Schizosaccharomyces pombe Retrotransposon Tf2 Mobilizes Primarily through Homologous cDNA Recombination

ELEANOR F. HOFF,[1] HENRY L. LEVIN,[2] AND JEF D. BOEKE[1]*

Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205,[1] and Laboratory of Eukaryotic Gene Regulation, National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892[2]

The Tf2 retrotransposon, found in the fission yeast *Schizosaccharomyces pombe*, is nearly identical to its sister element, Tf1, in its reverse transcriptase-RNase H and integrase domains but is very divergent in the *gag* domain, the protease, the 5′ untranslated region, and the U3 domain of the long terminal repeats. It has now been demonstrated that a *neo*-marked copy of Tf2 overexpressed from a heterologous promoter can mobilize into the *S. pombe* genome and produce true transposition events. However, the Tf2-*neo* mobilization frequency is 10- to 20-fold lower than that of Tf1-*neo*, and 70% of the Tf2-*neo* events are homologous recombination events generated independently of a functional Tf2 integrase. Thus, the Tf2 element is primarily dependent on homologous recombination with preexisting copies of Tf2 for its propagation. Finally, production of Tf2-*neo* proteins and cDNA was also analyzed; surprisingly, Tf2 was found to produce its reverse transcriptase as a single species in which it is fused to protease, unlike all other retroviruses and retrotransposons.

Transposable elements constitute up to 50% of the eukaryotic genome (51, 53). Though they can act as positive forces in the evolution of an organism, both by providing part of the chromosomal architecture (30) and by providing a source of mutagens (1), they also represent a burden on the host cell genome and a potential threat to host cell viability, should they "jump" into an essential region of the genome or mediate a rearrangement thereof. Different transposable elements have developed different means to balance their survival with that of the host cell by controlling their spread within the genome. Most balancing mechanisms described thus far appear to involve control of either transposon expression or target site selection (9). These mechanisms often rely on host-specific factors, such as transcription factors (10, 27), splicing factors (47), and chromatin organization (58). Characterization of different transposable elements can therefore reveal both new mechanisms for control of element mobility and new host-element interactions.

Long terminal repeat (LTR)-type retrotransposons (hereafter referred to simply as retrotransposons) have been isolated from many different eukaryotic organisms, including fruit flies, maize, and the yeast *Saccharomyces cerevisiae* (6, 23, 48–50). Retrotransposons resemble retroviruses in both genome structure and replication mechanism (9). Like retroviruses, they possess terminally redundant ends (LTRs), a primer binding site for the initiation of reverse transcription, and a polypurine tract that serves as the primer in second-strand synthesis of the cDNA copy of the element. A single mRNA encodes proteins homologous to the retroviral structural proteins capsid (CA) and, sometimes, nucleocapsid (NC) and to the retroviral enzymes protease (PR), reverse transcriptase-RNase H (RT), and integrase (IN). These proteins coassemble to form retrovirus-like particles in which reverse transcription of an RNA intermediate takes place. The resulting cDNA is then typically integrated into the host cell genome by the element-encoded IN, generating 4- to 6-bp target site duplications (TSDs).

Retrotransposition has been extensively studied in yeast. Several yeast retrotransposons exhibit target site bias during transposition. The Ty1 and Ty3 elements of *Saccharomyces cerevisiae* target integration to regions upstream of RNA polymerase III-transcribed genes (13, 16), while Ty5 exclusively targets telomeres and silenced chromatin (58). This targeting appears to rely on interactions with host cell factors such as RNA polymerase III, transcription factors, and chromatin components. Since the targeted regions lack open reading frames (ORFs) and elements for controlling expression of downstream genes, they represent one mechanism for balancing retrotransposon and host survival: preferential mobilization into a "safe haven" in the host cell genome (14, 55).

Retrotransposons are typically found in multiple copies in a host cell genome. In yeast, this has resulted in very-low-frequency homologous recombination of retrotransposon cDNA with preexisting genomic copies of its parent element, in addition to normal transposition (39). This integrase-independent pathway relies on host cell recombination machinery (43, 52). Because of these two different modes of entry into the host genome, we use the term mobilization to refer to total element movement when the use of the two pathways cannot be distinguished. Though recombination occurs at different percentages of total mobilization for each element examined (less than 10%), all yeast retrotransposons studied to date mobilize primarily through the integrase-dependent pathway (7, 21, 31, 58).

The Tf2 element is a retrotransposon isolated from the fission yeast *Schizosaccharomyces pombe* (35). It is closely related to the well-characterized *S. pombe* Tf1 element. Both are 4.9-kbp elements with a single ORF encoding the CA-, PR-, RT-, and IN-like domains. The cloned elements Tf2-1 and Tf1-107 are almost identical in nucleotide and amino acid sequence in the RT and IN regions and are identical in the extreme carboxyl terminus of the PR region; however, they are highly divergent in the 5′ untranslated region (5′ UTR) between the 5′ LTR and the ORF, which in retroviruses comprises all or part of a specific RNA packaging signal (for a review, see

* Corresponding author. Mailing address: Dept. of Molecular Biology and Genetics, Hunterian Bldg., Rm. 617, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205. Phone: (410) 955-2481. Fax: (410) 614-2987. E-mail: jboeke@jhmi.edu.

A.

B.

pEH 143 (Tf2-*neo*)
pEH 145-1 (Tf2GagFS-*neo*)
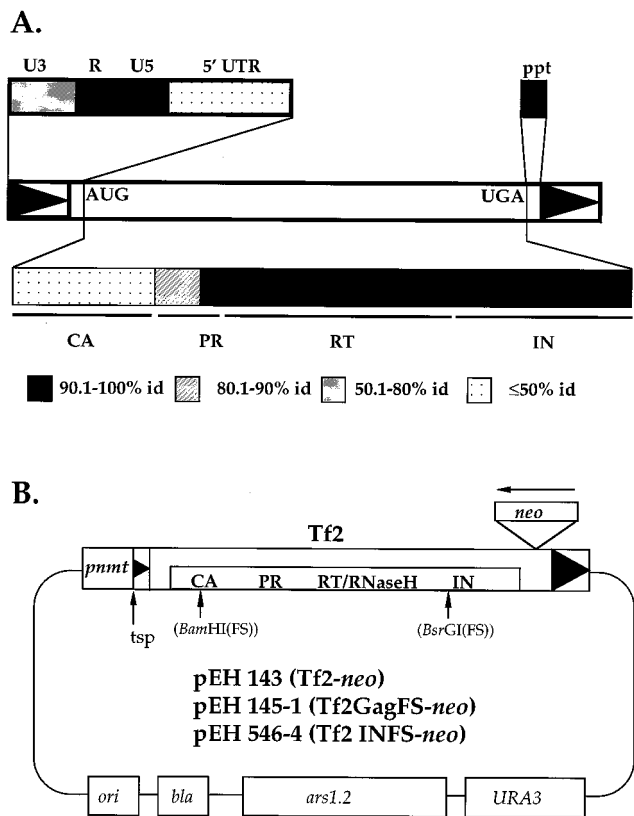pEH 546-4 (Tf2 INFS-*neo*)

FIG. 1. (A) Tf2 differs from Tf1 in both translated and untranslated sequences. Comparisons of Tf2 and Tf1 nucleotide and amino acid sequences revealed discrete regions of difference (35, 56), represented here by differently shaded boxes. Filled triangle, LTR; AUG, start of ORF; UGA, end of ORF; CA, capsid-like protein; ppt, polypurine tract (100% identity [id] between Tf2 and Tf1); thickly outlined filled boxes, nucleotide sequence; thinly outlined filled boxes, amino acid sequence. (B) Tf2-*neo* expression plasmids. Tf2 was placed under the control of the repressible *nmt* (no message in thiamine) promoter in a plasmid construct similar to that previously described for Tf1-*neo* (pHL414-2). Expression was induced by transferring cells from medium containing 10 μM thiamine to medium lacking thiamine, allowing intracellular levels of thiamine to drop below the repression threshold (50 pmol/10^7 cells). The *neo* gene, inserted in the 3′ end of the retrotransposon between the ORF and the 3′ LTR, in the orientation opposite that of *nmt1*-directed transcription, confers resistance to the drug G418 (Geneticin). This plasmid allows for selection of mobilized copies of the marked retrotransposon once a cycle of expression and transposition has been induced. tsp, transcription start point.

reference 46), as well as in the DNA encoding CA and about two-thirds of PR (56) (Fig. 1A). In addition, Tf2-1 and Tf1-107 are quite divergent in the U3 region of the LTR, a region containing *cis*-acting sequences involved in transcriptional regulation (5, 12, 18, 54). Only the Tf2 element is found in the form of full-length copies in the commonly used lab strains 972 and 975 (35).

Until now, analysis of the Tf2 element had been limited (35, 56), whereas the Tf1 element had been shown to be a functional retrotransposon, capable of generating true transposition events at a high frequency (31, 33). The presence of a Tf1-like ORF in Tf2 indicated that it too might be functional. Furthermore, based on Tf2's homology to Tf1, it appeared likely that the former shares Tf1's polyprotein processing (3, 34) and novel self-priming (31, 36) pathways. However, the differences between Tf2 and Tf1, particularly those in their 5′ UTR and the CA regions, raised the possibility that the Tf2 and Tf1 elements mobilize differently, in terms of either reg-

ulation or efficiency; these differences could occur at transcriptional, translational, or posttranslational steps. Our characterization of Tf2 has revealed that its mobilization is significantly different from that of Tf1. Most importantly, for its propagation, it relies heavily on homologous recombination of its cDNA. This mobilization phenotype, which we have termed integration site recycling, may represent a unique means of balancing retrotransposon survival with that of the host cell.

MATERIALS AND METHODS

Growth media. For nonselective growth, *S. pombe* was cultured in rich medium, YEC (0.5% yeast extract–0.2% Casamino Acids) plus uracil (250 μg/ml), adenine (500 μg/ml), and glucose (3%). For selective growth, *S. pombe* strains were grown in Edinburgh minimal medium (EMM), prepared according to the manufacturer's instructions (BIO 101, Vista, Calif.), supplemented with amino acids and pyrimidines, either individually or as part of a 5S dropout mix (with complete 5S containing Lys, His, Leu, uracil [U], and adenine [A], each at 250 μg/ml [final concentration]). Thiamine (10 μM) was added to the media when indicated. For selection against *URA3*-based plasmids, 5-fluoroorotic acid (FOA) was added to a final concentration of 0.1% to plates containing 250 μg of uracil per ml. For selection for the presence of the *neo* gene, Geneticin (G418) was added where indicated to a final actual concentration of 500 μg/ml.

Strains and plasmids. The strains used or constructed in this study are described in Table 1. Strains carrying a plasmid were grown and then transformed by the lithium acetate–Tris-EDTA method as described previously (25), with minor modifications as necessary. Plasmid pEH143-1 (Tf2-*neo* with an *nmt* promoter [*pnmt*]) was constructed in five steps. First, the 3′ end of Tf2 was subcloned as a 3.9-kbp *Pvu*II-*Xba*I fragment from pHL325-1 into *Eco*RV- and *Xba*I-cut pBSIIKS−. It was then subjected to site-specific mutagenesis with JB112 (5′-TACATAGAAGATCTTGGGGAGGG-3′), as described previously (29), to create a unique *Bgl*II site in the noncoding region between the end of the ORF and the beginning of the 3′ LTR, creating plasmid pEH129-1. The mutagenized Tf2 sequence was then subcloned as a 1.6-kbp *Bsr*GI-*Ehe*I fragment into the 10.5-kbp *Bsr*GI-*Bam*HI site of pHL411 (34), creating intermediate plasmid pEH130 and eliminating the *Bam*HI site. pEH130 also lacks the 2.2 kbp of *S. pombe* genomic sequence flanking the 3′ end of Tf2 in pEH125-1 and pHL416-38. Complete Tf2 with a *Bgl*II site and under the control of the *nmt* promoter was made by replacing the Tf1 sequence in pEH130 with Tf2 sequence (a 3.2-kbp *Xho*I-*Bsr*GI fragment) from pHL416-38, resulting in pEH133-1. pEH134-1 was created simultaneously by cloning the Tf2 sequence-containing *Xho*I-*Bsr*GI fragment from pEH128-1 (pHL416-38 cut with *Bam*HI, filled in, and religated) into *Xho*I- and *Bsr*GI-cut pEH130, resulting in a version of *pnmt*Tf2 with a *Bgl*II site and with a frameshift in the Gag (CA) region. pEH143-1 and pEH145-1 were generated by cloning the *neo* gene of Tn*903* as a *Bam*HI fragment from pGH54 into the *Bgl*II site of pEH133-1 or pEH134-1, respectively; in these plasmids, the orientation of *neo* is the reverse of *pnmt*-directed transcription (33). *pnmt*Tf2-*neo* with a frameshift in the IN domain was created by cutting pEH143-1 with *Bsr*GI, filling in with the Klenow fragment of DNA polymerase, and religating, resulting in pEH546-3.

RNA isolation and blotting. Total cellular RNA was typically isolated from 10 ml of an *S. pombe* culture grown to log phase (an optical density at 600 nm [OD$_{600}$] of 1.0) at 30°C and harvested by centrifugation. Cell pellets were disrupted by freeze-thawing and resuspended in 0.3 ml of RNA extraction buffer (0.1 M NaCl, 0.1 M Tris-Cl [pH 7.5], 0.03 M EDTA, and 1% Sarkosyl). Cold glass beads were added to the meniscus, 0.3 ml of PCIA (phenol-chloroform-isoamyl alcohol, 50:48:2) was added, and the mixture was vortexed for 10 s. After centrifugation in an Eppendorf microcentrifuge for 4 min at 4°C and 14,000 rpm, the aqueous phase was reextracted once with PCIA and once with chloroform-isoamyl alcohol (24:1), and then the RNA was precipitated. RNA pellets were washed with 70% ethanol and resuspended in 100 μl of diethyl pyrocarbonate-treated deionized water (DEPC-dH$_2$O) or DEPC-dH$_2$O containing 0.1 M EDTA.

For RNA blot analysis, 10 to 15 μg of each RNA preparation was ethanol precipitated and resuspended in 15 μl of RNA loading buffer (50% [vol/vol] formamide and 15% [vol/vol] formaldehyde in 1× MOPS [40 mM morpholinepropanesulfonic acid, 10 mM sodium acetate, 1 mM EDTA, pH 7.0]). Each sample was heated at 65°C for 10 min; then 3 μl of 6× tracking dye was added, and the sample was electrophoresed on a 1% agarose–formaldehyde gel in 1× MOPS for 4 h. The molecular size marker (0.4- to 9.0-kb RNA ladder; Gibco-BRL) lane was stained with ethidium bromide, while the rest of the gel was subjected to capillary transfer to a Genescreen Plus nylon membrane in 20× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) for at least 14 h. Blotting was performed in RNA hybridization buffer (10% [wt/vol] dextran sulfate, 0.33 M NaPO$_4$ [pH 7.0], 10 mM EDTA; pH 7.5) with probe (see below) added to a final concentration of 0.1 × 10^6 to 1.0 × 10^6 cpm/ml. Blots were washed in 0.2× SSC–0.5% sodium dodecyl sulfate (SDS) at 60°C and then exposed to film or subjected to phosphorimaging analysis with a STORM scanner and ImageQuant software (Molecular Dynamics, Synnyvale, Calif.).

For comparisons of RNA expression from the *nmt* promoter under inducing

TABLE 1. Yeast strains and plasmids used in this study

Strain[a]	Plasmid[b]	Construct	Source or reference	Notes
YHL912			J.D.B.	
SZP20-1	pHL416-38	Tf2	This study	
SZP57-1	pEH143-1	Tf2-*neo*	This study	*neo* in reverse orientation to Tf2
SZP59-1	pEH145-1	Tf2GagFS-*neo*	This study	*neo* in reverse orientation to Tf2
SZP64-1	pFL20	—[c]	34	Same strain as YHL1032
SZP66-1	pHL414-2	Tf1-*neo*	34	Same strain as YHL1089
SZP67-1	pHL415	Tf1PRFS (*Sac*I fill-in)-*neo*	34	Same strain as YHL1091
SZP68-1	pHL431-25	Tf1INFS (*Bsp*HI fill-in)-*neo*	This study	
SZP73-1	pEH546-4	Tf2INFS (*Bsr*GI fill-in)-*neo*	This study	

[a] YHL912 is the parent strain of the other strains listed.
[b] Except for pFL20, all plasmids are *URA3* based with constructs fused to the *nmt* promoter.
[c] See reference 37a. pFL20 is the *URA3 ars1-1* vector on which all other constructs were based.

and repressing conditions, 5- or 10-ml precultures were grown overnight in the absence of thiamine at 30°C; they were then diluted to an OD$_{600}$ of 0.2 and split, and thiamine was added to one tube to a final concentration of 10 μM. These cultures were then grown to an OD$_{600}$ of 1.0 to 3.0 and harvested as described above.

Probes. The DNA probes used were a 0.27-kbp Tf2-specific *Eco*RI-*Bam*HI probe, a 0.96-kbp *neo* fragment (a *Bam*HI fragment from pGH54), a 0.75-kbp *Bst*XI-*Bgl*II fragment from the 3′ end of Tf2 (taken from pEH129-1), and a 0.7-kbp *Bam*HI-*Hin*dIII fragment corresponding to part of the *act1* ORF (taken from pHL859). Probes were prepared by labelling the fragments with Redivue [α-^{32}P]dGTP (Amersham), by using the random hexamer labeling method of Feinberg and Vogelstein (17).

Protein preparations and blotting. Total soluble protein from cells expressing the Tf elements from the *nmt* promoter was prepared by growing a preculture of cells in EMM plus 5S and lacking U (EMM+5S-U) overnight at 30°C and then inoculating 10 ml of fresh EMM+5S-U to an OD$_{600}$ of 0.2 to 0.3. Each culture was grown to an OD$_{600}$ of 1.0 and harvested by centrifugation, and the pellets were subjected to freezing before extraction (5 min on dry ice-ethanol, or −70°C storage). Extraction was performed by resuspending each pellet in 0.1 ml of YLB (0.05 M Tris-Cl [pH 7.5], 1% SDS) plus additives (leupeptin, antipain, benzamidine, chymostatin, pepstatin A, each at 1 μg/ml; 10 U of aprotinin/ml; 0.014 M β-mercaptoethanol; 1 mM phenylmethylsulfonyl fluoride), adding cold glass beads to the meniscus, and vortexing the mixture at 4°C for 10 min. At ≥80% cell lysis (monitored by phase microscopy), the mixture was then boiled for 3 min, the supernatant was transferred to a new tube, and the beads were then washed with 50 μl of YLB plus additives. Pooled supernatants were microcentrifuged for 20 min at 4°C, the new supernatant was transferred to a new tube and respun for 5 min, and then the approximate protein concentration was measured by reading the A_{280} of a 1:200 dilution. On the basis of this measurement, about 150 to 300 μg (100 to 200 A_{280} units/ml) of each sample was mixed with 2× Laemmli buffer (20% glycerol, 125 mM Tris [pH 6.8], 5% SDS, 1.4 M β-mercaptoethanol), and the mixture was boiled for 1 min, loaded onto an SDS–10% polyacrylamide gel, and run in Tris (25 mM)–glycine (192 mM)–0.1% SDS buffer; 50 to 100 μg of high-molecular-weight prestained markers (Gibco-BRL) was also loaded for molecular sizing purposes. The proteins were then electroblotted onto an Immobilon-P (millipore) polyvinylidene difluoride membrane (60 min at 300 to 500 mA in Tris–glycine–15% methanol).

To detect Tf-specific antibodies (Abs), polyclonal antibodies (Abs) raised against Tf1 IN (Ab 657), Gag (CA; Ab 653) (34), and RT (Ab 1571; raised against a peptide, encoded in an *Eco*RI-*Bbs*I fragment, containing the entire RT domain of Tf1 [provided by E. Sweeny]) were incubated with the membrane at a 1:10,000 dilution in 5% milk–1× phosphate-buffered saline (PBS)–0.05% Tween 20 for 45 min to 3 h. After being washed in 0.05% milk–PBS–Tween 20, the blots were incubated with horseradish peroxidase-conjugated polyclonal Ab (goat anti-rabbit immunoglobulin G; Amersham), at a 1:10,000 dilution, for 30 to 60 min. After excess secondary Ab was washed off in 0.05% milk–PBS–Tween 20, signal was detected by using an Amersham ECL enhanced chemiluminescence kit.

Mobilization assay. To detect the Tf element expression-dependent movement of *neo* information from the *pnmt*Tf-*neo* plasmids into the *S. pombe* chromosomes, a mobilization assay was used. In the qualitative assay, cells from pertinent strains were patched in approximately 2-cm^2-sized patches on thiamine-containing medium (EMM+5S-U with 10 μM thiamine) and grown for 2 days at 32°C. They were then replica printed to nonrepressive medium and induced for 2 to 4 days at 32°C; in parallel, the thiamine dependence of the G418^r phenotype was tested by repressing on medium containing thiamine. The plates were then replica plated to EMM-5S-FOA-thiamine medium to select for the loss of the plasmid (2 days, 32°C) and then printed to YEC plus U-A-FOA-G418 to select for chromosomal Tf-*neo* mobilization events (i.e., growth). After 2 days as well as 3 days of incubation, growth on the G418 plates was scored and the plates were photographed. (Multiple transformants for each plasmid were initially tested to establish consistency of phenotype.)

For quantitation of the mobilization frequency, the assay was performed as described above until the end of the induction period, at which time the patches were scraped into 3 ml of sterile dH$_2$O. The OD$_{600}$ of a 1:10 dilution was measured, and appropriate dilutions of the cells were plated on EMM-5S-FOA-thiamine plates to allow for both a quantitation of Foa^R cells per milliliter of the original cell suspension and subsequent quantitation (by replica printing the FOA plates to FOA-G418 medium) of Foa^R G418^r cells per milliliter of the original cell suspension. The Tf-*neo* mobilization frequency was calculated with the following formula: (number of Foa^R G418^r cells per milliliter)/(total number of Foa^R cells per milliliter).

Genomic-DNA preparation and blotting. Total DNA was typically prepared from a 10-ml culture grown in rich medium (YEC-U-A plus 3% glucose; see above) for 24 to 36 h at 30°C, as previously described (35). For Southern analysis, 1 to 2 μg of DNA was digested with the appropriate restriction enzyme for at least 12 h, electrophoresed on a 1% agarose gel in 1× Tris-taurine-EDTA, and subjected to capillary transfer to a Genescreen Plus nylon membrane in 10× SSC after denaturation and neutralization were performed by standard methods.

Hybridization and washing. Filters were prehybridized in 4× Denhardt's solution–3× SSC for at least 3 h at 65°C. Hybridization was carried out with random-hexamer-labeled probes added at 5 × 10^5 to 1 × 10^6 cpm/ml of hybridization solution (6× SSC–4× Denhardt's solution–10 mM EDTA–0.5% [wt/vol] SDS) at 65°C for ≥16 h. Washes were performed in 200 ml of 2× SSC–1% (wt/vol) SDS, once for 30 min and once for 15 to 30 min, at 65°C. The filters were then subjected to autoradiography or phosphorimaging, using a STORM scanner and ImageQuant software.

Cloning of transposition sites. Putative Tf2-*neo* transpositions were cloned out of genomic DNA by doubly digesting DNA preparations with *Xba*I-*Spe*I or *Xba*I-*Nhe*I, electrophoresing the digests on an 0.8% agarose gel, and dividing the >4.4-kbp fraction into two pools; the DNA was then purified by using a Gene-Clean kit (Bio 101), and the fragments were ligated into *Xba*I-cut pBSIISK−. The ligation products were transformed into Electromax DH10b cells (Gibco-BRL) by electroporation, plated on Luria-Bertani medium containing carbenicillin at 50 μg/ml, and then replica printed to Luria-Bertani medium containing kanamycin at 25 μg/ml. DNA from Carb^R) Kan^R clones was analyzed by restriction digestion to determine if it contained Tf2-*neo*. DNA flanking Tf2-*neo* was sequenced by using primer JB236 (5′-AGAGTTCAGTTATTGTA-3′), which lies 3′ of the 5′ LTR, as well as the T3 and T7 primers. Flanking sequence was then used to search the Sanger Centre *S. pombe* sequence database, so that recombination into a preexisting Tf2 or LTR could be differentiated from an actual novel transposition event. Flanking sequence was also used to prepare primers for PCR of the empty site in the parent strain YHL912.

Analysis of gap-repaired plasmids by sampling genomic Tf2 elements. Colonies of strain YHL912 transformed with the 10.5-kbp *Bam*HI-*Bsr*GI fragment of pEH143-1 were tested for being simultaneously U$^+$ and G418^r while being separately Foa^R, indicating linkage of episomal Ura$^+$ phenotype to the *neo* gene. Plasmids were then rescued into *Escherichia coli* DH5α cells from several candidates satisfying these criteria, by using the STET/glass bead method of Robzyk and Kassir (48a). The rescued plasmids were then analyzed by digestion with *Bam*HI and *Bsr*GI and were shown to have the same overall structure as pEH143-1. A minimum of two bacterial transformants per original yeast DNA preparation were analyzed.

RESULTS

The Tf2-*neo* transcript is stably expressed from the *nmt* promoter. To characterize the expression and mobilization competence of Tf2, the *neo* gene from Tn*903* was inserted at a *Bgl*II site introduced downstream of the ORF; *neo* allows for selection of resistance to Geneticin (G418) in *S. pombe* in
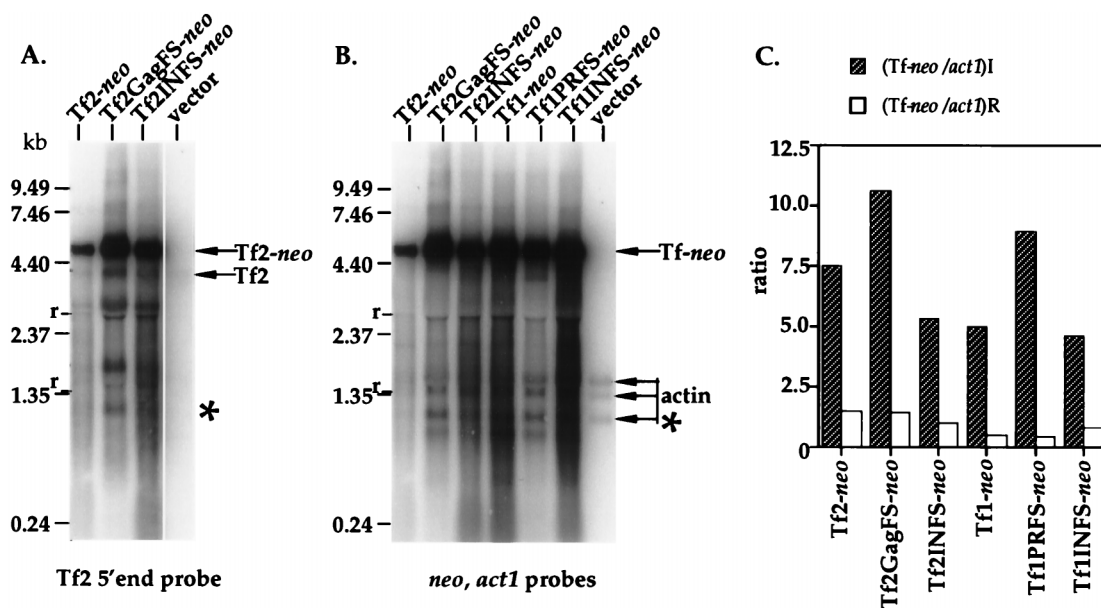
FIG. 2. RNA expression from *pnmt*Tf2-*neo* plasmids. (A) Tf2-*neo* RNA is stable and of the correct size. Expression of Tf2-*neo* in strains SZP57-1, 59-1, and 73-1 was compared to endogenous Tf2 RNA expression by RNA blot analysis with a ^{32}P-labelled Tf2-specific probe; the lanes are labeled with the Tf2-*neo* species. The signal intensity in this blot is not indicative of relative expression, since less sample was loaded for SZP57-1 (Tf2-*neo*). A single major band corresponding to Tf2-*neo* migrated to a position corresponding to a size about 1 kb larger than that of the endogenous Tf2 signal (seen in both the vector-only and Tf2-*neo* lanes). The secondary bands that comigrated with rRNA bands (marked "r") probably represent degraded RNAs. A quantitative comparison of the observed Tf2 and Tf2-*neo* signals shows that the Tf2-*neo* signal represents a 10-fold increase in total Tf2 message. (B and C) Tf2-*neo* RNA is expressed at levels similar to those of Tf1-*neo* RNA. (B) An RNA blot of samples from induced cultures of strains SZP57-1, 59-1, 73-1, 66-1, 67-1, 68-1, and 64-1, probed with *neo* and the *S. pombe* *act1* gene (41) simultaneously. The lanes are labeled with the Tf-*neo* species. The RNA load in the SZP57-1 (Tf2-*neo*) sample was much lower than those of the others; hence, the signal was less intense. (C) Tf-*neo* signals and *act1* signals from induced and repressed (grown in the presence of thiamine) samples were then quantitated by phosphorimaging; the Tf-*neo* signal was normalized to the *act1* signal, and normalized signals (measured in arbitrary units that are proportional to signal intensity) show that there is a similar level of expression. The *act1* probe detects three actin RNAs (41); the smallest one (∗) was used for normalization. This blot was stripped and reprobed for use in panel A (controlling appropriately for residual signal); thus, the smallest actin band is still apparent in panel A in an eightfold-longer exposure than that used for panel B. I, induced; R, repressed.

single copy in the Tf context (33). To control its expression, the Tf2-*neo* construct was cloned under the control of the *nmt1* promoter (*pnmt*) on an *S. cerevisiae* URA3-based plasmid (Fig. 1B, pEH143-1); *S. cerevisiae* URA3 complements *S. pombe* *ura4*. The *S. pombe* *nmt1* promoter allows high-level expression of genes cloned downstream of it when cells are grown in the absence of thiamine but can be repressed by 80-fold or more in the presence of ≥2 μM thiamine (4, 38). By using a similar strategy, Tf1-107 marked with *neo* (Tf1-*neo*) had previously been cloned under the control of the *nmt1* promoter on a URA3-marked plasmid and was shown to exhibit high-level expression of proteins and a high mobilization efficiency (34). Two mutant constructs, one with a frameshift mutation in the Tf2 Gag region (pEH145-1; GagFS) and the other with a frameshift mutation in IN (pEH546-3, INFS), were made as controls for experiments described below.

The size and stability of the Tf2 transcript initiated from the *nmt* promoter were determined by growing *S. pombe* strains (Table 1) harboring the *pnmt*Tf2-*neo* plasmids to mid-log phase (OD$_{600}$ = ~1.0) in the absence of thiamine to allow expression from that promoter. Total RNA extracted from these cells was subjected to RNA blot analysis, using a Tf2-specific probe (Fig. 2A). A single major species of the correct size was observed, as well as a band corresponding to endogenous Tf2 transcripts; as expected, the Tf2-*neo* band is at a position corresponding to a size about 1 kb larger than that of the unmarked Tf2 band. Quantitation of the Tf2-*neo* and Tf2 signals, using a phosphorimager, showed that the Tf2-*neo* mrNAs accumulated to a level 8- to 13-fold higher than the

sum of the levels of the endogenous Tf2 mRNAs (data not shown).

The same RNA samples, as well as total RNA from cells expressing Tf1-*neo* constructs, were probed with *neo* and an *S. pombe* *act1*-specific fragment in order to compare the expression levels of the Tf-*neo* constructs (Fig. 2B). When the Tf-*neo* signals were quantitated and normalized to the *act1* signal, the largest difference between the steady-state RNA levels of the different Tf-*neo* constructs was about twofold (Fig. 2C). When cells were grown in the presence of thiamine, RNA levels were reduced 5- to 20-fold (Fig. 2C).

Tf2-*neo* is competent for mobilization. Since the Tf2-*neo* element mRNA could be stably expressed in *S. pombe*, the mobilization phenotype of this element was assayed. Strains harboring the *pnmt*Tf2-*neo* plasmids were assayed for mobilization of *neo* information from the plasmid to the chromosome, after induction or repression of Tf2-*neo* expression, by selection for chromosomal acquisition of *neo* on medium containing FOA and G418. The FOA selects against the Tf2-*neo* plasmid, since URA3 confers FOA sensitivity (8); thus, only cells that have both lost the plasmid copy of *neo* and gained a chromosomal copy should be able to grow on medium containing both FOA and G418. Since acquisition of the chromosomal copy of *neo* should be dependent on mobilization of the Tf2-*neo* element, growth should also be observed only when an active copy of Tf2-*neo* has been expressed.

The results of these assays are shown in Fig. 3. At 32°C, Tf2-*neo* could mobilize; it generated G418^r papillae in a thiamine-dependent manner (Fig. 3A, panel I). On the negative-
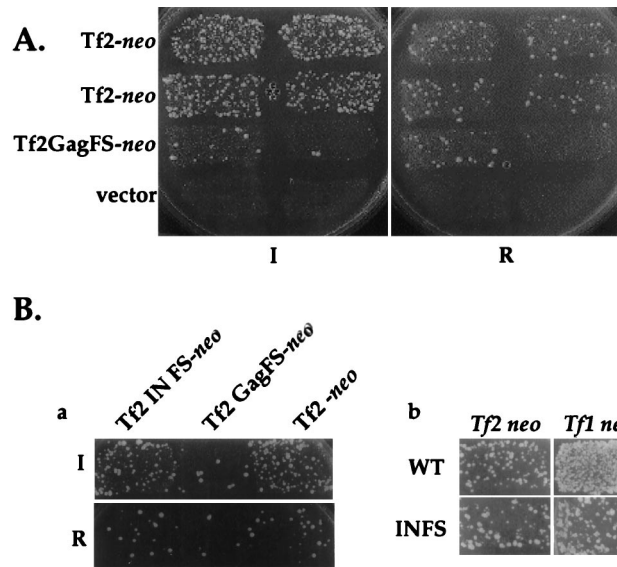
TABLE 2. Comparison of Tf-*neo* mobilization frequencies

Construct	Mobilization frequency (%)[a]	Frequency relative to:	
		Tf1-1-*neo*[b]	wt[c]
Tf2-*neo*	0.20 ± 0.04	0.06	1.00
Tf2GagFS-*neo*	≤0.01 ± 0.00	NA[d]	<0.05
Tf2INFS-*neo*	0.14 ± 0.03	0.04	0.67
Tf1-*neo*	3.40 ± 0.91	1.00	1.00
Tf1PRFS-*neo*	≤0.01 ± 0.00	NA	<0.003
Tf1INFS-*neo*	0.28 ± 0.10	0.08	0.082

[a] The Tf-*neo* mobilization frequency was determined for three experiments as described in Materials and Methods. Values are means ± standard errors of the means.
[b] Mean mobilization frequency relative to that of Tf1-*neo*.
[c] Mean mobilization frequency relative to that of the corresponding wt element.
[d] NA, not applicable.

FIG. 3. Tf2-*neo* mobilization. (A) Tf2-*neo* mobilization depends on expression of an intact Tf2-*neo* element. Results of a qualitative mobilization assay using patches of cells from SZP57-1,-2, 59-1,-2, and 64-1 are shown. Cells were induced in the mobilization assay for 3 days, then printed to minimal medium containing uracil and FOA and grown for 2 days; finally, they were transferred to rich YEC-glucose medium containing both FOA and G418 (YEC-FOA/G418). Growth on YEC-FOA/G418 medium is indicative of mobilization. Tf2-*neo* expression (induction [I]) generates clones that are resistant to G418 in the absence of the plasmid, as shown by papillation on YEC-FOA/G418. Induction of the negative control, Tf2GagFS-*neo*, generates Foa^R G418^r clones at a frequency no higher than that seen for patches of cells mock induced on repression-inducing medium containing thiamine (R). (B) Tf2-*neo* mobilization is not integrase dependent. Shown are the results of a mobilization assay comparing strains SZP57-1, 59-1, and 73-1, containing *nmt* plasmids with Tf2-*neo*, Tf2GagFS-*neo*, and Tf2INFS-*neo*, respectively (indicated in the figure). The frameshift in IN was introduced by filling in the *Bsr*GI site (Fig. 1B). Growth on G418-containing medium represents the relative mobilization frequency. Tf2INFS-*neo* mobilizes *neo* nearly as frequently as Tf2-*neo*. (a) The Tf2INFS-*neo* mobilization phenotype is not like the Tf1INFS-*neo* mobilization phenotype. I, induced; R, repressed. (b) Mobilization assay comparing IN frameshifts (INFS) of Tf2 and Tf1 to wt cognates.

control plate a background of papillae resulting from repression of Tf2-*neo* expression with thiamine (Fig. 3A, panel R) was observed, but this background was no higher than that seen with induction of a defective Gag frameshift (GagFS) version of Tf2-*neo*, and it thus represents either plasmid-containing colonies that escaped the FOA selection or plasmid recombination with the chromosome. In addition, since the number of G418^r papillae of the induced GagFS control did not differ from that of the repressed GagFS control, it appears that any expression of Tf2 proteins from the endogenous Tf2 elements is insufficient to complement the plasmid-expressed Tf2-*neo* RNA in *trans*.

However, the relative frequency of Tf2-*neo* mobilization is much lower than that of Tf1-*neo* mobilization (Fig. 3B, panel b). When a quantitative mobilization assay was performed, about 20-fold less mobilization by Tf2-*neo* than by Tf1-*neo* (Table 2) was observed. The observed difference between the Tf2-*neo* and Tf1-*neo* mobilization frequencies was not altered either by changing the temperature at which the induction step was performed to 22 or 37°C or by extending the induction time (data not shown).

Tf2-*neo* mobilization is mostly IN independent. The majority (about 94%) of Tf1-*neo* mobilization events are IN dependent (31), indicating the occurrence of true retrotransposition. This phenotype was observed when a Tf1-*neo* element with a frame-

shift in the IN domain was used. To determine genetically the percentage of Tf2-*neo* mobilization events that are IN dependent, a similar frameshift was introduced into the IN of Tf2, and the mobilization of this construct (Tf2INFS-*neo*) was compared to those of Tf2-*neo*, Tf1-*neo*, and Tf1INFS-*neo* (Fig. 3B). Surprisingly, the mobilization frequency of the Tf2INFS-*neo* mutant was about 70% of the wild-type (wt) Tf2-*neo* mobilization frequency whereas Tf1INFS-*neo* mobilized at only 8.2% of the wt Tf1-*neo* frequency (Table 2). Both the wt Tf2-*neo* and the Tf2INFS-*neo* mobilization frequencies were within twofold of that of the Tf1-INFS-*neo* construct also assayed in these experiments (Table 2), suggesting that these three Tf-*neo* elements follow similar mobilization pathways.

Tf2-*neo* targets endogenous Tf2 elements for recombination. The results obtained in these mobilization assays suggested that Tf2-*neo* might mobilize mostly, or even entirely, by an IN-independent pathway, presumably homologous recombination. Thus, it was important to determine the molecular nature of the Tf2-*neo* mobilization events. To facilitate the molecular analysis of the Tf2-*neo* elements mobilized into the YHL912 parent strain genome, a determination of endogenous Tf2 copy number and distribution in YHL912 was performed. This determination was performed by blot analysis of YHL912 genomic DNA digested with enzymes that cut Tf2-1 one or more times upstream of a 3′-end probe (Fig. 4A).

The results of the DNA blot analysis are shown in Fig. 4B. The highest degree of resolution of Tf2-hybridizing bands was obtained by using restriction enzyme *Bsr*GI. When the patterns and band intensities obtained with all of the different enzymes were compared, it was apparent that there are about 15 copies of Tf2 in the YHL912 genome. The existence of doublets in the *Bsr*GI digest was independently confirmed by the analysis of recombinants in which only one band in the doublet shifted (see below). This comparison also led to the detection of a Tf2 band at a position corresponding to a size of about 5.0 kbp, i.e., the length of a Tf2 element less one LTR, in all digests obtained by using enzymes that cut once in Tf2-1 (Fig. 4B and C, band T). A band was detected at the same position in a parallel blot of *Pst*I-cut DNA by a probe specific for the 5′ end of Tf2 (Fig. 4C), suggesting that this band represents a preexisting Tf2 tandem array. PCR experiments performed with YHL912 genomic DNA (data not shown), as well as analysis of Tf2 element sequences in GenBank (see below), support the existence of a tandem array of two Tf2 elements sharing a single LTR.

To analyze the molecular nature of the Tf2-*neo* mobilization events, and to determine if any of these events is a true trans-
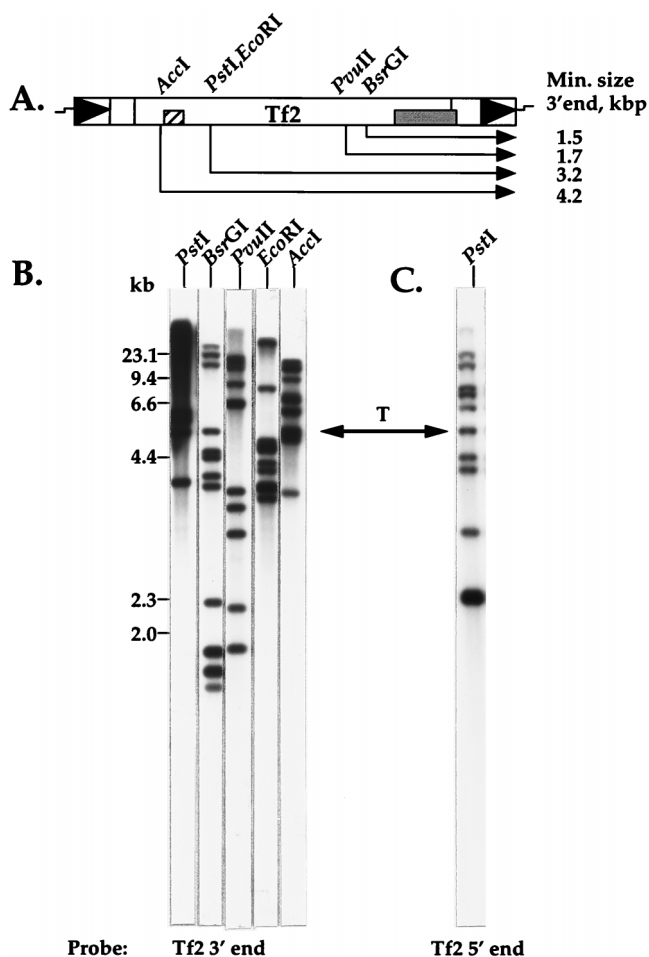
FIG. 4. DNA blot analysis of endogenous Tf2 elements. (A) Map of Tf2-1 showing minimum lengths of 3′-end fragments generated by cleavage at conserved restriction sites. Assuming a low degree of polymorphism, endogenous Tf2 elements cut with a particular restriction enzyme should generate fragments, plus flanking sequence, of a minimum predictable size that can be detected with the Tf2 3′-end probe. Dark box, Tf2 3′-end-specific fragment; hatched box, Tf2 5′-end fragment. (B and C) DNA blot analysis of digests of YHL912 DNA, using the Tf2 3′-end probe (B) and the 5′-end probe (C). (B) DNA from parent strain YHL912 was digested with the enzymes indicated in panel A, subjected to agarose gel electrophoresis, transferred to a Genescreen Plus membrane, and blotted with a ^{32}P-labeled Tf2 3′-end-specific probe. Five different fragment patterns are evident. A unit-sized band (T) of about 5.0 kb is detectable in YHL912 digests prepared with enzymes cutting once in Tf2 (*PstI*, *BsrGI*, and *AccI*). (C) This band is also present when the *PstI* digest is probed with the ^{32}P-labelled Tf2 5′-end-specific probe.
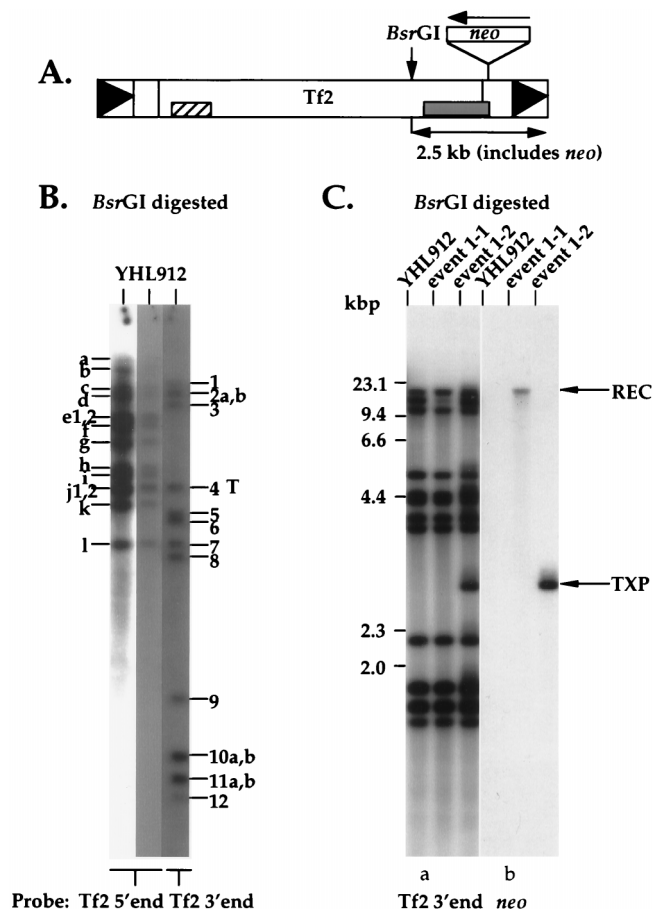
FIG. 5. DNA blot analysis of Tf2-*neo* and Tf2INFS-*neo* mobilization candidates: identification of putative transposition events. (A) Map of Tf2-*neo* depicting the *BsrGI* site and 5′-end and 3′-end probes. Hatched box, Tf2 5′-end probe; dark box, 3′-end probe. The *neo* probe encompasses the entire *neo* gene (open box). (B) DNA blot analysis of parent strain YHL912, using Tf2 probes. YHL912 was digested with *BsrGI* and hybridized with ^{32}P-labeled Tf2 5′-end probe (lanes 1 and 2) or 3′-end probe (lane 3). Hybridizing bands in the blot probed with the 5′-end probe are assigned letters a to l, with numbers being used when doublets are present. Hybridizing bands in the blot probed with the 3′-end probe are assigned numbers 1 to 12, with letters being used when doublets are present. T, tandem band (see the legend to Fig. 4). (C) DNA blot analysis comparing YHL912 and two Tf2-*neo* mobilization events, using the Tf2 3′-end probe (a) or the *neo* probe (b). Event 1-1 has been identified as a recombination event (REC) because it shifts an endogenous band (band 2a or b); event 1-2 has been identified as a possible transposition event (TXP) because it generates a new band with no shifts. The *neo* probe hybridized to the same bands.

position event, a similar DNA blot analysis was performed with DNA from cells that had undergone mobilization events generated by both Tf2-*neo* and Tf2INFS-*neo*. This analysis is depicted in Fig. 5A. When total genomic DNA from the parent strain YHL912 was cut with *BsrGI* and subjected to DNA blot analysis, using either a Tf2 3′-end probe or a Tf2 5′-end probe, two characteristic patterns of Tf2-specific bands were observed (Fig. 5B). Based on the analysis performed with the parent strain DNA, the effect of several different homologous recombination events or a true transposition event on these patterns was predicted (Table 3). Since both linear and putative circular cDNA species are generated when Tf2-*neo* or Tf1-*neo* expression is induced (see below), predictions of homologous recombination events involving either linear or circular cDNA species as donors and chromosomal elements as recipients were

made. Since the frameshift in the Tf2-*neo* IN domain was made by filling in the same *BsrGI* site used to cut the DNA, the possible outcomes for recombination with either a *BsrGI* site-containing donor cDNA or cDNA lacking that site were considered.

Taking the simplest event type, the use of the wt Tf2-*neo* linear cDNA as a substrate, as an example, the following analysis could be performed. If a forward transposition event occurs, a new Tf2-specific band should be detected with both Tf2 5′-end and 3′-end probes; however, if a homologous recombination event involving the linear cDNA and a single Tf2 element in the chromosome occurs, transferring the Tf2-*neo* sequence to an endogenous element, one of the endogenous Tf2 bands should shift upward by an amount corresponding to about 1 kb—the size of the *neo* gene—on a 3′-end probe blot, while the "new" 5′-end fragment should remain the same. An

TABLE 3. Distribution of Tf2-*neo* and Tf2INFS-*neo* mobilization events

Event class or subclass[a]	Tf2 wt events[b]			Tf2 IN frameshift events[c]		
	No. observed	Mobilization		No. observed	Mobilization	
		% of class	% of total[h]		% of class	% of total[h]
Transposition	7	ND[f]	33.3	0	ND	0.0
HR	14	ND	66.7	20	ND	100.0
LCR	8	57.1	38.1	8	40.0	40.0
Linear cDNA conversion	6	42.8	28.6	10	50.0	50.0
Difficult to interpret	0	0	0	2	10.0	10.0
Multiple events	1[e]	ND	0	0	ND	0
No observable event	0	ND	0	0	ND	0
HR targeting of tandem arrays	2 or 10[d]	ND	ND	5	ND	25.0
LCR	(0–8)[g]	ND	ND	4	80.0	20.0
Linear cDNA	2	ND	9.5	1	20.0	5.0

[a] HR, homologous recombination; LCR, LTR circular recombination.
[b] $n = 21$.
[c] $n = 20$.
[d] LCR events with Tf2-*neo* cDNA (wt) generate tandem arrays; thus, targeting to endogenous tandem arrays does not cause loss of the tandem band, making it impossible to determine by this analysis whether an event occurred at the tandem array or elsewhere. Because of the filled-in *Bsr*GI site in the IN frameshift, however, Tf2INFS-*neo* targeting of the tandem array can be determined when it causes the loss of the *Bsr*GI site and, hence, the loss of the tandem band.
[e] Event 6.
[f] ND, not determined.
[g] Zero to eight HR events may represent LCR type recombination with the tandem array.
[h] (No. observed/n) × 100.

example of a transposition event (event 1-2) and an example of a simple recombination event (event 1-1) are shown in Fig. 5C, panel a. (Note that if a Tf2-*neo* integrates fortuitously into one of the Tf2-containing *Bsr*GI fragments, a shift in the size of an endogenous band will also occur; however, this should be a shift down, and an additional, new band should always be evident.) This analysis can be confirmed by reblotting with the *neo* probe, as shown in Fig. 5C, panel b.

The results of analyzing mobilization candidates generated by wt Tf2-*neo* and Tf2INFS-*neo* are tabulated according to type in Table 3. Seven of 21 candidates generated by wt Tf2-*neo* showed banding patterns consistent with a possible transposition event; the rest showed banding patterns more consistent with gene conversion of an endogenous Tf2 element by linear Tf2-*neo* cDNA or a single crossover between a Tf2-*neo* circular cDNA species and an endogenous element. None of the 20 events generated by the Tf2INFS-*neo* mutant showed a banding pattern consistent with a true forward transposition event.
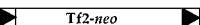
The tandem array of two Tf2 elements in the YHL912 genome was targeted for homologous recombination by 5 of the 20 Tf2 INFS mobilization events analyzed. Since the analysis of the Tf2 banding pattern in genomic blots of YHL912 DNA cut with different restriction enzymes (Fig. 4) led to an estimate of 15 Tf2 elements in the YHL912 background, the tandem array, representing about 13% of the possible Tf2 targets, was subjected to approximately 25% of the hits and thus might represent a recombinational "warm spot" in the *S. pombe* genome.

Tf2-*neo* can generate transposition events. What appeared in genomic DNA blot analyses to be Tf2-*neo* transposition events might actually have represented Tf2-*neo* cDNA recombination into endogenous solo LTRs or other sites in an IN-independent fashion, events that would have a restriction pattern indistinguishable from that of a real transposition event. To determine whether any of the candidates identified by
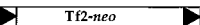
DNA blot analysis were the result of true transposition events—that is, that they represented the integration of a Tf2-*neo* sequence into a naive site (one lacking an endogenous Tf2 sequence), with hallmark TSDs flanking the element LTRs—candidate transposition events were cloned out. The genomic DNA flanking the element was sequenced to look for TSDs and to determine whether a corresponding empty site existed in the parent strain YHL912 (Fig. 6). Two of the seven candi-

A.

WT 3

GTAATGATAGTTAGTATA▶ Tf2-*neo* ▶GTATATTTACGAGACGAA

WT22

GGGCTAAATTGTATATTA▶ Tf2-*neo* ▶TATTATGAGAAATAAGAAG

B.

WT 3   GTAATGATAGTTAGTATATTTACGAGACGAA

WT22   GGGCTAAATTGTATATTATGAGAAATAAGAAG

FIG. 6. Cloned Tf2-*neo* transpositions. (A) Sequences of two transpositions with flanking genomic DNA. Tf2-*neo* candidate transpositions were cloned into pBSIISK−, and DNA flanking the event was sequenced, by using both the Tf2 and pBSII primers. WT22 shows the sequence obtained directly flanking either end of Tf2-*neo*, including TSDs (underlined). For WT3, the sequence flanking the 5′ LTR, including the putative TSD, was obtained; however, the sequence immediately flanking the 3′ LTR actually represents the predicted sequence at the site of insertion, based on identity of the 5′ flanking sequence to an *S. pombe* sequence in the Sanger Centre database. Thus, the indicated TSD and 3′ flanking sequence is predicted. (B) YHL912 sequences at Tf2-*neo* insertion sites. By using the *S. pombe* sequence database, the target site for each transposition event was found to lack Tf2 homology, having neither a Tf2 LTR nor a full-length Tf2; this was confirmed by PCR analysis of the genomic site of insertion (data not shown). Thus, WT22 and WT3 represent true transposition events.

date transposition events identified by Southern analysis, WT3 and WT22, were recovered as full-length clones with 5′-end and 3′-end flanking sequences. Restriction analysis diagnosed the presence of the Tf2-*neo* sequence, and sequence analysis with T3, T7, and/or Tf2-*neo*-specific primers directly revealed hallmark TSDs for WT22 and a presumed TSD (from one flanking sequence) for WT3. Subsequent BLAST searches of the Sanger Centre *S. pombe* sequence database, using the flanking DNA sequence, revealed the empty sites: a WT3 event flanking sequence is found on cosmid c9G1 from chromosome I (Chr I), and a WT22 event sequence is found on cosmid c17H9 (also from Chr I) (Fig. 6A); both sequences lack any Tf2 homology at the point of integration (Fig. 6B). The absence of preexisting LTRs or full-length elements at the point of integration in parent strain YHL912 (a derivative of strain 972) was directly confirmed for both events by PCR analysis of genomic DNA from this strain, using primers derived from the genomic sequence flanking each cloned Tf2-*neo* insertion; the products were of the expected size for a sequence lacking a preexisting Tf2 insertion (data not shown). Thus, Tf2-1 is capable of retrotransposition, albeit at a lower frequency than Tf1.

Analysis of mobilization steps: protein expression. Although Tf2-1 is capable of mobilization, its ability to generate true retrotransposition events was greatly reduced compared to that of Tf1 under the conditions tested. Thus, we attempted to elucidate the steps at which the Tf2 retrotransposition pathway is inhibited and to determine what directs Tf2 to an IN-independent pathway for mobilization. Since mRNA expression was similar for Tf2-*neo* and Tf1-*neo*, protein expression was examined.

Tf1-107 encodes a 1,330-amino-acid polyprotein, which is cleaved in a PR-dependent manner to release mature Tf1 CA (27 kDa), PR, RT, and IN (56 kDa) (34). By using anti-Tf1 IN polyclonal Abs, two intermediate species released during processing are detectable as well, a PR-RT-IN intermediate (125 kDa) and an RT-IN intermediate (110 kDa) (3). The nearly 100% identity of Tf2 to Tf1 in the IN domain allowed the use of this antiserum to detect the presence of Tf2-encoded proteins. To compare Tf2 and Tf1 protein expression, strains harboring the *pnmt*Tf2 and *pnmt*Tf1 (no *neo*) plasmids were grown under repressing and inducing conditions, and total soluble protein extracts prepared from these cells were separated by SDS-polyacrylamide gel electrophoresis and subjected to immunoblot analysis (Fig. 7A). The production of a Tf1-like IN species by Tf2 suggested that Tf2 expresses a full-length ORF and proteolytically processes it in a Tf1-like manner. The apparent difference in the amounts of IN produced by the two elements was confirmed by quantitative immunoblotting analysis of Tf1 and Tf2 IN species, which indicated that the Tf2 IN levels were reduced two- to fourfold relative to Tf1 IN levels (22).

Tf2 produces a PR-RT fusion protein. The use of an antiserum raised against Tf1 RT, which is greater than 99% identical to Tf2 RT, revealed differences in Tf2 and Tf1 protein processing. In the first direct detection of a Tf RT species, immunoblotting of Tf2-*neo* and Tf1-*neo* protein samples with anti-Tf1 RT antiserum demonstrated the presence of a mature Tf1 RT species of the expected size, 60 kDa (Fig. 7B, panel b). However, both this blot and a duplicate blot incubated with anti-Tf1 IN (Fig. 7B) showed that Tf2 does not make a detectable mature RT or RT-IN intermediate, molecules readily observed at 60 and 120 kDa in the Tf1 lanes (Fig. 7B, panel b). Instead, a PR-RT intermediate of 72 kDa (panel b) was observed. The 72-kDa species was also observed in the Tf1 lanes, but its intensity was lower than that of the mature RT. These

data suggest that Tf2 uses a proteolytic processing pathway different from that used by Tf1 (Fig. 7C).

The presence of a Tf1-like PR-RT-IN species (and no larger species) in the Tf2 protein samples indicated that Tf2 expresses its entire ORF and cleaves off its CA species appropriately; however, it could not be assumed that the free Tf2 CA species was stable. Since CA presumably plays a role in organizing the other retrotransposon proteins to make a replication intermediate, besides any other functions it might fulfill (44), it was of interest to determine whether Tf2 produces a stable CA species. Blotting Tf2 protein samples with anti-Tf1 CA (Fig. 7A, right panel) demonstrated that the anti-Tf1 CA Ab does not cross-react with Tf2-encoded proteins. Therefore, all of the Tf2 sequence upstream of the ORF was replaced with a hemagglutinin (HA) tag, making an N-terminal fusion, as was previously done for Tf1 (34); both the HA-Tf2 and HA-Tf1 constructs were then placed under the control of the *nmt* promoter (22). The HA-tagged proteins were expressed in *S. pombe*, and proteins of the expected size were observed in both cases. Quantitative immunoblotting experiments indicated that the Tf2 HA-CA proteins were two- to fourfold less abundant than the corresponding Tf1 species, as was also true for the IN species from both the HA-tagged species and the wt elements, indicating that the observed difference in the levels of Tf1 and Tf2 proteins is a posttranslational effect (22).

Tf2-*neo* produces fourfold less cDNA than Tf1-*neo*. Another crucial step in the retrotransposon life cycle that was examined with regard to phenotypic differences between Tf2 and Tf1 is the production of the full-length double-stranded cDNA genome. It has been demonstrated that overexpression of Tf1-*neo* from the *nmt* promoter produces easily detectable levels of full-length Tf1-*neo* cDNA intermediates (3). Detection and comparison of the levels of cDNA intermediates produced by Tf2-*neo* and Tf1-*neo* might therefore also offer insight into the reduced mobilization frequency of Tf2-*neo*.

A schematic of the cDNA blot analysis is shown in Fig. 8A. Digestion of the total DNA of cells induced for Tf-*neo* expression with a single-cut restriction enzyme leads to production from the linear cDNA of a 3′-end fragment of either 2.5 kb (*Bsr*GI) or 2.0 kb (*Bst*XI) that is detectable with the *neo* probe; this is easily differentiated from the Tf-*neo* plasmid band on the basis of size. The production of linear cDNA can be quantitatively compared by measuring the signal from the 3′-end fragment, normalizing it to the plasmid signal in each lane (the production of cDNA should be proportional to the number of plasmid molecules within each cell, internally controlling for loading differences), and then comparing the normalized numbers between samples.

The results of such an analysis are also presented in Fig. 8. A representative blot is shown in Fig. 8B. In addition to the predicted linear cDNA and plasmid band fragments, a third fragment, 3 to 4 kbp larger than the linear cDNA 3′-end fragment, was detected (marked "C"). In the case of Tf1, the source of this third fragment has been identified as a single-LTR-containing species of unknown structure, most likely a single-LTR circle or a tandem array (37). When production of cDNA was halted by the addition of thiamine to the medium of log-phase cells, the linear species had a shorter half-life than the larger species, suggesting that stability was conferred by the circular nature of the species (data not shown). The observation by DNA blot analysis of recombination events most simply explained by recombination with a circular species (Table 3) also supports the existence of circular recombination intermediates, and therefore the third species will be referred to as the circular cDNA species.

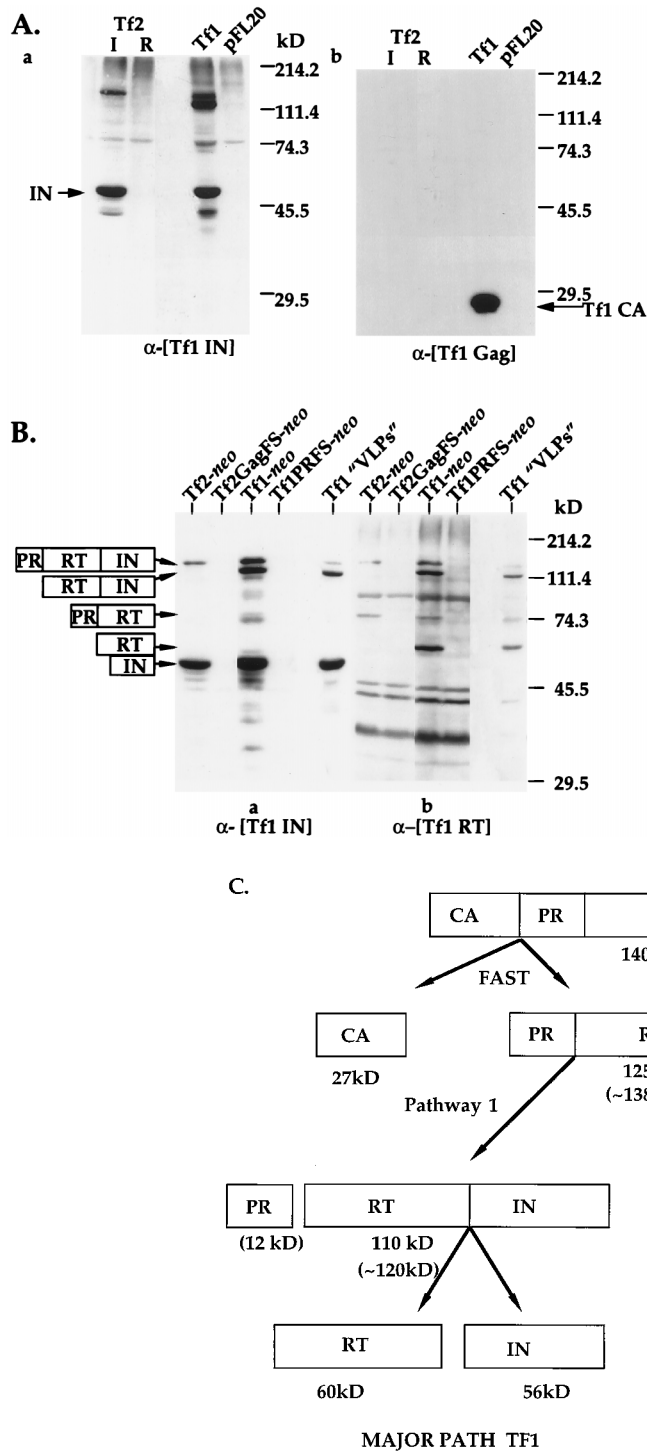Comparison of normalized cDNA values for Tf1-*neo* and

FIG. 7. Tf2 protein production and processing. (A) Tf2 produces and processes a polyprotein. Shown are the results of an immunoblot analysis of total soluble protein from strains expressing Tf2 or Tf1 from the *nmt* promoter. Tf1 and vector samples were taken from gradient-purified protein. Induced Tf2 samples show both a high-molecular-weight intermediate and a Tf1-like IN species (Tf2 IN) when blotted with antibody to Tf1 IN (α-[Tf1 IN]). Panel a was blotted with anti-Tf1 IN; panel b was blotted with anti-Tf1 Gag polyclonal Ab. (α-[Tf1 Gag]). I, induced, R, repressed. (B) Tf2-*neo* has a proteolytic processing defect. An immunoblot analysis of total soluble protein from strains expressing Tf2-*neo* or Tf1-*neo* constructs, using anti-Tf1 IN (a) or anti-Tf1 RT (b), is shown. Tf2 produces mature IN and PR-RT-IN but does not produce a detectable RT species or RT-IN species. A Tf2-*neo* species corresponding to the predicted size for a PR-RT intermediate is detectable in the RT panel; this same species is detectable in Tf1-*neo* protein samples as well, but at a much lower intensity than that for mature RT. Tf1 "VLPs", transposition intermediates (putative VLPs) isolated as described by Levin et al. (34). (C) Polyprotein processing pathways for Tf elements. Based on the protein species detected in panel B, two alternative pathways for Tf polyprotein processing can be envisioned. Both involve initial cleavage of the CA from the polyprotein, but they differ in the order of the second and third cleavages. Values in parentheses indicate molecular masses observed in this study, versus the published sizes (3).

Tf2-*neo* revealed that Tf1-*neo* produces about fourfold more cDNA than Tf2-*neo*. This is depicted graphically in Fig. 8C as the Tf1-*neo*/Tf2-*neo* cDNA species ratio for two experiments. Values for each of the cDNA species were compared, and a similar reduction in cDNA signal for Tf2-*neo* was observed.

Finally, cDNA production by the Tf2 IN frameshift was also compared to that of wt Tf2. The wt Tf2 and Tf2INFS mutant made similar amounts of total cDNA, as has been observed for

Tf1 (2), indicating that neither the absence nor the presence of Tf2 IN affects the production or stability of Tf2-*neo* cDNA (data not shown).

Comparison of Tf2-1 with other Tf2 elements. The differences in the Tf2-1 and Tf1-107 mobilization phenotypes led us to question whether the Tf2-1 clone is typical of the Tf2 group. The other element analyzed in the original sequence analysis of Tf2, Tf2-43, showed only a 1-base difference in the region of
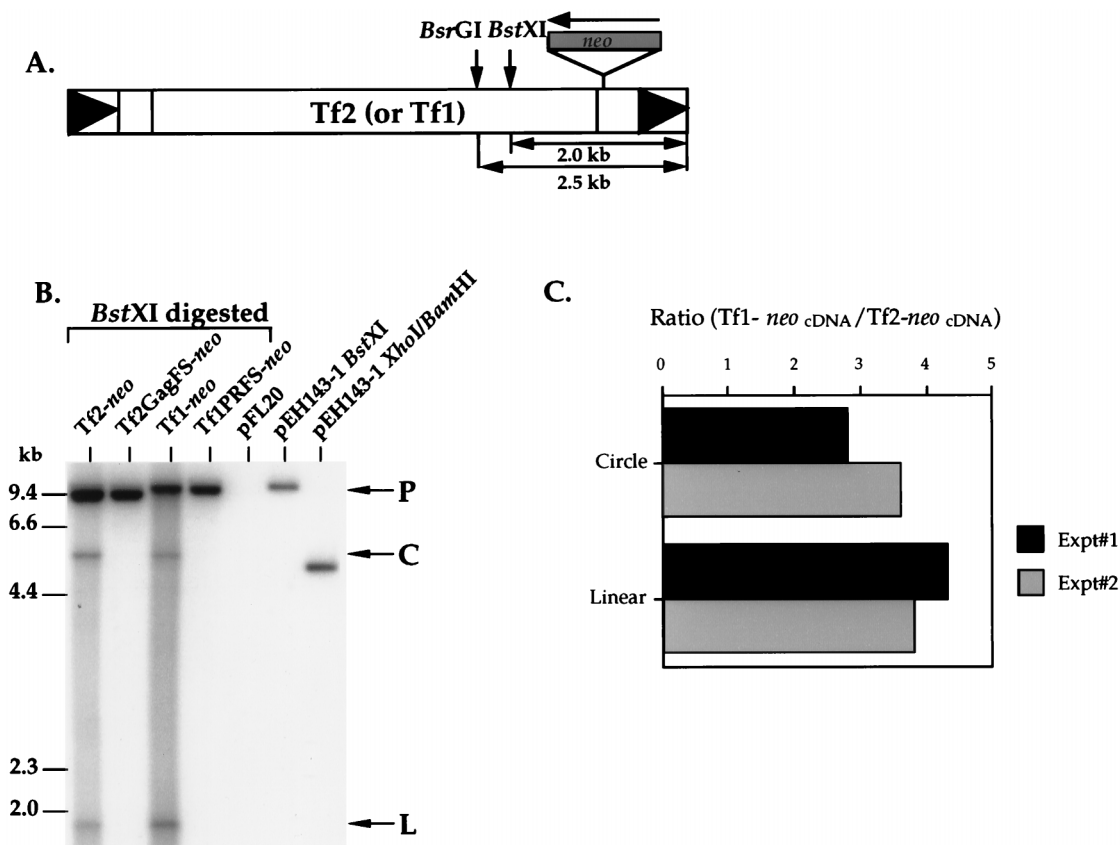
FIG. 8. Tf2-*neo* produces a steady-state level of cDNA about fourfold lower than that of Tf1-*neo* in log-phase cells. (A) Scheme of cDNA analysis by DNA blotting. Previously described by Atwood et al. (3), this scheme relies on easy detectability of cDNA in total-DNA preparations from induced cells harvested in log phase. The diagram shows the Tf-*neo* cDNA genome with *Bsr*GI and *Bst*XI sites; cutting with these enzymes produces a 2.5- or 2.0-kbp cDNA-specific fragment detectable with the *neo* probe. Boxed triangles represent LTRs. (B) DNA blotting of Tf2-*neo* and Tf1-*neo* samples cut with *Bst*XI. DNA preparations from cells expressing Tf2-*neo*, Tf2GagFS-*neo*, Tf1-*neo*, or Tf1PRFS-*neo* were cut with *Bst*XI, resolved by agarose gel electrophoresis, subjected to blotting, and then hybridized with the *neo* probe. Linear (L) and circular (C) cDNA species are apparent only when functional constructs are induced. P, plasmid band. pEH143-1 *Bst*XI and pEH143-1 *Xho*I/*Bam*HI are size standards. (C) Quantitative comparison of cDNA species, normalized to plasmid levels. *neo*-hybridizing species were quantitated by phosphorimage analysis, and the values obtained for the cDNA (L and C) species were normalized to those of the plasmid species (P) to enable the comparison of samples on the same blot. Normalized values are presented as a ratio of Tf1-*neo* to Tf2-*neo* for the linear cDNA and circular cDNA species.

overlap, occurring in the U3 region of the LTR (56). The 1,673 bp of Tf2-43 sequenced covers the entire region exhibiting the greatest degree of diversity between Tf2 and Tf1, suggesting that the Tf2 elements are conserved. However, given the estimated 15 Tf2 elements in strain YHL912, it was also possible that a subclass of Tf2 elements possessed a high transposition efficiency; if so, mobilization differences could be linked to crucial differences between Tf2-1 and one of these elements in either the coding or noncoding sequences.

Changes in the Tf2 mobilization phenotype would be easily detectable by the qualitative mobilization assay. Hypothesizing that differences in the *trans*-acting factors encoded by Tf2 and Tf1 are linked to Tf2's lower-level mobilization phenotype, we attempted to rescue the Tf2-1 mobilization phenotype by patching in coding sequences from other Tf2 elements (Fig. 9A). We first tried sampling genomic Tf2 sequences by simple gap repair of the *pnmt*Tf2-*neo* plasmid; this approach depended on the homologous repair of double-stranded DNA breaks observed in both *Saccharomyces cerevisiae* (45) and *S. pombe* (19, 20). pEH143-1 was "gapped" in the Tf2 region by excising the sequence extending from the *Bam*HI site in the middle of CA to the *Bsr*GI site at the 5′ end of the IN domain. The digestion products were transformed into YHL912, and repair and/or plasmid integration events were detected by se-

lecting for Ura[+] colonies and then selecting for G418[r] through replica printing; episomal Ura[+] phenotype was detected by replica printing to medium containing FOA. Only about 25% of the Ura[+] G418[r] candidates were also Foa[R], indicating a selection bias toward recombination products resolved into the chromosome. Thirty-one of the Ura[+] G418[r] candidates that were also Foa[R] were then tested in the mobilization assay and compared to Tf2-1; a subset of these is shown in Fig. 9B. Twenty-eight candidates had either the same or a worse mobilization phenotype than Tf2-1 (the latter phenotype indicating either incorrect repair or repair by defective Tf2 elements [see below]). Three candidates appeared to exhibit various degrees of rescue of the mobilization phenotype, but further genetic screens revealed that the G418[r] phenotype in these isolates was unlinked to the plasmid-borne copy of the element. In a similar experiment, repaired plasmids were rescued from 4 of 16 candidates tested by the mobilization assay and shown to have an overall structure like that of pEH143-1; the sequences on these plasmids conferred a mobilization phenotype either similar to (2 of 4) or worse than (2 of 4) that of Tf2-1 (reflecting the distribution of mobilization phenotypes in these 16 candidates) (data not shown). Thus, random sampling of the sequence between the *Bam*HI and *Bsr*GI sites did not
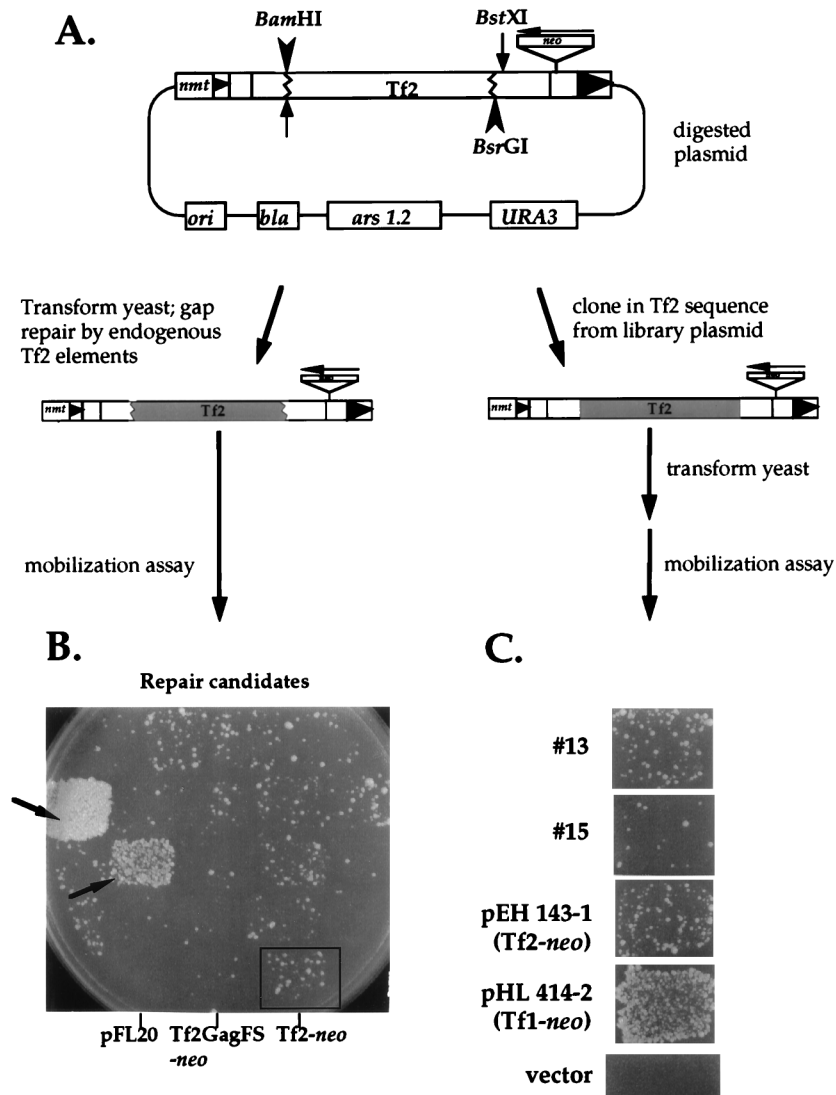
FIG. 9. Functional assays of other Tf2 elements. (A) Scheme used to assay for activity of other Tf2 elements. Either the Tf2-*neo* plasmid was cut with *Bam*HI and *Bsr*GI (arrowheads) and transformed directly into YHL912, to allow for homologous repair of the sequence between the *Bam*HI and *Bsr*GI sites by endogenous Tf2 elements (gray shading), or Tf2 sequences from library plasmid clones were subcloned into the plasmid's *Bam*HI-*Bst*XI sites (arrows). Jagged lines are used for gap-repaired Tf2 to indicate that we do not know the points of crossover in the repaired plasmids. (B) Mobilization assay of candidates from the gap repair transformation. Candidates were chosen for being Ura⁺, G418ʳ, and also Foaᴿ in the absence of selection for a Ura⁺ phenotype; a subset is shown (see text). Most candidates did not show an improved mobilization phenotype when compared to Tf2-1 (box), as judged by the degree of papillation on G418-containing medium. Two of three candidates isolated that were identified as having an apparently better mobilization phenotype than Tf2-1 are shown on this plate (arrows). In all three cases, this phenotype was subsequently found to be independent of the *URA3* plasmid. (C) Mobilization assay of strains transformed with plasmids carrying Tf2 sequences from random Tf2-containing library plasmids. Neither of the library clones tested (no. 13 and no. 15) exhibited an increased mobilization frequency associated with the sequence between the Tf2 *Bam*HI and *Bst*XI sites.

reveal any Tf2 elements with an improved mobilization phenotype.

Two random library clones containing Tf2 element sequences from strain 972 (35) were also tested for their ability to alter Tf2-1 mobilization. The sequences between the *Bam*HI and *Bst*XI sites of these clones were ligated into the same sites in the Tf2INFS-*neo* plasmid pEH546-4; this plasmid was used as the recipient to differentiate between the (presumably) wt sequence being cloned in and the recipient plasmid (Fig. 1B and 9A). The resulting hybrids were tested in the mobilization assay; one mobilized *neo* information into the genome at a level similar to that for Tf2-1, while the other appeared to possess a mutation rendering it inactive in the mobilization assay (Fig. 9C).

These functional analyses tested only a subset of the Tf2 sequence and assumed that differences in the ORF might determine the mobilization phenotype. Six full-length Tf2 sequences from strain 972 besides that of Tf2-1 have been submitted to GenBank. At the nucleotide level, very few differences are found in the sequences; their identities range from 98 to 100% (Table 4). One clone, SPAC19D5 Tf2, appears to be the 5′ member of the one-LTR tandem repeat (only a partial sequence of the downstream member was included in the cosmid sequence). Some of the nucleotide changes affect *cis*-acting domains of Tf2. There are some small internal microhomology-dependent deletions in the U3 region of the 5′ LTR of one of the elements (SPAC8E4 Tf2), including deletion of one of the 9-bp repeats upstream of the TATA

TABLE 4. Properties of other sequenced Tf2 elements from strain 972: DNA sequence analysis and comparison to Tf2-1

Cosmid	Nucleotide coordinates	Length (nucleotides)	% Nucleotide identity to Tf2-1	ORF length[a]	% Amino acid identity to Tf2-1	TSD
SPAC8E4	26021–30886	4,866[b]	98	1,333	99.4	No (5'GTAGC, 3'CTTTC)
SPCC1020	12250–17165	4,916	100	1,333	100.0	Yes (CTTAA)
SPAC26A3	27042–31956	4,915	99.9	1,333	99.9	Yes (CTTAA)
SPAC19D5	318–5233	4,916	99.8	1,333	99.8	Unknown; appears to be part of a one-LTR tandem array
SPAC27E2	21407–26322	4,916	100	1,333	100.0	Yes (TTTTA)
SPBC1E8	8941–13855	4,915	99	—[c]	NA[d]	Yes (TATAA)

[a] Amino acid residues.
[b] 5' LTR, 298 bp; internal deletions in LTR.
[c] C deleted at nucleotide 2601 (Tf2-1 position)—generates a frameshift; with C the ORF encodes 1,333 amino acids.
[d] NA, not applicable.

box which may have a role in transcription (35). This same element has two changes affecting the Tf self-priming structure (37), actually changing Tf2-specific nucleotides to Tf1-like nucleotides, and the SPAC26A3 Tf2 has a nucleotide missing in the element; however, these changes all occur in the hairpin of the self-priming structure, and deletion of this hairpin was not observed to have any effect on Tf1 mobilization (37).

Analysis of the translation products revealed some amino acid differences between Tf2-1 and four of the other Tf2 elements (Table 5). All of the observed differences fall within the coding sequence between the *Bam*HI and *Bsr*GI sites, indicating that for at least 42% of the genomic Tf2 elements other than Tf2-1 (6 of 14) we tested the appropriate region in our functional analyses. One notable change occurs in the YXDD motif of the RT domain: two of the elements, SPAC8E4 and SPAC19D5, have suffered a G-to-A change in the nucleotide sequence, resulting in a D-to-N amino acid change at the second D of the motif; this change alters the highly conserved RT active site and, when introduced into the Tf1-*neo* sequence, abrogates mobilization (32). Curiously, these two elements share two other amino acid changes as well, suggesting that the YXDN mutation was actually propagated, possibly through homologous recombination between genomic elements.

DISCUSSION

This study was undertaken to determine whether Tf2 is a functional retrotransposon and to compare and contrast its transposition capability with that of the related element Tf1. We have shown that Tf2 can be stably expressed from the *nmt* promoter and that it can mediate the occurrence of true transposition events in the host cell chromosome. However, the

TABLE 5. Properties of other sequenced Tf2 elements from strain 972: amino acid differences among Tf2 elements with full-length ORFs

Tf2-1 residue	Domain	Tf1 residue	Altered residue (other Tf2 copy or copies)
T 114	CA	M 106	P (SPAC8E4)
P 206	CA	T 203	L (SPAC26A3)
S 215	CA	N 212	F (SPAC8E4)
S 217	CA	F 214	P (SPAC8E4)
A 274	PR	A 271	T (SPAC8E4)[a]
T 540	RT	T 537	I (SPAC19, SPAC8E4)
A 556	RT	A 553	V (SPAC19, SPAC8E4)
D 567	RT	D 567	N (SPAC19, SPAC8E4)[b]

[a] DTGA→DTAT mutant; DTA is the PR catalytic triad.
[b] Alters RT active site (YXDD to YXDN).

frequency of Tf2-*neo* mobilization is about 20-fold lower than the Tf1-*neo* mobilization frequency. Furthermore, unlike what is observed with Tf1-*neo* or any other wt retrotransposon, the majority of Tf2-*neo* mobilization events result from homologous recombination between the Tf2-*neo* cDNA and endogenous Tf2 elements rather than from true transposition.

Tf2-*neo* mobilizes at a level similar to that seen with the Tf1-*neo* IN mutant. The Tf2-*neo* IN mutant, however, mobilizes at about the same frequency as wt Tf2-*neo* rather than dropping 15- to 100-fold, as has been observed for the corresponding Tf1-*neo* constructs (31, 37). The majority (70%) of the wt Tf2 mobilization events are homologous recombination events, indicating that integration is somehow blocked. Attempts to elucidate the reasons for Tf2's lower level of mobilization by analyzing mobilization components revealed that Tf2 has a different mode of proteolytic processing that results in accumulation of a PR-RT species and no detectable mature RT. Tf2 produces steady-state levels of IN that are two- to fourfold lower than those observed for Tf1, and it appears that its CA species is maintained at a similarly low level. Finally, Tf2 also produces fourfold less cDNA than Tf1. Based on both functional and sequence analyses of other Tf2 elements, we believe that these phenotypes are typical of these elements.

Tf2 proteolytic processing. Tf2 processes its polyprotein differently than Tf1. The Tf2 PR produces CA, IN, and a PR-RT species but does not release detectable mature RT. Tf1 makes the PR-RT species seen in Tf2 protein samples, but at a much lower level than it makes mature RT. Building on the previous model for Tf1 polyprotein cleavage (3), these data suggest that there are two alternative pathways for Tf polyprotein proteolytic processing: one in which the PR processes the CA-PR junction, the PR-RT junction, and then the RT-IN junction (favored by Tf1) and one in which the second cleavage occurs at the RT-IN junction (favored by Tf2). A key unknown is whether the PR-RT species produced in the second pathway is intrinsically refractory to cleavage by either Tf element PR or the Tf2 PR is simply unable to cleave the PR-RT junction in any context. It is also formally possible that Tf2 PR cleaves the PR-RT junction but an aberrant N terminus is generated and the resulting Tf2 RT and RT-IN proteins are highly unstable and, therefore, undetectable.

What is unique about Tf2 polyprotein processing is that in most retrotransposons, mutations that disrupt processing of the polyprotein usually eliminate the mobilization activity (11, 28, 40, 42, 57). Tf2 is the only retroelement we know of that produces its PR and RT as a single species and still maintains significant mobilization activity.

Tf2 recombination phenotype. Tf2-*neo* not only has a decreased mobilization frequency but also moves mostly by re-

combination of Tf2-*neo* cDNA with endogenous Tf2 elements. This recombination of cDNA with endogenous elements has been observed for other yeast retrotransposons (2, 24, 31, 39, 52), but never as the major mobilization pathway for the wt element in question. It is formally possible that the high proportion of cDNA recombination events occurs as a result of overexpression of the Tf2-*neo* element; however, yeast retrotransposon studies using overexpression systems have generally reflected the biology of the endogenous elements (15, 26).

When *neo* mobilization events generated by Tf2-*neo* and the Tf2INFS-*neo* mutant were subjected to genomic DNA blot analysis, none of the 20 IN frameshift events looked like transposition events, whereas 7 of the 21 Tf2-*neo* events looked like transposition events and two were confirmed as such. This indicated that although recombination is the primary pathway used by Tf2 cDNA for mobilization into the host genome, simple recombination into endogenous solo LTRs, an event that would produce a pattern indistinguishable from that of true transposition, does not occur at a significant frequency (<5%).

Affirming our observations by using Tf2-1, there are very few differences between this clone and six other full-length Tf2 elements whose sequences have been submitted to GenBank. Notably, none of the differences occur in IN, and only one Tf2 has a change in PR, although from our Tf2-1 data we might have expected that in one of these domains, Tf2-1 has a specific mutation that is deleterious to mobilization. We have also analyzed *S. pombe* wild strain NCYC132 Tf2 sequences amplified by PCR in the region between the transcription start point and the middle of PR; these also did not show any significant differences from the Tf2 elements in strain 972 (22a). This indicates that Tf2 sequences of at least one wild strain and lab strain 972 are highly conserved.

Part of our investigation involved comparing Tf2 to Tf1. Although the two Tf element transcripts are expressed at similar levels from the *nmt* promoter, the two elements differ markedly in their downstream phenotypes, from protein and cDNA expression to mobilization phenotype. These observations prompted the proposal of several hypotheses to explain the lower Tf2 mobilization levels, though not the lack of integrase dependence, including a simple reduction in the number of particles available to mobilize cDNA into the chromosomes. However, these observations also directed us to investigate Tf2 by using Tf1/Tf2 chimeras as diagnostic tools. Reciprocal swaps of the Tf1 and Tf2 IN domains clearly demonstrated that the Tf2 mobilization phenotype is not the result of differences in the IN domains. Instead, a far more complex picture of Tf mobilization has emerged, in work to be presented elsewhere (22a).

It is interesting to speculate how the Tf2 mobilization phenotype may have evolved. The presence of multiple full-length copies of the Tf2 element in all strains of *S. pombe* examined (including the lab strain 972, which lacks full-length Tf1 elements) clearly indicates that Tf2 is an evolutionarily successful retrotransposon. Perhaps a progenitor of the Tf2 element acquired a mutation(s) that caused the high recombination-low transposition phenotype and this then converted all of the other Tf2 elements to mutant elements with the same hyper-recombination phenotype.

The presence of solo Tf1 LTRs in strain 972 (22a, 35) indicates that many Tf1 and Tf2 elements coexisted in a direct ancestor of the Leupold strain background, as they still do in wild strains. If the Tf2 elements were engendered as integration defective de novo, they might have been mobilized in *trans* by Tf1 to allow the initial spread of Tf2 elements in the genome; Tf2 elements could then have replaced the Tf1 elements

through recombination. In support of this hypothesis, we have obtained genetic evidence that the Tf1 element can use the Tf2 RNA as a substrate for reverse transcription and mobilization into the genome, and vice versa (22a).

Finally, it is formally possible that Tf2 mobilization is developmentally or otherwise regulated and that Tf2 is capable of efficient transposition when in the presence of a host factor not appropriately available in mitotically dividing haploid cells grown under standard laboratory conditions. The differences between the Tf CA proteins might thus be symptomatic of different requirements for host factor interactions that could enhance virus-like particle formation or stability. Also, the differences between the U3 domains of the Tf1 and Tf2 LTRs, in which lie most of the transcription-regulatory elements, might cause the two elements to be expressed differently at different times during the cell cycle or during development. Transcripts from both sets of endogenous Tf elements have been observed in log-phase haploid cells from the wild strains (35), suggesting that if Tf2 mobilization is regulated differently from Tf1 mobilization, enhancement of expression or a post-translational requirement for a host factor or physiological state is the more likely explanation for the differences in mobilization phenotype.

Conclusions. The Tf2-1 element from Leupold strain 972, when marked and overexpressed in this strain background, exhibits reduced protein expression, cDNA generation, and mobilization compared with those of its sister element, Tf1. The majority of mobilization events that Tf2-1 can mediate involve homologous recombination into preexisting elements. The use of such a propagation pathway may aptly be named integration site recycling. This route of mobilization saves the host cell genome from a potentially lethal mutation resulting from fresh Tf2 integrations in or near an essential gene(s) while still enabling the Tf2 elements to evolve. As such, it represents a highly adapted relationship between the Tf2 retrotransposon and its fungal host, *S. pombe*, similar to what is seen with the Ty elements in *Saccharomyces cerevisiae*.

REFERENCES

1. **Ajioka, J. W., and D. L. Hartl.** 1989. Population dynamics of transposable elements, p. 939–958. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
2. **Atwood, A., J. Choi, and H. L. Levin.** 1998. The application of a homologous recombination assay revealed amino acid residues in an LTR-retrotransposon that were critical for integration. J. Virol. **72:**1324–1333.
3. **Atwood, A., J.-H. Lin, and H. L. Levin.** 1996. The retrotransposon Tf1 assembles virus-like particles that contain excess Gag relative to integrase because of a regulated degradation process. Mol. Cell. Biol. **16:**338–346.
4. **Basi, G., E. Schmid, and K. Maundrell.** 1993. TATA box mutations in the *Schizosaccharomyces pombe* nmt1 promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. Gene **123:**131–136.
5. **Bilanchone, V. W., J. A. Claypool, P. T. Kinsey, and S. B. Sandmeyer.** 1993. Positive and negative regulatory elements control expression of the yeast retrotransposon Ty3. Genetics **134:**685–700.
6. **Boeke, J. D., and V. G. Corces.** 1989. Transcription and reverse transcription of retrotransposons. Annu. Rev. Microbiol. **43:**403–434.
7. **Boeke, J. D., D. Eichinger, D. Castrillon, and G. R. Fink.** 1988. The *Saccharomyces cerevisiae* genome contains functional and nonfunctional copies of transposon Ty1. Mol. Cell. Biol. **8:**1432–1442.
8. **Boeke, J. D., F. LaCroute, and G. R. Fink.** 1984. A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol. Gen. Genet. **197:**345–346.
9. **Boeke, J. D., and J. P. Stoye.** 1997. Retrotransposons, endogenous retroviruses, and the evolution of retroelements, p. 343–435. *In* H. Varmus, S.

Hughes, and J. Coffin (ed.), Retroviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

10. **Boeke, J. D., C. A. Styles, and G. R. Fink.** 1986. *Saccharomyces cerevisiae SPT3* gene is required for transposition and transpositional recombination of chromosomal Ty elements. Mol. Cell. Biol. **6:**3575–3581.

11. **Braiterman, L. T., G. M. Monokian, D. J. Eichinger, S. L. Merbs, A. Gabriel, and J. D. Boeke.** 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: phenotypic analysis. Gene **139:**19–26.

12. **Casacuberta, J. M., S. Vernhettes, and M. A. Grandbastien.** 1995. Sequence variability within the tobacco retrotransposon Tnt1 population. EMBO J. **14:**2670–2678.

13. **Chalker, D. L., and S. B. Sandmeyer.** 1992. Ty3 integrates within the region of RNA polymerase III transcription initiation. Genes Dev. **6:**117–128.

14. **Craigie, R.** 1992. Hotspots and warm spots: integration specificity of retroelements. Trends Genet. **8:**187–189.

15. **Curcio, M. J., and D. J. Garfinkel.** 1991. Single-step selection for Ty1 element retrotransposition. Proc. Natl. Acad. Sci. USA **88:**936–940.

16. **Devine, S. E., and J. D. Boeke.** 1996. Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. Genes Dev. **10:**620–633.

17. **Feinberg, A. P., and B. Vogelstein.** 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **137:**266–267.

18. **Fulton, A. M., P. D. Rathjen, S. M. Kingsman, and A. J. Kingsman.** 1988. Upstream and downstream transcriptional control signals in the yeast retrotransposon, Ty. Nucleic Acids Res. **16:**5439–5458.

19. **Grallert, B., P. Nurse, and T. E. Patterson.** 1993. A study of integrative transformation in *Schizosaccharomyces pombe*. Mol. Gen. Genet. **238:**26–32.

20. **Grimm, C., and J. Kohli.** 1988. Observations on integrative transformation in *Schizosaccharomyces pombe*. Mol. Gen. Genet. **215:**87–93.

21. **Hansen, L. J., and S. B. Sandmeyer.** 1990. Characterization of a transpositionally active Ty3 element and identification of the Ty3 integrase protein. J. Virol. **64:**2599–2607.

22. **Hoff, E. F.** 1997. Ph.D. thesis. Johns Hopkins University School of Medicine, Baltimore, Md.

22a.**Hoff, E. F., and J. D. Boeke.** Unpublished data.

23. **Jin, Y. K., and J. L. Bennetzen.** 1989. Structure and coding properties of Bs1, a maize retrovirus-like transposon. Proc. Natl. Acad. Sci. USA **86:**6235–6239.

24. **Ke, N., and D. F. Voytas.** 1997. High frequency cDNA recombination of the *Saccharomyces* retrotransposon Ty5: the LTR mediates formation of tandem elements. Genetics **147:**545–556.

25. **Keeney, J. B., and J. D. Boeke.** 1994. Efficient targeted integration at *leu1-32* and *ura4-294* in *Schizosaccharomyces pombe*. Genetics **136:**849–856.

26. **Kinsey, P. T., and S. B. Sandmeyer.** 1995. Ty3 transposes in mating populations of yeast: a novel transposition assay for Ty3. Genetics **139:**81–94.

27. **Kirchner, J., C. M. Connolly, and S. B. Sandmeyer.** 1995. Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. Science **267:**1488–1491.

28. **Kirchner, J., and S. Sandmeyer.** 1993. Proteolytic processing of Ty3 proteins is required for transposition. J. Virol. **67:**19–28.

29. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82:**488–492.

30. **Le, M. H., D. Duricka, and G. H. Karpen.** 1995. Islands of complex DNA are widespread in *Drosophila* centric heterochromatin. Genetics **141:**283–303.

31. **Levin, H. L.** 1995. A novel mechanism of self-primed reverse transcription defines a new family of retroelements. Mol. Cell. Biol. **15:**3310–3317.

32. **Levin, H. L.** 1996. An unusual mechanism of self-primed reverse transcription requires the RNase H domain of reverse transcriptase to cleave an RNA duplex. Mol. Cell. Biol. **16:**5645–5654.

33. **Levin, H. L., and J. D. Boeke.** 1992. Demonstration of retrotransposition of the Tf1 element in fission yeast. EMBO J. **11:**1145–1153.

34. **Levin, H. L., D. C. Weaver, and J. D. Boeke.** 1993. Novel gene expression mechanism in a fission yeast retroelement: Tf1 proteins are derived from a single primary translation product. EMBO J. **12:**4885–4895.

35. **Levin, H. L., D. C. Weaver, and J. D. Boeke.** 1990. Two related families of retrotransposons from *Schizosaccharomyces pombe*. Mol. Cell. Biol. **10:**6791–6798.

36. **Lin, J.-H., and H. L. Levin.** 1997. Self-primed reverse transcription is a mechanism shared by several LTR-containing retrotransposons. RNA **3:**952–953.

37. **Lin, J.-H., and H. L. Levin.** 1997. A complex structure in the mRNA of Tf1 is recognized and cleaved to generate the primer of reverse transcription. Genes Dev. **11:**270–285.

37a.**Losson, R., and F. Lacroute.** 1983. Plasmids carrying the OMP decarboxylase structural and regulatory genes: transcription regulation in a foreign environment. Cell **32:**371–377.

38. **Maundrell, K.** 1990. nmt1 of fission yeast. A highly transcribed gene completely repressed by thiamine. J. Biol. Chem. **265:**10857–10864.

39. **Melamed, C., Y. Nevo, and M. Kupiec.** 1992. Involvement of cDNA in homologous recombination between Ty elements in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **12:**1613–1620.

40. **Merkulov, G. V., K. M. Swiderek, C. B. Brachmann, and J. D. Boeke.** 1996. A critical proteolytic cleavage site near the C terminus of the yeast retrotransposon Ty1 Gag protein. J. Virol. **70:**5548–5556.

41. **Mertins, P., and D. Gallwitz.** 1987. A single intronless action gene in the fission yeast *Schizosaccharomyces pombe*: nucleotide sequence and transcripts formed in homologous and heterologous yeast. Nucleic Acids Res. **15:**7369–7379.

42. **Monokian, G. M., L. T. Braiterman, and J. D. Boeke.** 1994. In-frame linker insertion mutagenesis of yeast transposon Ty1: mutations, transposition and dominance. Gene **139:**9–18.

43. **Nevo-Caspi, Y., and M. Kupiec.** 1996. Induction of Ty recombination in yeast by cDNA and transcription: role of the RAD1 and RAD52 genes. Genetics **144:**947–955.

44. **Orlinsky, K. J., J. Gu, M. Hoyt, S. Sandmeyer, and T. M. Menees.** 1996. Mutations in the Ty3 major homology region affect multiple steps in Ty3 retrotransposition. J. Virol. **70:**3440–3448.

45. **Orr-Weaver, T. L., J. W. Szostak, and R. J. Rothstein.** 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA **78:**6354–6358.

46. **Rein, A.** 1994. Retroviral RNA packaging: a review. Arch. Virol. Suppl. **9:**513–522.

47. **Rio, D. C.** 1990. Molecular mechanisms regulating *Drosophila* P element transposition. Annu. Rev. Genet. **24:**543–578.

48. **Robertson, H. M., and D. J. Lampe.** 1995. Distribution of transposable elements in arthropods. Annu. Rev. Entomol. **40:**333–357.

48a.**Robzyk, K., and Y. Kassir.** 1992. A simple and highly efficient procedure for rescuing autonomous plasmids from yeast. Nucleic Acids Res. **20:**3790.

49. **Sandmeyer, S. B.** 1992. Yeast retrotransposons. Curr. Opin. Genet. Dev. **2:**705–711.

50. **Sandmeyer, S. B., and T. M. Menees.** 1996. Morphogenesis at the retrotransposon-retrovirus interface: gypsy and copia families in yeast and Drosophila. Curr. Top. Microbiol. Immunol. **214:**261–296.

51. **SanMiguel, P., A. Tikhonov, Y. K. Jin, N. Motchoulskaia, D. Zakharov, A. Melake-Berhan, P. S. Springer, K. J. Edwards, M. Lee, Z. Avramova, and J. L. Bennetzen.** 1996. Nested retrotransposons in the intergenic regions of the maize genome. Science **274:**765–768.

52. **Sharon, G., T. J. Burkett, and J. D. Garfinkel.** 1994. Efficient homologous recombination of Ty1 element cDNA when integration is blocked. Mol. Cell. Biol. **14:**6540–6551.

53. **Smit, A. F.** 1996. The origin of interspersed repeats in the human genome. Curr. Opin. Genet. Dev. **6:**743–748.

54. **Suoniemi, A., A. Narvanto, and A. H. Schulman.** 1996. The BARE-1 retrotransposon is transcribed in barley from an LTR promoter active in transient assays. Plant Mol. Biol. **31:**295–306.

55. **Voytas, D. F.** 1996. Retroelements in genome organization. Science **274:**737–738.

56. **Weaver, D. C., G. V. Shpakovski, E. Caputo, H. L. Levin, and J. D. Boeke.** 1993. Sequence analysis of closely related retrotransposon families from fission yeast. Gene **131:**135–139.

57. **Youngren, S. D., J. D. Boeke, N. J. Sanders, and D. J. Garfinkel.** 1988. Functional organization of the retrotransposon Ty from *Saccharomyces cerevisiae*: Ty protease is required for transposition. Mol. Cell. Biol. **8:**1421–1431.

58. **Zou, S., N. Ke, J. M. Kim, and D. F. Voytas.** 1996. The *Saccharomyces* retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. Genes Dev. **10:**634–645.