MOLECULAR AND CELLULAR BIOLOGY, Apr. 1988, p. 1432–1442
0270-7306/88/041432-11$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 8, No. 4

The *Saccharomyces cerevisiae* Genome Contains Functional and Nonfunctional Copies of Transposon Ty1

JEF D. BOEKE,[1*] DANIEL EICHINGER,[1] DIEGO CASTRILLON,[2,3†] AND GERALD R. FINK[2,3]

Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, 725 N. Wolfe Street, Baltimore, Maryland 21205[1]; Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142[2]; and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139[3]

Received 10 August 1987/Accepted 23 December 1987

***Saccharomyces cerevisiae* Ty elements are transposons closely related to retroviruses. The DNA sequence of a functional Ty element (TyH3) is presented. The long terminal repeat sequences are different, suggesting that TyH3 is a recombinant Ty element. A chromosomal Ty element near the *LYS2* gene, Ty173, was found to be nonfunctional, even though it has no detectable insertions or deletions. The defect in Ty173 transposition is caused by a missense mutation giving rise to a Leu-to-Ile substitution in the *TYB* (*pol*) open reading frame. Several chromosomal Ty elements carry this lesion in their DNA, indicating that nonfunctional Ty elements are common in the yeast genome.**

The genome of the yeast *Saccharomyces cerevisiae* contains a family of approximately 35 transposons called Ty elements. Ty elements fall into two rather similar subfamilies called Ty1 and Ty2. These elements are about 6 kilobases (kb) in length and are bounded by long terminal repeat (LTR; also called δ) sequences of about 335 base pairs (5). An end-to-end transcript appears to be the major RNA product of Ty elements (9). This general structure, as well as DNA sequence analysis of two Ty1 elements, two Ty2 elements, and several structurally similar *Drosophila melanogaster* transposons, suggested a strong evolutionary link between Ty elements and the retroviruses of larger eucaryotes (for a review, see reference 1). A functional link between Ty elements and retroviruses was established when it was shown that Ty transposition proceeded via an RNA intermediate (2). Furthermore, reverse transcriptase and virus-like particles were shown to be associated with overexpression of certain Ty elements (13, 20), a condition which was shown to lead to a high frequency of Ty element transposition (2).

The analysis of Ty transposition was facilitated by fusion of the Ty element TyH3 to the inducible yeast *GAL1* promoter on a high-copy-number plasmid vector. The structure of the *GAL1*-Ty fusion is such that the promoter for the major Ty transcript is removed and replaced by the *GAL1* sequences; the structure of the *GAL1*-promoted Ty transcript is identical to that of native Ty transcripts. Galactose induction of transcription from plasmid pGTyH3 leads to frequent transposition of TyH3 itself and also of chromosomal Ty elements (1a, 2, 12). Strains containing pGTyH3 also grow poorly on transposition-inducing medium which is selective for the plasmid (Fig. 1).

Although members of the Ty1 family have a very similar overall structure (as is readily seen by comparison of the published DNA sequences of Ty elements), certain restriction endonuclease cleavage sites in Ty1 DNA are not well conserved. For example, the single *Hin*dIII site is found in about half of the randomly isolated Ty1 sequences, as is the *Xho*I site found in certain Ty1 LTRs. In addition, variants of Ty1 elements containing insertion, inversion, and deletion mutations have been observed (10, 17; J. D. Boeke and D. J. Garfinkel, unpublished results). The heterogeneity between Ty elements raises the question of whether they are all competent for transposition.

In this report we show that the Ty1 element Ty173 (27) is nonfunctional when assayed for transposition by using our p*GAL* induction system. Furthermore, such pGTy173 plasmids did not cause a growth defect in selectively grown colonies. Both of these differences between pGTy173 and pGTyH3 are due to a missense mutation in the Ty173 sequence.

The mutation that renders Ty173 nonfunctional falls within the second (*TYB*) open reading frame of the element. This open reading frame has weak but statistically significant sequence similarity at the amino acid level to the protease, integrase, reverse transcriptase, and RNase H domains of retroviral *pol* genes (6, 32; R. F. Doolittle, D.-F. Feng, M. S. Johnson, and M. A. McClure, submitted for publication). The missense mutation falls within the domain of *TYB*, which is similar to retroviral "integrase" domains, but very close to the region similar in sequence of retroviral "protease" domain (Doolittle et al., submitted). The mutation does not abolish the ability to form Ty-VLPs or reverse transcriptase activity, but does appear to affect proteolytic processing of *TYB* proteins.

TyH3 can be rendered nonfunctional (abolishing both Ty transposition and the growth defect) by specific alteration of a single nucleotide that results in the missense mutation present in Ty173. Approximately one-quarter of the genomic Ty1 elements have this missense mutation.

(A preliminary account of portions of this work was presented at a Banbury Conference on Eucaryotic Transposable Elements as Mutagenic Agents, 1987, Cold Spring Harbor, N.Y.)

MATERIALS AND METHODS

Strains, media, and plasmids. The *S. cerevisiae* and *Escherichia coli* strains used in these experiments have been

* Corresponding author.
† Present address: University of Texas Southwestern Medical School, Dallas, TX 75235.

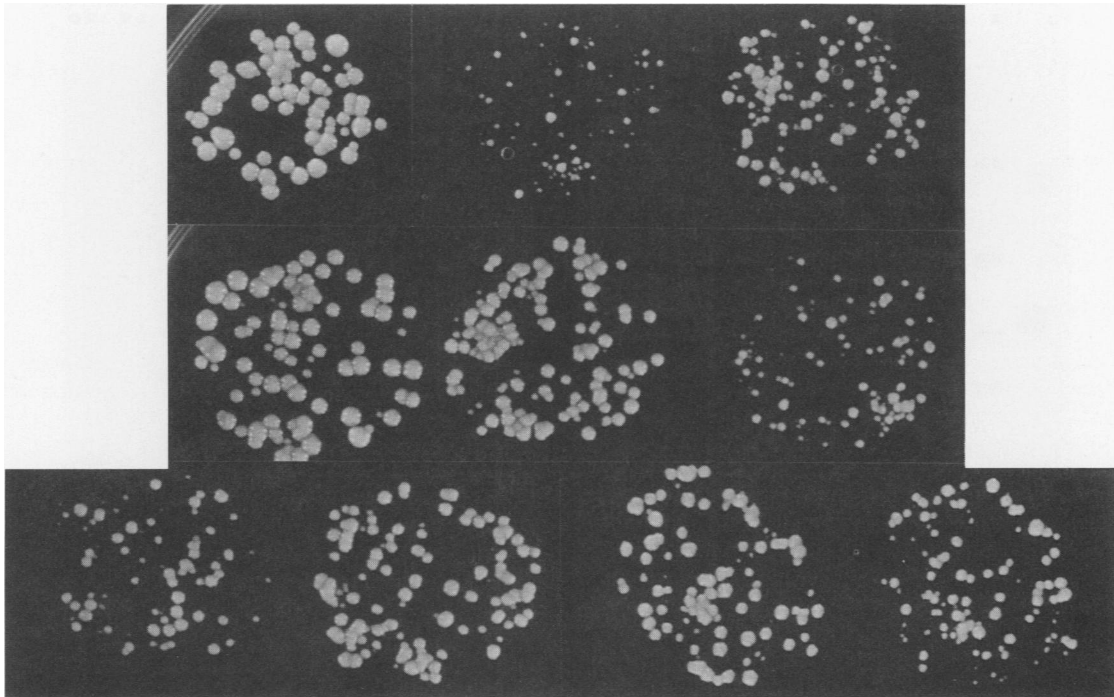


FIG. 1. Colony morphology of yeast cells bearing pGTy plasmids. Yeast strains (all are GRF167 transformants) containing pGTy plasmids and grown into colonies on SC-ura galactose plates. Top row, left to right: JB160 (pGAL vector), JB224 (pGTyH3), and JB442 (pGTyH3*lacO*); second row: JB160 (pGAL vector), JB161 (pGTy173), and JB162 (pGTy173-2); third row: DJ1346 (pJEF1346, a *Bst*EII-*Pst*I substitution plasmid), DJ1351 (pJEF1351, a *Bst*EII-*Sal*I substitution plasmid), DJ1 (pDE1-20, a *tyb-2098* point mutant plasmid), and DJ4 (pDE1-35, control plasmid for point mutant).

described in previous publications (2, 18, 27). Media were prepared as described by Sherman et al. (26).

Plasmids pGTyH3 and pGTyH3*lacO* have been described previously (2). Plasmid pGTy173 was constructed by ligating the internal *Xho*I fragment from Ty173 (27) into the unique *Xho*I site in the *GAL1* expression vector pCGE329, kindly provided by J. Mao and J. Schaum, Collaborative Research, Inc., Waltham, Mass. The hybrid plasmids described in Fig. 2 were constructed by two- or three-piece ligations as follows. pJEF1076, pJEF1071, and pJEF1339 were made by ligating the Ty-derived *Xho*I-*Bst*EII, *Xho*I-*Cla*I, and *Bst*EII-*Kpn*I fragments, respectively, from pGTy173, into pGTyH3*lacO* backbone fragments digested with the same pair of enzymes. pJEF1346/1347, pJEF1351, pJEF1357, and pJEF1363 were made as follows. The pGTyH3*lacO* backbone consisted of a *Bst*EII-*Kpn*I fragment. The corresponding *Bst*EII-*Kpn*I insert fragments from both Ty173 and TyH3 were eluted preparatively from an agarose gel slice and subsequently redigested with *Pst*I, *Sal*I, *Ava*I, or *Bgl*II. Pairwise combinations of fragments were ligated with backbone fragment to give rise to the hybrid Ty elements indicated in Fig. 2 and then used to transform competent *E. coli* cells. Control experiments indicated that omission of any one of the three fragments in the ligation resulted in 25-fold less transformants. Moreover, except for pJEF 1346/1347, the presence of *Mst*II or *Hha*I site polymorphisms was used to confirm that the appropriate Ty173 fragment had indeed been incorporated into the recombinant plasmid. For pJEF1346/1347, in which no such polymorphism was available, two independent constructs were used to assay transposition. Oligo-directed mutations (see below) were reintroduced into pGTyH3*lacO* as an *Eco*RI-*Sal*I fragment, which was ligated to *Sal*I-*Bam*HI and *Bam*HI-*Eco*RI backbone fragments isolated from pGTyH3*lacO*.

DNA manipulations. The DNA sequence of TyH3 was determined by using a combination of shotgun sequencing (21, 24) of randomly sheared fragments of TyH3 subcloned into the *Sma*I-digested, dephosphorylated M13mp8 vector (Amersham Corp.), dideoxy sequencing of fragments of TyH3 "force cloned" into mp18 and mp19 vectors (24), and Maxam-Gilbert sequencing (19) techniques.

Southern hybridization analysis of yeast chromosomal DNA with a *lacO* probe and with a probe derived from the 3′ end of TyH3 (the 3′ *Hin*dIII-*Bgl*II fragment) was carried out as described previously (2). Oligonucleotide mutagenesis was performed by the method of Kramer et al. (18) on a pGTyH3 *Eco*RI-*Sal*I fragment subcloned into M13mp8 by using the oligonucleotide 173-1 (CATGCGCAA<u>T</u>CATTC GATGA). The success of the mutagenesis was confirmed by dideoxy sequencing of 10 randomly chosen white plaques.

Southern hybridization experiments with oligonucleotide probes were carried out as follows. Dried agarose gels (0.8%), prepared by the methods of Studencki and Wallace (28), were prehybridized for 30 min at 65°C in 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.3% sodium dodecyl sulfate–0.1% sodium PP_i–100 μg of denatured salmon sperm DNA per ml. The gels were hybridized and washed under the conditions described previously (28). Oligonucleotides were labeled with [γ-^{32}P]ATP and polynucleotide kinase and were purified before use from a 10% polyacrylamide gel containing 50% urea. The oligonucleotides used as probes are described in Table 1.

Transposition assay. Cells (host strain GRF167 or BWG1-7a) transformed with the pGTy plasmid to be tested were

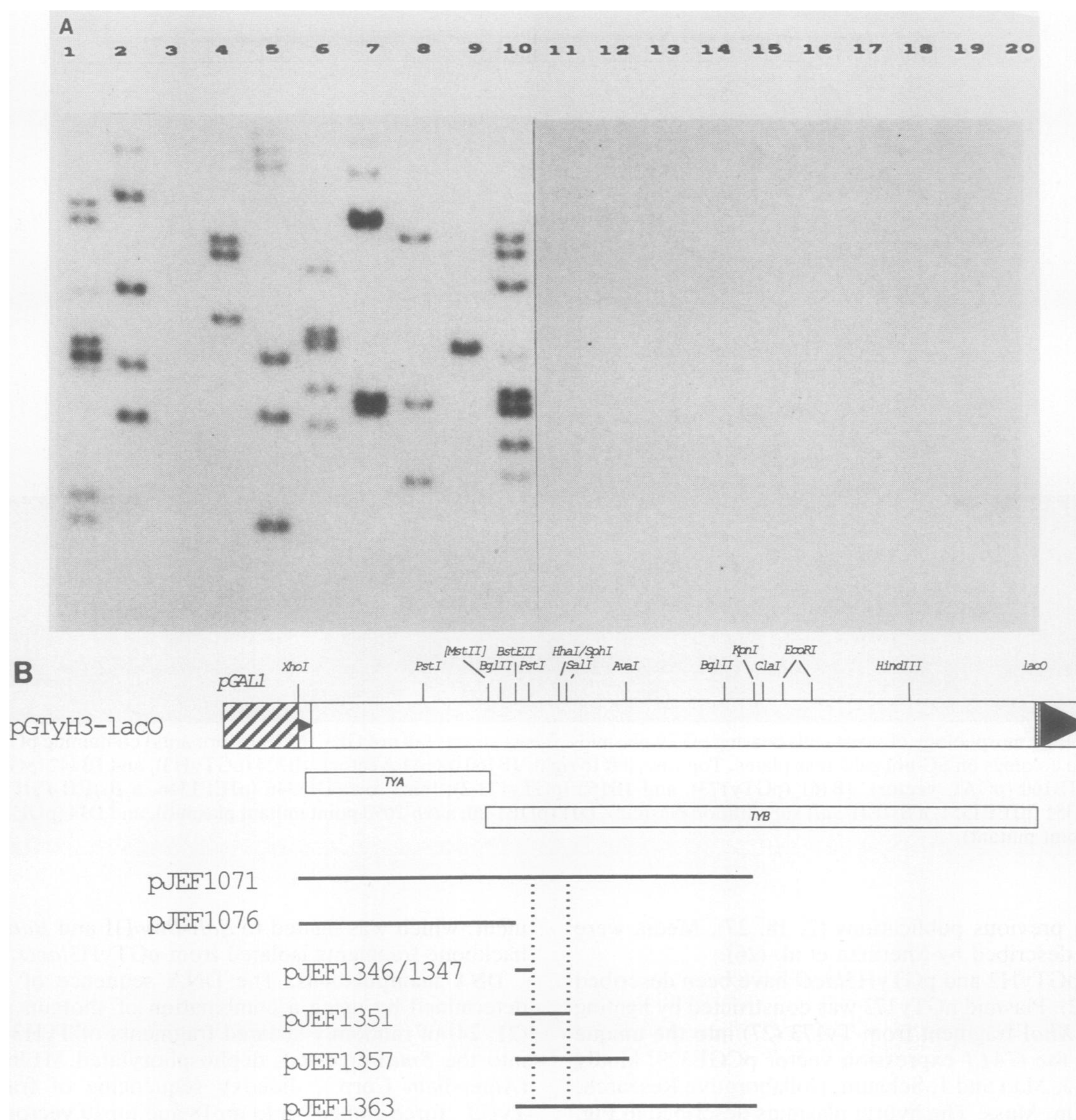


FIG. 2. Mapping the transposition defect in Ty173. (A) Southern blot analysis. A representative Southern blot analysis of Ty transposition is shown. Lanes: 1 to 10, strains carrying plasmid pJEF1346 (*Bst*EII-*Pst*I substitution plasmid); 11 to 20, strains carrying plasmid pJEF1351 (*Bst*EII-*Sal*I substitution plasmid). (B) Map. A map of the wild-type *GAL1*-Ty fusion, pGTyH3*lacO*, is shown. Symbols: ▨, *GAL1* promoter sequences; ▭, Ty sequences; ▧, *lacO* sequences; ▶, LTR (δ) sequences. The open reading frames, *TYA* and *TYB*, are indicated by labeled open boxes. The restriction sites are those referred to in the text and in Table 2; the *Mst*II site shown in brackets is found in Ty173 but not in TyH3. Only the *Hha*I and *Sph*I sites mentioned in the text (position 2103) are indicated; for more complete information on restriction sites, see Fig. 3. The horizontal lines below the box indicate the sequences which were replaced by Ty173 sequences in the recombinant plasmids indicated that are discussed in detail in Table 2. The vertical dashed lines delineate the region of Ty173 and Ty173-2 sequenced, within which the mutation lies.

grown as single colonies for 5 days at 22°C on SC-ura plates containing 2% galactose. At this time, independent colonies were restreaked onto SC-ura (26) plates containing 2% glucose and grown into single colonies at 30°C. The plasmids were then allowed to segregate from these cells (either by nonselective growth on YPD (26) plates and subsequent replica plating to SC-ura glucose plates or by replica plating to 5-fluoroorotic acid plates [3]). The plasmid-free segregants

TABLE 1. Oligonucleotides used

Oligonucleotide	Sequence[a]	Description
173-1	CATGCGCAATCATTCGATGA	*tyb-2098* (Leu-to-Ile) mutation alone
173-2	CATGTGCAATCATTCGATGA	*tyb-2098 tyb-2103* double mutant (like Ty173)
173-3	CATGTGCAAGCATTCGATGA	*tyb-2103* mutation alone
173-4	CATGCGCAAGCATTCGATGA	Wild type (like TyH3)

[a] Underlined positions are those differing from the wild-type TyH3 sequence.

were then analyzed by Southern hybridization analysis to determine the number of genomic fragments hybridizing to a *lacO* probe.

Western immunoblotting. Crude extracts (22) from galactose-induced strains JB442 and DJ1 were fractionated on sodium dodecyl sulfate–polyacrylamide gels and transferred to nitrocellulose (29). Filters were blocked with bovine serum albumin incubated with rabbit antiserum tyb-2 (the generous gift of D. Garfinkel; the serum was raised against a fusion protein containing *TYB* sequences extending from positions 2782 to 3947 [33]), washed, and incubated with ^{125}I-labeled protein A (Amersham Corp.).

RESULTS

TyH3 is a functional transposon. We have previously shown that TyH3 is a functional transposon by several criteria. When TyH3 is fused to the *GAL1* promoter and genetic markers are inserted within the *Bgl*II site near the 3′ end of the element, the TyH3 sequence and the inserted marker transpose into both plasmid and chromosomal targets at high frequency. For example, yeast cells containing plasmid pGTyH3*lacO* (carrying TyH3 marked with the synthetic *lacO* sequence) can be grown on galactose-containing (inducing) medium for several days: a procedure referred to as transposition-induction. Randomly selected colonies are streaked out on SC-ura medium and subsequently cured of the plasmid. Chromosomal DNA from such colonies typically contains 2 to 10 copies (mean, 3.2) of the *lacO* fragment as determined by Southern hybridization experiments (Fig. 2). In control experiments, glucose-grown cells do not give rise to chromosomal *lacO* sequences at a frequency detected by this assay.

The slow growth of colonies harboring the pGTyH3 plasmids is correlated with high transposition frequency: glucose-grown cells show no obvious growth phenotype, but the same cells show marked inhibition of colony-forming ability on galactose-containing medium that is also selective for the plasmid (i.e., SC-ura GAL medium). Insertion of a genetic marker such as *lacO* into the *Bgl*II site near the 3′ end of TyH3 slightly relieves this colony growth defect (Fig. 1), but these cells still grow much more slowly than cells containing control plasmids.

DNA sequence analysis of TyH3. To perform a meaningful comparison between TyH3, the functional Ty element, and other Ty elements studied, we determined the entire nucleotide sequence of TyH3 as described in Materials and Methods. Further details of the sequencing strategy are outlined in Fig. 3. As in the other sequenced Ty elements, two long open reading frames (*TYA* and *TYB*) are encoded by TyH3. The open reading frames *TYA* and *TYB* overlap by 38 nucleotides, as they do in the other sequenced Ty1 elements.

Comparison of published Ty1 sequences. The complete DNA sequences of four other Ty elements have been determined (6, 14, 32; P. J. Farabaugh, personal communication). However, in none of these cases are the Ty elements sequenced known to be functional. Two of these, Ty912 and Typ109, like TyH3 and Ty173, are members of the Ty1 family of transposons. The three Ty1 family members are remarkably similar in sequence and overall structure. Ty912 and TyH3 have the identical number of nucleotides; both can be related to the sequence of Typ109 by a small number of frameshifts. Nevertheless, there are a number of differences, mostly in the form of third-position and conservative coding changes, which account for the known restriction site polymorphisms between Ty912 and TyH3. Most of the differences are clustered in the 3′ segment of *TYB*, which encodes reverse transcriptase (13). In fact, 56% of the nucleotide sequence differences lie between positions 5070 and 5310 (only 4% of the sequence). The sequence differences between TyH3 and Ty912 are summarized in Table 2.

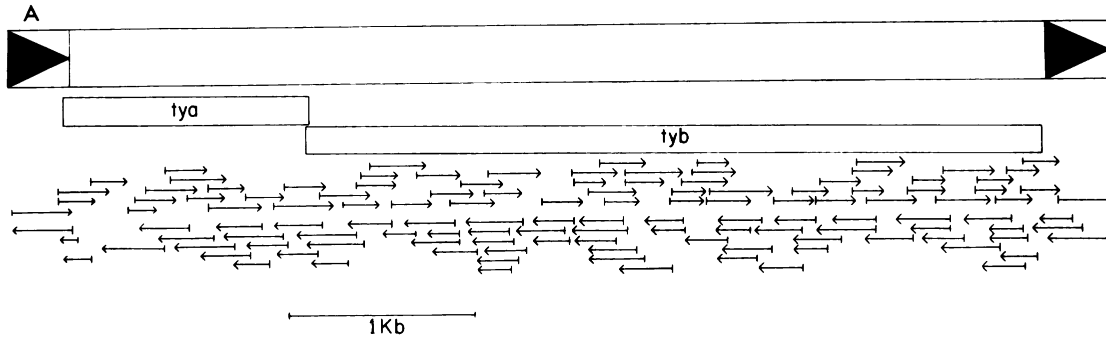
The 5′ and 3′ LTRs of TyH3 are not identical. The LTR or δ sequences of TyH3 differ at six positions (Fig. 3). Most (five of six) of these fall within the U3 region of the LTR, and so these heterogeneities are not relevant to the structure of the pGTyH3 plasmid, since pGTyH3 lacks U3 sequences in its 5′ LTR. The remaining heterogeneous base lies within the U5 sequences of TyH3 (2). The 3′ LTR is identical in sequence to the LTRs of Ty912. Other cases of Ty elements with nonidentical LTR sequences have been presented previously (see reference 23 for a review), but many (8 of 12) Ty elements have identical LTR sequences.

Ty173 is defective. A construct very similar to pGTyH3 was made starting with a second Ty element, Ty173. The main difference between the plasmids pGTyH3 and pGTy173 is the Ty elements used in their construction. The overall structure of the two elements is the same, except that pGTyH3 contains an intact 3′ LTR sequence and pGTy173 lacks 90 base pairs of 3′ LTR sequence. The coding region of Ty173 is unaffected by this difference (see Materials and Methods for details of the construction). The original sources of the Ty elements TyH3 and Ty173 are a His$^+$ revertant of the plasmid pNN162 (2, 25) and the spontaneous *lys2-173* mutation (27), respectively. Thus, both TyH3 and Ty173 have recently transposed.

In marked contrast with cells transformed with the pGTyH3 plasmid, the majority (11 of 12) of yeast cells transformed with pGTy173 (such as yeast transformant JB161) show little if any growth defect when transferred to galactose-containing medium. However, one transformant, JB162, shows a marked colony growth defect identical to that seen in cells transformed with pGTyH3 (strain JB224). The genetic nature of the difference between JB161 and JB162 resides in the plasmid; plasmid DNA isolated from JB162 (designated pGTy173-2 DNA) gives rise to yeast transformants with a growth defect. In addition, cured JB162 can be retransformed with pGTy173 and gives rise to transformants with the phenotype of JB161 (i.e., without a growth defect).

Because we had established that there was a genetic difference between the plasmids pGTy173 and pGTy173-2, we examined the structure of these plasmids in some detail. The overall structure of the plasmids is very similar; there is no evidence for any overall size difference between the two DNAs. However, pGTy173-2 differs from pGTy173 in at least two restriction sites which map within the Ty DNA segment of the plasmid. These are the *Mst*II site at position 1795 (present in Ty173) and a *Hha*I site, which overlaps a *Sph*I site (the latter two are lacking in Ty173) at position 2103. These restriction sites all lie within coding regions of the Ty element (Fig. 2).

Direct proof that a mutation in Ty173 reduced transposition was obtained from a functional analysis of hybrid pGTy plasmids consisting of segments of Ty173 DNA ligated into a pGTyH3*lacO* recipient plasmid. These hybrid pGTy plasmids are identical in overall structure to pGTyH3*lacO*, but contain regions of various sizes substituted by the corresponding Ty173 sequences. The detailed structures of these plasmids are given in Fig. 2. The plasmids were introduced into *S. cerevisiae* BWG1-7a, and the resulting transformants were tested for transposition by the *lacO* mobilization assay described in Materials and Methods. Since this assay detects



B

```
   1 TGTTGGAATAGAAATCAACTATCATCTACTAACTAGTATTTACATTACTAGTATATTATCATATACGGTGTTAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGA  150
         .        .      . SpeI  .     SpeI     .        .        .        .        .        .        .        .        .        .
                                                                                                                                              MetGlu
 151 TCAATGAATATAAACATATAAAATGATGATAATAATATTTATAGAATTGTGTAGAATTGCAGATTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAATATTATAGCCTTTATCAACAATGGAAT  300
         .        .        .        .        .        .     ^ .  ^^    .  ^  Xho^AvaI  .        .        .        .       START of tya
     SerGlnGlnLeuSerGlnHisSerProAsnSerHisGlySerAlaCysAlaSerValThrSerLysGluValHisThrAsnGlnAspProLeuAspValSerAlaSerLysThrGluGluCysGluLysAlaSerThrLysAlaAsnSer
 301 CCCAACAATTATCTCAACATTCACCCAATTCTCATGGTAGCGCCTGTGCTTCGGTTACTTCTAAGGAAGTCCACACAAATCAAGATCCGTTAGACGTTTCAGCTTCCAAAACAGAAGAATGTGAGAAGGCTTCCACTAAGGCTAACTCTC  450
         .        .     ^ .      HhaI     .        .        .        .        .        .        .        .        .        .        .
     GlnGlnThrThrThrProAlaSerSerAlaValProGluAsnProHisHisAlaSerProGlnProAlaSerValProProProGlnAsnGlyProTyrProGlnGlnCysMetMetThrGlnAsnGlnAlaAsnProSerGlyTrpSer
 451 AACAGACAACAACACCTGCTTCATCAGCTGTTCCAGAGAACCCCCATCATGCCTCTCCTCAACCTGCTTCAGTACCACCTCCACAGAATGGGCCGTACCCACAGCAGTGCATGATGACCCAAAACCAAGCCAATCCATCTGGTTGGTCAT  600
         .        .   PvuII.        .        .        .        .        .PflMI   .        .        .        .        .        .
     PheTyrGlyHisProSerMetIleProTyrThrProTyrGlnMetSerProMetTyrPheProProGlyProGlnSerGlnPheProGlnTyrProSerSerValGlyThrProLeuSerThrProSerProGluSerGlyAsnThrPhe
 601 TTTACGGACACCCATCTATGATTCCGTATACACCTTATCAAATGTCGCCTATGTACTTTCCACCTGGGCCACAATCACAGTTTCCGCAGTATCCATCATCAGTTGGAACGCCTCTGAGCACTCCATCACCTGAGTCAGGTAATACATTTA  750
         .        .        .        .        .        .        .        .        .        .        .        .        .        .        .
     ThrAspSerSerSerAlaAspSerAspMetThrSerThrLysLysTyrValArgProProProMetLeuThrSerProAsnAspPheProAsnTrpValLysThrTyrIleLysPheLeuGlnAsnSerAsnLeuGlyGlyIleIlePro
 751 CTGATTCATCCTCAGCGGACTCTGATATGACATCCACTAAAAAATATGTCAGACCACCACCAATGTTAACCTCACCTAATGACTTTCCAAATTGGGTTAAAACATACATCAAATTTTTACAAAACTCGAATCTCGGTGGTATTATTCCGA  900
         .        .        .        .        .        .      HpaI .        .        .        .        .        .        .        .
     ThrValAsnGlyLysProValArgGlnIleThrAspAspGluLeuThrPheLeuTyrAsnThrPheGlnIlePheAlaProSerGlnPheLeuProThrTrpValLysAspIleLeuSerValAspTyrThrAspIleMetLysIleLeu
 901 CAGTAAACGGAAAACCCGTACGTCAGATCACTGATGATGAACTCACCTTCTTGTATAACACTTTTCAAATATTTGCTCCCTCTCAATTCCTACCTACCTGGGTCAAAGACATCCTATCCGTTGATTATACGGATATCATGAAAATTCTTT  1050
         .        .        .        .        .        .        .        .        .        .        .        .        .        .        .
     SerLysSerIleGluLysMetGlnSerAspThrGlnGluAlaAsnAspIleValThrLeuAlaAsnLeuGlnTyrAsnGlySerThrProAlaAspAlaPheGluThrLysValThrAsnIleIleAspArgLeuAsnAsnAsnGlyIle
1051 CCAAAAGTATTGAAAAAATGCAATCTGATACCCAAGAGGCAAACGACATTGTGACCCTGGCAAATTTGCAATATAATGGCAGTACACCTGCAGATGCATTTGAAACAAAAGTCACAAACATTATCGACAGACTGAACAATAATGGCATTC  1200
         .        .        .        .        .        .        .        .      PstI  NsiI .        .        .        .        .
     HisIleAsnAsnLysValAlaCysGlnLeuIleMetArgGlyLeuSerGlyGluTyrLysPheLeuArgTyrThrArgHisArgHisLeuAsnMetThrValAlaGluLeuPheLeuAspIleHisAlaIleTyrGluGluGlnGlnGly
1201 ATATCAATAACAAGGTCGCATGCCAATTAATTATGAGAGGTCTATCTGGCGAATATAAATTTTTACGCTACACACGTCATCGACATCTAAATATGACAGTCGCTGAACTGTTCTTAGATATCCATGCTATTTATGAAGAACAACAGGGAT  1350
         .     SphI        .        .        .        .        .        .        .        .        .        .        .        .        .
     SerArgAsnSerLysProAsnTyrArgArgAsnProSerAspGluLysAsnAspSerArgSerTyrThrAsnThrThrLysProLysValIleAlaArgAsnProGlnLysThrAsnAsnSerLysSerLysThrAlaArgAlaHisAsn
1351 CGAGAAACAGTAAACCTAATTACAGGAGAAATCCGAGTGATGAGAAGAATGATTCTCGCAGCTATACGAATACAACCAAACCCAAAGTTATAGCTCGGAATCCTCAAAAAACAAATAATTCGAAATCGAAAACAGCCAGGGCTCACAATG  1500
         .        .        .        .        .        .        .        .        .        .        .        .      AsuII    .     BanII  .
     ValSerThrSerAsnAsnSerProSerThrAspAsnAspSerIleSerLysSerThrThrGluProIleGlnLeuAsnAsnLysHisAspLeuHisLeuArgProGluThrTyrEND
                                                                                        ThrIleSerThrThrPheIleLeuGlyGlnLysLeuThrGluSerThrValAsnHisThrAsnHisSerAspAsp
1501 TATCCACATCTAATAACTCTCCCAGCACGGACAACGATTCCATCAGTAAATCAACTACTGAACCGATTCAATTGAACAATAAGCACGACCTTCATCTTAGGCCAGAAACTTACTGAATCTACAGTAAATCATACTAATCATTCTGATGAT  1650
         .        .        .        .        .        .        .        .     START of tyb .        .    END of tya  .        .        .
     GluLeuProGlyHisLeuLeuLeuAspSerGlyAlaSerArgThrLeuIleArgSerAlaHisHisIleHisSerAlaSerSerAsnProAspIleAsnValValAspAlaGlnLysArgAsnIleProIleAsnAlaIleGlyAspLeu
1651 GAACTCCCTGGACACCTCCTTCTCGATTCAGGAGCATCACGAACCCTTATAAGATCTGCTCATCACATACACTCAGCATCATCTAATCCTGACATAAACGTAGTTGATGCTCAAAAAAGAAATATACCAATTAACGCTATTGGTGACCTA  1800
         .        .        .        .        . BglII   .        .        .        .        .        .        .        .        . BstEII .
     GlnPheHisPheGlnAspAsnThrLysThrSerIleLysValLeuHisThrProAsnIleAlaTyrAspLeuLeuSerLeuAsnGluLeuAlaAlaValAspIleThrAlaCysPheThrLysAsnValLeuGluArgSerAspGlyThr
1801 CAATTTCACTTCCAGGACAACACCAAAACATCAATAAAGGTATTGCACACTCCTAACATAGCCTATGACTTACTCAGTTTGAATGAATTGGCTGCAGTAGATATCACAGCATGCTTTACCAAAAACGTCTTAGAACGGTCTGACGGCACT  1950
         .        .        .        .        .        .        .        .        . PstI   .      SphI      .        .        .        .
     ValLeuAlaProIleValLysTyrGlyAspPheTyrTrpValSerLysLysTyrLeuLeuProSerAsnIleSerValProThrIleAsnAsnValHisThrSerGluSerThrArgLysTyrProTyrProPheIleHisArgMetLeu
1951 GTACTTGCACCTATCGTAAAATATGGAGACTTTTACTGGGTATCTAAAAAGTACTTGCTTCCATCAAATATCTCCGTACCCACCATCAATAATGTCCATACAAGTGAAAGTACACGCAAATATCCTTATCCTTTCATTCATCGAATGCTT  2100
         .        .        .        .      ScaI       .        .        .        .        .        .        .        .        .      FspI
     AlaHisAlaAsnAlaGlnThrIleArgTyrSerLeuLysAsnAsnThrIleThrTyrPheAsnGluSerAspValAspTrpSerSerAlaIleAspTyrGlnCysProAspCysLeuIleGlyLysSerThrLysHisArgHisIleLys
2101 GCGCATGCCAATGCACAGACAATTCGATACTCACTTAAAAATAACACCATCACGTATTTTAACGAATCAGATGTCGACTGGTCTAGTGCTATTGACTATCAATGTCCTGATTGTTTAATCGGCAAAAGCACCAAACACAGACATATCAAA  2250
 HhaI SphI  .        .        .        .        .        .        . SalI   .        .        .        .        .        .        .
     GlySerArgLeuLysTyrGlnAsnSerTyrGluProPheGlnTyrLeuHisThrAspIlePheGlyProValHisAsnLeuProAsnSerAlaProSerTyrPheIleSerPheThrAspGluThrThrLysPheArgTrpValTyrPro
2251 GGTTCACGACTAAAATACCAAAATTCATACGAACCCTTTCAATACCTACATACTGACATATTTGGTCCAGTTCACAACCTACCAAATAGTGCACCATCCTATTTCATCTCATTTACTGATGAGACAACAAAATTCCGTTGGGTTTATCCA  2400
         .        .        .        .        .        .        .        .      ApaLI     .        .        .        .        .
     LeuHisAspArgArgGluAspSerIleLeuAspValPheThrThrIleLeuAlaPheIleLysAsnGlnPheGlnAlaSerValLeuValIleGlnMetAspArgGlySerGluTyrThrAsnArgThrLeuHisLysPheLeuGluLys
2401 TTACACGACCGTCGCGAGGACTCTATCCTCGATGTTTTTACTACGATACTAGCTTTTATTAAAAACCAGTTTCAGGCCAGTGTCTTGGTTATACAAATGGACCGTGGTTCTGAGTATACTAACAGAACTCTCCATAAATTCCTTGAAAAA  2550
       . NruI   .        .        .        .        .        .        .        .        .        .        .        .        .
     AsnGlyIleThrProCysTyrThrThrThrAlaAspSerArgAlaHisGlyValAlaGluArgLeuAsnArgThrLeuLeuAspAspCysArgThrGlnLeuGlnCysSerGlyLeuProAsnHisLeuTrpPheSerAlaIleGluPhe
2551 AATGGTATAACTCCATGCTATACAACCACAGCGGATTCCCGAGCACATGGAGTCGCTGAACGGCTAAACCGTACCTTATTAGATGACTGCCGTACTCAACTGCAATGTAGTGGTTTACCGAACCATTTATGGTTCTCTGCAATCGAATTT  2700
         .        .        .      AvaI       .        .        .        .        .        .        .        .        .        .
     SerThrIleValArgAsnSerLeuAlaSerProLysSerLysLysSerAlaArgGlnHisAlaGlyLeuAlaGlyLeuAspIleSerThrLeuLeuProPheGlyGlnProValIleValAsnAspHisAsnProAsnSerLysIleHis
2701 TCTACTATTGTGAGAAATTCACTAGCTTCACCTAAAAGCAAAAAATCTGCAAGACAACATGCTGGCTTGGCAGGACTTGATATCAGTACTTTGTTACCTTTCGGTCAACCTGTTATCGTCAATGATCACAACCCTAACTCCAAAATACAT  2850
         .        .        .        .        .        .        .        .      ScaI .        .        .      BclI   .        .        .
     ProArgGlyIleProGlyTyrAlaLeuHisProSerArgAsnSerTyrGlyTyrIleIleTyrLeuProSerLeuLysLysThrValAspThrThrAsnTyrValIleLeuGlnGlyLysGluSerArgLeuAspGlnPheAsnTyrAsp
2851 CCTCGTGGCATCCCAGGCTACGCTCTACATCCGTCTCGAAACTCTTATGGATATATCATCTATCTTCCATCCTTAAAGAAGACAGTAGATACAACTAACTATGTTATTCTTCAGGGCAAGGAATCCAGATTAGATCAATTCAATTACGAC  3000
         .        .        .        .        .        .        .        .        .        .        .        .        .        .        .
```

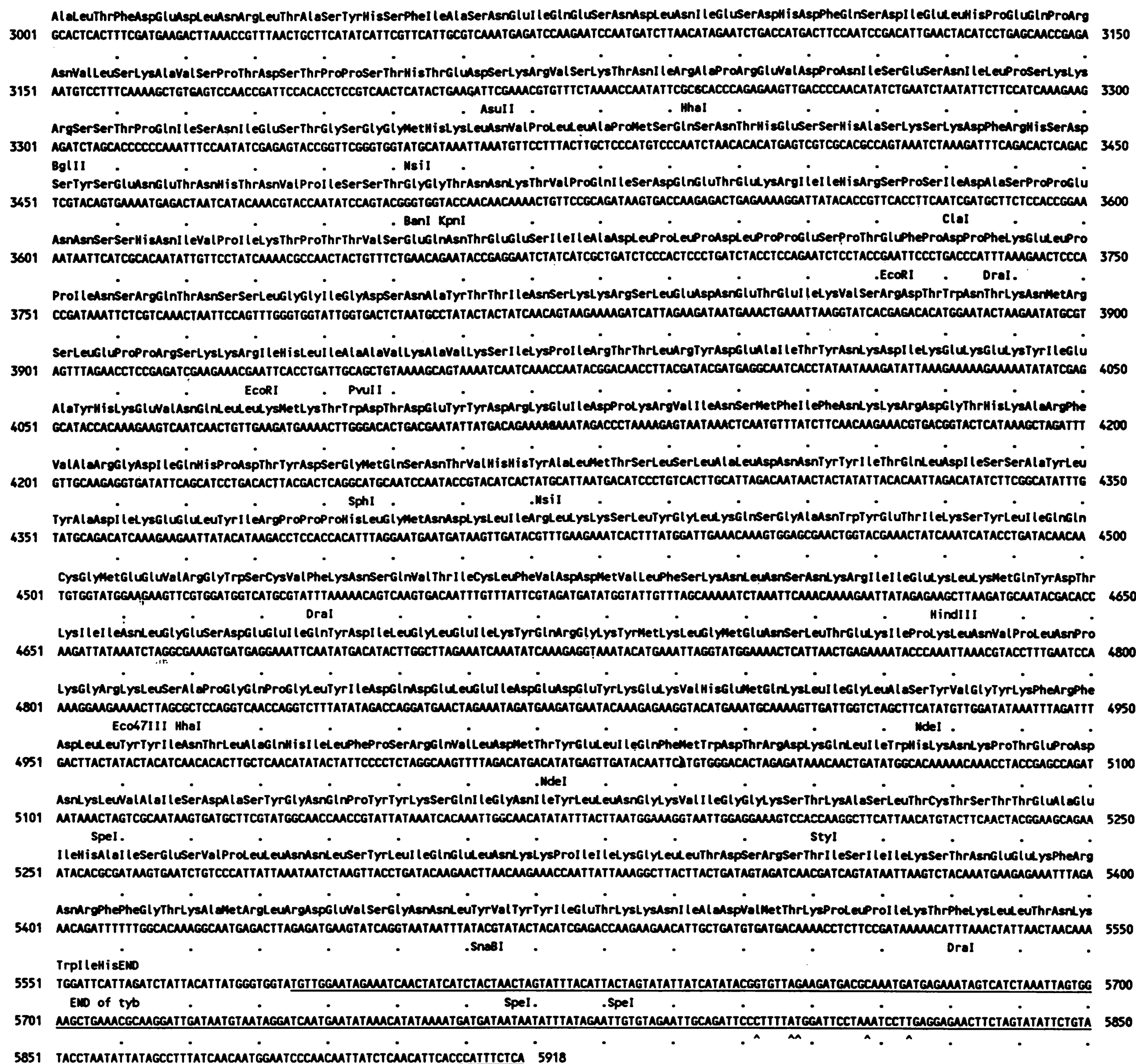


FIG. 3. DNA sequence of TyH3. (A) Sequencing strategy. Symbols: ▭, Ty sequence; ▶, LTR (δ) sequences. The boxes labeled *TYA* and *TYB* are the two long open reading frames found in all sequenced Ty elements. The arrows indicate the DNA sequences read to produce the sequence in panel B. (B) DNA and predicted protein sequences. LTR (δ) sequences are underlined. Bases which are different in the 5′ and 3′ LTR sequences are indicated by a ^ symbol.

TABLE 2. Sequence difference between TyH3 and Ty912

Region	Nucleotide interval	No. of base differences	% Nucleotide sequence identity	No. of amino acid differences	% Amino acid sequence identity
LTRs[a]	1–334, 5585–5918	6	91.5	1	92.9
TYA	294–1613	7	99.5	4	99.1
TYB[b]	1614–5563 (1576-5563)	51 (53)	98.7 (98.6)	12 (14)	99.1 (99.0)
Hot spot[a]	5070–5310	45	81.2	10	87.5
Total	1–5918	64	98.9	16 (18)	99.1 (99.0)

[a] Region of TyH3 most different from Ty912 (see text).

[b] *TYA* and *TYB* overlap for 38 nucleotides. The precise makeup of the *TYB* protein is unknown. Two possible cases are considered. The numbers without parentheses are derived by assuming that *TYB* starts at the end of the overlap region; the numbers in parentheses are derived by assuming that *TYB* starts at the beginning of the overlap region (there are two single-base changes in the overlap region; one results in a coding change in *TYB* only, the other in both *TYA* and *TYB*).

TABLE 3. Transposition of wild-type and mutant *lacO*-marked Ty elements

Plasmid	Type of substitution made[a]	Total colonies tested	Total bands hybridizing[b]	Transposition frequency[c]
Wild type				
pGTyH3*lacO* (pJEF896)	None	18	54	3.0
pJEF1076	*Xho*I–*Bst*EII (238–1902)	20	59	3.0
pJEF1346	*Bst*EII–*Pst*I (1792–1892)	12	51	4.3
pJEF1347	*Bst*EII–*Pst*I (1792–1892)	9	31	3.4
pDE1-35	Reconstructed pJEF896	9	23	2.6
Mutant				
pJEF1071	*Xho*I–*Cla*I (238–3580)	33	0	<0.030
pJEF1339	*Bst*EII–*Kpn*I (1792–3505)	19	0	<0.053
pJEF1351	*Bst*EII–*Sal*I (1792–2173)	13	0	<0.077
pJEF1357	*Bst*EII–*Ava*I (1792–2588)	13	3	0.23
pJEF1363	*Bst*EII–*Bgl*II (1792–3301)	17	1	0.059
pDE1-20, pDE1-22, pDE1-23	Oligo-directed Leu-to-Ile	23	2	0.087

[a] Fragment of Ty173 substituted for the homologous fragment in pGTyH3*lacO* (pJEF896). Numbers in parentheses indicate sequence coordinates.
[b] Fragments of genomic yeast DNA digested with *Hin*dIII which hybridize to a *lacO* probe.
[c] Total bands hybridizing/total colonies tested. The number reflects the number of transpositions of the *lacO*-marked Ty element into genomic DNA over a 5-day period at 22°C on SC-ura medium containing 2% galactose. The average frequency for the wild-type plasmids is 3.2; the average frequency for the mutant plasmids is 0.051.

only the transposition of the marked Ty element itself, the Southern blots were rehybridized with a Ty-specific probe to determine whether unmarked (chromosomal) Ty elements could transpose in cells containing wild-type or mutant plasmids. In all cases, the LacO phenotype and the pattern of total Ty elements was consistent; constructs which led to *lacO* mobilization (transposition of the marked Ty element) also gave rise to many new bands hybridizing to Ty probes; constructs which did not support *lacO* mobilization had a wild-type Ty pattern, showing that the mutation also affected the transposition of chromosomal (i.e., unmarked) Ty elements. The results of these experiments are presented in Fig. 2 and Table 3. The results show that sequences in Ty173 lying between the *Pst*I site at position 1897 and the *Sal*I site at position 2220 destroy the ability of TyH3 to transpose.

The relevant *Pst*I-*Sal*I fragments from both Ty173 and Ty173-2 were inserted into M13mp18 and M13mp19 and sequenced. Although the Ty173-2 sequence was identical to the TyH3 sequence within this segment, the Ty173 sequence differed at three positions from the TyH3 sequence (Fig. 3 and 4). Two of the sequence changes fell at the third position of the codons and did not alter coding (G to A at position 1939 and G to A at position 2103). The third difference resulted in substitution of a codon specifying leucine in the wild type to one encoding isoleucine in the mutant; this mutation lies within *TYB* (C to A at position 2098; hereafter referred to as *tyb-2098*). The mutation at position 2103 destroys recognition sites for the three enzymes *Hha*I, *Sph*I, and *Fsp*I and is very near the mutation causing the coding change. Because this second mutation is so tightly linked to the first, it provides a useful marker for the presence of the inactivating *tyb-2098* mutation (see below).

Oligonucleotide-directed mutagenesis of nucleotide 2098 in TyH3. The mapping and sequencing experiments outlined above suggested that the *tyb-2098* mutation causing the Leu-to-Ile coding change was responsible for the mutant phenotype, although it was conceivable that the third-position changes might have played some role. The role of the *tyb-2098* mutation was established by using a mutagenic oligonucleotide to construct a plasmid which differed from pGTyH3 only at the appropriate leucine codon (i.e., C to A at position 2098).

The *Eco*RI-*Sal*I fragment of pGTyH3 containing the region of interest was subcloned between the *Eco*RI and *Sal*I sites in M13mp8, giving rise to bacteriophage strain fDE20. fDE20 was then mutagenized by the method of Kramer et al. (18) with the mutator oligonucleotide 173-1 (Table 1). A total of 10 mutagenized white plaques (selected at random) were sequenced; 6 of these carried the mutation of interest. *Eco*RI-*Sal*I fragment DNA from three independently isolated mutant plaques and from unmutagenized fDE20 was ligated back into pGTyH3*lacO* between the appropriate *Sal*I and *Eco*RI sites as described in Materials and Methods, giving rise to mutagenized plasmids pDE1-20, pDE1-22, and pDE1-23 and unmutagenized plasmid pDE1-35. The four plasmids were transformed into *S. cerevisiae* GRF167. The presence of the *lacO* sequences and the mutant sequence in the yeast transformants was confirmed by recovering the plasmids from the yeast cells and by restriction and hybridization analysis of the recovered plasmids.

The four transformants were tested for their ability to mobilize *lacO* sequences into the genome. The three transformants carrying the mutated (*tyb-2098*) plasmid, like the hybrid elements described above, are unable to mobilize *lacO* at the normal frequency; the frequency of transposition of mutant *lacO*-marked Ty elements is about 60-fold lower than with a wild-type *lacO*-marked element (Table 3).

The Ty173 mutation is recessive. We wished to determine whether the *tyb-2098* mutant could be complemented by the wild-type (TyH3) gene product or whether the mutant gene product might instead have a *trans*-dominant effect on the

```
                             2100
                             |      SphI
                             | FspI
TyH3 &     CAT CGA ATG CTT GCG CAT GCC
Ty912      His Arg Met Leu Ala His Ala

Ty173      CAT CGA ATG ATT GCA CAT GCC
           His Arg Met Ile Ala His Ala

Ty1-17     CAT CGA ATG CTT GGA CAT GCT
& Ty917    His Arg Met Leu Gly His Ala
```

FIG. 4. Ty173 mutation. The sequences of five Ty elements in the region of the Ty173 mutation are compared. Bases and amino acids which are heterogeneous are underlined. TyH3 and Ty912 are identical in this region (there is an inversion of two base pairs in the published sequence of Ty912 (6; Farabaugh, personal communication). Ty1-17 and Ty917 are identical to each other in this region (32; Farabaugh, personal communication).

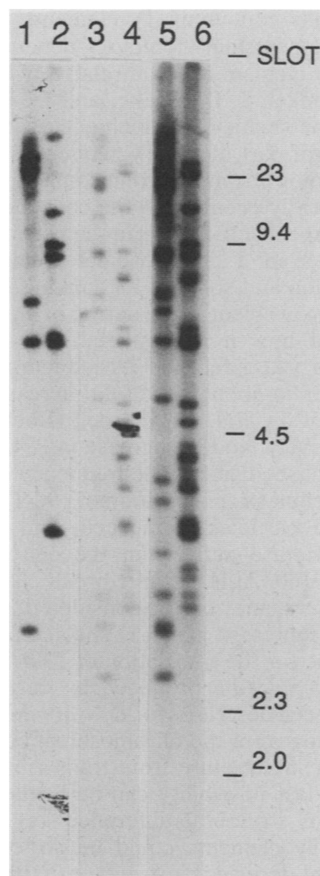


FIG. 5. Prevalence of the Ty173 mutation among chromosomal Ty elements. Yeast genomic DNA (strain BWG1-7a) was digested with *BclI* (lanes 1, 3, and 5) or *SalI* (lanes 2, 4, and 6) and electrophoretically separated on a 0.8% agarose gel. The dried gel panels were hybridized to oligonucleotide probes 173-1 to 173-4 and to a Tyl-specific probe prepared by the method of Feinberg and Vogelstein (11). Although the single mutant oligo probes 173-1 and 173-3 bound no genomic sequences under conditions of appropriate stringency, 173-2 (the doubly mutant oligo probe; lanes 1 and 2) and 173-4 (the wild-type oligo probe; lanes 3 and 4) bound different subsets of the fragments bound by the Tyl probe (lanes 5 and 6). Size markers are indicated in kilobases.

wild-type gene product. This experiment was carried out by introducing two complementary pGTy plasmids into the same cell. We introduced the plasmid pX1, which is pGTyH3*lacO* in which the *URA3* selectable marker in the plasmid has been replaced with the yeast *TRP1* selectable marker, into a strain (DJ1) already carrying pDE1-20 (pGTyH3*lacO* bearing the *tyb-2098* mutation described above, with *URA3* as the plasmid selectable marker) by selecting Ura^+ Trp^+ transformants. Transposition of the *lacO* marker into genomic DNA was assayed as described previously and was found to occur at about half the normal frequency (the mean number of *lacO*-hybridizing bands was 1.3 per genome). Thus, the *tyb-2098* mutation does not inhibit the action of wild-type TyH3 gene products.

A significant fraction of chromosomal Tyl elements carry the *tyb-2098* mutation. We wished to determine what fraction of chromosomal Tyl elements in *S. cerevisiae* strains carried the *tyb-2098* mutation. Sequence analysis of a randomly isolated Ty element (J. D. Boeke, unpublished data) lacking the *HhaI* site at position 2103 revealed that this element also carried the inactivating *tyb-2098* mutation. The presence of this mutation in a second element suggests that the *HhaI* site-inactivating mutation (*tyb-2103*) and missense (*tyb-2098*) mutation are usually linked owing to their extreme proximity.

Four oligonucleotides (Table 1) were synthesized for use as hybridization probes to determine the fraction of Ty elements carrying the *tyb-2098* mutation directly. All of the oligonucleotides are of the same length and sense; they represent the four possible combinations of the known alleles at positions 2098 and 2103 (Table 1). The oligonucleotides were end labeled with [γ-^{32}P]ATP and used as hybridization probes against genomic Southern blots (Fig. 5). As can be seen from these results, this hybridization protocol permits the *tyb-2098* mutation to be assayed in total genomic Southern blots. Strain BWG1-7a contains approximately 25 to 30 Tyl elements (all of the probes used in this experiment were Tyl specific). These Tyl elements fall into two classes. Seven of the chromosomal Tyl elements in BWG1-7a contain both the *tyb-2103* mutation and the *tyb-2098* mutation (i.e., they are similar to Ty173). The remaining Ty elements correspond to the TyH3 sequence; none of the Ty elements carry only the *tyb-2103* or the *tyb-2098* mutation by themselves. Similar results were obtained with strain GRF167 (not shown).

Ty173 itself is a nonfunctional chromosomal Ty element. It could be argued that chromosomal Ty elements bearing the *tyb-2098* mutation might contain a second mutation in *TYB* which suppressed the effects of the first and that this suppressor was separated from *tyb-2098* in the cutting and pasting experiments discussed above. To rule out this possibility, we examined the ability of intact Ty173 to support the transposition of chromosomal Ty elements (because it lacks a complete 3' LTR, pGTy173 itself cannot transpose). We compared the ability of pGTy173 and pGTy173-2 to support chromosomal Ty transposition and found that pGTy173-2 supported about 10-fold more transposition than pGTy173 did (Table 4).

Proteolysis defect resulting from Ty173 mutation. Strain DJ1, while contains the oligonucleotide-directed Ty173 mutation described above, was compared with isogenic strain JB442, which carries pGTyH3*lacO*, in terms of reverse transcriptase activity and the ability to form particles as described by Garfinkel et al. (13). No significant differences were noted in reverse transcriptase activity or the presence of Ty-VLPs (data not shown). The ability to produce p90-TYB, an antigen produced by transposition-induced cells carrying pGTyH3, was examined by Western blotting (Fig. 6). Although a protein of approximately 90 kilodaltons can be seen in these cells, its amount is greatly reduced; a considerable amount of an apparent precursor molecule of approximately 97 kilodaltons can be seen in the Ty173 mutant cells.

A variety of mutations in the protease region of *TYB* were studied by Western blotting (33) and were shown to be defective in producing p90-TYB as well as other Ty-encoded

TABLE 4. Transposition in pGTy173 and pGTy173-2 transformants

Plasmid	No. of new bands hybridizing to Tyl probe	No. of bands disappearing	No. of colonies tested	No. of new bands per colony tested (mean)
pGTy173	1	0	16	0.06
pGTy173-2	15	1	18	0.78

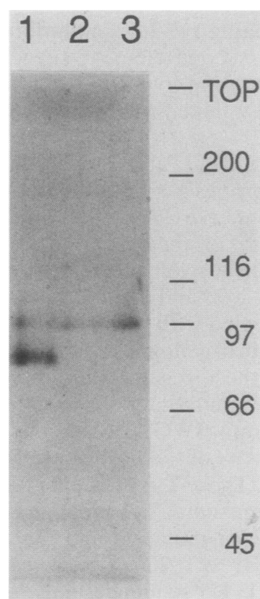


FIG. 6. Western blotting of *TYB* proteins from TyH3 and the Leu-to-Ile point mutant. Protein extracts from galactose-induced cultures of strain GRF167 carrying the indicated pGTy plasmids were electrophoresed on sodium dodecyl sulfate-polyacrylamide gels, blotted to nitrocellulose, and reacted with a *TYB*-specific antibody kindly provided by D. Garfinkel. Immune conjugates were identified by using ^{125}I-labeled protein A and autoradiography. Lanes: 1, pGTyH3 (wild type); 2, pDE1-20 (oligo-directed *tyb-2098* point mutant); 3, pJEF1293 (a known protease mutant made in pGTyH3 [33]). Size markers are indicated in kilodaltons.

products. One such known protease mutant of plasmid pGTyH3 has a Western blot pattern similar to that seen in the *tyb-2098* mutant (Fig. 6). It should be pointed out, however, that this experiment does not distinguish between a defect in the protease enzyme and a defect in the protease substrate (the *TYB* protein itself). Recently performed sequence alignments (Doolittle, et al., submitted) suggest that the *tyb-2098* mutation actually lies within the region of *TYB* whose sequence is similar to integrase; the mutation lies within a putative metal-binding domain conserved among many if not all retroviral and retrotransposon integrases (15). Since this metal-binding domain lies near to the N-terminal extremity of the region of sequence similarity, it may indeed fall near to a protease cleavage site.

DISCUSSION

Our results show that yeast cells contain many nonfunctional Ty elements. When functional Ty elements (typified by TyH3) are fused to the *GAL1* promoter, high-frequency transposition of both the donor (*GAL1*-fused) Ty and other, chromosomal, Ty elements is seen. In contrast, nonfunctional Ty elements (such as Ty173), when fused to the *GAL1* promoter in the same way, fail to give rise to high-frequency transposition of either plasmid-borne or chromosomal Ty elements. The defect responsible for the failure of Ty173 to transpose is a mutation leading to a substitution of isoleucine for leucine in the *tyb* open reading frame (*tyb-2098*). This subtle coding changes results in a dramatic decrease in Ty transposition frequency.

When plasmids containing Ty elements are transformed into yeast cells, certain homology-dependent rearrangements of the Ty element occur at relatively high frequencies (30; D. J. Garfinkel, J. D. Boeke, and G. R. Fink, unpublished data). One such event which occurs rather frequently is replacement of part of the Ty sequences on the plasmid with their homologs from a chromosomal Ty element. Such an event apparently occurred in the conversion of pGTy173 to pGTy173-2; the inactive copy in the plasmid was replaced by sequences from a chromosomal element lacking the *tyb-2098* mutation and some of its flanking DNA.

An interesting problem in the nature of Ty transposition is the question of how it is regulated. The introduction of pGTyH3 into a yeast cell and its induction by growth on galactose results in about a twofold increase in the overall concentration of Ty RNA in the cell (Boeke, unpublished data); this relatively modest increase in RNA concentration has dramatic consequences. The frequency of Ty transposition, the intracellular concentration of p90-TYB, the number of Ty viruslike particles in the cell, and the quantity of reverse transcriptase activity in the cell all increase by factors of 20 to 100. At least three broad classes of explanations for the discrepancy can be considered. (i) There may be a translational repressor which is specific for Ty RNA and possibly specific for the translation of *TYB*. (ii) An inhibitor of transposition may be produced by certain Ty element copies, as has been described for *Drosophila* P elements. (iii) Ty RNA deriving from the chromosomal copies of Ty may somehow differ in structure from transcripts deriving from pGTyH3. This last possibility can be further split into two subcategories of structural difference: (a) the RNA from chromosomal Ty elements could be structurally different from the plasmid-derived transcript, or (b) transcripts encoding functional and nonfunctional proteins could derive from different copies of Ty elements which are transcribed unequally. The above possibilities are explored below.

(i) A compelling, although indirect argument can be made against the existence of a translational repressor. Mutations introduced in vitro into many different sites scattered throughout pGTyH3 result in the loss of transposition of both the donor (*GAL1*-fused) Ty itself and chromosomal Ty elements (J. D. Boeke and G. Monokian, unpublished data). This result is inconsistent with the notion of a translational repressor which operates by binding specifically to a site on Ty RNA.

(ii) The existence of a transposition inhibitor is formally possible. For example, certain copies of Ty may specify the synthesis of gene products which inhibit rather than facilitate transposition. Such mutations would be dominant; perhaps twofold overproduction of wild-type Ty gene products could overcome such inhibition. Although this argument cannot presently be refuted, it is clear that Ty173 does not produce a *trans*-dominant inhibitor. The inhibitor argument also fails to explain the great discrepancy in amount of p90-TYB observed in transposition-induced and uninduced cells (13).

(iiia) The 5′ ends of pGTyH3-derived transcripts have been shown to be indistinguishable (by primer extension) from those encoded by chromosomal Ty elements, although pGTyH3 transcripts are produced by the heterologous *GAL1* promoter (A. Bystrom, J. D. Boeke, and G. R. Fink, unpublished data). The 3′ ends of *GAL1*-driven and LTR-driven Ty transcripts are unlikely to be different, because both pGTyH3 and chromosomal Ty elements contain intact 3′ δ (transcriptional terminator) sequences. Thus it seems unlikely that the structure of the termini of Ty transcripts could account for the observed discrepancy. An as yet

undiscovered form of covalent modification of one class of transcripts is a formal possibility, but seems unlikely and is unprecedented.

(iiib) Many of the chromosomal transcripts may encode nonfunctional, nonprocessable, or unstable gene products. If many or most of these mutant Ty elements produce either a mutant protease or *TYB* proteins which are unstable or poorly processed, it would explain the apparent lack of p90-TYB and, consequently, transposition in uninduced cells.

In support of possibility (iiib) above, we have shown here that one chromosomal Ty element, Ty173, whose structure is normal at a gross level (i.e., no significant insertions or deletions) contains a missense mutation (resulting in a Leu-to-Ile coding change) which renders its *TYB* protein(s) nonfunctional. Hybridization analysis with oligonucleotide probes for this mutation shows that about one-quarter of the Ty1 elements in the genome carry this inactivating mutation, as well as an associated third-position change which destroys a *Hha*I restriction site but does not alter coding (the *tyb-2103* mutation) 5 base pairs away. The fact that this pair of mutations is found in a discrete subset of Ty1 elements suggests that the mutation arose once and subsequently spread to other Ty elements in the genome. Of course, the mutation may also have arisen independently several times by the templating effect of a quasipalindromic sequence (8), although no good candidate sequence for such a mechanism could be identified. In either case, the presence of multiple copies of this mutant element in more than one strain suggests that it may confer a selective advantage to the host cell. Naturally, there may in addition be as yet undiscovered inactivating mutations in other Ty elements.

Since the transcription rate of a given Ty element is probably dependent on the chromosomal environment in which it finds itself (the LTRs being juxtaposed to different flanking regions in every case), it is formally possible to explain the extremely low transposition rates in wild-type cells by proposing that (i) the yeast genome contains both functional and nonfunctional copies of Ty elements and (ii) the functional Ty elements, by and large, are poorly transcribed or silent, whereas the nonfunctional elements produce a disproportionate amount of the Ty mRNA.

It is interesting in this regard that TyH3, which is clearly a functional Ty element, has six differences between its 5′ and 3′ LTR sequences. Transposition events derived from a single type of RNA template molecule would be predicted to contain identical LTR sequences, whereas transpositions resulting in dissimilar LTRs probably require that information is ultimately derived from two different RNA templates (see references 4, 7, 16, and 31 for discussions of retroviral recombination and replication). Thus, the heterogeneities in the LTR sequences of TyH3 suggest that TyH3 is a recombinant Ty which arose during the transposition event which generated it. The TyH3 sequence may be derived from transcripts of two nonfunctional Ty elements.

Cells may shuffle functional and nonfunctional transposon-coding sequences behind weak and strong LTR promoters, respectively, by using the cellular homologous recombination system. Such recombination events would modulate the level of transposition; a cell with an optimal level of transposition would obviously have a selective advantage. Interestingly, the *tyb-2098* mutation described above lies near the center of the Ty element, where it is more likely to participate in gene conversion events with other Ty elements. Perhaps mutations falling within this region are predisposed to spread throughout the genome. Predictions of this model are currently being tested.

ACKNOWLEDGMENTS

We thank C. Riley for preparing the synthetic oligonucleotides; R. F. Doolittle for communicating results prior to publication; and G. Natsoulis, K. Durbin, C. Dowling, H. Kazazian, and D. Garfinkel for helpful discussions.

This work was supported in part by Public Health Service grants GM35010 to G.R.F. and GM36481 to J.D.B. from the National Institutes of Health and in part by support from the Searle Scholars Program/The Chicago Community Trust to J.D.B. G.R.F. is an American Cancer Society Professor. D.E. was supported by Public Health Service training grant CA09139 from the National Institutes of Health.

LITERATURE CITED

1. **Boeke, J. D.** 1988. Retrotransposons, p. 59–103. *In* E. Domingo, J. Holland, and P. Ahlquist (ed.), RNA genetics, vol. 2. Retroviruses, viroids, and RNA recombination. CRC Press, Boca Raton, Fla.

1a. **Boeke, J. D., and D. J. Garfinkel.** 1988. Yeast Ty elements as retroviruses, p. 15–39. *In* M. J. Leibowitz and Y. Koltin (ed.), Viruses of fungi and simple eukaryotes. Marcel Dekker, Inc., New York.

2. **Boeke, J. D., D. J. Garfinkel, C. A. Styles, and G. R. Fink.** 1985. Ty elements transpose through an RNA intermediate. Cell **40:** 491–500.

3. **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.

4. **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.

5. **Cameron, R. R., E. Y. Loh, and R. W. Davis.** 1979. Evidence for transposition of dispersed repetitive DNA families in yeast. Cell **16:**739–751.

6. **Clare, J., and P. J. Farabaugh.** 1985. Nucleotide sequence of a yeast Ty1 element: evidence for an unusual mechanism of gene expression. Proc. Natl. Acad. Sci. USA **82:**2829–2833.

7. **Coffin, J. M.** 1979. Structure, replication, and recombination of retrovirus genomes: some unifying hypotheses. J. Gen. Virol. **42:**1–26.

8. **DeBoer, G. G., and L. S. Ripley.** 1984. Demonstration of the production of frameshift and base-substitution mutations by quasipalindromic sequences. Proc. Natl. Acad. Sci. USA **81:** 5528–5531.

9. **Elder, R. T., E. Y. Loh, and R. W. Davis.** 1983. RNA from the yeast transposable element Ty1 has both ends in the direct repeats, a structure similiar to retrovirus RNA. Proc. Natl. Acad. Sci. USA **80:**2432–2436.

10. **Errede, B., M. Company, and R. Swanstrom.** 1986. An anomalous Ty1 structure attributed to an error in reverse transcription. Mol. Cell. Biol. **6:**1334–1338.

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

12. **Fink, G. R., J. D. Boeke, and D. J. Garfinkel.** 1986. The mechanism and consequences of retrotransposition. Trends Genet. **2:**118–123.

13. **Garfinkel, D. J., J. D. Boeke, and G. R. Fink.** 1985. Ty element transposition: reverse transcriptase and virus-like particles. Cell **42:**507–517.

14. **Hauber, J., P. Nelbock-Hochstetter, and H. Feldman.** 1985. Nucleotide sequence and characteristics of a Ty element from yeast. Nucleic Acids Res. **13:**2745–2758.

15. **Johnson, M. S., M. A. McClure, D.-F. Feng, J. Gray, and R. F. Doolittle.** 1986. Computer analysis of retroviral *pol* genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. Proc. Natl. Acad. Sci. USA **83:**7648–7652.

16. **Junghans, R. P., L. R. Boone, and A. M. Skalka.** 1982. Retroviral DNA H structures: displacement-assimilation model of

recombination. Cell **30**:53–62.

17. **Kingsman, A. J., R. L. Gimlich, L. Clarke, A. C. Chinault, and J. Carbon.** 1981. Sequence variation in dispersed repetitive sequences in *Saccharomyces cerevisiae*. J. Mol. Biol. **145**:619–632.
18. **Kramer, W., V. Drutsa, H. W. Jansen, B. Kramer, M. Pflugfelder, and H. J. Fritz.** 1984. The gapped duplex DNA approach to oligonucleotide-directed mutation construction. Nucleic Acids Res. **12**:9441–9456.
19. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**: 499–560.
20. **Mellor, J., M. H. Malim, K. Gull, M. F. Tuite, S. McCready, T. Dibbayawan, S. M. Kingsman, and A. J. Kingsman.** 1985. Reverse transcriptase activity and Ty RNA are associated with virus-like particles in yeast. Nature (London) **318**:583–586.
21. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
22. **Ohashi, A., J. Gibson, I. Gregor, and G. Schatz.** 1982. Import of proteins into mitochondria. J. Biol. Chem. **257**:13042–13047.
23. **Roeder, G. S., and G. R. Fink.** 1983. Transposable elements in yeast, p. 299–326. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
24. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
25. **Scherer, S., C. Mann, and R. W. Davis.** 1982. Reversion of a promoter deletion in yeast. Nature (London) **298**:815–819.
26. **Sherman, F., G. R. Fink, and C. W. Lawrence (ed.).** 1978. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Simchen, G., F. Winston, C. A. Styles, and G. R. Fink.** 1984. Ty-mediated gene expression of the LYS2 and HIS4 genes of Saccharomyces cerevisiae is controlled by the same SPT genes. Proc. Natl. Acad. Sci. USA **81**:2431–2434.
28. **Studencki, A. B., and R. B. Wallace.** 1984. Allele-specific hybridization using oligonucleotide probes of very high specific activity: discrimination of the human beta A- and beta S-globin genes. DNA **3**:7–15.
29. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
30. **Tschumper, G., and J. Carbon.** 1986. High frequency excision of Ty elements during transformation of yeast. Nucleic Acids Res. **14**:2989–3001.
31. **Varmus, H. E., and R. Swanstrom.** 1984. Replication of retroviruses, p. 369–512. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Warmington, J. R., R. B. Waring, C. S. Newlon, K. J. Indge, and S. G. Oliver.** 1985. Nucleotide sequence characterization of Ty1-17, a class II transposon from yeast. Nucleic Acids Res. **13**:6679–6693.
33. **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**:1432–1448.