Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III

Scott E. Devine and Jef D. Boeke¹

Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205 USA

Retroviruses and their relatives, the LTR-containing retrotransposons, integrate newly replicated cDNA copies of their genomes into the genomes of their hosts using element-encoded integrases. Although target site selection is not well understood for this general class of elements, it is becoming clear that some elements target their integration events to very specific regions of their host genomes. Evidence is accumulating that the yeast retrotransposon Ty1 behaves in this manner. Ty1 is found frequently adjacent to tRNA genes in the yeast genome and experimental evidence implicates these regions as preferred integration sites. To determine the basis for Ty1 targeting, we developed an *in vivo* integration assay using a Ty1 donor plasmid and a second target plasmid that could be used to measure the relative frequency of Ty1 integration into sequences cloned from various regions of the yeast genome. Targets containing genes transcribed by RNA polymerase III (Pol III) were up to several hundredfold more active as integration targets than "cold" sequences lacking such genes. High-frequency targeting was dependent on Pol III transcription, and integration was "region specific," occurring exclusively upstream of the transcription start sites of these genes. Thus, Ty1 has evolved a powerful targeting mechanism, requiring Pol III transcription to integrate its DNA at very specific locations within the yeast genome.

[**Key Words:** Integration; retrotransposon; Ty1; targeting; transcription; RNA polymerase III]

Received November 17, 1995; revised version accepted January 19, 1996.

Ty1 is a long terminal repeat (LTR)-containing retrotransposable element in *Saccharomyces cerevisiae* that resembles vertebrate retroviruses structurally and functionally (Boeke and Sandmeyer 1991). It is one of five retrotransposons in this yeast (Ty1–Ty5), and is a representative of the Ty1-copia family of retroelements (Boeke and Sandmeyer 1991; Voytas and Boeke 1992). Like other retrotransposons and retroviruses, Ty1 transposes through an RNA intermediate (Boeke et al. 1985), and produces intracellular virus-like particles (VLPs) (Garfinkel et al. 1985). VLPs contain several element-encoded proteins required for Ty1 replication and integration into the host genome, including reverse transcriptase (RT) (Garfinkel et al. 1985) and integrase (IN) (Eichinger and Boeke 1988). RT is required for the replication of Ty1 genomic RNA into double-stranded cDNA, whereas IN mediates the insertion of the newly replicated cDNA into the host genome.

A remarkable aspect of Ty1 is that integrated copies are frequently found near tRNA genes in the yeast genome. In fact, Cameron et al. (1979) first reported Ty1 as a class of repetitive DNA associated with tRNA genes. The *S. cerevisiae* genome project has provided convinc-

ing additional evidence for this association, with solo LTRs and full Ty1 elements found almost exclusively next to tRNA genes on those chromosomes sequenced to date (Oliver et al. 1992; Dujon et al. 1994; Feldman et al. 1994; Johnston et al. 1994; Bussey et al. 1995). Experimentally induced Ty1 transposition events are also targeted frequently to integration hot spots located near tRNA genes. In a study of >100 insertions generated experimentally in yeast chromosome III (chr III), integration events were clustered at hot spots located near tRNA genes and solo LTRs (Ji et al. 1993); these hot spots were separated by "cold" regions of 10–60 kb lacking transposition events suggesting that Ty1 integration is targeted to regions containing tRNA genes and LTRs (Ji et al. 1993). Although an association exists between such regions and integration hot spots, the mechanism responsible for Ty1 target site selection remains unclear.

Interestingly, the yeast retrotransposon Ty3 targets selectively tRNA genes and other genes transcribed by RNA polymerase III (Pol III) for integration (Sandmeyer et al. 1988; Chalker and Sandmeyer 1990, 1992). Genome sequencing efforts have established clearly that Ty3 and its solo LTRs are associated almost exclusively with tRNA genes in the yeast genome (Oliver et al. 1992; Dujon et al. 1994; Feldman et al. 1994; Johnston et al.

¹Corresponding author.

1994; Bussey et al. 1995), and experimentally induced transposition events are also targeted to such genes (Chalker and Sandmeyer 1990, 1992). Ty3 integrates precisely at, or within a few bases of, the RNA Pol III transcription start sites of these genes, and integration is dependent on the presence of the Pol III transcription factors TFIIB and TFIIC at the target gene (Kirchner et al. 1995). Because Ty3 targeting occurs *in vitro* in the absence of active transcription with only a limited subset of Pol III transcription factors, protein–protein interactions between the Ty3 integration machinery and a subset of RNA Pol III transcription factors are likely to be responsible for targeting (Kirchner et al. 1995).

Chromosome sequencing projects have revealed that the remaining retrotransposons of yeast, Ty2, Ty4, and Ty5, also show evidence for targeted rather than random integration (Oliver et al. 1992; Dujon et al. 1994; Feldman et al. 1994; Johnston et al. 1994; Bussey et al. 1995). The genomic distributions of Ty2 and Ty4 most closely resemble Ty1, suggesting that these transposons have similar targeting mechanisms. The distribution of Ty5, on the other hand, is unique, indicating that this transposon is likely to have a very different targeting mechanism. Ty5 was identified originally in the telomeric regions of yeast chr *III* (Oliver et al. 1992; Voytas and Boeke 1992), and endogenous Ty5 insertions are found in subtelomeric regions or near the silent mating locus HMR (Zou et al. 1995). This distribution suggested that Ty5 may target its integration events to silenced chromatin, and recent experimental evidence supports such Ty5 targeting (Zou et al. 1996).

Certain poly(A) retrotransposons of the non-LTR class (lacking LTRs) also show evidence for targeted integration at specific regions of the genome. The DRE element of *Dictyostelium* is a particularly relevant example. DRE is found associated exclusively with tRNA genes in the *Dictyostelium* genome, and is found generally 50 bp upstream of the Pol III transcription start sites of these genes (Marschalek et al. 1989). Although other non-LTR retroelements, such as human LINE-1, display no known integration preferences (beyond a possible bias for purine-rich stretches), the R1 and R2 elements of insects integrate sequence-specifically within rDNA (Jakubczak et al. 1991). Thus, targeted integration is observed for a variety of retrotransposons.

Mammalian and avian retroviruses have also been examined for integration patterns within their respective hosts (for review, see Sandmeyer et al. 1990). Although there are studies supporting random integration models for retroviruses (Withers-Ward et al. 1994), there are also those supporting nonrandom integration models (King et al. 1985), as well as those demonstrating regions of highly preferred integration (Shih et al. 1988). Nevertheless, all studies to date indicate the presence of multiple integration sites within mammalian or avian genomes. Recently, a host-encoded factor In1, which interacts with human immunodeficiency virus (HIV) integrase and resembles the yeast transcription factor Snf5, was identified (Kalpana et al. 1994). The interaction of a transcription factor with this integrase suggests that HIV in-

tegration may be directed by host factors to specific genomic regions.

To further investigate Ty1 integration targeting, we have developed a plasmid-based genetic assay in yeast to test various sequences as targets for Ty1 integration. We isolated the most preferred integration hotspot of yeast chr *III* along with its flanking regions and tested it as an integration target using this *in vivo* assay. Our results indicate that preferred integration at the hot-spot site is dependent on the presence of a tRNA gene, *SUF16*, which makes the target >100-fold more active for Ty1 integration than otherwise identical targets lacking this gene. Moreover, hot-spot activity was dependent on RNA Pol III transcription at the target tRNA gene. Other genes transcribed by RNA Pol III such as *SNR6* (encoding U6 snRNA) also behaved as highly active integration targets with integration frequencies up to several hundred-fold higher than “cold” DNA lacking such genes. Integration events were targeted exclusively to the regions upstream of the RNA Pol III transcription start sites of these genes. Thus, a powerful integration targeting mechanism, requiring RNA Pol III transcription, is responsible for the nonrandom distribution of Ty1 elements in the yeast genome.

Results

A plasmid assay for Ty1 integration specificity

An *in vivo* assay was developed using a Ty1 donor plasmid and a second target plasmid to evaluate the relative frequencies of Ty1 integration into various target sequences (Fig. 1; see Materials and methods). The donor Gal–Ty1–*neo* plasmid (Fig. 1A) places the expression of Ty1 genomic RNA transcripts under the control of the *GAL1* promoter. Galactose induction leads to the production of element-encoded, intracellular VLPs, which are transposition intermediates (Boeke et al. 1985; Garfinkel et al. 1985; Eichinger and Boeke 1988) (Fig. 1B), and transposition is increased consequently 20- to 100-fold compared with cells grown on glucose (Boeke et al. 1985) (Fig. 1D). The Ty1 element originating from the donor plasmid is marked with the *neo* gene such that newly transposed copies of Ty1–*neo* confer G418 resistance to recipient cells. Ty1–*neo* may integrate into either the yeast genome or into the target plasmid (Fig. 1C). Any desired “test” sequence can be evaluated as a Ty1–*neo* target by cloning such sequences into the target plasmid and measuring the relative frequency of Ty1–*neo* integration *in vivo*. The donor plasmid, which also confers G418^R, is shuffled out by growth on 5-fluoroorotic acid (5-FOA) such that only cells containing new Ty1–*neo* integration events are recovered on G418 medium. Target plasmids are then rescued by transformation of *Escherichia coli*, and the number of Ty1–*neo* recombinants (Amp^R/Kan^R colonies) divided by the total number of targets recovered (Amp^R colonies) provides the integration frequency for a given target. The precise sites of Ty1–*neo* integration are determined directly by

Figure 1. Two-plasmid genetic assay for Ty1 integration targeting. (A) The Ty1-*neo* donor plasmid carries *CEN* and *URA3* sequences as well as the Gal-Ty1-*neo* transposon, but lacks a functional *bla* (Amp^R) gene. The Gal promoter directs the expression of Ty1-*neo* transcripts. The LTRs are indicated by boxes containing black triangles, the arrow indicates the start site for transcription, and the hatched box represents the *neo* gene. (B) Upon galactose induction, Ty1-*neo* transcripts are expressed (wavy line), which serve as both translation templates and replication intermediates. The transcripts are translated to produce the element-encoded proteins necessary to assemble Ty1 virus-like particles, and genomic Ty1-*neo* transcripts are replicated into cDNA. (IN) integrase; (RT) reverse transcriptase; (PR) protease; (black circles) TYA; (wavy line) Ty1-*neo* transcript. Ty1-*neo* cDNAs are then integrated into the host genome or the target plasmid by Ty1 IN. (C) The target plasmid (derived from pRS323) carries 2 μ , *HIS3* and *bla* (Amp^R) sequences, as well as the test target sequence, which is cloned into the multicloning site of pRS323. A Ty1-*neo* insertion is depicted. (D) Transposition is galactose dependent. The patches shown were carried through one round of Ty1-*neo* transposition with the pSD530 donor plasmid and a target plasmid. The donor plasmid was shuffled out by growth on 5-FOA, and the patches were replica plated to medium containing G418 (shown). The targets are as follows: upper left, pSD546 (target A); upper right, pSD547 (target B); lower left, pSD521 (target D); lower right, pSD551 (target E). G418^R yeast patches arise with growth on galactose but not glucose. The cells in the G418^R patches have received at least one new Ty1-*neo* element insertion in their genomes or the target plasmid. Plasmids are then rescued into *E. coli* and the recombinant frequency is determined.

sequencing the integration junctions of independent Ty1-*neo* insertions, using Ty1-specific primers located near the transposon termini.

The *SUF16* hot spot on yeast chromosome III

In a previous study of >100 Ty1 insertions generated in yeast chr III, a region near the *SUF16* (tRNA^{Gly}) locus was identified as a major hot spot for integration (Ji et al. 1993). In fact, 32 of 116 chromosome insertions were mapped to within 10 kb of the *SUF16* locus, making it the most preferred region of Ty1 integration on this chromosome (Ji et al. 1993). Because the resolution of pulsed-field mapping was only accurate to within ~10 kb, it is possible that some of these insertions were not closely associated with *SUF16*. Nevertheless, ≥ 14 were determined to be associated with *SUF16* by sequence analysis (Ji et al. 1993). We first isolated and characterized a clone from a yeast genomic library (see Materials and methods) that contained an 8-kb insert carrying the *SUF16* gene and its flanking regions, extending from yeast chr III nucleotides 136,155 through 144,333 (Oliver et al. 1992). A 3-kb subclone containing the *SUF16* gene and its flanking regions was the focus of our initial studies (Fig. 2A). This 3-kb region also contained a Ty1 solo LTR directly upstream of *SUF16*, and two open reading frames (ORFs), *YCR15c* and *YCR16w*, proposed to repre-

sent protein-encoding genes transcribed by RNA polymerase II (Pol II).

High-frequency integration into the *SUF16* hot spot on a plasmid target

The 3-kb hot-spot target (Fig. 2A) was transformed into the Ty1-*neo* donor host ySD10 and 44 independent transformants were carried through one round of transposition induction. Ura⁻ cells having lost the donor plasmid were selected by growth on 5-FOA and cells containing new transposon insertions were identified subsequently by replica plating to medium containing G418 (see Fig. 1D). G418^R cells arose in a galactose-dependent manner, as expected, indicating that at least one new Ty1-*neo* transposition event had occurred in these cells (see Fig. 1D). DNA was prepared from individual yeast patches and transformed into *E. coli*. Recombinant target plasmids containing a Ty1-*neo* insertion were readily recovered from these G418^R patches, and 37 of 44 (84.1%) of the yeast transformants carrying the 3-kb hot-spot target gave rise to at least one recoverable Ty1-*neo*/target recombinant (Fig. 2A). Multiple recombinants were recovered from single patches, and a total of 472 recombinants were recovered from the 44 G418^R patches. These recombinants arose from a total of 4.1×10^5 Amp^R target plasmids recovered, yielding a recombinant frequency of 1.2×10^{-3} . The recombinant fre-

Figure 2. Analysis of the *SUF16* integration hot spot using the two-plasmid assay. (A) The 3-kb *SUF16* hot-spot target is shown. The *SUF16* tRNA gene is in the center (white box with an arrow showing the transcription start site and direction of transcription). It is flanked upstream by a solo Ty1 LTR (white box with a black triangle) and the *YCR16w* ORF (hatched box with arrow showing proposed Pol II transcription orientation), and downstream by the *YCR15c* ORF (hatched box with arrow). (B) The *suf16Δ* target is shown. A 70-bp *CfoI* restriction fragment was eliminated to remove the *SUF16* gene. (C) The *suf16^{G56}* target is shown. The invariant C56 within the B-box of the Pol III promoter was changed to G (shown in the expanded view beneath). The *CfoI* sites used to generate *suf16Δ* (target B above) are also depicted in the expanded view. The results testing each target in the two-plasmid assay are tabulated below. The table lists the following data: (a) the target, corresponding to the same target lettering above; (b) the number of independent ySD10 donor transformants carried through a single round of transposition with the target; (c) the number of transformants giving rise to at least one recoverable Ty1-*neo*/target recombinant; (d) the percentage (c/b×100); (e) the number of Ty1-*neo*/target recombinants recovered from all transformants with the target (Amp^R/Kan^R colonies); (f) the total number of targets recovered from all transformants (Amp^R colonies); (g) the recombinant frequency for the target (e/f).

quency varied in independent experiments by approximately two- to threefold, and was reproducible in two different strain backgrounds (data not shown).

Since the tRNA gene and the immediately adjacent solo LTR were likely candidates for causing hot-spot activity, targets were constructed to test the possible contributions of these sequences to this activity. Indeed, upon deleting only a 70-bp *CfoI* fragment, which almost precisely contained the *SUF16* gene, the integration frequency dropped to $<6.7 \times 10^{-6}$ (Fig. 2B), corresponding to a reduction of >172 -fold relative to the intact hot spot (Fig. 2A). Moreover, zero recombinants were recovered from 57 independent patches carried through Ty1-*neo* transposition (Fig. 2B; cf. 37 of 44 for the hotspot, Fig. 2A). These results indicate that the *SUF16* tRNA gene is a critical component of the hot spot.

An intact RNA Pol III promoter is required for *SUF16* hot-spot activity

To investigate whether transcriptional activity of *SUF16* is required for hot-spot activity, a single point mutation (C → G) was constructed at position 56 within the B-box of the Pol III promoter (Fig. 2C). The C at position 56 is invariant among yeast tRNA genes (Geiduschek and Tocchini-Valentini 1988), and has been shown to be involved in binding the Pol III transcription factor TFIIC (Baker et al. 1986). Moreover, alteration of this position to G has been shown to severely reduce transcription from a variety of yeast tRNA genes (Koski et al. 1980; Newman et al. 1983; Chalker and Sandmeyer 1992; Hull

et al. 1994). The 3-kb hot-spot target altered at this single nucleotide position had an integration frequency of 1.7×10^{-5} (Fig. 2C), corresponding to a reduction of 66-fold relative to the intact hot spot (Fig. 2A). Thus, hot-spot activity near the *SUF16* locus on yeast chr III requires Pol III transcription at that locus.

To examine the *SUF16* hot spot and its flanking regions more closely, we constructed a 0.5-kb *SalI/HindIII* fragment containing a minimal hot-spot region including both the LTR and tRNA gene. Like the full 3-kb hot spot, the 0.5-kb fragment had a recombinant frequency that was very close to 1×10^{-3} (0.94×10^{-3} ; Fig. 3, target D). Thus, the hot-spot activity was largely or completely contained within this fragment. Because solo LTRs are associated frequently with integration hot spots (Ji et al. 1993), a target was then constructed that contained only the LTR from the 0.5-kb *SUF16* region. This LTR is ~80% identical to the LTRs of functional Ty1 elements (Boeke and Sandmeyer 1991). Solo LTRs are thought to be derived from full Ty elements by recombination between the two terminal LTRs followed by the elimination of intervening sequences (Winston et al. 1984). Interestingly, the LTR alone acted as a very cold target for Ty1 integration, with a frequency of $<1.6 \times 10^{-6}$, which is at least 719-fold lower than the intact 3-kb hot spot (see Fig. 2C). In fact, this target was among the coldest of the targets tested, and we consider it to represent baseline levels of integration. Thus, LTRs are not sufficient for integration hot-spot activity, and the correlation between LTRs and hot spots presumably reflects integration events previously targeted to tRNA genes.

a. Column headings all as defined for Fig. 2.

Figure 3. Ty1-*neo* targeting and Pol III transcription analysis of the minimal *SUF16* hot spot. The 0.5-kb *Sal*I-*Hind*III minimal hot-spot region (target D) containing the solo LTR and the *SUF16* gene is shown below the full 3-kb hot spot (target A). Various targets containing portions of the minimal hot spot are shown. These include the solo LTR target (E), the *SUF16* target (F), and the A-box deletion mutant (G). The results of transposition are tabulated beneath with the same heading definitions as in Fig. 2. Pol III transcription analysis in vitro is shown for each of the templates (inset). The lanes are lettered the same as the targets. The lane labeled vector included pRS323 as the template, and the lane labeled no DNA was carried out in the absence of any added template. Transcripts of 67 nucleotides (nt), corresponding to the size estimated for *SUF16* transcripts, are noted (arrow).

To test directly whether the tRNA gene alone could support hot-spot activity, a target was constructed that contained only the *SUF16* tRNA gene and 45 bp of upstream sequence, but lacked LTR sequences. Although the frequency of integration at this target was about four-fold lower than the intact 3-kb hot spot, it was at least 181-fold more active than the LTR alone (cf. Fig. 3, targets E and F). Thus, hot-spot activity is specified by the *SUF16* gene. The LTR provides only a small stimulatory effect.

Finally, a target was constructed that contained only the *SUF16* B-box and the immediate downstream sequences. Therefore, this target lacked an A-box and a transcription start site (Fig. 3, target G). Similar templates containing isolated B-boxes from tRNA genes have been shown to be sufficient for TFIIC binding, but support only barely detectable levels of Pol III transcription (Chalker and Sandmeyer 1992; Geiduschek and Tocchini-Valentini 1988). Generally, an alternative near-consensus A-box is used to generate such transcripts, and the start sites for transcription are altered (Geiduschek and Tocchini-Valentini 1988). Like the solo LTR target, the A-box mutant showed very low levels of Ty1 targeting (Fig. 3G). Thus, although this target might be expected to bind TFIIC and associated factors, it did not act as an efficient Pol III transcription template (see below), and it did not act as a hot spot.

Hot-spot activity is linked to Pol III transcription

Originally the *SUF16* locus was identified genetically as a functional suppressor tRNA (Gaber and Culbertson

1982), therefore the plasmid-borne copy we studied was likely to be expressed. To confirm that the *SUF16* gene on our targets could serve as a template for RNA Pol III transcription, targets were tested using yeast transcription extracts in vitro. The full 3-kb hot spot supported Pol III transcription and generated transcripts of the size expected for *SUF16* (Fig. 3 inset, lane A). In contrast, the 3-kb *suf16*Δ target (Fig. 2B) did not support Pol III expression, as would be expected given that the *SUF16* locus was deleted (data not shown), and the *suf16*^{G56} mutant only weakly supported Pol III transcription (with transcription levels <5% of wild type; data not shown). We conclude that *SUF16* is the only Pol III-transcribed gene within the 3-kb region. Targets D and F, which contained the unaltered *SUF16* gene and also acted as hot-spot targets, supported Pol III transcription, whereas the LTR alone (target E) and the A-box deletion mutant (target G) did not (Fig. 3). Thus, there was an absolute correlation between hot-spot activity and Pol III transcription. Of all targets derived from the *SUF16* region, only those capable of supporting Pol III transcription behaved as hotspots for Ty1 integration.

Other Pol III-transcribed genes act as hot spots for Ty1 integration

Because Ty1 targeting to *SUF16* required Pol III transcription, it was likely that other genes transcribed by RNA Pol III might also serve as active integration targets. Therefore, to test this hypothesis, a number of additional Pol III-transcribed genes, including other tRNA

genes as well as 5S rDNA and the *SNR6* gene (encoding U6 RNA), were tested as integration targets (Fig. 4). Although a range of integration frequencies was observed, all of the Pol III-transcribed genes behaved as active integration targets, and most served as highly active integration targets compared to cold DNA such as the solo LTR (Fig. 3E) and the *suf16Δ* target (see Fig. 2B). Indeed, the *SNR6* gene was as active as the *SUF16* 3-kb hot spot itself; 94% of the patches tested gave rise to recombinants, and the overall recombinant frequency was 1.5×10^{-3} (Fig. 4, target J). The *SUP2* tRNA gene (Fig. 4, target H), and an additional target from yeast chr III containing two tRNA genes (data not shown) also behaved as active targets, with recombinant frequencies $>1 \times 10^{-4}$ (Fig. 4, target H; data not shown). The frequencies for these tRNA genes were similar to the minimal *SUF16* target shown in Figure 3F. Although the 5S rDNA target had a relatively lower frequency of integration (~ 35 -fold higher than the LTR target), it also served as an active Ty1 integration target. Examination of integration specificity (below) confirmed its similarity to the other Pol III-transcribed genes.

SNR6 targets

Although its promoter includes an unusual combination

of Pol II and Pol III *cis* elements including a TATA box, an internal A-box, and a downstream B-box, *SNR6* is transcribed by RNA Pol III in vitro and in vivo (Moenne et al. 1990; Roberts et al. 1995). An intact TATA box is not required for Pol III transcription of *SNR6* (Gerlach et al. 1995), and TFIIC plays a dominant role in assembling the Pol III transcription apparatus in vivo (Burnol et al. 1993; Roberts et al. 1995). The TATA box, however, is important in directing Pol III transcription of *SNR6* in vitro in the absence of the transcription factor TFIIC, and can influence the selection of transcription start sites in vivo (Gerlach et al. 1995). To examine the potential effects of these sequences on Ty1-*neo* targeting, mutant *SNR6* targets were constructed and tested.

Sequences upstream of the *SNR6* TATA box have been shown previously to have no effect on transcription in vivo and in vitro (Eschenlauer et al. 1993). Deletion of most of the upstream LTR sequences (Fig. 4K) had only a modest (2.2-fold) effect on the Ty1-*neo* recombinant frequency. Further deletion of more proximal upstream sequences had an additional (5.1-fold) effect (Fig. 4L). This decrease is similar to that obtained with the *SUF16* hot spot upon deletion of its flanking sequences (see Fig. 3F). To examine whether the TATA box played a role in Ty1-*neo* integration targeting, a TATAAA to TGTA mutation (Gerlach et al. 1995) was constructed and

a. Column headings as defined for Fig. 2.

Figure 4. Analysis of other Pol III-transcribed targets. Three additional Pol III-transcribed targets and mutants are shown. (H) *SUP2* tRNA target; (I) 5S rDNA; (J) *SNR6* encoding U6 RNA. Note the solo LTR (white box with black triangle), a TATA box just upstream of the transcription start site for *SNR6* (black box), the *SNR6* gene (white box), and the B-box (black box downstream of *SNR6*). The remaining targets are *SNR6* mutants: (K) *Cla*I deletion mutant removing most of the LTR; (L) *Eco*RI truncation mutant; (M) TATA box mutant (TATA to TGTA); (N) B-box deletion mutant. The results of transposition are tabulated beneath with the same headings as defined for Fig. 2.

tested. Although this mutation was shown previously to abolish TFIIC-independent Pol III transcription in vitro (which depends on protein–DNA contacts between TFIIB and the TATA box), in vivo transcription was not affected, nor was the selection of the Pol III transcription start site (Gerlach et al. 1995). This mutation did not decrease Ty1–*neo* targeting (Fig. 4, target M), and the TATA box does not appear to play a role in determining hot-spot activity at *SNR6*. Next, a B-box mutation deleting the B-box and downstream sequences was tested. The B-box has been shown previously to bind TFIIC and is required for efficient Pol III transcription of *SNR6* in vivo (Brow and Guthrie 1990). Deletion of the B-box led to a 24-fold decrease in Ty1 integration frequency (278-fold lower than the intact *SNR6* target), reducing the integration frequency to baseline levels (Fig. 4N). Therefore, hot-spot activity at *SNR6* was dependent on the presence of the B-box and thus the target's ability to bind TFIIC and support Pol III transcription.

Integration occurs in regions upstream of Pol III-transcribed genes

Individual Ty1–*neo* recombinants were examined by restriction mapping and sequence analysis to determine the sites of integration within the tested targets. The integration sites of 28 independent Ty1–*neo* elements within the 3-kb *SUF16* hot spot were determined (Fig. 5A). To ensure that each Ty1–*neo* recombinant represented an independent integration event, only a single recombinant was analyzed from each patch. The integration events were found exclusively in the region immediately upstream of the *SUF16* gene and were not found elsewhere in the target. The insertion sites began at position –85 and extended through position –650 relative to the Pol III transcription start site; 22 of 28 (79%) were within 300 bp of the start site and 100% of all recovered insertions were within 700 bp (Fig. 5A; data not shown). Ty1–*neo* elements were found almost equally in both orientations, with 13 of 28 (46%) in the forward orientation (the same orientation as *SUF16* transcription) and 15 of 28 (54%) in the reverse orientation. Multiple independent integrations occurred at the same sites in some cases, implying that local sequence may play a role in the final selection of the integration site. Indeed, a loose AT-rich “anticonsensus” sequence has been identified for Ty1 integration both in vivo and in vitro (Natsoulis et al. 1989; Ji et al. 1993).

Several constructs containing the *SUF16* gene with various amounts of flanking DNA showed the same phenomenon. Ty1 insertions always occurred “region-specifically” upstream of the tRNA gene. An extreme example of this phenomenon is shown in Figure 5F, where a *SUF16* target containing only 45 bp of natural upstream sequence adjacent to vector sequences directed integration into these vector sequences. Interestingly, insertions occurred with a different pattern than that occurring with the wild-type flanking sequences (Fig. 5A), further supporting the idea that, although Pol III

Figure 5. Ty1–*neo* integration sites in Pol III-transcribed targets. (A) The precise integration sites were determined for a number of independent Ty1 insertions in four different targets: (A) The full 3-kb *SUF16* hot spot; (F) *SUF16* only (Fig. 3, target F); (I) 5S rDNA (Fig. 4, target I); and (J) the full *SNR6* target (Fig. 4, target J). Note that each arrow represents a single independent Ty1 integration event. The numbers above the arrows indicate the distance relative to the transcription start sites of the genes. Arrows pointing downward represent Ty1–*neo* insertions transcribed in the same direction as the target gene, and those facing upward represent Ty1–*neo* insertions in the opposite orientation. The white boxes represent the target genes and an arrow shows the transcription start sites.

transcription strongly targets insertion upstream of the transcription start site, final target selection at the nucleotide level is somewhat flexible and influenced by the local sequence. Interestingly, the few residual Ty1–*neo* insertions occurring in the 3-kb *suf16*^{G56} promoter mutant (see Fig. 2C) were found in the upstream region (data not shown). Thus, although the recombinant frequency was reduced drastically by this mutation, targeting still occurred for those few remaining integration events. Finally, other target genes transcribed by Pol III showed the same behavior as the tRNA genes, such as region-specific integration upstream of the transcription start sites of these genes, over a range of positions extending from –73 to –491 (Fig. 5I,J). Thus, the basic features of Ty1 targeting are conserved among Pol III-transcribed genes.

Target site duplications

Ty1 creates a 5-bp target site duplication upon integration that occurs as a consequence of staggered cuts by Ty1 integrase, followed by strand joining with the transposon and subsequent host repair (Farabaugh and Fink 1980). In all cases in which both ends of a Ty1 insertion were sequenced, a target site duplication was observed. A 5-bp target site duplication was observed in 8 of 9 cases, and an unprecedented 6-bp target site duplication was observed in the remaining case. Target site duplications flanking these transposons indicate that recombination was mediated by Ty1 integrase rather than occurring by an alternative mechanism.

Hot-spot targeting occurs in vivo but not with integration reactions in vitro using isolated virus-like particles

Ty1 VLPs are transposition intermediates that are abundant in cells undergoing active transposition (Garfinkel et al. 1985). VLPs contain RT (Garfinkel et al. 1985) and IN activities (Eichinger and Boeke 1988), as well as Ty1 genomic RNA and cDNA. The final step of Ty1 transposition can be carried out entirely in vitro using isolated VLPs and target DNA, resulting in the integration of Ty1 cDNAs into the target (Eichinger and Boeke 1988). Artificial transposons may also be added to such reactions, and these transposons will be integrated efficiently into DNA targets as well (Eichinger and Boeke 1990; Braiterman and Boeke 1994a,b; Devine and Boeke 1994). Although hot-spot targeting with the 3-kb hot-spot target occurred in vivo (Figs. 2A and 5A), integration into the same stretch of cloned DNA in vitro was random using transposon substrates and isolated VLPs (Fig. 6) (Devine and Boeke 1994). Thus, in vitro and in vivo reactions have very different features. In vitro reactions are efficient and random, whereas in vivo reactions are only highly efficient when targeted to hot spots. Integration in vivo appears to be blocked elsewhere as insertion events were never found in other regions of the target plasmids, and targets lacking hot spots such as the solo LTR (see Fig. 3E) or the *suf16Δ* target (see Fig. 2B) did not behave as active targets at all. Thus, host factors, which might include chromatin components or Pol III transcription factors, regulate the frequency and location of Ty1 integration. Biochemical reconstitution of hot-spot activity in vitro may be possible by supplementing in vitro integration reactions with specific host factors, and this approach may facilitate the identification of the host factors responsible for hot-spot activity in vivo.

Discussion

Pol III transcription in vivo greatly stimulates Ty1 integration into target DNA

A full Ty1 or delta element is present upstream of 66% of the 53 tRNA genes located on those yeast chromosomes sequenced and published to date (chromosomes I, II, III, VIII, and XI; Oliver et al. 1992; Dujon et al. 1994;

Figure 6. Comparison of Ty1 integration in vivo and in vitro. Integration events are shown for the 3-kb hot-spot region both in vivo and in vitro. The 28 integration events identified in the 3-kb hot spot (from Fig. 5A) are plotted above the x-axis, whereas 27 integration events that were generated previously in the same region in vitro using a naked DNA target are plotted below (from Devine and Boeke 1994). Note that the in vivo events are targeted to a narrow region of the target, whereas the in vitro events are scattered throughout the target.

Feldman et al. 1994; Johnston et al. 1994; Bussey et al. 1995; data not shown]. We developed a two-plasmid genetic assay in vivo to investigate the basis for this association of Ty1 with tRNA genes in the yeast genome. The plasmid assay was, in principle, similar to an assay developed previously to examine Ty3 integration (Chalker and Sandmeyer 1992); however, our assay had the additional feature of providing precise integration frequencies for each target. The targets tested could be classified as either active hot-spot targets receiving Ty1 integration events at a relatively high frequency ($>10^{-4}$), or as much less active cold targets with very low recombinant frequencies ($<5 \times 10^{-6}$), yielding few or no recombinants. Targets containing an actively transcribed tRNA gene behaved as highly active integration targets, and displayed integration frequencies up to several hundredfold higher than cold DNA targets lacking such genes. Several distinct Pol III-transcribed genes behaved in a similar manner, including tRNA genes, 5S rDNA, and the *SNR6* (U6) gene, indicating that not only tRNA genes—but all genes transcribed by Pol III—behave as active Ty1 integration targets. High frequency targeting was dependent on the target's ability to be transcribed actively by RNA Pol III. Any mutation expected to decrease Pol III transcription led to a concomitant decrease in targeting. The LTR and Pol II-transcribed genes flanking the *SUF16* hot spot did not act as efficient targets alone, further supporting the notion that the mechanism is Pol III specific. The presence of this powerful targeting mechanism provides an explanation for the association of Ty1 with a large number of tRNA genes—and other genes transcribed by RNA Pol III—in the yeast genome.

The intact *SUF16* hot spot cloned from yeast chr III and the *SNR6* gene from yeast chr II behaved as the most

active targets, with recombinant frequencies $>1 \times 10^{-3}$, suggesting that these regions of the genome would be highly active integration targets in strains undergoing Ty1 transposition. In support of this notion, both genes contain solo LTRs (remnants of previous integration events) in their upstream regions (Oliver et al. 1992; Feldman et al. 1994). Most of the remaining hot-spot targets, including the *SUP2* tRNA gene, had integration frequencies in the 10^{-3} to 10^{-4} range (~ 5 - to 10-fold lower than intact *SUF16* and *SNR6*), indicating that, although Ty1 is targeted to tRNA genes scattered throughout the genome, the precise efficiency of targeting is likely to vary among different tRNA genes. The 5S rDNA target was on the lower end of the range of hot-spot activity with a recombinant frequency of 5.6×10^{-5} . Nevertheless, this target was >35 -fold more active than the coldest targets tested (which yielded no recombinants), and the 5' regional clustering of integration sites closely resembled the other Pol III target genes. Moreover, as 5S rDNA is contained within large tandem rDNA repeats containing 100–200 5S transcription units on chr *XII* (Piper et al. 1984), it would be expected to serve as a highly active target. In support of this notion, there have been several reports of Ty sequences in 5S rDNA: (1) a solo LTR is found 144 bp upstream of at least one copy of 5S rDNA in certain strains (Piper et al. 1984) (European Molecular Biology Laboratory accession no. X00602), (2) a complete Ty1 insertion is present adjacent to a copy of 5S in another strain (Vincent and Petes 1986), and (3) Ty1 is found in rDNA-containing satellite DNA (Oyen and Gabrielsen 1983). In contrast, Ty3 insertions have not been found in rDNA, although 5S rDNA has been shown to serve as an active target for Ty3 integration (Chalker and Sandmeyer 1990). It is noteworthy that the non-LTR retrotransposons R1Bm and R2Bm (and R1 and R2 elements from other organisms) also target their integration events to rDNA (Jakubczak et al. 1991), perhaps indicating a functional similarity between these apparently unrelated transposons. At any rate, rDNA appears to be a commonly chosen target for retrotransposon integration.

Integration occurred region specifically upstream of the RNA Pol III transcription start sites of the active targets, over a range of sites extending from nucleotide positions -73 through -650 , and this feature was conserved among all Pol III-transcribed target genes. Hence, in addition to activating a target for high-efficiency integration, the targeting mechanism also dictates the region within which integration occurs. Efficient upstream targeting occurred even when the upstream sequences were replaced by vector sequences (Fig. 5F), indicating that specific upstream sequences were not required for active targeting. However, the integration patterns were somewhat different for different targets and upstream sequences. Thus, the final selection of target site is influenced by at least two factors: (1) the powerful targeting mechanism involving Pol III transcription that activates initially the regions upstream of the gene for integration, and (2) a secondary component related to the sequence or structure upstream that dictates the exact

nucleotide positions at which integration will occur. Interestingly, the region immediately upstream of the transcription initiation site was resistant to integration events, and the nearest events were ~ 70 bp upstream. This resistant region corresponds precisely to the region protected by TFIIB at Pol III promoters (Kassavetis et al. 1989), and this region may be blocked by the presence of this factor.

Frequency and specificity were tightly linked aspects of the targeting mechanism. If integration occurred, it occurred in the upstream region, suggesting that integration may be generally blocked in the absence of the targeting mechanism (perhaps by chromatin structure). In contrast, integration is *stimulated* by the presence of nucleosomes with retroviral integration reactions in vitro (Pryciak et al. 1992), suggesting that yeast retrotransposons and mammalian retroviruses might have very different requirements for integration. In this context it is significant that the *SUF16* region behaved as a hot spot on a small circular plasmid, and did not require its normal chromosomal environment for this behavior. Such plasmids assemble into chromatin in vivo (Dean et al. 1989), and therefore, the DNA segments carried on these plasmids were likely to resemble their chromosomal counterparts with respect to chromatin structure. Our studies indicate that specific alterations to target plasmids containing Pol III-transcribed genes result in drastic effects on Ty1 transposition frequency. These differences in frequency cannot be accounted for by differences in the copy number of the different target plasmids, because the recovery of the various target plasmids into *E. coli* by transformation never varied by more than threefold, whereas effects on transposition were as large as several orders of magnitude. A plasmid-based system facilitated target modification, and will be useful for future reconstitution experiments in vitro.

Ty1 vs. Ty3 targeting

The retroelement Ty3 of yeast also targets tRNA genes and other genes transcribed by RNA Pol III for integration, with insertions occurring at or within a few base pairs of the transcription start sites of these genes (Sandmeyer et al. 1988; Chalker and Sandmeyer 1990, 1992, 1993). Ty1, on the other hand, integrates region specifically over a range of hundreds of base pairs upstream of such genes (Fig. 5). Ty3 targeting has been linked to the presence of the Pol III transcription factors TFIIC and TFIIB at the target gene, suggesting that protein–protein interactions between Ty3 and these factors mediate targeting (Kirchner et al. 1995). Pol III transcription itself does not appear to be required for Ty3 targeting in vitro. In vivo, actively transcribed targets always act as good Ty3 targets, and mutations leading to poor transcription consistently debilitate targeting. Ty1 transposition in vivo is similar with regard to the requirement for target gene transcription. Rather than reflecting an explicit requirement for transcription itself, however, Ty1 may also require only a subset of Pol III transcription factors for active targeting. Another similarity is that both ele-

ments target 5S rDNA less efficiently than tRNA genes (this study; Chalker and Sandmeyer 1992). However, these elements differ dramatically with regard to the precision of integration; Ty3 integrates at the transcription start sites of Pol III-transcribed genes, whereas Ty1 shows regional specificity, integrating within a window of 100–700 bp upstream of the same class of genes (Fig. 5; Table 1). Clearly, the molecular mechanisms of these two types of integration are likely to be quite different.

One possible explanation for the similar targeting strategies used by Ty1 and Ty3 would be that both arose from a common evolutionary progenitor that possessed this powerful targeting mechanism. On the other hand, a progenitor hypothesis is not supported by sequence and structural comparisons. Ty3 belongs to the Ty3–gypsy family of retrotransposons with a genomic organization that closely resembles the vertebrate retroviruses (Hansen et al. 1988), whereas Ty1 belongs to the Ty1–copia family with a slightly different genomic organization (Boeke and Sandmeyer 1991). Although both Ty1 and Ty3 encode functional integrases responsible for carrying out integration, Ty1 integrase is 71.5 kd (Moore and Garfinkel 1994), whereas Ty3 integrase is 58–61 kd (Hansen and Sandmeyer 1990) and the two integrases share little sequence homology outside of the core domain. These differences argue strongly against the existence of a progenitor transposon, and indicate that the similar targeting mechanisms used by these two elements arose by convergent evolution guided by selective pressure to avoid integrative damage.

Ty3 has never been found to integrate at sites other than those containing Pol III-transcribed genes, whereas Ty1 has been observed to integrate at integration hot spots (Figs. 2–5) (Ji et al. 1993), as well as upstream or within a variety of Pol II-transcribed genes (for review,

see Natsoulis et al. 1989; Wilke et al. 1989; Boeke and Sandmeyer 1991). Thus, Ty3 appears to be limited to a single mode of integration whereas Ty1 has two modes. This may simply reflect a weaker or less penetrating mechanism of targeting for Ty1 that leads to effective Pol III targeting but allows low frequency nontargeted integration elsewhere. The bacterial transposon Tn7 shows a similar duality of integration. Targeted integration at the *attTn7* site occurs at a high frequency in the presence of the transposon-encoded protein TnsD, whereas nontargeted, random integration occurs at a lower frequency in the presence of only TnsE (Bainton et al. 1993).

The Ty3 mechanism appears to be accomplished by “active targeting,” whereby the integration machinery is recruited to the site of integration by Pol III transcription factors (Kirchner et al. 1995). Although Ty1 may function similarly, an alternative possibility is that the latter transposon uses a “restricted access” mechanism, whereby integration only occurs at positions or regions lacking nucleosomes. This model assumes that chromatin prohibits Ty1 integration throughout the genome and that Pol III transcription removes selectively the barrier in the upstream regions of target genes to create a “window of integration.” Integration is known to occur at high frequencies *in vitro* with only naked DNA as a target, therefore a window of naked DNA would be expected to serve as a good target (Eichinger and Boeke 1988, 1990; Devine and Boeke 1994). However, in contrast to *in vivo* integration, *in vitro* integration is very random (Devine and Boeke 1994) (Fig. 6). Also, as some Pol II promoters have been shown to have nucleosome-free regions that might likewise be expected to serve as integration windows in chromatin, restricted access alone is unlikely to account for selective integration near tRNA genes. Thus, targeting may combine features of both restricted access and active targeting models. A combined model predicts that chromatin blocks integration at most sites throughout the genome, but that a targeting mechanism actively recruits and tethers the integration complex to the Pol III transcription complex to facilitate integration upstream of the target gene. Irrespective of the specific mechanism, it is clear that host factors play a critical role in regulating the frequency and specificity of Ty1 integration.

Ty targeting

Ty5 is found primarily at the telomeric regions of yeast chromosomes, where only inactive copies were observed initially (Oliver et al. 1992; Voytas and Boeke 1992). The proximity to the telomere suggested the possibility that Ty5, like several other transposons, including Tart and HetA, may preferentially target new integration events to regions of the genome containing silenced heterochromatin. In the case of Tart and HetA, this targeting has been proposed to have led to a role in telomere formation in *Drosophila* (Levis et al. 1993; Sheen and Levis 1994). Functional studies with an active copy of Ty5 have now confirmed that this transposon indeed targets silenced

Table 1. Features of yeast Ty retrotransposon targeting

Retro-transposon	Family	Target	Comment
Ty1	Ty1–copia	Pol III-transcribed genes ^a	region-specific ^b
Ty2	Ty1–copia	tRNA genes ^c	
Ty3	Ty3–gypsy	Pol III-transcribed genes ^d	position-specific ^e
Ty4	Ty1–copia	tRNA genes ^c	
Ty5	Ty1–copia	silenced chromatin ^f	region-specific ^g

^aExperimentally determined (this paper).

^bIntegrates at a range of sites within a few hundred base pairs of a specific target site (this paper; Ji et al. 1993).

^cInferred from the genomic distribution of the element revealed by sequencing (Oliver et al. 1992; Dujon et al. 1994; Feldman et al. 1994; Johnston et al. 1994; Bussey et al. 1995).

^dExperimentally determined (Chalker and Sandmeyer 1992).

^eIntegrates at or near a specific nucleotide position (Chalker and Sandmeyer 1992).

^fExperimentally determined (Zou et al. 1996).

^gIntegrates at a range of sites within a few hundred bp of a specific target site (Zou et al. 1996).

chromatin, with de novo insertion events preferentially targeting the telomeric and silenced mating regions of yeast chr *III* (Zou et al. 1996). Interestingly, Ty1 insertions have been observed occasionally near the telomeres (e.g., there is a Ty1 solo LTR on yeast chr *I* in the W' region; Bussey et al. 1995), and Ty1 is capable of inserting into the silent mating cassettes (Mastrangelo et al. 1992). Moreover, the regions upstream of tRNA genes have been shown to suppress the expression of Pol II promoters placed directly upstream, suggesting the possibility that these regions possess features resembling silenced chromatin (Hull et al. 1994). Taken together with the fact that Ty1 and Ty5 are both members of the Ty1-copia family of retrotransposons, these elements may be related more closely than a quick glance would indicate. The targeting mechanisms may involve host factors common to Pol III-transcribed genes and silenced sites. It is tempting to speculate that, at lower frequencies, Ty1 may target silenced chromatin, and Ty5 may target tRNA genes.

One general model that would explain all Ty targeting would involve the element-encoded integrases and one or more specific host factors. In each case, the integrase would interact with a host factor, which would escort the integration complex to the integration site. Although somewhat speculative, this model is supported by the fact that targeted integration of at least three of these transposons, Ty1 (this study), Ty3 (Chalker and Sandmeyer 1992; Kirchner et al. 1995), and Ty5 (Zou et al. 1996), require specific host factors. Indeed, all retrotransposons in yeast now appear to have evolved very specific targeting mechanisms (Table 1). This likely reflects a selection for those transposons that avoid integrative damage to the host genome. It is remarkable that these transposons appear to have accomplished their goals with at least partially different host factors. The eventual mastery of specific combinations of element and host-encoded factors will allow a better understanding of genome evolution, and may facilitate the advanced manipulation of eukaryotic genomes.

Materials and methods

Cloning the *SUF16* region of yeast chr *III*

Approximately 200,000 clones of a yeast genomic library constructed in pRS200 (C. Connelly and P. Hieter, unpubl.) were screened with a 168-bp *CfoI* restriction fragment derived from pRG2-2 corresponding to the region directly downstream of *SUF16* (Gaber and Culbertson 1982). The insert of one of the isolated clones (pSD523) spanned nucleotides 136,155 through 144,333 of yeast chr *III* (Oliver et al. 1992), and this clone was chosen for further studies.

Targeting assay

The Ty1 donor plasmid, pSD530, was constructed in two steps, and consists of a plasmid backbone (pSD528) derived from pRS316 [*URA3 CEN*] (Sikorski and Hieter 1989), which was modified to eliminate a large portion of the *bla* gene. In the second step, the *neo*-marked Gal-Ty1 element from pJEF1933

was inserted into the unique *Bam*HI site. Similar Gal-Ty1 constructs placing Ty1 transcription under the control of the Gal promoter have been used to induce Ty1 transposition 20- to 100-fold by growth on galactose (Boeke et al. 1985, 1988). The *neo* marker within the Ty1 element confers G418 resistance to yeast cells carrying newly transposed Ty1-*neo* elements (Boeke et al. 1988). Thus, after transposition induction and subsequent loss of the donor plasmid, *Ura*⁻ cells displaying G418 resistance have undergone transposition.

The target plasmid for all of the studies was the 2 μ plasmid pRS323 (Christianson et al. 1992), which was engineered to contain various inserts within the multicloning site. The pRS323 backbone contains a *HIS3* marker and a functional *bla* gene. The donor plasmid lacked the *bla* gene (and could not confer Amp resistance) and the target lacked the *neo* (Kan^R) marker. Thus, Amp^R/Kan^R clones could only arise by the transposition of Ty1-*neo* into the target plasmid. Therefore, transposition events occurring into the pRS323 target could be rescued by transformation into *E. coli* and identified by growth on medium containing ampicillin and kanamycin.

A donor strain, ySD10, containing the donor plasmid pSD530 was prepared by transformation of the yeast strain YPH499 (MATa *ura3-52 trp1 Δ 63 his3 Δ 200 lys2-801 ade2-101 leu2 Δ 1*) (Sikorski and Hieter 1989) by the lithium acetate method (Kaiser et al. 1994) with subsequent growth on SC medium lacking uracil. The transformant strain displayed galactose-dependent transposition as judged by the generation of G418^R yeast with growth on galactose but not glucose (Fig. 1D). Each of the test targets was then transformed into ySD10 and carried through the transposition protocol described below.

Approximately 20–60 independent transformants were patched onto plates containing SC medium lacking uracil and histidine (to select the two plasmids) with 2% glucose as the carbon source, and incubated for 1–2 days at 30°C. The cell patches were then replica-plated to SC medium containing 2% galactose (lacking uracil and histidine) to induce Ty1 transposition, and incubated at 22°C for 4 days. Patches were next replica-plated to SC medium containing 2% glucose (lacking uracil and histidine) and incubated at 30°C overnight. The donor plasmid was shuffled out by growth on YPD medium at 30°C overnight followed by growth on SC-His medium containing 1 gram/liter 5-FOA (Boeke et al. 1984) at 30°C for 1–2 days. Finally, patches were replica-plated to SC-His medium containing 500 μ M/ml of G418, and incubated at 30°C for 1–2 days to select cells with new copies of the marked Ty1 element.

Recombinant target plasmids containing Ty1-*neo* insertions were recovered in *E. coli* by preparing total DNA from yeast patches by the glass bead-phenol method (Kaiser et al. 1994) and transforming DH10B cells (Bethesda Research Laboratories) by electroporation using a Bio-Rad electroporation apparatus. Recombinant frequencies were measured by plating the transformation mixture on LB medium containing 50 μ g/ml of ampicillin, and LB medium containing both 50 μ g/ml of ampicillin and 25 μ g/ml of kanamycin. The recombinant frequency was determined by dividing the number of Amp^R/Kan^R colonies (recombinants) by the number of Amp^R colonies (the total number of targets recovered). For each target analyzed the frequency was determined by pooling the data from independent patches and recovering >10⁵ Amp^R targets. The number of Amp^R plasmids rescued from individual yeast patches was similar for each target, indicating that targets were propagated similarly.

Target plasmids

The 3-kb *SUF16* hot-spot target (pSD546, Fig. 2A) contained sequences from the *XhoI* site at position 141,200 of yeast chr *III*

through position 144,333 (Oliver et al. 1992) (from pSD523) cloned into the *XhoI*–*NotI* sites of pRS323. The Δ *suf16* mutant construct (pSD547, Fig. 2B) was generated by first constructing a 0.5-kb *HindIII*–*SalI* fragment lacking the 70-bp *SUF16* *CfoI* fragment by three-piece ligation in pBluescript II to create the plasmid pSD522. The deletion was confirmed by sequencing, and the *HindIII*–*SalI* fragment was then inserted back into the *HindIII*–*SalI* sites of pSD546. The G56 mutant 3-kb target (pSD566, Fig. 2C) was constructed by PCR using an oligonucleotide containing the appropriate mismatch. The majority of the PCR-derived sequences were replaced with a *XhoI*–*EcoRV* fragment derived from pSD546, and the remaining region was sequenced to confirm the presence of the G at position 56 and no other alterations. The 0.5-kb minimal hot spot (pSD521, Fig. 3D) was constructed by inserting the 0.5-kb *HindIII*–*SalI* fragment from pSD523 into pRS323. The solo LTR target (pSD552, Fig. 3E) was constructed by inserting a (klenow-filled) *AccI* fragment from pSD521 into the *SmaI* site of pRS323. The minimal *SUF16* gene target (pSD548, Fig. 3F) was constructed by inserting a (klenow-filled) *AccI*–*HindIII* fragment from pSD521 into the *SmaI* site of pRS323. The A-box mutant (pSD551, Fig. 3G) was constructed by inserting a (klenow-filled) *Apal*–*HindIII* fragment from pSD521 into the *SmaI* site of pRS323. *SUP2* (pSD557, Fig. 4H) was constructed by inserting a 1-kb *BamHI*–*EcoRI* fragment from pDLC300 (Chalker and Sandmeyer 1992) into the *BamHI*–*EcoRI* sites of pRS323. The 5S rDNA (pSD558, Fig. 4I) was constructed by inserting a 0.35-kb *BamHI*–*EcoRI* fragment from pDLC301 (Chalker and Sandmeyer 1992) into pRS323. *SNR6* constructs were derived from pEP6 (Brow and Guthrie 1988). The *SNR6* gene along with its flanking regions (including an upstream Ty solo LTR) was cloned as a 1.6-kb *EcoRI*–*HpaI* fragment into the *EcoRI*–*SmaI* sites of pRS323 to create target pSD565 (Fig. 4J). A 0.6-kb *ClaI* fragment was then deleted, eliminating most of the LTR sequence to create target pSD572 (Fig. 4K). An *EcoRI* site was then introduced 15 bp upstream of the *SNR6* TATA box using a PCR adapter spanning the *EcoRI* site through the unique *NruI* site within the *SNR6* coding region. The resulting target construct, pSD575, lacked all sequences upstream of the *EcoRI* site (Fig. 4L). Sequence analysis confirmed that this segment was not mutated during PCR. The TATA box was mutated using a similar approach as for pSD575, except that a G was substituted for the A at position 2 in the TATA consensus sequence in the oligonucleotide used to make the PCR adapter, to create pSD578 (Fig. 4M). The sequence was again confirmed. Finally, an *EcoRI*–*RsaI* fragment from pSD575, which deletes the B-box and downstream sequences, was cloned into the *EcoRI*–*SmaI* sites of pRS323 to generate pSD576 (Fig. 4N).

Sequencing

The insertion sites of Ty1–*neo* recombinants were determined by sequence analysis either extending primers homologous to the subterminal regions of Ty1–*neo* outward into the adjacent sequences, or using specific primers within the *SUF16* region. A single recombinant clone was selected from each patch to ensure that each recombinant sequenced represented an independent transposition event. Ty1–*neo* sequencing primers were SD125 (JB939) 5′-CCTTAGAAGTAACCGAAGCAC-3′ and SD126 (JB940) 5′-GATCTATTACATTATGGGTG-3′. Sequencing was performed with Sequenase (for manual sequencing) or with *Taq* polymerase (for automated sequencing) using an ABI Prism kit and analyzed on an ABI373A sequencer. Sequences were aligned and analyzed with the aid of the Wisconsin GCG Gap program or with Sequencher by Genecodes.

RNA Pol III assay

RNA Pol III extracts were prepared from the yeast strain YPH499 essentially as described by Evans and Engelke (1990). Transcription reactions were carried out in a volume of 20 μ l containing 100 ng of template plasmid. Reactions were incubated for 30 min at 30°C, terminated by adding EDTA to 25 mM, phenol/chloroform extracted once, and ethanol precipitated. Dried pellets were resuspended in 50% deionized formamide/50% TE buffer containing 0.01% bromophenol blue dye, heated, and analyzed on a 6% polyacrylamide/8 M urea gel. A sequencing ladder was used to estimate the transcript sizes. Dried gels were exposed to XAR5 film at –70°C with an intensifying screen.

Acknowledgments

We thank the members of the Boeke and Hieter laboratory, Hanlee Ji and Shari Corin for helpful discussions and reagents. We also thank Jeff Corden, Shari Corin, Dan Voytas, and Cathy Wolkow for critical review of the manuscript, Mike Culbertson for pRG2-2, Christine Guthrie for pEP6, and Suzanne Sandmeyer for pDLC300 and pDLC301 as well as for helpful advice. This work was supported by postdoctoral fellowship PF4040 from the American Cancer Society (S.E.D.), grants GM36481 and CA16519 from the National Institutes of Health (J.D.B.), and a Faculty Research Award from the American Cancer Society (J.D.B.).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Bainton, R.J., K.M. Kubo, J.N. Feng, and N.L. Craig. 1993. Tn7 transposition: Target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. *Cell* 72: 931–943.
- Baker, R.E., O. Gabrielsen, and B.D. Hall. 1986. Effects of tRNA^{Tyr} point mutations on the binding of yeast RNA polymerase III transcription factor C. *J. Biol. Chem.* 262: 5275–5282.
- Boeke, J.D. and S.B. Sandmeyer. 1991. Yeast transposable elements. In *The molecular and cellular biology of the yeast Saccharomyces: Genome dynamics, protein synthesis, and energetics* (ed. J.R. Broach, E.W. Jones, and J. Pringle), pp. 193–261. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Boeke, J.D., H. Xu, and G. R. Fink. 1988. A general method for the chromosomal amplification of genes in yeast. *Science* 239: 280–282.
- 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. Genetics* 197: 345–346.
- 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.
- Braiterman, L.T. and J.D. Boeke. 1994a. In vitro integration of retrotransposon Ty1: A direct physical assay. *Mol. Cell. Biol.* 14: 5719–5730.
- . 1994b. Ty1 in vitro integration: Effects of mutations in cis and in trans. *Mol. Cell. Biol.* 14: 5731–5740.
- Brow, D.A. and C. Guthrie. 1988. Spliceosomal RNA U6 is re-

- markedly conserved from yeast to mammals. *Nature* **334**: 213–218.
- . 1990. Transcription of a yeast U6 snRNA gene requires a polymerase III promoter element in a novel position. *Genes & Dev.* **4**: 1345–1356.
- Burnol, A.F., F. Margottin, J. Huet, G. Almouzni, M.N. Prioleau, M. Mechali, and A. Sentenac. 1993. TFIIC relieves repression of U6 snRNA transcription by chromatin. *Nature* **362**: 475–477.
- Bussey, H., D.B. Kaback, W. Zhong, D.T. Vo, M.W. Clark, N. Fortin, J. Hall, B.F. Ouellette, T. Keng, A.B. Barton et al. 1995. The nucleotide sequence of chromosome I from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **92**: 3809–3813.
- Cameron, J.R., E.Y. Loh, and R.W. Davis. 1979. Evidence for transposition of dispersed repetitive DNA families in yeast. *Cell* **16**: 739–751.
- Chalker, D.L. and S.B. Sandmeyer. 1990. Transfer RNA genes are genomic targets for de novo transposition of the yeast retrotransposon Ty3. *Genetics* **126**: 837–850.
- . 1992. Ty3 integrates within the region of RNA polymerase III transcription initiation. *Genes & Dev.* **6**: 117–128.
- . 1993. Sites of RNA polymerase III transcription initiation and Ty3 integration at the U6 gene are positioned by the TATA box. *Proc. Natl. Acad. Sci.* **90**: 4927–4931.
- Christianson, T.W., R.S. Sikorski, M. Dante, J.H. Schero, and P. Hieter. 1992. Multifunctional yeast high-copy shuttle vectors. *Gene* **110**: 119–122.
- Dean, A., D.S. Pederson, and R.T. Simpson. 1989. Isolation of yeast plasmid chromatin. *Methods Enzymol.* **170**: 26–41.
- Devine, S.E. and J.D. Boeke. 1994. Efficient integration of artificial transposons into plasmid targets in vitro: A useful tool for DNA mapping, sequencing and genetic analysis. *Nucleic Acids Res.* **22**: 3765–3772.
- Dujon, B., D. Alexandraki, B. Andre, W. Ansorge, V. Baladron, J.P. Ballesta, A. Banrevi, P.A. Bolle, M. Bolotin-Fukuhara, P. Bossier, et al. 1994. Complete DNA sequence of yeast chromosome XI. *Nature* **369**: 371–378.
- Eichinger, D.J. and J.D. Boeke. 1988. The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: Cell-free transposition. *Cell* **54**: 955–966.
- . 1990. A specific terminal structure is required for Ty1 transposition. *Genes & Dev.* **4**: 324–330.
- Eschenlauer, J.B., M.W. Kaiser, V.L. Gerlach, and D.A. Brow. 1993. Architecture of a yeast U6 gene promoter. *Mol. Cell. Biol.* **13**: 3015–3026.
- Evans, C. and D.R. Engelke. 1990. Yeast extracts for transfer RNA gene transcription and processing. *Methods Enzymol.* **181**: 439–450.
- Farabaugh, P.J. and G.R. Fink. 1980. Insertion of the eukaryotic transposable element Ty1 creates a 5-base pair duplication. *Nature* **286**: 352–356.
- Feldman, H., M. Aigle, G. Aljinovic, B. Andre, M.C. Baclet, C. Barthe, A. Baur, A.M. Becam, N. Bateau, E. Boles, et al. 1994. Complete DNA sequence of yeast chromosome II. *EMBO J.* **13**: 5795–5809.
- Gaber, R.F. and M.R. Culbertson. 1982. The yeast frameshift suppressor gene SUF16-1 encodes an altered glycine tRNA containing the four-base anticodon 3'-CCCG-5'. *Gene* **19**: 163–172.
- Garfinkel, D.J., J.D. Boeke, and G.R. Fink. 1985. Ty element transposition: Reverse transcriptase and virus-like particles. *Cell* **42**: 507–517.
- Geiduschek, E.P. and G.P. Tocchini-Valentini. 1988. Transcription by RNA polymerase III. *Ann. Rev. Biochem.* **57**: 873–914.
- Gerlach, V.L., S.K. Whitehall, E.P. Geiduschek, and D.A. Brow. 1995. TFIIB placement on a yeast U6 RNA gene in vivo is directed primarily by TFIIC rather than by sequence-specific DNA contacts. *Mol. Cell. Biol.* **15**: 1455–1466.
- Hansen, L.J., and S.B. Sandmeyer. 1990. Characterization of a transpositionally active Ty3 element and identification of the Ty3 integrase protein. *J. Virol.* **64**: 2599–2607.
- Hansen, L.J., D.L. Chalker, and S.B. Sandmeyer. 1988. Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses. *Mol. Cell. Biol.* **8**: 5245–5256.
- Hull, M.W., J. Erickson, M. Johnston, and D.R. Engelke. 1994. tRNA genes as transcriptional repressor elements. *Mol. Cell. Biol.* **14**: 1266–1277.
- Jakubczak, J.L., W.D. Burke, and T.D. Eickbush. 1991. Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *Proc. Natl. Acad. Sci.* **88**: 3295–3299.
- Ji, H., D.P. Moore, M.A. Blomberg, L.T. Braiterman, D.F. Voytas, G. Natsoulis, and J.D. Boeke. 1993. Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. *Cell* **73**: 1007–1018.
- Johnston, M., S. Andrews, R. Brinkman, J. Cooper, H. Ding, J. Dover, Z. Du, A. Favello, L. Fulton, S. Gattung, et al. 1994. Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VIII. *Science* **265**: 2077–2082.
- Kaiser, C., S. Michaelis, and A. Mitchell. 1994. *Methods in yeast genetics, A Cold Spring Harbor Laboratory course manual*, 1994 ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Kalpana, G.V., S. Marmon, W. Wang, G.R. Crabtree, and S.P. Goff. 1994. Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* **266**: 2002–2006.
- Kassavetis, G.A., D.L. Riggs, R. Negri, L.H. Nguyen, and E.P. Geiduschek. 1989. Transcription factor IIIB generates extended DNA interactions in RNA polymerase III transcription complexes on tRNA genes. *Mol. Cell. Biol.* **9**: 2551–2566.
- King, W., M.D. Patel, L.I. Lobel, S.P. Goff, and M.C. Nguyen-Huu. 1985. Insertion mutagenesis of embryonal carcinoma cells by retroviruses. *Science* **228**: 554–558.
- Kirchner, J., C.M. Connolly, and S.B. Sandmeyer. 1995. Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. *Science* **267**: 1443–1444.
- Koski, R.A., S.G. Clarkson, J. Kurjan, B.D. Hall, and M. Smith. 1980. Mutations of the yeast SUP4 tRNA^{Tyr} locus: Transcription of the mutant genes in vitro. *Cell* **22**: 415–425.
- Levis, R.W., R. Ganesan, K. Houtchens, L.A. Tolar, and F.M. Sheen. 1993. Transposons in place of telomeric repeats at a *Drosophila* telomere. *Cell* **75**: 1083–1093.
- Marschalek, R., T. Brechner, E. Amon-Bohm, and T. Dinger-mann. 1989. Transfer RNA genes: Landmarks for integration of mobile genetic elements in *Dictyostelium discoideum*. *Science* **244**: 1493–1496.
- Mastrangelo, M.F., K.G. Weinstock, B.K. Shafer, A.M. Hedge, D.J. Garfinkel, and J.N. Strathern. 1992. Disruption of a silencer domain by a retrotransposon. *Genetics* **131**: 519–529.
- Moenne, A., S. Camier, G. Anderson, F. Margottin, J. Beggs, and A. Sentenac. 1990. The U6 gene of *Saccharomyces cerevisiae* is transcribed by RNA polymerase C (III) in vivo and in vitro. *EMBO J.* **9**: 271–277.
- Moore, S.P. and D.J. Garfinkel. 1994. Expression and partial purification of enzymatically active recombinant Ty1 integrase in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **91**: 1843–1847.
- Natsoulis, G., W. Thomas, M.C. Roghmann, F. Winston, and

- J.D. Boeke. 1989. Ty1 transposition in *Saccharomyces cerevisiae* is nonrandom. *Genetics* **123**: 269–279.
- Newman, A.J., R.C. Ogden, and J. Abelson. 1983. tRNA gene transcription in yeast: Effects of specified base substitutions in the intragenic promoter. *Cell* **35**: 117–125.
- Oliver, S.G., Q.J. van der Aart, M.L. Agostoni-Carbone, M. Aigle, L. Alberghina, D. Alexandraki, G. Antoine, R. Anwar, J.P.G. Ballesta, P. Benit et al. 1992. The complete DNA sequence of yeast chromosome III. *Nature* **357**: 38–46.
- Oyen, T.B. and O.S. Gabrielsen. 1983. Non-random distribution of the Ty1 elements within nuclear DNA of *Saccharomyces cerevisiae*. *FEBS Lett.* **161**: 201–206.
- Piper, P.W., A. Lockheart, and N. Patel. 1984. A minor class of 5S rRNA genes in *Saccharomyces cerevisiae* X2180-1B, one member of which lies adjacent to a Ty transposable element. *Nucleic Acids Res.* **12**: 4083–4096.
- Pryciak, P.M., A. Sil, and H.E. Varmus. 1992. Retroviral integration into minichromosomes in vitro. *EMBO J.* **11**: 291–303.
- Roberts, S., T. Colbert, and S. Hahn. 1995. TFIIC determines RNA polymerase III specificity at the TATA-containing yeast U6 promoter. *Genes & Dev.* **9**: 832–842.
- Sandmeyer, S.B., V.W. Bilanchone, D.J. Clark, P. Marcos, G.F. Carle, and G.M. Brodeur. 1988. Sigma elements are position-specific for many different tRNA genes. *Nucleic Acids Res.* **16**: 1499–1515.
- Sandmeyer, S.B., L.J. Hansen, and D.L. Chalker. 1990. Integration specificity of retrotransposons and retroviruses. *Ann. Rev. Genet.* **24**: 491–518.
- Sheen, F.M. and R.W. Levis. 1994. Transposition of the LINE-like retrotransposon TART to *Drosophila* chromosome termini. *Proc. Natl. Acad. Sci.* **91**: 12510–12514.
- Shih, C.C., J.P. Stoye, and J.M. Coffin. 1988. Highly preferred targets for retrovirus integration. *Cell* **53**: 531–537.
- Sikorski, R.S. and P. Hieter. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19–27.
- Vincent, A. and T.D. Petes. 1986. Isolation and characterization of a Ty element inserted into the ribosomal DNA of the yeast *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **14**: 2939–2949.
- Voytas, D.F. and J.D. Boeke. 1992. Yeast retrotransposon revealed. *Nature* **358**: 717.
- Wilke, C.M., S.H. Heidler, N. Brown, and S.W. Liebman. 1989. Analysis of yeast retrotransposon Ty insertions at the CAN1 locus. *Genetics* **123**: 655–665.
- Winston, F., D.T. Chaleff, B. Valent, and G.R. Fink. 1984. Mutations affecting Ty-mediated expression of the HIS4 gene of *Saccharomyces cerevisiae*. *Genetics* **107**: 179–197.
- Withers-Ward, E.S., Y. Kitamura, J.P. Barnes, and J.M. Coffin. 1994. Distribution of targets for avian retrovirus DNA integration in vivo. *Genes & Dev.* **8**: 1473–1487.
- Zou, S., D.A. Wright, and D.F. Voytas. 1995. The *Saccharomyces* Ty5 retrotransposon family is associated with origins of DNA replication at the telomeres and the silent mating locus HMR. *Proc. Natl. Acad. Sci.* **92**: 920–924.
- Zou, S., N. Ke, J.M. Kim, and D.F. Voytas. 1996. The *Saccharomyces* retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. *Genes & Dev.* (this issue).

Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III.

S E Devine and J D Boeke

Genes Dev. 1996, **10**:

Access the most recent version at doi:[10.1101/gad.10.5.620](https://doi.org/10.1101/gad.10.5.620)

References

This article cites 62 articles, 31 of which can be accessed free at:
<http://genesdev.cshlp.org/content/10/5/620.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Focused on your science.

USC
SCIENTIFIC
SIDE BY SCIENCE