An active retrotransposon in *Candida albicans*

Nicholas J. Holton, Timothy J. D. Goodwin, Margaret I. Butler and Russell T. M. Poulter*

Department of Biochemistry, University of Otago, PO Box 56, Dunedin, New Zealand

ABSTRACT

Tca2 is a Ty1/*copia*-type retrotransposon from the pathogenic yeast *Candida albicans*. It was originally identified as an abundant, linear, extrachromosomal, double-stranded DNA molecule. Here we show that Tca2 is widespread in *C.albicans*, but that the abundance of extrachromosomal Tca2 DNA varies greatly among different strains and is strongly dependent on the growth temperature. The relative levels of Tca2 RNA vary in a similar pattern to the extrachromosomal DNA, raising the possibility that the variations in extrachromosomal DNA levels are introduced predominantly at the level of transcription. We have also analysed the retrotranspositional activity of the element by developing a transposition assay involving a cloned Tca2 element tagged with a selectable marker gene that is activated by passage through an RNA intermediate. We show that the marked Tca2 is transpositionally active as transposed copies of the marked element became integrated at a variety of new positions in the genome and an intron in the donor element was precisely removed in the newly transposed copies. This is the first report of transposition in *C.albicans*.

INTRODUCTION

Retrotransposons are mobile genetic elements capable of autonomous transposition via RNA intermediates (1). They are widespread in eukaryotes and often make up a large proportion of their genomes. In maize, for example, it has been estimated that at least 50% of the genome is derived from retrotransposons (2). LTR retrotransposons carry two long terminal repeats (LTRs), typically 250–600 bp in length, flanking an internal protein-coding domain 5–7 kb long. The LTRs contain signals for promoting and terminating transcription and are also required in the reverse transcription process. The internal domain usually contains two open reading frames (ORFs) known as *gag* and *pol*. The *gag* ORF encodes proteins that make up the major structural component of a cytoplasmic particle in which reverse transcription occurs. The *pol* ORF encodes enzymes (reverse transcriptase, ribonuclease H, protease and integrase) which are involved in making a double-stranded DNA copy of the retrotransposon mRNA and inserting it into the host genome.

Several LTR retrotransposons present in yeast have been shown to be active, for example Ty1, Ty2 and Ty3 of *Saccharomyces cerevisiae*, Ty5 of *Saccharomyces paradoxus* and Tf1 and Tf2 of *Schizosaccharomyces pombe* (3–6). Using these model systems, it has been possible both to dissect the functional components of these elements, and to examine what influence the host organisms have on their transposition (7–11). For instance, several retrotransposons have been shown to be regulated by host transcription factors (12).

The chromosomal integration site of many retrotransposons appears to involve some degree of target site selection. Ty3 integrates within a few base pairs of RNA polymerase III (Pol III)-transcribed genes, via its interaction with Pol III transcription factors. The integration of Ty5 is typically near telomeres or the silent mating loci (5). The targeting of other well-studied retroelements is generally less precise. For instance, Ty1 often integrates near Pol III-transcribed genes, but it can integrate near and within RNA polymerase II (Pol II)-transcribed genes and within pre-existing LTRs (13). Tf1 integrates within long intergenic regions between Pol II-transcribed genes (6).

We previously described an unusual retrotransposon from the pathogenic yeast, *Candida albicans* (14). This element, Tca2, is 6426 bp long and contains two long ORFs, similar to the *gag* and *pol* ORFs of other retrotransposons, which are flanked by 280-bp LTRs. Tca2 was originally identified as an abundant, extrachromosomal, linear, double-stranded DNA molecule (known as pCal) present at an estimated 50 copies per cell. No other retrotransposons are known to produce such abundant extrachromosomal DNA copies.

In this paper we describe various aspects of the Tca2 element in some detail. Most importantly we show that Tca2 is a retrotranspositionally active element. This is the first report of transposition in *C.albicans*.

MATERIALS AND METHODS

Strains and culture conditions

The *C.albicans* strains analysed were hOG759 (*ade1 met2*), hOG1042 (*ade2*) and hOG1051 (an *ade2 ura3* derivative of hOG1042) (all derived from iB65) (14), ATCC10261 (15), F16932 [a prototrophic clinical isolate from Poulter,R. (Bath, UK), unpublished], SA40 (16), SC5314 (17), SGY269 (*ura3*) (18), hOG963 (an *ade2 ura3* derivative of SGY269, Poulter,R., unpublished) and CAI-8 (*ade2 ura3*) (19). The strains were grown at either 27 or 37°C in YPD medium (1% yeast extract, 2% peptone, 2% glucose) or YNB (0.7% yeast nitrogen base,

*To whom correspondence should be addressed. Tel: +64 3 4797856; Fax: +64 3 4797866; Email: russell@sanger.otago.ac.nz

0.5% glucose) supplemented with adenine and uridine, where appropriate.

Enzymes

Restriction enzymes, Expand high fidelity PCR system, RNase A, Klenow, DNase I and Proteinase K were purchased from Roche, GmbH, Germany. T4 DNA ligase was supplied by Promega, Madison, WI. Zymolyase 100T was purchased from Seikagaku Corporation, Tokyo, Japan. Strataclean was supplied by Stratagene, La Jolla, CA.

Candida albicans nucleic acid isolations

For DNA isolations, cells were grown at 27 or 37°C to an A_{600} of ~1.0. DNA for Southern blots and PCR was prepared as described by Philippsen *et al.* (20). Small scale DNA preparations for inverse PCR were performed using High Pure PCR Template Preparation Kits purchased from Roche, GmbH, Germany.

RNA extractions were performed from cells grown in YPD medium at either 27 or 37°C to an A_{600} of ~1.0, using the method of Schmitt *et al.* (21).

Southern and northern blotting

DNA was electrophoresed through ~1% agarose and transferred to Hybond-N+ nylon membranes (Amersham), by alkaline transfer, essentially as described by Sambrook *et al.* (22). Probes were radiolabelled with $[\alpha^{32}P]dCTP$ by random-primed labelling using Hexanucleotide Mix from Roche. Hybridizations were performed at 65°C. The hybridization buffer was similar to that of Church and Gilbert (23) (0.36 M Na_2HPO_4, 0.12 M NaH_2PO_4, 1 mM EDTA, 7% SDS). Membranes were prehybridised in this buffer for 2 h, the denatured probe was then added in 5 ml of fresh buffer and hybridisation was allowed to proceed for 16–20 h. Post-hybridisation washes consisted of two rinses in 2× SSC at room temperature followed by stringency washes in 0.2× SSC, 0.1% SDS at the hybridisation temperature. Finally membranes were rinsed in 2× SSC then exposed to Kodak X-Omat AR film at –80°C using an intensifying screen. Membranes were stripped for reprobing by rinsing in H_2O for 1 min, followed by two washes in 0.2 M NaOH, 0.1% SDS at 37°C and then a final rinse in 2× SSC.

Northern blotting was performed essentially as described by Brown and Mackey (24). After hybridization the membranes were washed twice, for 5 min per wash, in 2× SSC at room temperature, washed twice, for 5 min per wash, in 0.2× SSC, 0.1% SDS at room temperature, and washed twice, for 15 min per wash, in 0.2× SSC, 0.1% SDS at 65°C. Finally, the membranes were rinsed in 2× SSC and exposed to X-ray film at –80°C.

The films from the Southern and northern blots were scanned using a Bio-Rad GS-670 imaging densitometer. The brightness/contrast of the whole scan was adjusted for printing using Adobe Photoshop 3.0.

Recombinant DNA manipulations

Bacterial plasmids were prepared using an alkaline lysis/polyethylene glycol precipitation method from Applied BioSystems. Sequencing was performed using a combination of subcloning and specifically designed oligonucleotide primers. Sequences were determined on an ABI 377 DNA Sequencer and edited using SeqEd 1.0.3. PCRs were performed on an Autogene II programmable cycling water bath from Grant Instruments, Cambridge. Primers were synthesised on an ABI 380B oligonucleotide synthesiser or purchased from Genset Singapore Biotech., Singapore.

Transformation of *C.albicans*

Candida albicans transformations were performed using a modification of a PEG-mediated transformation protocol (25). Modifications were such that β-glucuronidase was replaced by 1 mg/ml Zymolyase 100T, β-mercaptoethanol replaced DTT, and the cells were plated directly onto osmotically supported medium (YNB + 3×10^{-3} M uridine + 1 M Sorbitol) without an agar overlay.

Construction of a tagged Tca2 element

A cloned copy of the Tca2 element was tagged with a *URA3* gene under the control of its own promoter. This *URA3* gene (26) was disrupted by the insertion of a small (84 bp) intron from the *C.albicans* peptide transporter gene (27) close to the start of the ORF. The intron, with an enhanced polypyrimidine tract (27), was synthesised from two complementary oligonucleotides. The intron was inserted, in both orientations, into the unique *Hin*dII site of a 967-bp *URA3* gene with its own promoter. The *URA3*/intron construct was inserted as a *Pst*I fragment into Tca2 at a synthetic *Nsi*I site, generated by site-specific mutagenesis (28), near the end of the *pol* ORF [at position 6084 of the GenBank sequence AF007776 (14)]. The *URA3* gene was inserted in both orientations with respect to Tca2. These constructs were used to test the ability of the intron to be spliced from the *URA3* gene. The construct containing the *URA3* gene in the reverse orientation to Tca2 and the intron in the same orientation as Tca2 was used for subsequent experiments. The *C.albicans ADE2* gene (29) was introduced into this vector outside the 3′ LTR of Tca2, to produce the plasmid pRUIA. This plasmid was linearised at the unique *Xba*I site in the *ADE2* gene then stably integrated into the *ADE2* loci of strains hOG963 and hOG1051. Stable *ade2+*, *ura3–* transformants (named H963R and H1051R, respectively), which contain integrated copies of the marked element, were obtained.

In a separate experiment, the *URA3*-tagged Tca2 was also introduced into CAI-8 (*ade2 ura3*), by first adding a *C.albicans* origin of replication (25) into pRUIA. Transformants were selected on medium without adenine, and grew as unstable, irregular, white and red colonies. This suggests that the *ADE2* gene was carried by a extrachromosomal plasmid. Subsequently, stable, white, *ade2+*, *ura3–* colonies were isolated from the original unstable transformants. The stability suggests that the plasmid had integrated into the host genome. One of these stable *ade2+*, *ura3–* isolates, W3, was used in further analyses.

Tca2 transposition assay

Candida albicans H963R, H1051R and W3 were grown overnight in YPD at 27 or 37°C. Approximately 10^7 cells of these cultures were then plated onto YNB media containing 4% casamino acids and grown at 27 or 37°C, for 4–7 days. Potential retrotransposition events were indicated by the appearance of *ura3+* colonies.

Sequencing of Tca2 insertion sites

The sequences of the insertion sites were determined by inverse PCR. Genomic DNA was first digested with *Nla*III or *Taq*I. The enzymes were then inactivated by heating to 65°C for 20 min or using Strataclean, respectively. Ligations were performed with DNA at a concentration of ~20 ng/µl and T4 DNA ligase, according to the manufacturer's instructions. Inverse PCR primers were designed to the 3′ *URA3*/Tca2 junction (ATGAATTATCATTGAAGAACTGCATTAAC) and within the *URA3* gene adjacent to the site (underlined) at which the intron was inserted, so as to bind effectively only in the absence of the intron (ATTCTAGACAAAGATATGACAGT<u>CA</u>AC). PCR reactions were then performed using 200 ng of ligated DNA and *Taq* polymerase, in a total volume of 100 µl. The sequence of the PCR products was obtained by direct sequencing using the sequencing primer (TCGTAATGCTCAATCAGGGGAGTG).

Sequence analyses

The locations of the tagged Tca2 elements in the genome were determined by comparing the sequences generated by inverse PCR with the *C.albicans* genomic sequences, generated by a *C.albicans* genome sequencing project at the Stanford Genome Technology Center (http://www-sequence.stanford.edu/group/candida). The database used was assembly 6. The annotations and ORF assignments were as determined in the genomic database. To produce a description of the *C.albicans* intergenic regions we analysed sequences from 24 contigs from assembly 4 (contig4-3084 to contig4-3108), which contained 593 intergenic regions. Intergenic regions were determined by translating each contig, using only nonoverlapping ORFs, and calculating the distance between each ORF. A minimum ORF size of 300 bp was used to decrease the likelihood of including noncoding ORFs in the analysis. Motif searches were performed using MEME (30). Searches for tRNAs were performed using tRNAscan-SE (31).

RESULTS

Levels of extrachromosomal Tca2 DNA

The retrotransposon Tca2 was originally identified as an abundant, extrachromosomal DNA molecule, in the *C.albicans* strain hOG1042 (14). Such extrachromosomal DNA had not been reported in any other *C.albicans* strain, suggesting that this abundance of Tca2 extrachromosomal DNA was peculiar to this strain. To investigate this phenomenon further we analysed a variety of strains for the presence of extrachromosomal Tca2 DNA by Southern blotting. The strains examined included hOG1042 and its close relative hOG759, two recent clinical isolates (SA40 and F16932), and three common laboratory strains (SGY269, SC5314 and ATCC10261). Figure 1A shows a Southern blot containing approximately equal amounts of uncut genomic DNA isolated from 27 and 37°C cultures of four of these strains, hybridised to a probe derived from the internal region of Tca2. Bands hybridising to the Tca2 probe were detected in all of the *C.albicans* strains examined, showing that this element is widespread in the species. The full-length extrachromosomal Tca2 DNA is evident as a band running at ~6.5 kb with the large molecular weight DNA bands (<20 kb) representing hybridisation to chromosomal copies of
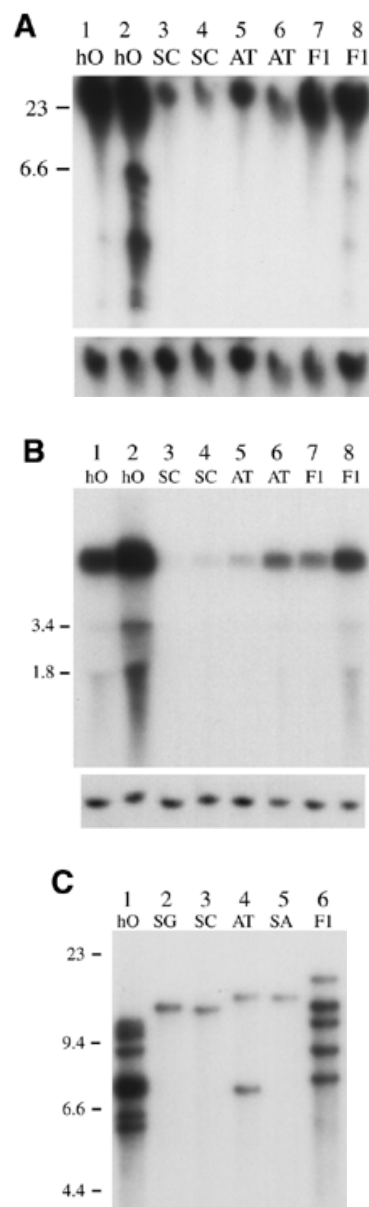
Figure 1. (**A**) Comparative levels of Tca2 extrachromosomal DNA. Southern blot analysis was performed on DNA from four *C.albicans* strains grown at 27 or 37°C (odd and even numbered lanes, respectively). Hybridisation was performed using a probe synthesised from the 3′ *Asp*718/*Eco*RI fragment of Tca2. The same blot stripped and reprobed with the *C.albicans* actin gene is shown below. (**B**) Comparative levels of Tca2 RNA. A northern blot containing RNA from 27 and 37°C cultures (odd numbered and even numbered lanes, respectively) of four different *C.albicans* strains, hybridised with the 3′ *Asp*718/*Eco*RI fragment of Tca2. The same blot stripped and reprobed with the *C.albicans* actin gene is shown below. (**C**) Copy number of Tca2. *Eco*RI-digested genomic DNA from a variety of strains was probed with the 3′ *Asp*718/*Eco*RI fragment of Tca2. In all panels the key to strains is as follows: hO, hOG1042; SG, SGY269; SC, SC5314; AT, ATCC10261; SA, SA40; F1, F16932.

Tca2. In addition, hybridisation to lower molecular weight DNA (3–4 kb) was observed. These bands probably represent incomplete reverse transcripts, similar to those observed for Ty1 (32). The highest level of Tca2 extrachromosomal DNA was observed in hOG1042. The other strains contain considerably

less extrachromosomal Tca2 DNA. No extrachromosomal Tca2 DNA was detected in ATCC10261. All strains which contain detectable levels of extrachromosomal Tca2 also show a temperature-dependent variation in levels, with substantially more extrachromosomal DNA being detected at 37 than at 27°C.

Strain- and temperature-dependent variation of Tca2 transcript levels

The transcriptional activity of Tca2 was analysed by hybridising northern blots, containing RNA isolated from 27 and 37°C cultures of a variety of *C.albicans* strains, to a probe derived from the internal region of Tca2 (Fig. 1B). In all the strains examined one or more bands hybridising to the Tca2 probe were detected. The most prominent band in each strain corresponds to RNA molecules of sufficient size to be full-length transcripts. Interestingly, in each strain, these full-length transcripts are much more abundant at 37 than at 27°C, similar to the pattern observed for the extrachromosomal DNA levels. In addition, there also appears to be a good correlation between the amounts of Tca2 RNA in different strains and the amounts of extrachromosomal DNA. Given the likelihood that the Tca2 RNA is an intermediate in the synthesis of the extrachromosomal DNA, this raises the possibility that most of the variations in the levels of extrachromosomal Tca2 DNA are a consequence of similar variations in the RNA levels. It is also possible, however, that some degree of transcription of the extrachromosomal DNA molecules also contributes to the correlation between extrachromosomal DNA and RNA levels.

### Variation in chromosomal Tca2 copy numbers

Interstrain variation of both Tca2 extrachromosomal DNA and RNA levels may simply reflect the number of copies of Tca2 in these strains; for instance, strains carrying more copies of Tca2 might produce more transcripts, as a result of the cumulative contribution of each element. To investigate this possibility, Southern blots containing *Eco*RI-digested DNA from a variety of *C.albicans* strains were hybridised to probes derived from the internal region of Tca2 to determine the number of integrated Tca2 copies per strain (an example of these Southern blots is shown in Fig. 1C). The DNA for these analyses was isolated from cells grown at 27°C to minimise the interference from the extrachromosomal reverse transcripts, since only low levels of extrachromosomal DNA are present at this temperature. In four of the *C.albicans* strains the element is present at just one or two copies per cell. In the other two strains, however, it exists at a higher copy number: at least five in F16932, and 10 or more in hOG1042. The banding patterns obtained with the Tca2 probe can be seen to vary considerably from strain to strain. This finding suggests that Tca2 has been active in these strains since they diverged. The number of copies of Tca2 in each strain correlates well with the chromosomal band intensity seen in the Southern blot containing uncut genomic DNA (Fig. 1A). Interestingly, the two strains with the highest Tca2 copy numbers (F16932 and hOG1042) also have the highest levels of Tca2 transcripts. Similarly, SC5314, with apparently just a single copy of Tca2, has a lower level of Tca2 RNA than ATCC10261 (two copies). These findings show that there is some correlation between Tca2 RNA levels and Tca2 copy numbers, although additional factors also appear to influence the levels of RNA (see below).
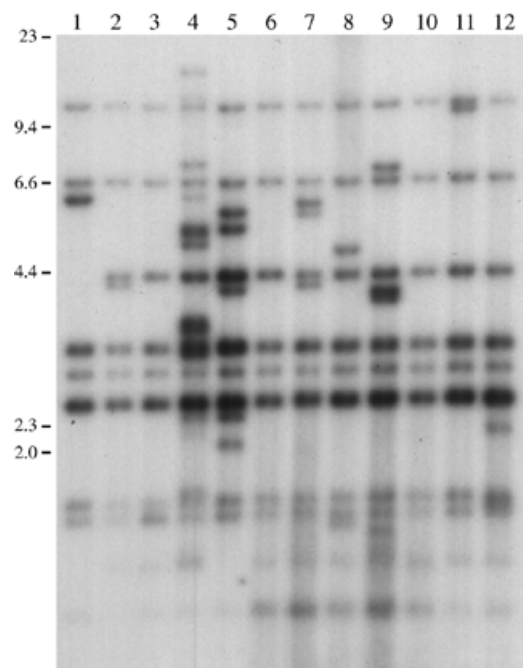
Figure 2. Tca2 is transpositionally active. A blot of *Eco*RI-digested DNA from hOG1042 and related strains, hybridised to the Tca2 3′ *Asp*718/*Eco*RI probe. Lane 1, hOG759; lane 2, hOG762; lane 3, hOG1042; lanes 4 and 5, two strains derived from hOG1042 after one month's selection for slow growth on plates; lanes 6–12, seven strains derived from hOG1042 after one month's growth in liquid culture.

Tca2 is capable of transposition

The abundance of full-length, double-stranded, extrachromosomal copies of Tca2 in some strains suggests that some Tca2 elements are capable of performing all the steps of the replication cycle, up to and including the generation of full-length double-stranded DNA molecules. It is not clear, however, whether the abundant extrachromosomal forms contribute to high levels of Tca2 transposition in these strains, or if they are a result of some blockage of the integration or degradation processes. To investigate these possibilities three *C.albicans* strains, that vary in the number of integrated copies of Tca2 they contain, and in the amount of extrachromosomal DNA they produce, were grown continuously for ~35 days at 37°C. Serial subcultures were made every second day. After the growth period, genomic DNA was prepared from 10–12 independent isolates derived from each of the three strains, and from each of the parent strains. Several of the new isolates were deliberately selected as slow-growing derivatives. The number of copies of Tca2 in each strain was then estimated by Southern blotting using a probe derived from the internal region of Tca2 (Fig. 2). No new bands were detected in strains derived from ATCC10261 or SA40, although SA40 is known to produce full-length double-stranded DNA copies of Tca2 (data not shown). In contrast, most of the derivatives of hOG1042 had additional bands, usually 2 or 3, although 9 extra bands were recorded in one derivative. None of the hOG1042 derivatives lacked a band present in hOG1042. The new bands are, therefore, not due to rearrangements of the existing Tca2 insertions, but represent additional copies of Tca2. These
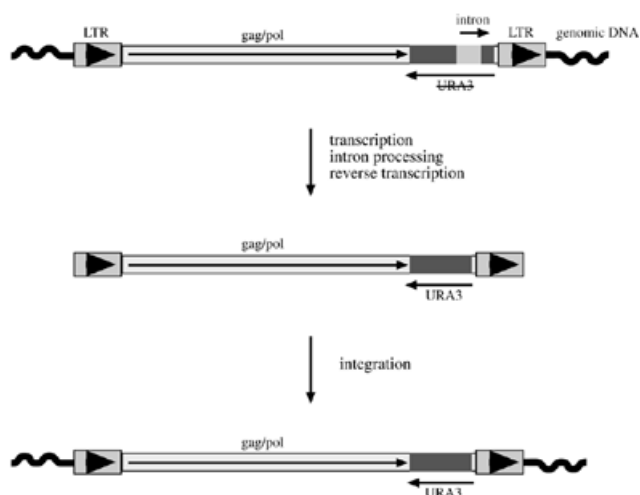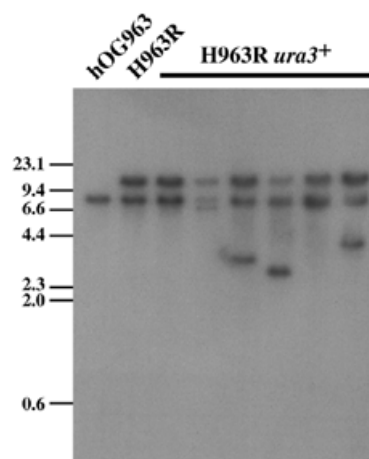
Figure 3. The retrotransposition assay. The *URA3* transgene is non-functional until after the antisense intron is spliced out of the RNA and the spliced RNA is copied into DNA.

Figure 4. Southern blot analysis of Tca2 transposition clones from H963R. Genomic DNA was digested with *Eco*RV and probed with the coding region of the *URA3* gene. The band at ~8.5 kb represents the native disrupted *URA3* gene. The band at ~15 kb represents the integrated *URA3*-tagged Tca2. Additional bands, not present in H963R, are from the transposition of the tagged Tca2. No additional bands are present in lanes 3 and 7 suggesting the occurrence of recombination events.

findings suggest that Tca2 is capable of transposition, and that transposition occurs at a sufficiently high frequency that it can be detected without selection. It is also clear that Tca2 is much more active in hOG1042 than in the other two strains.

Tca2 retrotransposition assay

To study several aspects of the Tca2 system, including its retrotranspositional activity and possible target site preference, as well as strain-dependent influences on Tca2 replication, we developed a retrotransposition assay using a Tca2 element tagged with a *URA3* gene under the control of its own promoter. The *URA3* gene was disrupted by the insertion of a small intron which was inserted in either forward or reverse orientations (with respect to the *URA3* gene) to allow analysis of the intron processing. The intron-containing *URA3* gene was inserted into the extreme 3′ end of the Tca2 *pol* ORF, in both orientations with respect to Tca2 transcription. Transformation of these constructs into *ura3*⁻ strains of *C.albicans* initially showed that the *URA3* gene was functional in either orientation with respect to Tca2, but only when the intron was in the correct orientation with respect to the *URA3* gene itself. Thus, the *URA3* gene was expressed when embedded within Tca2 sequences, and the intron was processed normally.

The retrotransposition assay was conducted using the construct in which the *URA3* gene was inserted in the opposite orientation with respect to Tca2 and was disrupted by the insertion of an antisense intron. Since the intron is, therefore, in the same orientation as Tca2 transcription it can potentially be spliced out of full-length retrotransposon transcripts of the marked element, but not from transcripts generated off the *URA3* promoter. Therefore, in principle, the *URA3* gene can only lose the intron when the element goes through a cycle of retrotransposition (Fig. 3). The vector carrying the marked Tca2 element was introduced into the genomes of hOG1051 and hOG963 on the basis of it complementing the *ade2* auxotrophies of these strains, as described in Materials and Methods. The initial transformants were selected on the basis

of the *ADE2* gene carried on the vector. The linearised element was introduced into hOG1051 (a *ura3 ade2* derivative of hOG1042) at the defective *ade2* locus, producing the strain H1051R, and also into the *ade2* locus of strain hOG963 (a *ura3 ade2* derivative of SGY269) producing strain H963R. Southern analysis was performed on these strains to ensure the presence of the tagged Tca2 and the location of the integration. H1051R has two copies of the *URA3*-tagged Tca2 whereas H963R contains a single copy (not shown). Potential retrotransposition events were subsequently identified by the generation of *ura3*⁺ colonies from the *ade2*⁺ *ura3*⁻ strains. This assay is similar in principle to the assays previously developed to study retrotransposition of the Ty elements in *S.cerevisiae* (1,33,34).

A high frequency of *ura3*⁺ colonies (10^{-5}–10^{-6} *ura3*⁺ colonies per cell plated) was obtained by growing the strains containing the marked Tca2 elements in rich medium at 37°C overnight and plating appropriate volumes of the resulting cultures onto medium lacking uridine. Analyses of these *ura3*⁺ colonies showed that their *ura3*⁺ phenotypes were stable at both 27 and 37°C. Southern analysis showed that many of these colonies contained additional integrated copies of the *URA3*-tagged Tca2 element not found in the parent strains (Fig. 4). Inverse PCR products containing the intron insertion site, the 3′ LTRs and the genomic DNA flanking the new tagged Tca2 elements were then obtained. Sequencing of these PCR products revealed that the intron had been precisely removed from each element, indicating that the elements had transposed through RNA intermediates. In addition, each element was flanked by *C.albicans* genomic DNA, confirming that the elements were integrated into the host genome. This was true not only for strains which showed extra copies of the tagged Tca2 on Southern blots, but also for those in which no new copy could be detected (see below). The tagged Tca2 elements are therefore capable of retrotransposition.

Our analyses, however, indicated that not all the *ura3*+ colonies were due to retrotransposition. Retrotransposition involves both reverse transcriptase activity giving the RNA-derived double-stranded DNA, and the action of an integrase, inserting the DNA into a new location in the host genome. A number of *ura3*+ colonies were found, on analysis, to have no additional copies of the Tca2 retroelement. These colonies were most likely produced by one of two possible pathways. The first possibility is gene conversion between the genomic mutant *ura3* allele and the *ura3* (intron-carrying) transgene, generating a functional *URA3* gene from the two non-functional ones. This would not involve reverse transcription. Secondly, they could be generated by gene conversion or homologous recombination between the *URA3*-tagged DNA product of reverse transcription and one of the mutant genomic *ura3* genes or the intron-carrying donor Tca2 element. It is not possible with our data to determine which of these pathways has been used in these particular *ura3*+ strains. All of the 16 independent *ura3*+ colonies analysed from H1051R were found to have no additional copies of Tca2. Southern analysis and junction sequencing of the *ura3*+ colonies derived from another strain, H963R, revealed that approximately half represented the integration of Tca2 at a variety of genomic positions, as shown in Figure 4. All of the *ura3*+ colonies derived from W3 (an *ade2*+, *ura3* derivative of CAI-8) appeared to be the result of the integration of new copies of Tca2.

H1051R carries (homozygously) a *ura3* point mutation. It gave rise to numerous *ura3*+ derivatives, none of which carried an extra copy of Tca2. It seems probable that at least some of these derivatives are due to gene conversion or homologous recombination between one of the mutant genomic *ura3* alleles and the introduced *ura3* transgene. It is therefore uncertain whether any of these *ura3*+ derivatives arose from retrotransposition, although retrotransposition was observed with native Tca2 elements in the parental strain, hOG1042. H963R carries (homozygously) an *ADE2* gene insertion into *ura3* (18). The proximity of the disruption to the intron inserted in the Tca2 *URA3* gene may reduce the incidence of gene conversion or homologous recombination. About 50% of the *ura3*+ derivatives of H963R carried extra tagged copies of Tca2. Strain W3 carries (homozygously) a deletion of its *ura3* genes and their surrounding genomic contexts. This eliminates any possibility of gene conversion or homologous recombination with the native *ura3* gene in this strain. Reflecting this expectation, all *ura3*+ derivatives of W3 carry additional tagged Tca2 elements.

Retrotransposition of an element leads to the integration of an additional copy at a new site. Usually this involves the element-encoded integrase, but integration might be achieved by homologous recombination between the ends of the double-stranded DNA and some homologous genomic sequence. In the present analysis, such homologous targets would be genomic Tca2 sequences and 'solo LTRs'. Few of the H963R events analysed by inverse PCR were in pre-existing Tca2 sequences. Of the 14 W3 events analysed, one was the result of integration via homologous recombination into the left LTR of a pre-existing complete Tca2 element. This suggests that homologous recombination is found significantly less often in this strain than integrase-mediated integration.

Tca2 is preferentially targeted to promoter regions

The strains that resulted from the integration of the tagged Tca2 at positions not ascribed to recombination events were analysed further to determine the integration site preference of Tca2. Genomic junctions of a large number of tagged Tca2 insertions were sequenced in strains derived from H963R and W3.

As can be see in Table 1 and Figure 5, the majority of insertions occur in the intergenic regions between ORFs. Within these regions there seems to be a preference for the promoter regions, with 92% of insertions closer to start codons than stop codons of ORFs. This pattern of targeting is similar to that seen for Tf1 from *S.pombe* (6), although several differences should be noted. Tca2 insertions were observed within ORFs, whereas Tf1 was not observed to integrate into ORFs. Tf1 appears to have a preference for longer intergenic regions whereas no such preference was observed for Tca2. The average length of intergenic regions was estimated for each of the promoter/terminator combinations from a selection of *C.albicans* genomic sequences: divergent promoters, 1088 bp; terminator-terminator, 521 bp; and promoter-terminator, 770 bp. From the data presented in Table 1, there does not appear to be any strong bias towards integration in long intergenic regions. In addition, no integration events were observed within terminator-terminator intergenic regions, further supporting the notion that Tca2 is targeted to regions near promoters.

Other retrotransposons, such as Ty1 and Ty3, are known to integrate near tRNAs or pre-existing LTRs (8,9,35). Analysis of sequences adjacent to the integration sites of Tca2 revealed no such correlation, with only one insertion observed within 500 bp of a tRNA (H963RU53). No insertions were observed within 1 kb of a pre-existing Tca2 LTR (excepting the events determined to be the result of homologous recombination with Tca2 LTRs), indicating no preference for integrating near such elements. Interestingly, two of the insertions (H963RU52 and H963RUAA5) occur within 395 bp of each other, suggesting that this is a hotspot for Tca2 integration. Multiple insertions were not observed at any other site, although further hotspots may be revealed by the analysis of additional integration events.

We compared the sequences adjacent to the integration sites of Tca2 in an attempt to identify sequence elements potentially involved in determining the target site preference of Tca2. Alignment of the integration site sequences did not reveal any strong consensus that could be involved in determining the target site. Analysis of the 5-bp duplication caused by retrotransposon integration, however, revealed some sequence bias: a preference for A and T residues in the three central positions of the target site duplications (Fig. 5B). This is a similar pattern to that observed previously for *S.cerevisiae* and *C.albicans* LTR retrotransposons (36,37). Motif searches 500 bp either side of the integration sites, using MEME (30), revealed the presence of a polypurine motif of ~15 bases (with a consensus of RAARRARAAAAAARW) near all of the integration sites. This motif was found to be highly abundant in *C.albicans*, predominantly within intergenic regions. However, the motif, while abundant in the genome, occurred at varying distances from the integration site and may be of little relevance to the targeting of Tca2.

Table 1. Location of Tca2 insertion sites

Clone	Name of 5′ ORF[a]	Tca2 to ORF[b] (bp)	Name of 3′ ORF	Tca2 to ORF (bp)	Size of intergenic region (bp)	Orientation gene Tca2 gene[c]
H963RU3	MOT2	−111	YMR123	380	269	← → →
H963RU6	YDL45	478	RNA11	153	631	← → ←
H963RU8	hyp	250	YER166	310	560	← → ←
H963RU10	YFR44	1	if4e-like	786	787	→ → →
H963RU18	hyp	251	DRS1	1160	1411	← ← →
H963RU30	YHR47	202	NOP4	112	314	→ ← →
H963RU35	REPEAT SEQUENCE					
H963RU43	CET1	1130	ADH5	646	1776	→ → →
H963RU46	hyp	202	DAO2	492	694	← ← ←
H963RU50	Not seq		YBR287	147		? → →
H963RU52	hyp	27	MSS51	228	255	← ← ←
H963RU53	YOR161	158	Tca8 partial or tRNA^Thr			← ←
H963RU59	ECM3	414	YDL72	−14	400	→ ← →
H963RU62	End of contig		hyp	355		? ← ←
H963RU63	YMR130	789	DFG5	94	883	→ ← →
H963RU65	SMD3	950	HAP1	939	1889	← → →
H963RU75	SAC6	−74	VPS18	567	493	← → →
H963RU80	RAD17	76	hyp	261	337	← ← ←
H963RUAA1	FET3	21	FET99	1585	1606	← ← ←
H963RUAA5	hyp	−368	MSS51	623	255	← ← ←
H963RUAA8	YR388	−316	hyp	1351	1035	← ← ←
H963RUAB1	YBR175	466	YOL129	237	703	← ← →
H963RUAB3	hyp	69	hyp	143	212	← ← →
H963RUAB6	SPE2	223	TFB3	74	297	→ ← →
W3U1	REPEAT SEQUENCE					
W3U2	THR4	331	PTH1	99	430	← ← →
W3U3	REPEAT SEQUENCE					
W3U4	PET123	228	hyp	119	347	→ ← →
W3U5	hyp	2182	hyp	107	2289	← ← →

[a]Annotations assigned at the *C.albicans* sequence database are shown in capitals. ORFs with partial homology, but not annotated, are shown with lower case letters. When no significant homology could be found ORFs are labelled as hypothetical (hyp).
[b]Distance, in base pairs, taken from the Tca2 insertion site to the nearest base of the 3′ or 5′ coding region.
[c]Outer arrows indicate the direction of the ORFs either side of the Tca2 insertion. The middle arrow indicates the orientation of the inserted Tca2.

## Strain-dependent factors influence Tca2 transcription

Northern analysis of H963R and H1051R was performed to determine the influence the host strain has on the expression of Tca2. Hybridisations performed with a *URA3* probe allow identification of the tagged Tca2 transcripts. Substantially more Tca2 transcript was observed in H1051R compared to that observed in H963R (Fig. 6). In addition, the temperature dependence that was observed for wild-type Tca2 was also seen for the tagged copies. The tagged Tca2 was introduced into the same locus (*ade2*) of both strains, and there is no sequence variation between the introduced elements, thus the relative abundance of RNA produced by a particular element is strongly influenced by the host strain.

Expression of Tca2 is influenced by genomic location

During analysis of the *ura3*[+] clones derived from retrotransposition events we observed that some strains grew more slowly on uridine-deficient medium than others (not shown). When the medium was supplemented with uridine all the strains analysed grew at a similar rate. This was thought to be due to the influence of the insertion sites on the levels of expression of Tca2 and the associated *URA3* gene. To investigate the effect of the Tca2 insertion sites on the expression of the tagged elements, we compared the levels of Tca2 RNA produced by strains carrying single marked elements at different loci. To do this it was first necessary to remove the original Tca2 tagged with the intron-containing *URA3*. This
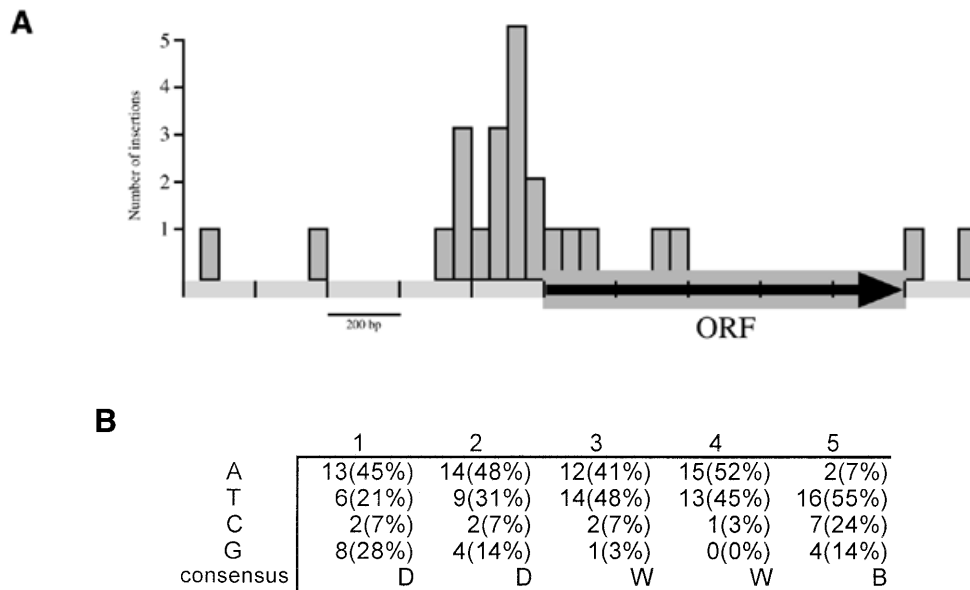
Figure 5. Analysis of Tca2 insertions. (**A**) Schematic diagram of the insertion sites of Tca2. The locations of the insertion sites in relation to the closest ORFs are shown. The heights of the bars indicate the number of insertions identified within the 50-bp intervals that they cover. (**B**) Base composition of the 5 bp Tca2 target site duplication. Positions shown 5′ to 3′.
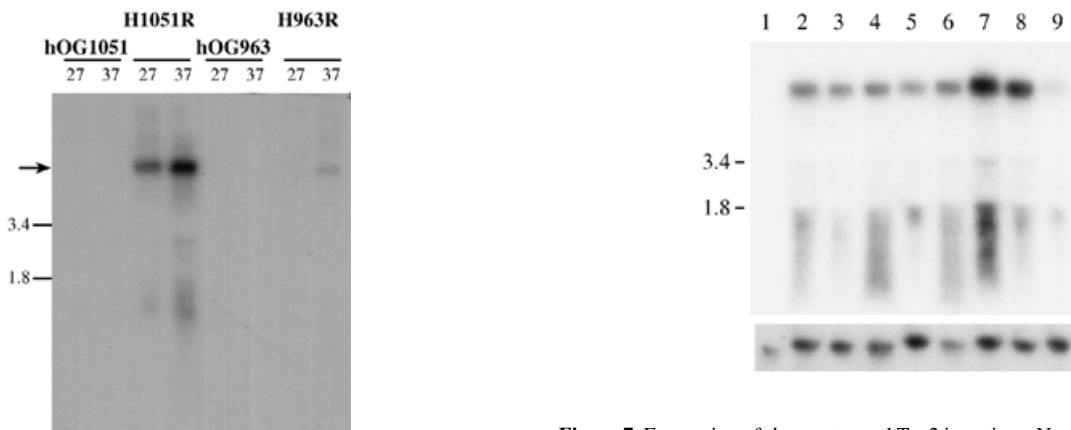
Figure 6. Expression of *URA3*-tagged Tca2. Northern analysis was performed on H1051R and H963R, and their parental strains hOG1051 and hOG963, respectively. RNA was extracted from strains grown at 27 and 37°C and probed with the *URA3* gene.

Figure 7. Expression of *de novo* tagged Tca2 insertions. Northern analysis was performed on RNA extracted from transposition clones from which the intron-containing tagged Tca2 had previously been removed. RNA was probed with the *URA3* gene. The same blot stripped and reprobed with the actin gene is shown below. Lanes: 1, hOG963; 2, H963R; 3, H963RU30r; 4, H963RU43r; 5, H963RU65r; 6, H963RU75r; 7, H963RU80r; 8, H963RUAA1r; 9, H963RUAA8r.

was achieved by selecting *ura3⁺* cells which had spontaneously become *ade2* auxotrophs as a result of intrachromosomal recombination between the two *ade2* genes that flank the marked Tca2 donor elements. Such recombinations result in the loss of the donor elements from the cells. The loss of the donor elements and the presence of single marked Tca2 elements in each strain was confirmed by Southern analyses (not shown). The original transgene was removed from eight strains which contained marked transposed Tca2 elements at different loci. Northern analysis using the *URA3* gene as a probe indicated that five of the eight Tca2 insertion sites analysed have almost identical RNA levels (Fig. 7). Strains

H963RU80r and H963RUAA1r both have a higher level of marked Tca2 transcripts than the other strains. Interestingly, their insertions are both very close to the 5′ ends of ORFs (74 and 21 bp) and the Tca2 element is in the same orientation as the ORF, suggesting that the expression of Tca2 in these strains may be elevated by the promoters into which the elements have been inserted. Strain H963RUAA8r has a substantially lower level of Tca2 transcripts than the other strains. The insertion in this strain is within an ORF, 316 bp from the start codon and in the same orientation as the ORF. These findings show that the transcriptional activity of Tca2 can be influenced by its genomic context.

DISCUSSION

The retrotransposon Tca2 was originally discovered due to its appearance as an abundant, extrachromosomal DNA molecule in the *C.albicans* strain hOG1042 (14). We have now shown that Tca2 is widespread in *C.albicans*, but that different *C.albicans* strains vary greatly in the amount of extrachromosomal Tca2 DNA that they produce and that Tca2 DNA levels are also strongly temperature dependent.

The abundance of extrachromosomal Tca2 DNA in different strains correlates well with the abundance of Tca2 RNA. This is most likely because the RNA is an intermediate in the synthesis of the extrachromosomal DNA, although it is possible that transcription of the extrachromosomal DNA also contributes to the correlation.

The correlation between extrachromosomal DNA and RNA levels is not perfect, however. For instance, no extrachromosomal copies of Tca2 were detected in ATCC10261, despite this strain having an abundance of apparently full-length Tca2 transcripts, whereas extrachromosomal Tca2 DNA molecules were detected in SC5314 (not shown) despite this strain apparently having lower levels of Tca2 RNA than ATCC10261. This lack of a perfect correlation between RNA and extrachromosomal DNA suggests that the efficiency of the reverse transcription reaction may differ among strains. In some cases this might be due to sequence differences among the elements in different strains. For example, sequence analysis of genomic copies of Tca2 in hOG759 show variations within their primer binding sites which result in different lengths of complementarity to the primer tRNA fragments (not shown). The variations may also, of course, result from strain-dependent influences on the efficiency of the process, for instance the abundance of the primer tRNA fragment may vary in different strains.

The levels of Tca2 RNA found in different strains appear to be related to the numbers of Tca2 elements in the strains, with strains having more copies generally having higher levels of RNA. However, the introduction of identical tagged elements into two different strains (hOG1051 and hOG963), at the same genomic location, resulted in very different levels of transcript. Similarly, marked Tca2 elements located at different sites in the genome were found to give rise to different levels of RNA. These findings indicate that not every Tca2 element produces the same level of RNA. The variations are probably produced both by strain-dependent factors, such as differences in the abundance of important transcription factors, and differences in the sequence and genomic context of the elements themselves.

While the relative amounts of transcript differ between strains, all strains analysed show a similar temperature-dependent expression pattern. Temperature-dependent regulation of transcription has also been observed for the *C.albicans* retrotransposon Tca1, although, in contrast to Tca2, higher levels of Tca1 expression were observed at 25 than at 37°C (38). Retrotransposition efficiency of Ty1 in *S.cerevisiae* is sensitive to increased temperature, while Tf1 transposition in *S.pombe* is unaffected by temperature increases from 22 to 36°C (39). Temperature-dependent expression and a controlled rate of Tca2 retrotransposition may be important in the evolution and virulence of *C.albicans*.

Homologous recombination of retrotransposons, either via LTRs or complete elements, has been observed for other retroelements (40,41). Such homologous recombination of retroelements has been shown to be independent of the integrase, suggesting that recombination is not directed by the integrase (40–42). It is clear that in W3, the *C.albicans* strain most easily analysed, homologous recombination is much less frequent than integrase-mediated integration.

Each retrotransposon family of *S.cerevisiae* and *S.pombe* shows a distinct integration site preference. While targeting of Ty3 is highly specific (35), the other elements appear to be targeted, in a less stringent fashion, to particular genomic environments, such as intergenic regions (6). This imprecise targeting, such as found here for Tca2, suggests that the integration is not occurring through the recognition of specific sequence elements, to which the integrase is binding. While the integration pattern of Tca2 and other retroelements appears imprecise, there is, however, some degree of targeting, since integration is clearly not randomly distributed throughout the genome. Ty3 has been shown to interact with Pol III transcription factors, thereby providing a mechanism for precise targeting to Pol III-transcribed genes (35). Interactions between the integration machinery and host-derived DNA binding proteins are thought to be involved in the targeting of other retrotransposons. Such interactions presumably allow the integrases to become associated with genomic DNA, and therefore the targeting is dependent on the position at which the host proteins bind. Targeting of Tca2 is mostly within a few hundred base pairs of the transcription start site of Pol II-transcribed genes, suggesting a possible interaction between the integrase and a Pol II-associated transcription factor(s). The Tca2 integrase may interact directly with Pol II-associated transcription factors or the relaxation of the DNA caused by their binding might be important. Such a hypothesis could explain why integrations were observed within ORFs, where presumably no transcription factors are bound. In these cases integration could occur in association with transcription, with Pol II being responsible for relaxing the DNA.

It is clear from the 'foot-printing' analyses of Smith *et al.* (43), that Ty1 insertion into *S.cerevisiae* ORFs can occur even though Ty1 more usually targets sites near tRNA genes. During growth of a Ty1-mutagenised population, those strains carrying insertions which are deleterious form an ever smaller proportion of the total. Those strains remaining exhibit a diminished set of insertion sites, as a result of selection against the strains carrying the deleterious insertions rather than as a result of any target specificity of the particular element. We do not consider that the integration bias we observed in the *C.albicans* Tca2 system is attributable to such selection, as *C.albicans* strains are all diploid, in contrast to the *S.cerevisiae* strains used by Smith *et al.* (43). Selection is much less likely to operate against the integrations that will be present heterozygously in the diploid *C.albicans*.

In this paper we have demonstrated for the first time the activity of a *C.albicans* transposable element. This has been shown in hOG1042 by observing the unselected accumulation of additional Tca2 copies. Retrotransposition has also been shown in two other strains (H963R, W3) by the development of a retrotransposition assay. The evidence for the activity of the marked element employed in this assay is the appearance of

new marked copies at a variety of loci, and the precise excision of an intron from the transposed copies.

The development of this assay system should enable a further dissection of the functional elements of Tca2: specific sites could be mutagenised to determine, for example, the key amino acid sequences involved in retrotransposition, and the mechanism by which the termination codon at the boundary of the *gag* and *pol* ORFs is bypassed. In addition to the study of Tca2, this system could be used to perform efficient random insertional mutagenesis of *C.albicans*. The distribution of Tca2 insertions observed here, suggests that gene expression and function might be affected for most insertions. Given the diploid nature of the *C.albicans* genome, however, it will also be necessary to develop strategies to bring these insertions to homozygosity or to develop haploid or partially haploid recipient strains. This may be achievable, given the recent reports suggesting the possibility of a sexual cycle for *C.albicans* (44,45).

ACKNOWLEDGEMENTS

REFERENCES

1. Boeke,J.D., Garfinkel,D.J., Styles,C.A. and Fink,G.R. (1985) Ty elements transpose through an RNA intermediate. *Cell*, **40**, 491–500.
2. SanMiguel,P., Tikhonov,A., Jin,Y.-K., Motchoulskaia,N., Zakharov,D., Melake-Berhan,A., Springer,P.S., Edwards,K.J., Lee,M., Avramova,Z. and Bennetzen,J.L. (1996) Nested retrotransposons in the intergenic regions of the maize genome. *Science*, **274**, 765–768.
3. Curcio,M.J., Sanders,N.J. and Garfinkel,D.J. (1988) Transpositional competence and transcription of endogenous Ty elements in *Saccharomyces cerevisiae*: implications for regulation of transposition. *Mol. Cell. Biol.*, **8**, 3571–3581.
4. Hansen,L.J. and Sandmeyer,S.B. (1990) Characterization of a transpositionally active Ty3 element and identification of the Ty3 integrase protein. *J. Virol.*, **64**, 2599–2607.
5. Zou,S., Ke,N., Kim,J.M. and Voytas,D.F. (1996) The *Saccharomyces* retrotransposon Ty5 integrates preferentially into regions of silent chromatin at the telomeres and mating loci. *Genes Dev.*, **10**, 634–645.
6. Behrens,R., Hayles,J. and Nurse,P. (2000) Fission yeast retrotransposon Tf1 integration is targeted to 5′ ends of open reading frames. *Nucleic Acids Res.*, **28**, 4709–4716.
7. Chapman,K.B., Bystrom,A.S. and Boeke,J.D. (1992) Initiator methionine tRNA is essential for Ty1 transposition in yeast. *Proc. Natl Acad. Sci. USA*, **89**, 3236–3240.
8. Kirchner,J., Connolly,C.M. and Sandmeyer,S.B. (1995) Requirement of RNA polymerase III transcription factors for *in vitro* position-specific integration of a retrovirus like element. *Science*, **267**, 1488–1491.
9. Devine,S.E. and Boeke,J.D. (1996) Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. *Genes Dev.*, **10**, 620–633.
10. Natsoulis,G., Thomas,W., Roghmann,M.C., Winston,F. and Boeke,J.D. (1989) Ty1 transposition in *Saccharomyces cerevisiae* is nonrandom. *Genetics*, **123**, 269–279.
11. Tumer,N.E., Parikh,B.A., Li,P. and Dinman,J.D.(1998) The pokeweed antiviral protein specifically inhibits Ty1-directed +1 ribosomal frameshifting and retrotransposition in *Saccharomyces cerevisiae*. *J. Virol.*, **72**, 1036–1042.
12. Morillon,A., Springer,M., and Lesage,P. (2000) Activation of Kss1 invasive-filamentous growth pathway induces Ty1 transcription and retrotransposition in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **20**, 5766–5776.
13. Ji,H., Moore,D.P., Blomberg,M.A., Braiterman,L.T., Voytas,D.F., Natsoulis,G. and Boeke,J.D. (1993) Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences. *Cell*, **73**, 1007–1018.
14. Matthews,G.D., Goodwin,T.J., Butler,M.I., Berryman,T.A. and Poulter,R.T. (1997) pCal, a highly unusual Ty1/copia retrotransposon from the pathogenic yeast *Candida albicans*. *J. Bacteriol.*, **179**, 7118–7128.
15. Mackinnon,J.E. and Artagaveytia-Allende,R.C. (1945) The so-called genus *Candida* Berkhout, 1923. *J. Bacteriol.*, **49**, 317–334.
16. Agatensi,L., Franchi,F., Mondello,F., Bevilacqua,R.L., Ceddia,T., De Bernardis,F. and Cassone,A. (1991) Vaginopathic and proteolytic *Candida* species in outpatients attending a gynaecology clinic. *J. Clin. Pathol.*, **44**, 826–830.
17. Gillum,A.M., Tsay,E.Y.H. and Kirsch,D.R. (1984) Isolation of the *Candida albicans* gene for orotidine-5′-phosphate decarboxylase by complementation of *S. cerevisiae ura3* and *E. coli pyrF* mutations. *Mol. Gen. Genet.*, **198**, 179–182.
18. Kelly,R., Miller,S.M., Kurtz,M.B. and Kirsch,D.R. (1987) Directed mutagenesis in *Candida albicans*: one-step gene disruption to isolate *ura3* mutants. *Mol. Cell. Biol.*, **7**, 199–207.
19. Fonzi,W.A. and Irwin,M.Y. (1993) Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics*, **134**, 717–728.
20. Philippsen,P., Stotz,A. and Scherf,C. (1991) DNA of *Saccharomyces cerevisiae*. *Methods Enzymol.*, **194**, 169–182.
21. Schmitt,M.E., Brown,T.A. and Trumpower,B.L. (1990) A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **18**, 3091–3092.
22. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Analysis of genomic DNA by Southern hybridisation. In Fore,N., Nolan.C. and Ferguson,M. (eds), *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
23. Church,G.M. and Gilbert,W. (1984) Genomic sequencing. *Proc. Natl Acad. Sci. USA*, **81**, 1991–1995.
24. Brown,T. and Mackey,K. (1997) Analysis of RNA by northern and slot blot hybridisation. In Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds), *Current Protocols in Molecular Biology*. John Wiley and Sons, Inc., NY.
25. Cannon,R.D., Jenkinson,H.F. and Shepherd,M.G. (1990) Isolation and nucleotide sequence of an autonomously replicating sequence (ARS) element functional in *Candida albicans* and *Saccharomyces cerevisiae*. *Mol. Gen. Genet.*, **221**, 210–218.
26. Losberger,C. and Ernst,J.F. (1989) Sequence and transcript analysis of the *C. albicans* URA3 gene encoding orotidine-5′-phosphate decarboxylase. *Curr. Genet.*, **16**, 153–158.
27. Basrai,M.A., Lubkowitz,M.A., Perry,J.R., Miller,D., Krainer,E., Naider,F. and Becker,J.M. (1995) Cloning of a *Candida albicans* peptide transport gene. *Microbiology*, **141**, 1147–1156.
28. Plant,E.P. (1999) Retrotransposons in *Candida albicans*. PhD thesis, University of Otago, Dunedin, New Zealand.
29. Schmuke,J.J., Davisson,V.J., Bonar,S.L., Gheesling Mulli,K. and Dotson,S.B. (1997) Sequence analysis of the *Candida albicans ADE2* gene and physical separation of the two functionally distinct domains of the phosphoribosylaminoimidazole carboxylase. *Yeast*, **13**, 769–776.
30. Bailey,T.L. and Elkan,C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In Altman,R., Brutlag,D., Karp,P., Lathrop,R. and Searls,D. (eds), *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*. AAAI Press, Menlo Park, CA.
31. Lowe,T.M. and Eddy,S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.*, **25**, 955–964.
32. Pochart,P., Agoutin,B., Rousset,S., Chanet,R., Doroszkiewicz,V. and Heyman,T. (1993) Biochemical and electron microscope analyses of the DNA reverse transcripts present in the virus-like particles of the yeast transposon Ty1. Identification of a second origin of Ty1DNA plus strand synthesis. *Nucleic Acids Res.*, **21**, 3513–3520.
33. Curcio,M.J. and Garfinkel,D.J. (1991) Single-step selection for Ty1 element retrotransposition. *Proc. Natl Acad. Sci. USA*, **88**, 936–940.
34. Kawakami,K., Pande,S., Faiola,B., Moore,D.P., Boeke,J.D., Farabaugh,,P.J., Strathern,J.N., Nakamura,Y. and Garfinkel,D.J. (1993) A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in *Saccharomyces cerevisiae*. *Genetics*, **35**, 309–320.

35. Yieh,L., Kassavetis,G., Geiduschek,E.P. and Sandmeyer,S.B. (2000) The Brf and TATA-binding protein subunits of the RNA polymerase III transcription factor IIIB mediate position-specific integration of the *gypsy*-like element, Ty3. *J. Biol. Chem.*, **275**, 29800–29807.

36. Kim,J.M., Vanguri,S., Boeke,J.D., Gabriel,A. and Voytas,D.F. (1998) Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res.*, **8**, 464–478.

37. Goodwin,T.J. and Poulter,R.T. (2000) Multiple LTR-retrotransposon families in the asexual yeast *Candida albicans*. *Genome Res.*, **10**, 174–191.

38. Chen,J.Y. and Fonzi,W.A. (1992) A temperature-regulated, retrotransposon-like element from *Candida albicans*. *J. Bacteriol.*, **174**, 5624–5632.

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

40. Hoff,E.F., Levin,H.L. and Boeke,J.D. (1998) *Schizosaccharomyces pombe* retrotransposon Tf2 mobilizes primarily through homologous cDNA recombination. *Mol. Cell. Biol.*, **18**, 6839–6852.

41. Ke,N. and Voytas,D.F. (1999) cDNA of the yeast retrotransposon Ty5 preferentially recombines with substrates in silent chromatin. *Mol. Cell. Biol.*, **19**, 484–494.

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

43. Smith,V., Chou,K.N., Lashkari,D., Botstein,D. and Brown,P.O. (1996) Functional analysis of the genes of yeast chromosome V by genetic footprinting. *Science*, **274**, 2069–2074.

44. Tzung,K.W., Williams,R.M., Scherer,S., Federspiel,N., Jones,T., Hansen,N., Bivolarevic,V., Huizar,L., Komp,C., Surzycki,R., Tamse,R., Davis,R.W. and Agabian,N. (2001) Genomic evidence for a complete sexual cycle in *Candida albicans*. *Proc. Natl Acad. Sci. USA*, **98**, 3249–3253.

45. Hull,C.M., Raisner,R.M. and Johnson,A.D. (2000) Evidence for mating of the "asexual" yeast *Candida albicans* in a mammalian host. *Science*, **289**, 307–310.