagenesis. pHL281 contains a 1.7 kb *Apa*I−*Bam*HI fragment that contains the 3′ end of Tf1-107 with the *Bgl*II site into which the *neo* fragment was inserted. pHL234 (inset) is pCHY20 with two inserts that contain 5′ and 3′ portions of Tf1-107. pHL297 was produced by replacing the *Apa*I−*Bam*HI piece of pHL234 with the *Apa*I−*Bam*HI fragment from pHL283 that contains the *neo* insert. Because there is an additional *Apa*I site in the *URA3* section of pHL234, pHL297 transformants that were the result of partial rather than complete *Apa*I digestion were selected on LB-kanamycin plates to guarantee the presence of *neo*.

started at the same sequence as the natural Tf1 transcript. The purpose of preserving the normal Tf1 transcription start was to guarantee that the RNA could be reverse transcribed properly. To determine whether the transcription start site in pHL297 corresponds to the natural Tf1 start, primer extension reactions with a Tf1 specific oligonucleotide were performed with RNA isolated from strains that only possessed plasmid copies of *fbp*−Tf1 or endogenous copies of Tf1. Figure 4B shows primer extension assays that demonstrated that in strain YHL462 (a derivative of strain NCYC 132, containing many endogenous copies of Tf1), transcripts from endogenous copies of Tf1 initiate at the same position as the RNA produced by 570 and 573, two strains in which the only Tf1 copies derive from the *fbp*−Tf1 plasmid pHL297. A small fraction of the Tf1 transcripts produced by YHL570 and YHL573 are slightly larger than the native Tf1 transcripts which may mean that these mRNA copies initiate just within the *fbp* sequence. The control strain YHL438 contained a pFL20 plasmid without Tf1 and produced no Tf1 specific signal. There was, however, a non-Tf1 extension product observed in YHL438 that was seen below the Tf1 signal produced by all strains.

Transposition assay

The method used to detect transposition events is described in detail in Materials and methods. Strains containing *fbp*-fused Tf1-107 plasmids were grown as single colonies on SC-ura plates to select for the presence of the plasmids. Cells from individual colonies were then allowed to go through ~13 generations of growth on medium selective for the plasmid. Aliquots of the saturated cultures were spread on 5-fluoroorotic acid (FOA) plates to a density of ~10 000 colonies per plate. The presence of FOA allows only the cells that have lost the *URA3* plasmid to grow (Boeke *et al.*, 1984). As ~0.5% of the cells in an SC-ura culture have lost the plasmid, no curing step was required before the cells were plated onto the FOA plates. The FOA resistant colonies were then replica plated to YE-G418 to determine the number of colonies that had retained the *neo* marker but lost *URA3*.

Gene conversion of endogenous Tf1 elements by the Tf1−*neo* copy on the plasmid cannot explain the observed events, because the strain used contained no Tf1 copies. However, gene conversion of other Tf1 elements with limited homology, such as Tf2, or of Tf1 LTRs could give rise to chromosomal Tf1−*neo* copies. To rule out this type of mechanism, we repeated the transposition assay with a version of Tf1−*neo* that had a frameshift mutation in the protease domain. This frameshift mutation should eliminate translation of RT and IN. This control strain YHL675, is isogenic to YHL569 and YHL570 except that it contained the frameshift mutation. As indicated in Table I, the

frameshift mutation reduced the appearance of G418^r colonies to an undetectable level.

Physical characterization of transposition candidates

In addition to the genetic evidence for transposition, DNA blots were used to determine whether Tf1 as well as *neo* sequences had been transferred to the genomes of the G418 resistant cells. Genomic DNA was prepared from 12 transposition candidates of YHL570 from the first experiment in Table I (at least eight of these were independent). Because the Tf1−*neo* contained no sites of *Bam*HI, the DNAs from the 12 candidates were digested with this enzyme and hybridized to a *neo* probe. As seen in Figure 5A, each of the candidates contained a hybridization signal of varying size while the parent strain produced no detectable bands. All but one of the candidates clearly had a single restriction fragment that contained all of the *neo* hybridizing sequence, while one candidate (#1) produced two bands possibly because of an incomplete restriction digest. Candidate #1 genomic DNA produced a single band hybridizing to *neo* in a second DNA blot (data not shown). Each of the candidates apparently received a single copy of *neo* because there are no *Bam*HI sites in the *neo* sequence.

Figure 5B is a DNA blot identical to that in Figure 5A except that the filter was hybridized to a Tf1 specific probe (Figure 4A). The pattern of bands from this blot matched the signals from the previous blot, indicating that each of the candidates also acquired a single copy of Tf1 sequence and the copy of Tf1 and *neo* reside on the same restriction fragment. However, one exceptional candidate (#2) showed no Tf1 signal suggesting that the *neo* gene in this case might have entered the genome by a recombination pathway other than transposition or alternatively, that a deletion of this region occurred during transposition. Nevertheless, the DNA blot results from 11 of 12 candidates suggest that their ability to grow on G418 was due to simple transposition of Tf1−*neo*. Although the band for candidate #3 is light in this exposure, it is clearly seen in a dark exposure of this blot.

Isolation and characterization of target sites before and after transposition

The genomic Tf1−*neo* fragments and flanking sequences from candidates YHL663 and YHL665 were cloned into pBSII as *Spe*I−*Nhe*I fragments by selecting for resistance to kanamycin in *E.coli*. Neither Tf1 nor *neo* contains a site for either enzyme. The structure of the recombinant plasmids was determined by restriction digestion and determination of the sequences flanking the LTRs as described in Materials and methods. In both candidates, the *neo* genes were found inserted in an intact Tf1 element as in pHL297 except that the 5′ LTRs were present in an intact form (i.e. not fused to the *fbp* promoter). Like previously cloned genomic copies of Tf1 and Tf2, the transposed Tf1−*neo* elements were flanked by 5 bp duplications. As expected, these were different in sequence from the 5 bp duplications flanking Tf1-107.

The presence of intact 5′ and 3′ LTRs and 5 bp duplications were consistent with Tf1−*neo* transposition events. To confirm this hypothesis, we used sequence flanking Tf1−*neo* from each candidate to design oligos that were used in PCR reactions with genomic DNA from the parent strain. For both candidates, the PCR reaction with the candidate DNA produced no obvious bands while the reactions with the parent DNA produced a band of the size predicted for true transposition events. The PCR products from the third candidate, YHL665 are shown in Figure 6. We cloned the insertion sites of these two candidates from patient strain YHL232 so that parental genomic sequences could be compared to the sequences of the candidates. Figure 7 compares the sequence flanking the Tf1−*neo* insertions in the two candidates to the parental sequences. In both candidates, the only changes in the insertion sites were the duplication of the target site and the presence of the Tf1−*neo* element. Hence, both of these strains acquired the *neo* genes as the result of true transposition events.

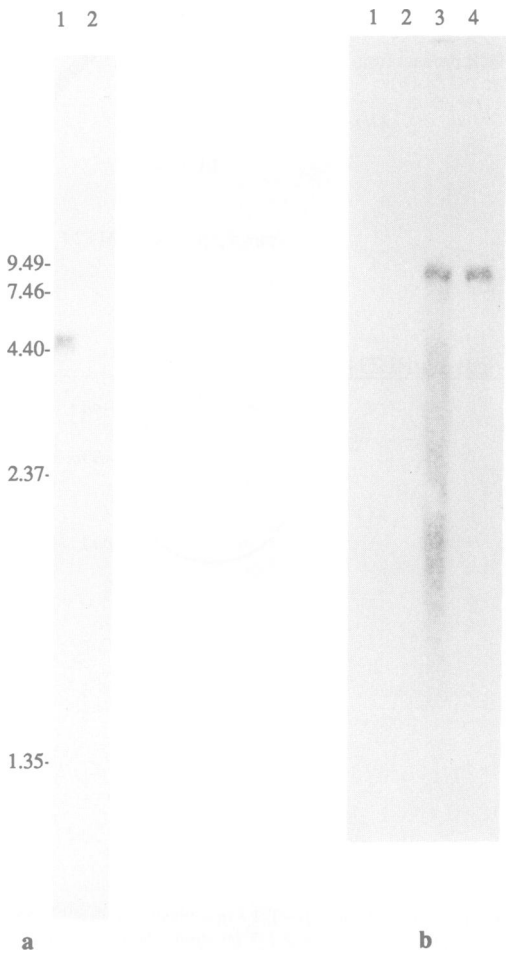

Fig. 3. (**a**) Total RNA was isolated from YHL448 (lane 1), a strain with Tf1 fused to the *fbp* promoter and YHL351 (lane 2), which contains Tf1 with its own promoter. Cells were grown in 3% glucose (see b). The total amount of RNA loaded in each lane was equivalent as determined by ethidium bromide staining. The RNA blot was hybridized with the Tf1 specific *Eco*RI probe shown in Figure 4A. The positions of 0.24−9.5 kb molecular weight standards (Bethesda Research Laboratory) are indicated to the left of the gel. (**b**) Total RNA was isolated from YHL458 (lanes 1 and 2), a strain with the pCYH20 vector that carries no Tf1 insert and YHL450 (lanes 3 and 4), which contains the *fbp*−Tf1 fusion. RNA loaded in lanes 1 and 3 came from cultures grown in 3% glycerol plus 0.1% glucose while lanes 2 and 4 contain RNA from cells grown in 3% glucose. The Tf1 specific *Eco*RI fragment probe was used.

Tf1 transposes at both high and low temperatures

Retrotransposition efficiency of Ty1 in *S.cerevisiae* is extremely sensitive to increased temperature (Paquin and Williamson, 1984, 1986; Boeke *et al.*, 1986). To measure

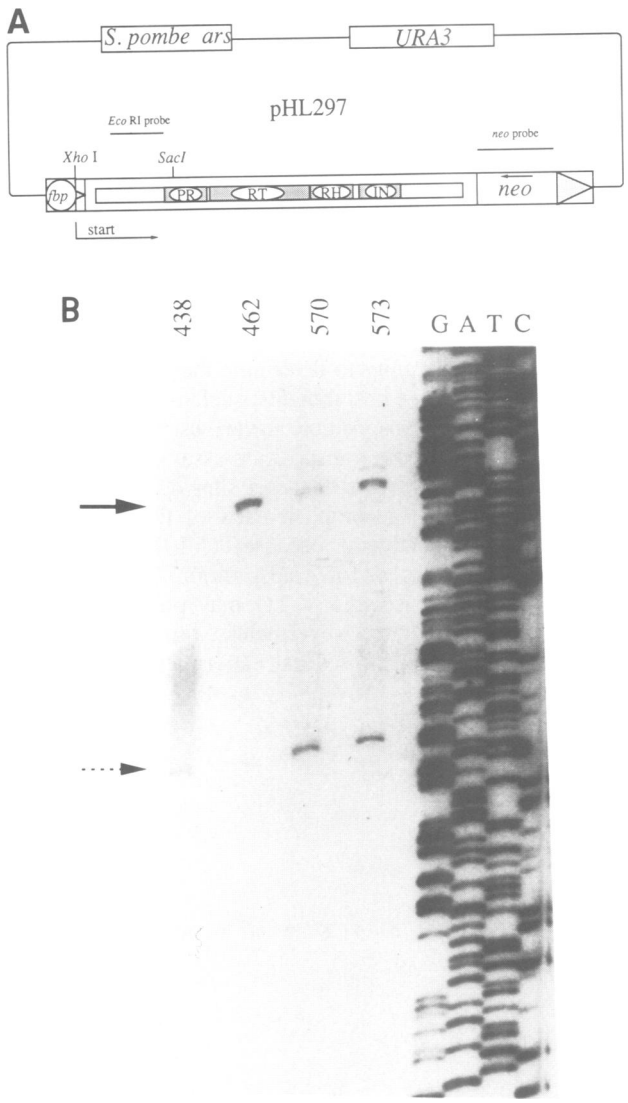

Fig. 4. (**A**) The diagram of pHL297 shows the position of the pFL20 that contains the *S.pombe ars*. The position of the *S.cerevisiae URA3* gene is also shown. The circle labeled *fbp* indicates the position of the promoter with relation to the natural transcription start site of Tf1 indicated by the left side of the triangle. The site of transcriptional initiation of the *fbp* fusion plasmid pHL297, is indicated by the arrow labeled start. The stippled bars show the positions of Tf1 amino acid sequences that are homologous to the retroviral domains PR, RT, RH, and IN. The stippled triangles represent the 3′ Tf1 LTR and the R/U5 portions of the 5′ LTR. The horizontal lines indicate the position of hybridization probes used for the blots in this paper. The *Sac*I site indicates the position at which a frameshift mutation was introduced. (**B**) These lanes are the products of primer extension reactions with oligo JB181 used to determine the start site of Tf1 transcription for YHL438, YHL462, YHL570 and YHL573 (solid arrow). The position of the dashed arrow marks a non-Tf1 specific product observed in all strains tested. The sequence reactions were produced with ^{32}P kinased JB181 on a Tf1 plasmid template to serve as molecular weight standards.

transposition frequency for Tf1 at different temperatures, we repeated the assay using cultures of YHL573, YHL675 and YHL677 grown at either 22°C, 30°C or 36°C. YHL573 was an independent transformant of plasmid pHL297 while YHL675 and YHL677 were independent transformants of the frameshift plasmid pHL326. G418 resistant colonies arose at frequencies of 0.06−0.2% regardless of the growth temperature. Mean frequencies were 0.14% at 22°, 0.10% at 30° and 0.10% at 36°C in YHL573. Under the same conditions two independent cultures of both YHL675 and YHL677 produced no G418 resistant colonies (<0.01%) at all three temperatures.

Effect of glycerol regulation of the fbp promoter on transposition

The effect of growth in glycerol on transposition was determined to compare with the RNA blot data that indicated that Tf1 RNA levels did not increase on glycerol. Growth in glycerol failed to stimulate transposition (0.17% on glucose versus 0.12% on glycerol). This result is consistent with the RNA blot in Figure 3b that shows no obvious difference in transcript levels. A possible explanation for the apparent lack of Tf1 induction from the *fbp* fusion plasmids is that there could be sequence in the 5′ end of Tf1 that serves as a transcriptional enhancer which interferes with the glucose repression.

Discussion

A system for studying transposition in *S.pombe* has been developed. This system combines the use of a strong promoter to direct the expression of Tf1 element proteins and the insertion of the selectable *neo* marker into a nonessential region of Tf1, allowing detection of relatively rare transposition events. The *fbp* promoter overproduces Tf1 transcripts that initiate at the natural Tf1 start site. pHL297, an *fbp* plasmid containing Tf1−*neo*, caused strains YHL569 and YHL570 to acquire G418 resistance at a frequency of ~0.2−4.1%. The frequency was reduced by at least 20-fold as the result of a frameshift mutation in the protease domain of pHL326. The simplest conclusion from this experiment is that the process that caused cells to gain resistance to G418 required the expression of functional Tf1 protein, as would be predicted for retrotransposition. The lack of G418 resistant cells produced by pHL326 demonstrated that homologous recombination between pHL297 and sequences within the genome was not responsible for the ability of pHL297 to convert cells to G418 resistance.

Genomic DNA from 12 YHL570 transposition candidates was analyzed to determine whether Tf1−*neo* sequences had actually been incorporated into the genome. Eleven of the 12 candidates each had *neo* and Tf1 sequences on the same fragment of genomic DNA, as predicted for true transposition events. One of the candidates clearly lacked any Tf1 signal, suggesting that either i) this particular strain acquired the *neo* marker by a mechanism other than transposition or ii) this element suffered a deletion during transposition. Because copies of the related element Tf2 (but not Tf1) are found in the chromosomes of strain YHL570 (Levin *et al.*, 1990), homologous recombination between the Tf1−*neo* on pHL297 and the endogenous Tf2 copies could have produced the candidate that became G418^r but lacked Tf1 sequences. In support of the second possibility, internal deletion of Ty1 sequences during transposition has been reported (Eichinger and Boeke, 1988; Curcio *et al.*, 1990).

Although DNA blot data suggested that the candidate strains had acquired G418 resistance via transposition events, direct examination of the flanking sequence was undertaken to provide definitive evidence. Structural analysis of the cloned DNA flanking the Tf1−*neo* element of two transposition candidates with the same regions in the parent strain,

showed that the ability to grow on G418 correlated with the simple insertion of Tf1−*neo*. The duplication of the five bases flanking the Tf1−*neo* suggests that the integration event is initiated by a 5 bp staggered cut mediated by an integrase of the type encoded by retroviruses and retrotransposons.

Table I. Transposition assay results

Experiment 1 Strain	FOA Res.	G418 Res.	%G418 Res./FOA Res.[a]
YHL569	2,000 ± 100	83	4.1
YHL570	2,000 ± 100	38	1.9
Experiment 2 Strain	**FOA Res.**	**G418 Res.**	**%G418 Res./FOA Res.**
YHL570A	7,000 ± 500	13	0.18
YHL570B	5,600 ± 200	14	0.25
YHL675A	5,200 ± 200	0	<0.02
YHL675B	4,000 ± 200	0	<0.02

[a]The transposition frequencies observed in experiment 1 are high compared to subsequent experiments for unknown reasons. The data in experiment 2 are typical of the results obtained since (Levin and Boeke, 1992).

The temperature sensitivity of Ty1 retrotransposition in *S.cerevisiae* has previously been reported (Paquin and Williamson, 1984, 1986; Boeke *et al.*, 1986). The ability to monitor Tf1 transposition at 36°C is experimentally useful. It suggests that some aspect of Tf1 transposition is substantially more heat stable than its counterpart in Ty1 transposition. Should this heat resistant step involve integration or reverse transcription, biochemical characterization may be facilitated by stability at higher temperatures. In addition, the ability of Tf1 to transpose at high temperatures provides an opportunity to determine the effect of various temperature sensitive host mutants, such as *cdc* mutants, on the frequency of transposition at various temperatures.

The results from the transposition assays along with the physical characterizations of target sites demonstrates that Tf1-107 is able to accomplish each of the steps in the retrotransposition pathway. Because Tf1-107 is the first Tf1 element copy assayed for transposition, the possibility remains that other isolates of Tf1 may possess higher or lower levels of activity. Nevertheless, this active copy of Tf1 forms the basis for several experimental approaches

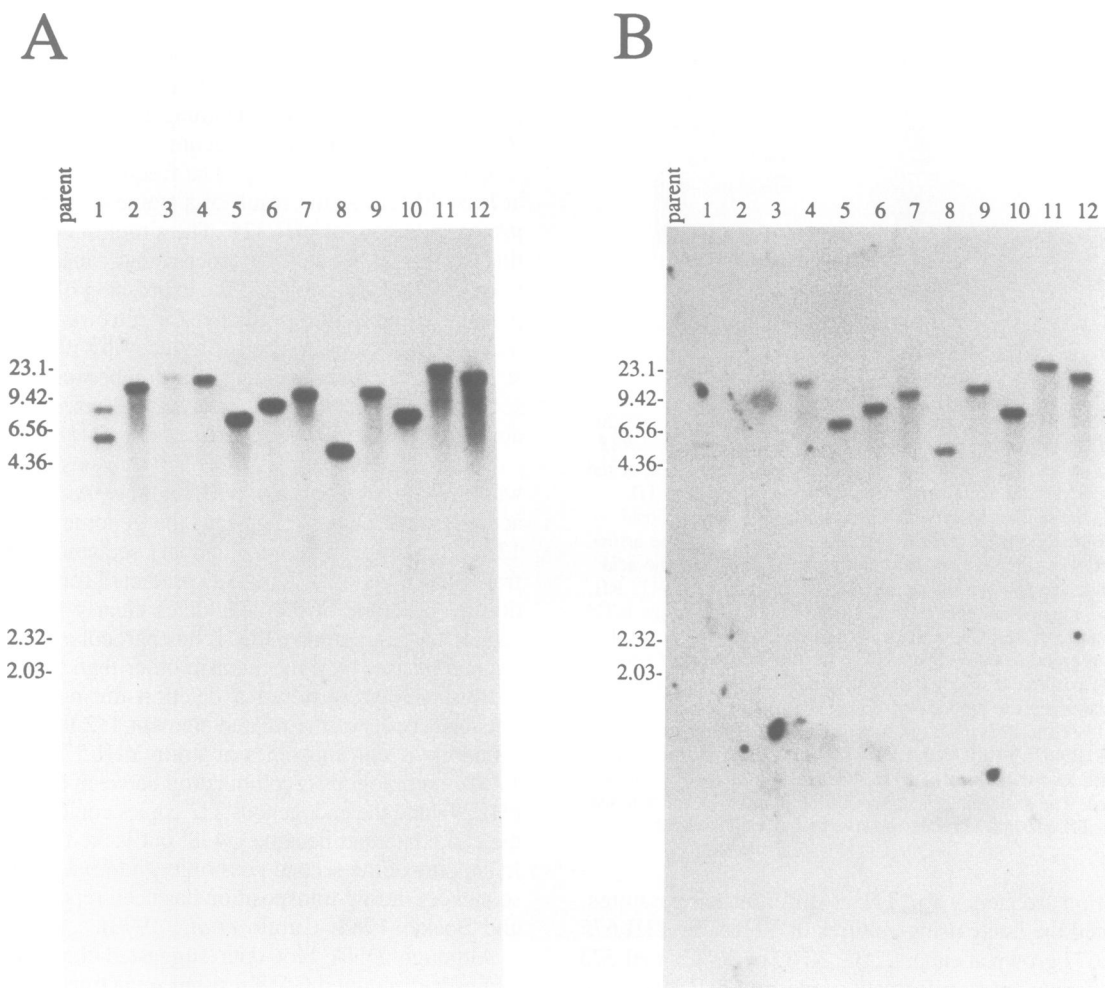

Fig. 5. (**A**) Genomic DNA was isolated from 12 transposition candidates and the parent, YHL232, after plasmid pHL297 was lost. The DNA was digested with *Bam*HI, which cuts neither Tf1 nor *neo*. The DNA blot was hybridized with the *neo* probe shown in Figure 4A. The position of molecular weights are shown to the left of the gel. (**B**) This blot is identical to that shown in A except that the hybridization probe used was the Tf1-specific *Eco*RI fragment shown in Figure 4A.

aimed at characterizing the mechanisms of each step in transposition, including the contribution of host functions.

The frequency of transposition observed in this system is perhaps 100-fold lower than transposition of Ty1−*neo* from a *GAL*−Ty1 plasmid in *S.cerevisae*. This difference could be due to many different factors, such as expression efficiency, marker effects or relative efficiency of this particular copy. However, we note that marked Ty3 elements, which like Tf1 belong to the *gypsy* family of retrotransposons, transpose at a similar frequency to Tf1−*neo*. Since many members of the *gypsy* family have apparent target site specificity, lower transposition frequency could also be a consequence of less available target sites in the genome. Since, in the seven Tf1/Tf2 insertion sites sequenced, there is no evidence that these elements prefer a specific target sequence, nor that they insert upstream of pol III transcription units, such target specificity would have to be of a novel type.

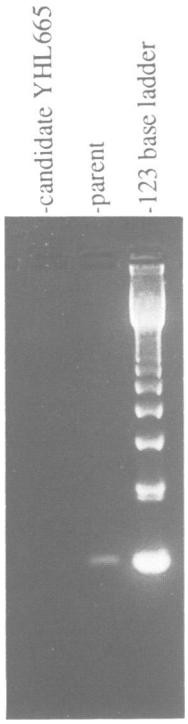

Fig. 6. The lane marked YHL665 contained the PCR products of a reaction containing DNA from this candidate with oligos JB140 and JB141 that hybridized to positions flanking the target site. The parent lane reaction contained the same oligos but the genomic DNA was derived from the parent strain, YHL232. The expected molecular weight for the amplified pretransposition DNA is 127 bp. The molecular weight markers used were a 123 base pair ladder.

Materials and methods

Plasmid constructions

The transposition plasmid pHL297, was produced by ligating specially modified versions of the two halves of Tf1 into the *Xho*I and *Bam*HI sites of the expression plasmid pCHY20. pCHY20 is identical to pCHY21 (Hoffman and Winston, 1989) except that in pCHY20 the polylinker was placed at the start site of transcription (C.Hoffman and F.Winston, personal communication). The 5′ half of Tf1 in this construct contains an *Xho*I site at the transcription start site. The 3′ portion contains a bacterial *neo* gene in a nonessential region of the element. As shown in Figure 2, the *Xho*I site introduced into Tf1 at the site of transcription initiation was created by PCR with oligonucleotides JB68 (CCAAGGTCTGGTAGGAAG) and JB70 (GCCTCGAGTTCAGTTATGAGCTATAT). The PCR product was cleaved with *Xho*I and *Sac*I and the product was ligated into pCHY20 together with the *Sac*I−*Bam*HI fragments of Tf1 to produce pHL234. A section of Tf1 from the *Apa*I site to the *Bam*HI site was cloned into M13mp19, creating phage 3′ pst9, and site specific mutagenesis (Kunkel, 1985) was used to create a unique *Bgl*II site at position #4568 just downstream of the Tf1 ORF. Oligonucleotide JB112 (TACATAGAAGATCTTGGG-GAGGG) was used to create the *Bgl*II site. The *Apa*I−*Bam*HI fragment was then subcloned into pBSII, creating pHL281. The mutagenized *Apa*I−*Bam*HI fragment was cut with *Bgl*II and the *neo* fragment from pGH54 (Joyce and Grindley, 1984; Garfinkel *et al.*, 1988) was inserted, giving rise to pHL283. To create pHL297, the *Apa*I−*Bam*HI fragment containing the 3′ half of Tf1 marked with *neo* was excised from pHL283 and ligated into pHL234 cut with *Apa*I and *Bam*HI. pHL326 is identical to pHL297 except that it has a frameshift mutation in the protease region. The frameshift was constructed by inserting a double-stranded oligonucleotide JB127: pGCCGTGCGGCCGCACGGCAGCTTCGACGGCACGCCG-GCGTGCCGp, which contains a *Not*I site, into the *Sac*I site.

Growth media

The liquid medium used for the transposition assay was SC-ura (synthetic complete medium lacking uracil; Sherman *et al.*, 1986), but differed from this formula in that 3% glucose was used as the carbon source unless otherwise specified. The SC-ura plates contained 2% agar in addition to the other components. The YE-G418 plates contained 0.5% yeast extract, 3% glucose, 2% agar, and 0.5 mg/ml Geneticin (GIBCO; the concentration was corrected for purity). The FOA plates were SC-ura plates that contained 50 μg/ml uracil and 1 mg/ml 5-fluoro-orotic acid (SCM Specialty Chemicals).

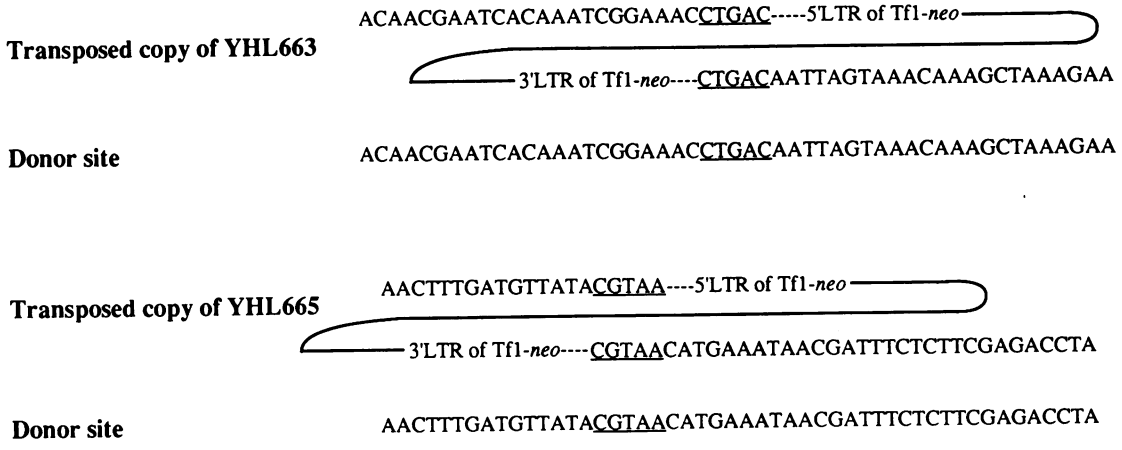

Fig. 7. The results of DNA sequencing reactions of the parent and candidate target site region. The curved line broken with two short lines represents the body of the transposed Tf1. The underlined nucleotides indicate the position of the five base target site that is duplicated upon insertion.

Table II. Strains

Strain	Genotype	Source	Plasmid description
972	h^-	[c]	
975	h^+	[c]	
NCYC 132	wt	[c]	
YHL232	*ura4 D-18*, h^-	(Grimm *et al.*, 1988)	
YHL334	*ura4*$^-$	[d]	
YHL351	YHL232/pHL117	this paper	Tf1 inserted in pFL20
YHL438	YHL232/pDW231	[e]	truncated Tf1-107 fused to an epitope tag
YHL448	YHL232/pHL235	this paper	same as pHL234 except independent PCR reaction
YHL450	YHL232/pHL234	this paper	*fbp* promoter fused to PCR copy of Tf1 (Figure 1)
YHL458	YHL232/pFL20	this paper	*S.pombe ars* plasmid
YHL462	YHL334/pCYH20	this paper	*fbp* fusion vector
YHL569	YHL232/pHL297	this paper	*fbp* fusion of Tf1−*neo* (Figure 1)
YHL570[a]	YHL232/pHL297	this paper	*fbp* fusion of Tf1−*neo*
YHL573	YHL232/pHL298	this paper	differs from pHL297 by having an independent PCR product
YHL663	YHL232 Tf1−*neo* transposition #1	this paper	
YHL665	YHL232 Tf1−*neo* transposition #3	this paper	
YHL675	YHL232/pHL326	this paper	identical to pHL297 except pHL326 has frameshift mutation in protease
YHL677[b]	YHL232/pHL326	this paper	identical to pHL297 except pHL326 has frameshift mutation in protease

[a]YHL570 was an independent transformant from YHL569
[b]YHL677 was an independent transformant from YHL675
[c]The references for these strains are listed in Zimmer *et al.* (1987).
[d]This strain was isolated in this lab as a spontaneous *ura4*$^-$ mutant of NCYC 132.
[e]D.Weaver, H.Levin and J.Boeke, unpublished.

Primer extension and S1 nuclease protection assays

When indicated, 500 μg of RNA were pretreated with 50 units of DNase (RQ1 from Promega) in 3 ml of buffer (50 mM Tris−Cl, pH 7.2, 10 mM $MgSO_4$) for 1 h at 37°C. Some RNA samples (10 μg) were also pretreated in 1 μg/ml RNase in TE (10 mM Tris pH 7.9, 1 mM EDTA) at 37°C for 1 h. The RNase and DNase treatments were followed by two phenol extractions. The first step of the primer extension assay was the hybridization reaction, containing 10 μg of genomic RNA in 15 μl of reaction buffer (0.1 M Tris−HCl (pH 8.1), 1.0 mM EDTA, 17.3 mM NaCl). Approximately 0.5 pmol of ^{32}P-labeled oligonucleotide was added to the mixture that was then incubated at 65°C for 2 min followed by a 42°C incubation for 60 min. After the hybridization, 9 μl of extension buffer (55.5 μM each dNTP, 28 mM $MgCl_2$, and 2.8 mM DTT) and 2 units of AMV reverse transcriptase (Boehringer Mannheim) was added to the reactions, and then incubated at 37°C for 1.5 h. The extension products were ethanol precipitated and loaded onto sequencing gels. The S1 nuclease protection assays were performed according to Ausubel *et al.* (1987). The primers used for these assays were JB54.3 (ATCACAAGAGTTCAGTTA) and JB181 (GTTGTAGTGATGGACGC).

Transposition assay

Individual colonies of *S.pombe* strains (see Table II) transformed with either pHL297 or pHL326 were used to inoculate the growth of patches of 1 square cm grown for 3 days on SC-ura plates. The patches were used to inoculate 10 ml liquid cultures that were adjusted to obtain uniform initial cell densities in each culture with an OD_{600} of 0.05−0.1 in SC-ura medium. The cultures were grown to saturation (~1 × 10^8 cells/ml) at 30°C (unless otherwise specified). 0.1 ml of a 100-fold dilution was spread on FOA plates, giving rise to 5000−50 000 colonies. The FOA^r colonies were then printed to YE-G418 plates. Each G418^r colony was scored as a transposition candidate. DNA blots used to characterize the candidate strains were performed using a previously described phenol extraction technique (Levin *et al.*, 1990).

Probes

The *Eco*RI probe of Tf1 shown in Figure 4A derives from nucleotides 564−1312. Because this region of sequence is only found in Tf1 and not Tf2 (Levin *et al.*, 1990), genomic DNA extracted from cells derived from the laboratory strains 972 and 975 does not hybridize with the *Eco*RI probe. The *neo* probe used is also shown in Figure 4A and is the 1.0 kb *Bam*HI fragment from the *neo* gene (Joyce and Grindley, 1984; Garfinkel *et al.*, 1988).

Isolation of candidate Tf1−neo insertions and target sites

Genomic DNA was isolated from candidate transposition strains YHL663 and YHL665. 5−10 μg of total genomic DNA was cut with *Nhe*I and *Spe*I and loaded onto a 0.6% agarose gel. Fragments >4 kb long were ligated into pBSII KS− (Stratagene), cut with *Xba*I and transformed into *E.coli*, selecting for growth on LB plates containing 50 μg/ml kanamycin. After the two halves of the Tf1−*neo* elements were subcloned into independent pBSII constructs, JB44 and JB45, oligonucleotides that hybridize to the LTR were used to sequence the genomic DNA that flanked the insertions (Levin *et al.*, 1990). Oligonucleotides that hybridized to the insertion flanks for use in PCR reactions with genomic DNA of the parent, G418^s strain YHL232 were synthesized. JB138 (ACAACGAATCACAAATC), JB139 (TTACATACTTATTATTT), JB140 (AACTTTGATGTTATA) and JB141 (AAAGGATTTGAGTTATTT) were used to produce PCR fragments corresponding to pretransposition sequences that were then gel purified and ligated into pBSII and sequenced.

Acknowledgements

We thank C.Hoffman and F.Winston for sharing strains and data prior to publication and P.Hieter, F.Spencer and D.Nathans for helpful comments on this manuscript. Supported in part by a postdoctoral fellowship (H.L.L.) and grant FRA-366 (J.D.B.) from the American Cancer Society, NIH training grant CA09139 (H.L.L.) and Public Health Service grant CA16519 from the National Institutes of Health (J.D.B.).

References

Ausubel,F., Brent,R., Kingston,R.E., Moore,D.D., Smith,J.A., Seidman,J.G. and Struhl,K. (eds) (1987) *Current Protocols in Molecular Biology*. J. Wiley and Sons, New York.
Boeke,J.D., Xu,H. and Fink,G.R. (1988) *Science*, **239**, 280−282.
Boeke,J.D., Lacroute,F. and Fink,G.R. (1984) *Mol. Gen. Genet.*, **197**, 345−346.
Boeke,J.D., Garfinkel,D.J., Styles,C.A. and Fink,G.R. (1985) *Cell*, **40**, 491−500.
Boeke,J.D., Styles,C.A. and Fink,G.R. (1986) *Mol. Cell. Biol.*, **6**, 3575−3581.
Boeke,J.D., Eichinger,D.J., Castrillon,D. and Fink,G.R. (1988) *Mol. Cell. Biol.*, **8**, 1432−1442.
Brierley,C. and Flavell,A.J. (1990) *Nucleic Acids Res.*, **18**, 2947−2951.
Chalker,D.L. and Sandmeyer,S.B. (1990) *Genetics*, **126**, 837−850.
Curcio,M.J., Hedge,A.M., Boeke,J.D. and Garfinkel,D.J. (1990) *Mol. Gen. Genet.*, **220**, 213−221.
Curcio,J. and Garfinkel,D.J. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 936−940.
Eichinger,D.J. and Boeke,J.D. (1988) *Cell*, **54**, 955−966.

Garfinkel,D.J., Mastrangelo,M.F., Sanders,N.J., Shafer,B.K. and Strathern,J.N. (1988) *Genetics*, **120**, 95–108.
Grimm,C., Kohli,J., Murray,J. and Maundrell,K. (1988) *Mol. Gen. Genet.*, **215**, 81–86.
Hansen,L.J. and Sandmeyer,S.B. (1990) *J. Virol.*, **64**, 2599–2607.
Heidmann,O. and Heidmann,T. (1991) *Cell*, **64**, 159–170.
Heidmann,T., Heidmann,O. and Nicolase,J.F. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2219–2223.
Heyer,W., Sipiczki,M. and Kohli,J. (1986) *Mol. Cell. Biol.*, **6**, 80–89.
Hoffman,C.S. and Winston,F. (1989) *Gene*, **84**, 473–479.
Jensen,S. and Heidmann,T. (1991) *EMBO J.*, **10**, 1927–1937.
Joyce,C.M. and Grindley,N. (1984) *J. Bacteriol.*, **158**, 636–643.
Kunkel,T.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 488–492.
Levin,H.L., Weaver,D.C. and Boeke,J.D. (1990) *Mol. Cell. Biol.*, **10**, 6791–6798.
Levin,H.L. and Boeke,J.D. (1992) In Leibowitz,M.J., Koltin,Y. and Rubio,V. (eds), *The Proceedings of the International Workshop on Viruses of Fungi and Simple Eukaryotes*. University of Delaware Press, DE, in press.
Miller,K., Rosenbaum,J., Zbrzezna,V. and Pogo,A.O. (1989) *Nucleic Acids Res.*, **17**, 2134.
Paquin,C.E. and Williamson,V.M. (1984) *Science*, **226**, 53–55.
Paquin,C.E. and Williamson,V.M. (1986) *Mol. Cell. Biol.*, **6**, 70–79.
Pelisson,A., Finnegan,D.J. and Bucheton,A. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 4907–4910.
Sandmeyer,S.B., Hansen,L. and Chalker,D.L. (1990) *Annu. Rev. Genet.*, **24**, 491–518.
Sherman,F., Fink,G.R. and Hicks,J.B. (eds) (1986) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Voytas,D.F., Konieczny,A., Cummings,M.P. and Ausubel,F.M. (1990) *Genetics*, **126**, 713–721.
Yoshioka,K., Honma,H., Zushi,M., Knodo,S., Togashi,S., Miyaek,T. and Shiba,T. (1990) *EMBO J.*, **9**, 535–541.
Zimmer,M., Welser,F., Oraler,G. and Wolf,K. (1987) *Curr. Genet.*, **12**, 329–336.

Received on September 20, 1991; revised on November 26, 1991