Nucleic Acids Research, 1996, Vol. 24, No. 23 **4825–4831**

The *Saccharomyces* retrotransposon Ty5 influences the organization of chromosome ends

Sige Zou, Jin M. Kim and Daniel F. Voytas*

Department of Zoology and Genetics, 2208 Molecular Biology Building, Iowa State University, Ames, IA 50011, USA

Received June 11, 1996; Revised and Accepted October 15, 1996

DDBJ/EMBL/GenBank accession nos U19263, U19264

ABSTRACT

Retrotransposons are ubiquitous components of eukaryotic genomes suggesting that they have played a significant role in genome organization. In *Saccharomyces cerevisiae*, eight of 10 endogenous insertions of the Ty5 retrotransposon family are located within 15 kb of chromosome ends, and two are located near the subtelomeric *HMR* locus. This genomic organization is the consequence of targeted transposition, as 14 of 15 newly transposed Ty5 elements map to telomeric regions on 10 different chromosomes. Nine of these insertions are within 0.8 kb and three are within 1.5 kb of the autonomously replicating consensus sequence in the subtelomeric X repeat. This suggests that the X repeat plays an important role in directing Ty5 integration. Analysis of endogenous insertions from *S.cerevisiae* and its close relative *S.paradoxus* revealed that only one of 12 insertions has target site duplications, indicating that recombination occurs between elements. This is further supported by the observation that Ty5 insertions mark boundaries of sequence duplications and rearrangements in these species. These data suggest that transposable elements like Ty5 can shape the organization of chromosome ends through both transposition and recombination.

INTRODUCTION

Telomeres are specific protein–DNA structures found at the termini of eukaryotic chromosomes (1). Telomere sequences typically consist of tandem arrays of simple repeats synthesized by telomerase, a cellular reverse transcriptase (2). Most organisms have short and precise telomeric repeat sequences that are evolutionarily conserved. The *Saccharomyces cerevisiae* telomeric sequences, however, are atypically heterogeneous and consist of arrays of $TG_{1–3}$ (3). A variety of middle repetitive sequences, called subtelomeric repeat sequences, are found associated with telomeres. These repeats are highly polymorphic and not well-conserved among eukaryotes.

In *S.cerevisiae*, subtelomeric repeat sequences have been studied in great detail (4–7). They are comprised of two major groups, called Y′ elements and X repeats. Y′ elements are immediately internal to the telomeres. There are two major classes that differ by size, called Y′-long (6.7 kb) and Y′-short (5.2 kb). The size differences are due to a series of small insertions and/or deletions. Y′ elements in the same class are highly conserved and typically share ~99% nucleotide identity. Y′ elements are found on most chromosomes and are highly polymorphic among different strains. At the end of any particular chromosome, Y′ elements are present in zero to four tandemly arranged copies.

X repeats are centromere-proximal when present in conjunction with Y′ elements (5). Junction sequences between X repeats and Y′ elements are normally short stretches of telomere sequences $TG_{1–3}$. On some chromosomes, such as chromosome III (chr III), no Y′ elements are present and X repeats are found immediately internal to the telomere. X repeats consist of a 473 bp core X sequence as well as varying numbers of short subtelomeric repeats (STR-A, STR-B, STR-C and STR-D) that range in size from 45 to 140 bp (5,6). Boundaries of X repeats vary dramatically due to the presence of STRs. The STRs are not present at all chromosome ends, while core X sequences are found in all but one subtelomeric region and share ~80% nucleotide identity.

The widespread and polymorphic distribution of X repeats and Y′ elements suggests that these subtelomeric sequences are in constant flux. Some Y′ elements have a large open reading frame (ORF) with weak homology to viral helicases, suggesting that Y′ elements may be related to transposable elements (4). However, transposition of Y′ elements has never been documented. On the other hand, Y′ recombination occurs frequently and depends on the *RAD52* gene, which is required for homologous recombination (4,8). Y′ elements preferentially recombine with members of the same size class, which results in a preponderance of one size class in any given strain. Recombination of Y′ elements can also result in exchange of sequences between ends of chromosomes. For example, gene conversion can replace sequences at one end of a chromosome with those from another end. Therefore, Y′ recombination can clearly reshape the organization of the chromosome ends. The presence of X repeats at the ends of almost all chromosomes suggests that they may also participate in homologous recombination. However, movement of X repeats by recombination has not been documented.

Ty5 is a retrotransposon identified from *S.cerevisiae* and its close relative *S.paradoxus* (9,10). The copy number and distribution of Ty5 insertions are polymorphic in different *Saccharomyces* species and strains. Characterization of eight endogenous Ty5 insertions in *S.cerevisiae* showed that two are associated with the silent mating locus *HMR* and six are located in subtelomeric

*To whom correspondence should be addressed. Tel: +1 515 294 1963; Fax: +1 515 294 0345; Email: voytas@iastate.edu

regions. None of these elements are transposition competent. In *S.paradoxus*, two of three characterized Ty5 insertions are likely subtelomeric, based on their association with a subtelomeric X repeat (10). A Ty5 transposition assay was developed in *S.cerevisiae* using an *S.paradoxus* element (11). Of 19 newly transposed Ty5 elements on chr III, four were inserted near the left telomere and 14 were inserted near transcriptional silencers at the *HMR* and *HML* loci. These regions are bound in a unique type of chromatin, called silent chromatin, which represses the transcription of adjacent genes (12). The target bias for Ty5 suggests that this element recognizes silent chromatin during integration.

Transposable elements have been found associated with telomeres in other organisms. In *Drosophila melanogaster*, for example, two families of non-LTR retrotransposons have been identified that serve as telomeres (13,14). The telomeric location of Ty5 suggests that it may play a role in genome organization, and in particular, contribute to the dynamic nature of chromosome ends through transposition and/or recombination. To test this, we characterized the genomic location of additional Ty5 transposition events. We also used the complete nucleotide sequence of the *S.cerevisiae* genome to compare the organization of endogenous elements between *S.cerevisiae* and *S.paradoxus*.

MATERIALS AND METHODS

Strains

The following yeast strains were used in this study: *S.paradoxus* NRRL Y-17217 (Northern Regional Research Laboratory); wild-type *S.cerevisiae* strains SK1 and S288C for the characterization of endogenous Ty5 insertions (J.D. Boeke, Johns Hopkins University); *S.cerevisiae* W303-1A (*MAT***a** *ade2-1 can1-100 his3-11 leu2-3 trp1-1 ura3-1*) for the Ty5 transposition assay (A. Myers, Iowa State University). The *E.coli* strain XL1-Blue (Stratagene) was used for recombinant DNA manipulations. Transformation of *E.coli* and yeast strains was performed by electroporation as described (15).

Mapping newly transposed Ty5 insertions

Sequences flanking newly transposed Ty5 elements were amplified by inverse PCR method as described (16). Briefly, ~100 ng of genomic DNA was digested with *Msp*I and self-ligated in a 50 µl ligation mixture. Sequences flanking Ty5 insertions were amplified from 2 µl of the ligation mixture with Ty5-6p LTR-specific oligonucleotides DVO219 (5′-TACTGTCGGATCGGAGGT-TT-3′) and DVO220 (5′-CTGTGTACAAGAGTAGTACC-3′). PCR products were sequenced with oligo- nucleotides DVO214 (5′-CCCTCGAGCATTTACATAACATATAGAAAG-3′) or DVO243 (5′-CCTTGTCTAAAACATTACTG-3′). Ty5 integration sites were determined by comparing these sequences to the *S.cerevisiae* genome database.

DNA manipulations and analysis

Yeast genomic DNA and chromosomes were prepared as described (15). The genomic DNA was digested with restriction enzymes and separated by agarose gel electrophoresis. Yeast chromosomes were separated by pulsed-field gel electrophoresis, and chromosome identity determined by their mobility (5). Gels containing genomic DNA or chromosomes were transferred to nylon membranes by alkaline transfer. Filters were hybridized with DNA fragments that had been radio-labeled by random-priming (Promega). Hybridization probes included Ty5 internal sequences (probe A and B in Fig. 3A), the long terminal repeat (LTR) (Fig. 3A), as well as sequences flanking Ty5 insertions. The LTR was amplified from Ty5-5p with oligonucleotides DVO182 (5′-GGGTAATGTTTC-AGT-3′) and DVO116 (5′-TAGTAAGTTTATTGGACC-3′). Sequence flanking the 5′-end of Ty5-12p element was amplified with DVO200 (5′-CATTACCCATATCATGCT-3′) and the reverse primer, which is complementary to the vector. DNA sequences were determined with the fmol sequencing kit (Promega), or by the Nucleic Acid Facility of Iowa State University. Sequence analysis was performed using the GCG computer programs (17). LTR sequences were identified from the complete nucleotide sequence of *S.cerevisiae* using the program BLAST (18). Sequences were considered that had >65% nucleotide identity to the Ty5-1 left LTR.

RESULTS

Ty5 preferentially transposes to subtelomeric regions in *S.cerevisiae*

We have previously characterized 19 newly transposed Ty5 elements on chr III. Of these, 18 occurred in regions of silent chromatin, including 14 at the *HMR* and *HML* mating loci and four at the left telomere (11). We hypothesized that this integration pattern is due to a mechanism that targets Ty5 to silent chromatin. For chromosomes other than chr III, the only known regions of silent chromatin are at the telomeres. We wanted to investigate whether Ty5 insertions on other chromosomes are near the telomeres, which would support a general role for Ty5 in the genomic organization of chromosome ends.

Fifteen strains were randomly chosen from a collection of strains with newly transposed Ty5 elements (11). Sequences flanking these elements were amplified by inverse PCR and used directly for DNA sequencing. These insertions were found to reside on ten different chromosomes (Fig. 1). One insertion, W3, is on chr XI 152 kb from the end of the chromosome. The remaining 14 insertions are all subtelomeric and are within 15 kb of chromosome ends. The insertions show no orientation specificity with respect to the ends of the chromosomes; eight insertions are in the same 5′ to 3′ orientation as the chromosome sequence, and seven are in opposite orientation.

Ty5 preferentially inserts near the X repeat

We have previously used the autonomously replicating consensus sequences (ACS) in the X repeat as a reference point for Ty5 insertions. In this study, nine of the 15 Ty5 insertions are within 0.8 kb on either side of the ACS in the X repeat; three additional insertions are within 1.5 kb. Eight of the 11 telomeres that have Ty5 insertions also have Y′ elements. These Y′ elements separate the X repeat from the telomere by >5 kb. In all eight cases, Ty5 insertions are clustered within 1.5 kb of the ACS in the X repeat and are consequently several kilobases from the $TG_{1–3}$ telomeric repeats. This suggests that the X repeat serves as a nucleation site for factors important for Ty5 targeting.

Features of sequences flanking endogenous Ty5 insertions in the subtelomeric regions of *S.cerevisiae*

The clear preference for Ty5 to integrate near the telomeres indicates an active role for Ty5 in shaping the organization of chromosome ends. We took advantage of the complete genome

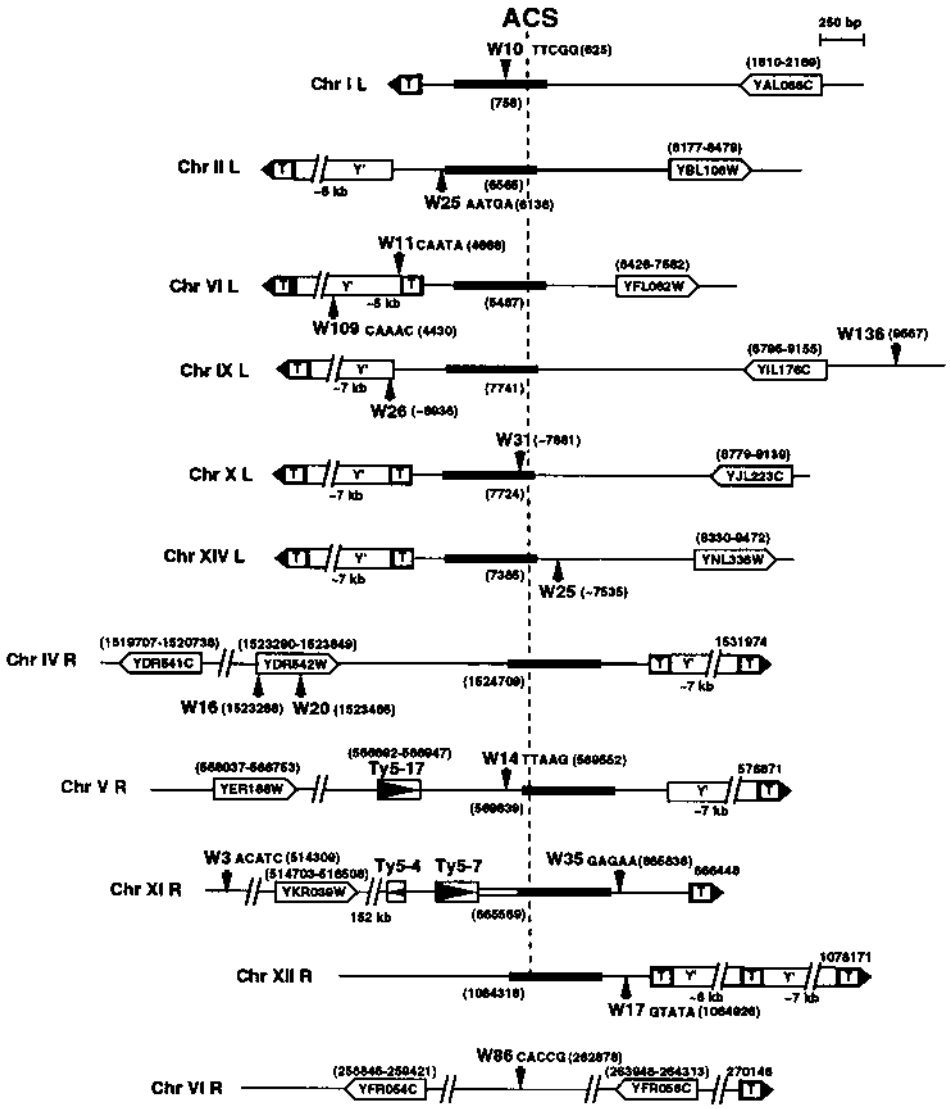

Figure 1. Location of *de novo* Ty5 transposition events. Left and right arms of chromosomes are labeled as L and R, respectively. T indicates telomeric repeat sequences, and the narrow boxes indicate X repeats, which are aligned at the ACS (the autonomously replicating consensus sequences). The sizes of Y′ elements are provided. Open boxes with arrowheads depict endogenous Ty5 LTRs, and open boxes with labels indicate open reading frames. The Ty5 insertions are labeled as W followed by a number, which refers to the strain from which they were isolated. Arrows pointing down indicate insertions in the same 5′ to 3′ orientation as the chromosome sequences. Arrows pointing up represent insertions in the opposite orientation. Numbers in parentheses designate base positions of open reading frames and Ty5 insertion sites. The base position of the ACS is provided adjacent to the dashed line, which marks the position of this sequence.

sequence of *S.cerevisiae* to evaluate whether endogenous Ty5 insertions have played a passive role in genome organization through recombination. We identified all Ty5 insertions in S288C, the strain used for the yeast genome project (Fig. 2A). Ten Ty5 insertions were found, including eight previously identified on chr III, VII, VIII and XI (10). The chr VII insertion (designated Ty5-15) had previously been characterized only by Southern hybridization analysis (10). Two new insertions were identified on chr V, designated Ty5-16 and Ty5-17. The chr V insertions are near the right telomere but are in opposite orientation. Ty5-17 is within 600 bp of a X ACS, and Ty5-16 is within 2.7 kb.

Target site sequences were characterized for eight Ty5 insertions with full-sized LTRs. None of these insertions have the perfect 5 bp target site duplications characteristic of newly transposed Ty5 elements (11), although Ty5-16 has flanking target sequences with four identical nucleotides out of five (Table 1). The 5′ target site of Ty5-17 is the same as the 3′ target site of Ty5-16. However, the 3′ target site of Ty5-17 is different from the 5′ target sites of these two insertions, suggesting that a gene conversion event, or two sequential reciprocal recombination events, occurred between these elements.

Some Ty5 elements mark boundaries of duplicated sequences in the *S.cerevisiae* genome. Genome sequencing efforts have identified extensive duplications between the telomeric regions of chr III and XI (19–21). Four Ty5 insertions are present in these duplicated regions (Fig. 2B). The chromosome ends, including the X repeat, are similar between the chr III left telomere and both telomeres of chr XI (region **a**). The similarity ends at the Ty5-1 insertion on chr III. Downstream of region **a**, the chr XI left-end has a unique 50 bases and both chr XI ends share a second duplicated sequence (region **b**). For the chr XI left-end, the **b** region terminates in sequences that have been duplicated from the

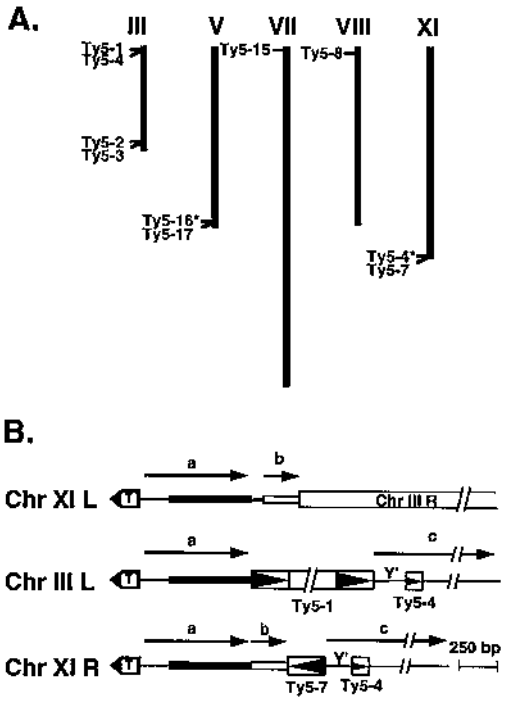

Figure 2. *Saccharomyces cerevisiae* Ty5 elements. (**A**) Chromosome location of endogenous Ty5 elements. Chromosomes are drawn to scale with the left end on top. The asterisk reflects elements on the Crick strand of the chromosome sequences. Base positions for insertions are: Ty5-1, 1172–4314; Ty5-2, 290646–290891; Ty5-3, 291015–291252; Ty5-4 on chr III, 4471–4572; Ty5-4 on chr XI, 664909–664808; Ty5-7, 665062–665300; Ty5-8, 7993–8224; Ty5-15, 863–1079; Ty5-16, 562209–562459; Ty5-17, 564300–564533. (**B**) Sequence rearrangements between chr III and chr XI of *S.cerevisiae*. Symbols are as in Figure 1. Duplicated sequence domains are indicated by arrows and designated as a, b or c. The Y′ element sequence in region c is labeled. The open box labeled chr III R represents an additional sequence duplication between chr XI L and chr III R.

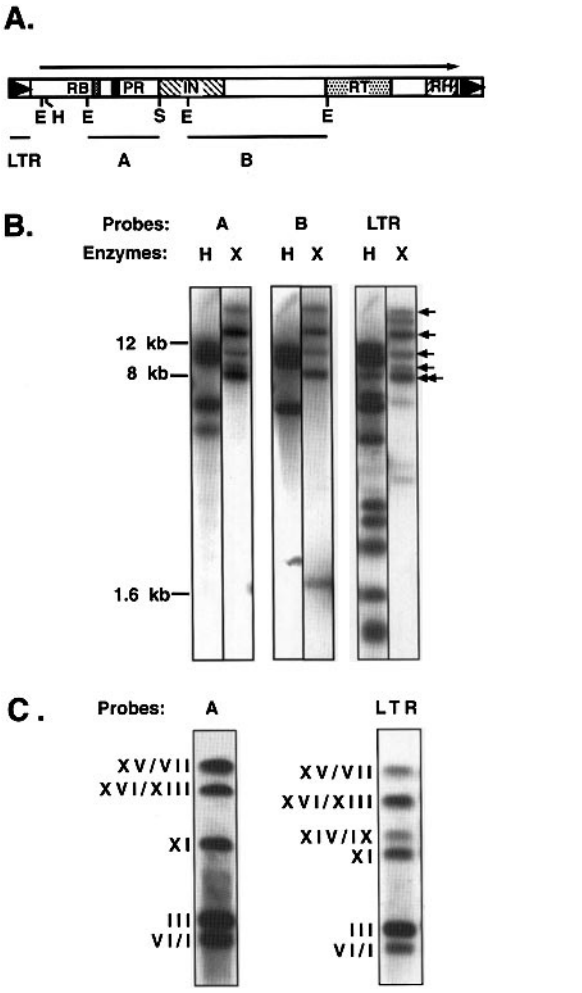

Figure 3. Copy number of Ty5 elements in strain NRRL Y-17217. (**A**) Genomic organization of Ty5-6p. Open boxes with arrows indicate the LTRs. Boxes within the internal domain depict conserved amino acid sequence domains: RB, RNA binding domain; PR, protease; IN, integrase; RT, reverse transcriptase; RH, RNase H. The arrow over the element indicates the open reading frame. E, H and S denote restriction endonuclease sites for *Eco*RI, *Hin*dIII and *Sma*I. The lines under the element depict probes used for hybridization analyses. (**B**) Southern hybridization analyses of strain NRRL Y-17217. DNA was digested with *Hin*dIII (H) or *Xho*I (X). Molecular length markers are indicated, and arrows denote *Xho*I restriction fragments that hybridize to probe A. Arrowheads indicate bands that hybridize to both internal and LTR probes. (**C**) Chromosome distribution of Ty5 in NRRL Y-17217. Filters were prepared from pulsed-field gels and hybridized with probe A and the LTR probe. Chromosome designations for hybridizing chromosomes are indicated.

right-end of chr III. The right-end of chr XI, however, has a Ty5 insertion at the end of the **b** region (Ty5-7). This insertion has different target sites from Ty5-1 and is in the opposite orientation, clearly indicating that they are different insertions. Centromere proximal to Ty5-1 and Ty5-7 are several kb of duplicated sequences (region **c**), including a Ty5 insertion (Ty5-4). It has previously been noted that the beginning of the **c** region contains 140 bp of a Y′ element (22). The location of Ty5-1 and Ty5-7 at the boundaries of rearrangements suggests that these elements have played a role in these events.

Distribution of Ty5 insertions in *S.paradoxus*

Characterization of Ty5 elements showed that transposition-competent insertions are not present in *S.cerevisiae*, but are present in its close relative *S.paradoxus*. The close relationship between these species suggests that they may be a good model further understanding the role of Ty5 in genome organization. We focused our characterization efforts on the *S.paradoxus* strain NRRL Y-17217, which harbors the most Ty5 insertions, at least one of which is transpositionally active (10,11). We first estimated copy number in this strain by Southern hybridization analysis using restriction enzymes that do not cut within the element or cut only once. Filters were hybridized with probes specific to either Ty5 internal sequences or Ty5 LTRs (Fig. 3). An example of this analysis is presented in Figure 3B, using *Hin*dIII (one internal site) and *Xho*I (no cut sites). Based on this analysis, at least six restriction fragments hybridized to both internal and LTR sequences in NRRL Y-17217. In the *Hin*dIII lanes, there are at least seven more restriction fragments that hybridized to the LTR. Some of these extra LTRs may be solo LTRs, which are derived from recombination between LTRs of full-length Ty5 elements.

The chromosomes of NRRL Y-17217 were separated by pulsed-field gel electrophoresis and transferred to nylon filters. The chromosome location of Ty5 insertions was analyzed by hybridizing the filter with either Ty5 internal or LTR sequences (Fig. 3C). Ty5 elements with internal sequences were located on at least five chromosomes, namely chr VI or I, III, XI, XVI or XIII and XIV or VII. Ty5 LTRs were located on these same chromosomes as well as chr XIV or IX.

Table 1. Sequences of Ty5 target sites

Species	Insertion (chr)	Left target sequence	Right target sequence	Target nucleotide identity	LTR nucleotide identity (%)[a]
S.cerevisiae	Ty5-1 (III)	TTTCA	TATCC	3/5	86, 89
	Ty5-2 (III)	TTCCT	TAAAA	1/5	73
	Ty5-3 (III)	ATCGC	TTTGC	3/5	55
	Ty5-7 (XI)	CGTGG	TACCG	1/5	82
	Ty5-8 (VIII)	GTATA	ATATG	3/5	77
	Ty5-15 (VII)	TTTCA	CCCAA	1/5	81
	Ty5-16 (V)	GTTAT	GTTCT	4/5	91
	Ty5-17 (V)	GTTCT	TTACA	2/5	73
S.paradoxus	Ty5-6p (XI)	TCGTA	TCGTA	5/5	100
	Ty5-5p (III/XI)	TGTCA	CTATC	0/5	100, 98
	Ty5-10p (III/XI)	AGTAT	TATAA	2/5	98
	Ty5-12p (III/XI)	AGTAT	TTTTC	1/5	97
	Ty5-14p (ND)	–	TGTCA	–	98

[a]Values are derived from comparisons with the Ty5-6p right LTR. Multiple values refer to comparisons with the left and right LTRs, respectively.

Duplication and rearrangement of sequences flanking Ty5 insertions in *S.paradoxus*

We previously isolated five of the approximately 13 Ty5 insertions present in *S.paradoxus* strain NRRL Y-17217 (10). To investigate the relationship between Ty5 and genome organization, sequences flanking these insertions were determined. The target sites of several insertions were analyzed, and only the Ty5-6p insertion was found to have target site duplications (Table 1). It is interesting to point out that the 5 bp target sequence at the 3′-end of Ty5-14p is the same as the 5′ target sequence of Ty5-5p. This suggests that Ty5-5p and Ty5-14p recombined and exchanged targets, and Ty5-14p subsequently suffered a deletion of its 5′ region (Fig. 4). Evidence for such a reciprocal translocation is also supported by sequence differences among the LTRs of these elements. The 3′ LTR of Ty5-5p has four bp that differ from the 5′ LTR. The Ty5-5p 3′ LTR, however, is identical at these four nucleotide positions to the 5′ LTR of Ty5-14p, arguing strongly that a recombination event had occurred between these elements.

Flanking sequences of several *S.paradoxus* insertions were compared to the *S.cerevisiae* genome database or used in Southern hybridization analysis. As previously reported, 5′ and 3′ sequences flanking Ty5-6p share ~90% nucleotide identity to sequences on *S.cerevisiae* chr XI (10). The 5′ flanking sequence hybridized to chr XI of *S.paradoxus* as well as *S.cerevisiae* (Fig. 5A), indicating that Ty5-6p is located on *S.paradoxus* chr XI and its flanking sequences are conserved between the two species. No evidence for a Ty5 insertion, however, was found at the corresponding region on *S.cerevisiae* chr XI, suggesting Ty5-6p transposed to this site after species divergence.

Analysis of flanking sequences of some *S.paradoxus* insertions support a role for Ty5 in genome rearrangements. Several Ty5 elements were flanked by sequences unique to *S.paradoxus*. For example, the 5′ flanking sequence of Ty5-5p shows no significant homology to any *S.cerevisiae* sequences, while the 3′ flanking sequence shares high homology with the subtelomeric X repeat. Southern analysis indicated that the unique 5′ flanking sequence hybridizes to *S.paradoxus* chr III and XI, suggesting that this

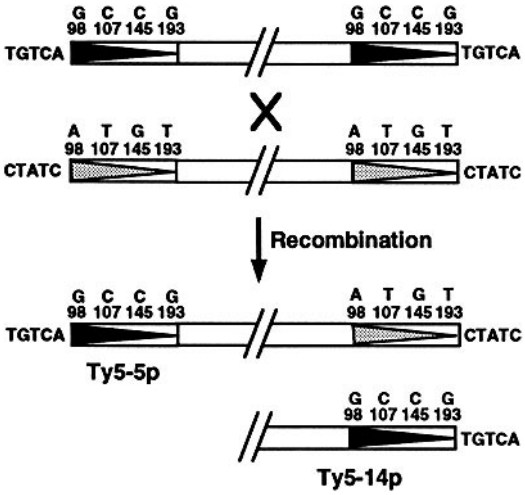

Figure 4. Proposed model for the reciprocal translocation that generated Ty5-5p and Ty5-14p. Two parental elements are shown with LTRs that differ by black or speckled arrowheads. Nucleotide differences between these LTRs are shown, with numbers indicating base positions. Sequences of target sites on either side of the elements are provided.

sequence is duplicated between these chromosomes (Fig. 5B). Since Ty5-5p is associated with a subtelomeric X repeat, the duplication may have occurred between the ends of chr III and XI. Sequence analysis of Ty5-14p indicated that it has a 5′ deletion, which includes the 5′ LTR. The 3′ flanking sequence has no significant similarity to any *S.cerevisiae* sequences.

Flanking sequences of some elements suggest that Ty5 insertions mark sites that have been rearranged between *S.paradoxus* and *S.cerevisiae*. For example, the 5′ sequence of Ty5-12p hybridized to *S.cerevisiae* chr V (Fig. 5C). This sequence, however, hybridized to *S.paradoxus* chr III and XI, indicating that it has been duplicated and rearranged between these species. Consistent with the hybridization analysis, the 5′ flanking sequence shares 90% nucleotide identity with a subtelomeric

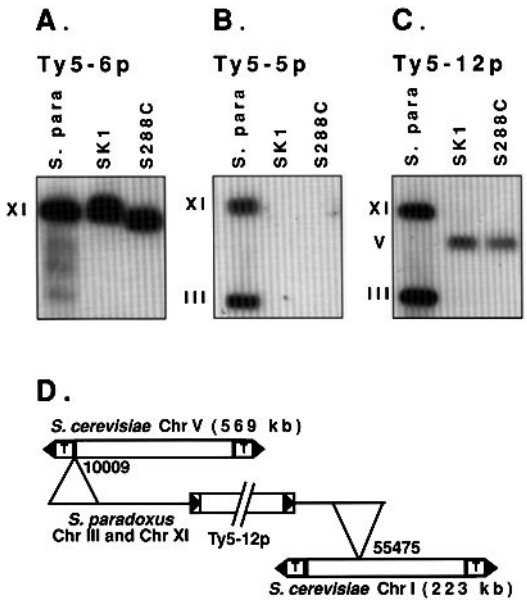

Figure 5. Organization of sequences flanking *S.paradoxus* Ty5 insertions in *S.paradoxus* and *S.cerevisiae*. S. para designates *S.paradoxus* and SK1 and S288C are wild-type strains of *S.cerevisiae*. Filters prepared from pulsed-field gels were hybridized with 5′ flanking sequences of Ty5-6p (**A**), Ty5-5p (**B**) and Ty5-12p (**C**). (**D**) Rearrangements of sequences flanking Ty5-12p. The 5′ flanking sequence of Ty5-12p is located on chr III and chr XI of *S.paradoxus*, but is located on chr V of *S.cerevisiae*. The number denotes the base position corresponding to the *S.cerevisiae* chr V nucleotide sequence. The 3′ flanking sequence shows homology to the middle of *S.cerevisiae* chr I, with the number indicating the base position within the chromosome.

region of *S.cerevisiae* chr V (Fig. 5D). The 3′ sequence of Ty5-12p shares 90% identity to the middle of the left arm of *S.cerevisiae* chr I, providing evidence for additional rearrangements. Part of the flanking sequences of insertion Ty5-11p were also determined. The 5′ sequence shares 84% identity to the subtelomeric region of *S.cerevisiae* chr XVI, but the 3′ sequence shares 84% identity to the subtelomeric region of *S.cerevisiae* chr XIII. Although the chromosome location of this element was not determined by hybridization analysis, these results indicate a sequence rearrangement between *S.paradoxus* and *S.cerevisiae* and implicate a role for Ty5 elements in genome rearrangements.

DISCUSSION

Ty5 is a subtelomeric repeat

From the available *S.cerevisiae* genomic sequence, we have identified 10 endogenous Ty5 insertions. Of these, seven are located within 10 kb of the ends of chromosomes, and one is within 15 kb. Two insertions are near the *HMR* locus. If *HMR* (~20 kb from the end of chr III) is considered a subtelomeric region, then all the endogenous Ty5 insertions are in the vicinity of chromosome ends. In addition, characterization of Ty5 in different *Saccharomyces* species and strains indicated that Ty5 distribution is polymorphic and varies in copy number and chromosome location (10). We have previously mapped the locations of 19 newly transposed Ty5 insertions on chr III (11). One insertion was in the middle of the right arm of chr III; 14 were at the silent mating loci, *HML* and *HMR*; four were within 2 kb of the left telomere. Here we mapped the locations of 15 newly transposed Ty5 insertions on 10 chromosomes other than chr III, which are not known to have silent chromatin except at the telomeres. Fourteen of fifteen are located within 15 kb from the ends of these chromosomes. These results indicate that Ty5 is a subtelomeric repetitive element based on its location and polymorphic nature.

At the subtelomeric regions of *S.cerevisiae*, the X repeats and Y′ elements are the two most abundant repetitive sequences (7). Y′ elements are immediately adjacent to the telomere sequences and are found at the subtelomeric regions of most but not all chromosomes (4). Internal to Y′ elements are X repeats, which are found at the ends of all but one chromosome. Chromosome ends, therefore, have a relatively rigid organization; internal to the telomere sequences are Y′ elements, followed by X repeats. In contrast, the location of Ty5 is very flexible. Insertions can be found within the telomere sequences, between X repeats and the telomere sequences or centromere-proximal to X repeats or Y′ elements. Unlike Y′ elements, Ty5 insertions can occur in either orientation with respect to the chromosome end. Ty5 transposition, therefore, has a regional specificity as opposed to a site specificity.

A potential role for the X repeat in directing Ty5 transposition

In our previous study of Ty5 insertions on chr III, 14 of 19 elements were clustered near the E and I transcriptional silencers that flank *HML* and *HMR*. We have recently shown that the assembly of silent chromatin mediated by these silencers is critical for Ty5 targeting (S. Zou and D. F. Voytas, unpublished). Of the 14 telomeric insertions identified in this study, and the four telomeric insertions on chr III, 12 are located within 0.8 kb of the X repeats. Particularly notable are insertions on chromosome ends with Y′ elements, which are all >5 kb from the chromosome ends. The $TG_{1–3}$ telomeric sequences can assemble silent chromatin in the absence of X repeats (23). However, the clustering of Ty5 insertions near X sequences suggests that some unique feature of the X repeat directs Ty5 integration. The X repeats have binding sites for the origin recognition complex and the transcription factor ABF1. Future experiments will test the role of these binding sites in directing Ty5 transposition.

Transposable elements and the origin of subtelomeric repeats

Telomere repeat sequences are generated by reverse transcription, which is carried out by telomerase (2). Telomerase is the only known reverse transcriptase that is not associated with retroelements, and may have originated from a retrotransposon or a retrovirus. The Y′ elements have some features of transposable elements; however, transposition of Y′ elements has never been demonstrated (4). Although X repeats are conserved among *Saccharomyces* species, their origin is largely unknown (5). In contrast to the other repeat sequences, the Ty5 elements are typical LTR retrotransposons and actively transpose to subtelomeric regions. This provides direct evidence that subtelomeric repeats can originate from transposable elements. The link between transposable elements and telomeres is further substantiated by the observation that the HeT and TART transposable elements of *Drosophila melanogaster* serve as telomeres (13,14). Transposable elements, therefore, may generally contribute to the structure of chromosome ends.

Ty5 and the organization of subtelomeric regions

The Ty5 elements are unique among yeast subtelomeric repeats in that they can shape chromosome ends actively through transposition and passively through recombination. The copy number of Ty5 varies extensively in strains of both *S.cerevisiae* and *S.paradoxus* (10). Although functional Ty5 elements do not exist in *S.cerevisiae*, they still influence genome organization through recombination, similar to the Y′ elements. In *S.paradoxus*, however, there are transposition-competent elements that can actively participate in restructuring chromosomes. The close relationship between the two yeast species makes them ideal models to more precisely evaluate the extent to which Ty5 has influenced genome organization, especially in light of the complete genome sequence of *S.cerevisiae*.

Analysis of *de novo* transposition events clearly demonstrated that Ty5 generates 5 bp target site duplications (11). Characterization of endogenous insertions, however, showed that only one is flanked by such duplications. There are two possibilities to explain this phenomenon. First, the absence of target site duplications may be due to random mutation. We reason that if this is the case, the LTR sequences among different Ty5 insertions should be degenerate to a similar extent as the target sites. The LTRs of Ty5-5p and Ty5-12p share >98% nucleotide identity with the transpositionally functional Ty5-6p LTR, suggesting that these insertions are not ancient and their target sites should not have mutated dramatically. Nonetheless, the target sites of Ty5-5p share no similarity and the target sites of Ty5-12p have only one nucleotide in common. It is difficult to argue that mutation alone could be responsible for the extreme differences in target site sequences.

A second possibility is that recombination between elements resulted in the lack of target site duplications. There are several examples that directly support this model. The 5 bp at the 5′ target site of Ty5-5p are the same as those at the 3′ target site of Ty5-14p, suggesting these two insertions recombined, resulting in the exchange of target sites. We can not tell whether the 3′ target site of Ty5-5p is the same as the 5′ target site of Ty5-14p, because Ty5-14p has suffered a deletion of its 5′ LTR. Additional support, however, is provided by four nucleotide differences between the LTRs of Ty5-5p. These four nucleotides in the Ty5-5p 5′ LTR are shared with the 3′ LTR of Ty5-14p, suggesting that these LTRs originated from the same element. Evidence for recombination is also found among the *S.cerevisiae* elements. The 5′ target sequence of Ty5-17 shares only two nucleotides with its 3′ target sequence, but is the same as the 3′ target sequence of Ty5-16; the target sites of Ty5-16 differ by only one nucleotide. A possible explanation for this observation is that gene conversion occurred between Ty5-16 and Ty5-17. In this process, the 5′ target site of Ty5-17 was replaced by the target sequence of Ty5-16, but the other site of Ty5-17 and both target sites of Ty5-16 remained unchanged. The 5′ target site of Ty5-16 subsequently mutated, resulting in one nucleotide difference. Experiments need to be conducted to test whether subtelomeric Ty5 elements can recombine among themselves and exchange their newly acquired target sites.

Recombination between repetitive sequences has likely played an important role in restructuring chromosomes. We have obtained evidence that Ty5 has been involved in recombination events, and these events have reorganized chromosomes in *S.cerevisiae* and *S.paradoxus*. For example, the 5′ flanking sequence of Ty5-12p is located on chr V of *S.cerevisiae* but is duplicated on chr III and chr XI in *S.paradoxus*. The 5′ flanking sequence of Ty5-5p is duplicated between chr III and chr XI in *S.paradoxus*, but is completely absent from *S.cerevisiae*. Similarly, duplicated sequences between chr III and XI in *S.cerevisiae* have boundaries that are marked by Ty5 insertions. Recombination between some subtelomeric repeats, such as Y′ elements, has been well characterized (4). Taken together, these observations clearly support the role of repetitive sequences, including transposable elements, in influencing the organization of chromosome ends. Further characterization of the genomes of closely related species such as *S.paradoxus* and *S.cerevisiae* will likely offer additional perspective on the extent to which transposable elements have shaped chromosome architecture.

ACKNOWLEDGEMENTS

We would like to thank M. Cummings for critical reading of the manuscript and David Wright for technical assistance. This work was supported by a grant from the American Cancer Society and an American Cancer Society Junior Faculty Research Award to D.F.V. This is journal paper no. J-16941 of the Iowa Agriculture and Home Economics Experiment Station (Ames); Project no. 3120.

REFERENCES

1 Zakian,V.A. (1995) *Science*, **270**, 1601–1607.
2 Blackburn,E. (1992) *Annu. Rev. Biochem.*, **61**, 113–129.
3 Shampay,J., Szostak,J. and Blackburn,E. (1984) *Nature*, **310**, 154–157.
4 Louis,E.J. and Haber,J.E. (1992) *Genetics*, **131**, 559–574.
5 Louis,E.J., Naumova,E.S., Lee,A., Naumov,G. and Haber,J.E. (1994) *Genetics*, **136**, 789–802.
6 Fiona,E.P., Tonie,C.H. and Louis,E.J. (1995) *Yeast*, **11**, 371–382.
7 Chan,C.S.M. and Tye,G.-K. (1983) *Cell*, **33**, 563–573.
8 Lundblad,V. and Blackburn,E.H. (1993) *Cell*, **73**, 347–360.
9 Voytas,D.F. and Boeke,J.D. (1992) *Nature*, **358**, 717.
10 Zou,S., Wright,D.A. and Voytas,D.F. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 920–924.
11 Zou,S., Ke,N., Kim,J.M. and Voytas,D.F. (1996) *Genes Dev.*, **10**, 634–645.
12 Laurenson,P. and Rine,J. (1992) *Microbiol. Rev.*, **56**, 543–560.
13 Biessmann,H., Mason,J.M., Ferry,K., d'Hulst,M., Valgeirsdottir,K., Traverse,K.L. and Pardue,M.L. (1990) *Cell*, **61**, 663–673.
14 Levis,R.W., Ganesan,R., Houtchens,K., Tolar,L.A. and Sheen,F.M. (1993) *Cell*, **75**, 1083–1093.
15 Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1987) *Current Protocols in Molecular Biology*. Greene/Wiley Interscience, New York.
16 Ochman,H., Gerber,A.S. and Hartl,D.L. (1988) *Genetics*, **120**, 621–623.
17 Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
18 Altschul,S., Gish,W., Miller,W., Myers,E. and Lipman,D. (1990) *J. Mol. Biol.*, **215**, 403–410.
19 Gromadka,R., Gora,M., Zielenkiewicz,U., Slonimski,P. and Rytka,J. (1996) *Yeast*, **12**, 583–591.
20 Oliver,S.G., van der Aart,Q.J.M., Agostoni-Carbone,M.L., Aigle,M., Alberghina,L., Alexandraki,D., Antoine,G., Anwar,R., *et al.* (1992) *Nature*, **357**, 38–46.
21 Dujon,B., Alexandraki,D., Andre,B., Ansorge,W., Baladron,V., Ballesta,J.P., Banrevi,A., Bolle,P.A., Bolotin-Fukuhara,M., Bossier,P. *et al.* (1994) *Nature*, **369**, 371–378.
22 Louis,E. (1995) *Yeast*, **11**, 1553–1573.
23 Gottschling,D.E., Aparicio,O.M., Billington,B.L. and Zakian,V.A. (1990) *Cell*, **63**, 751–762.