JOURNAL OF BACTERIOLOGY, May 1994, p. 2477–2482
0021-9193/94/$04.00+0

Vol. 176, No. 9

Ylt1, a Highly Repetitive Retrotransposon in the Genome of the Dimorphic Fungus *Yarrowia lipolytica*

NATASCHA SCHMID-BERGER, BEATRICE SCHMID, AND GEROLD BARTH*

Department of Microbiology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

Received 22 December 1993/Accepted 25 February 1994

A highly repetitive composite element, Ylt1, was detected in the genome of the dimorphic fungus *Yarrowia lipolytica*. Ylt1 resembles retrotransposons found in other eukaryotes. It is about 9.4 kb long and can transpose in the genome. The Ylt1 element is bounded by a long terminal repeat (LTR), the zeta element. Several copies of zeta were isolated and sequenced. The sequence of this element is well conserved. It is 714 bp long and is bounded by nucleotides 5′-TG ... CA-3′, which are part of a short inverted repeat, a feature conserved in the LTRs of retroviruses and retrotransposons. Sequence analysis revealed motifs commonly found in LTR elements, like signals for the start and termination of transcription. The zeta element exists as part of retrotransposon Ylt1, as well as a solo element in the genome. Ylt1 and solo zeta elements are flanked by a 4-bp directly repeated genomic sequence. The copy numbers of Ylt1 and solo zeta are dependent on the strain examined, but at least 35 copies of the composite Ylt1 element and more than 30 copies of the solo zeta element per haploid genome have been observed.

Retrovirus-like elements have been detected in five species of higher fungi (3, 7, 16, 19). These transposable elements have a life cycle similar to that of retroviruses; however, they do not produce infectious virus particles. There exist two main classes of these mobile genetic elements, which are distinguished by the presence or absence of long terminal repeats (LTRs) at their ends (designated retrotransposon or retroposon, respectively) (15). The LINE-like transposon detected in an African strain of *Neurospora crassa* (16) is the sole described example of a retroposon in fungi. Retrotransposons in the yeast *Saccharomyces cerevisiae* are well studied. They are divided into two distinct subclasses on the basis of sequence comparisons (29), i.e., the Ty1/copia class and the Ty3/gypsy class. Four retrotransposons of *S. cerevisiae* (Ty1, Ty2, Ty4, and Ty5) (for reviews, see references 3 and 27), as well as the Tca1 retrotransposon of *Candida albicans* (7), are members of the first class related to the copia element of *Drosophila melanogaster*. Ty3 of *S. cerevisiae* (3), Tf1 and Tf2 of *Schizosaccharomyces pombe* (18, 28), and CfT-1 of the phytopathogenic filamentous fungus *Cladosporium fulvum* (19) are in the second class, related to the gypsy element of *D. melanogaster*. The LTRs have three functionally distinct regions (U3-R-U5) which are involved in initiation and termination of transcription and in transposition (25, 26). LTRs are often found as solo elements divorced from the retrotransposon. These solo elements arise by homologous recombination between the LTRs of the retrotransposon, with subsequent deletion of the internal region (5, 21). Although LTRs differ considerably in length in higher eukaryotes (from several hundred base pairs to more than 2 kb) (2, 15), only short ones (from 332 to 427 bp) have been detected in fungi (3, 7, 14, 18, 19). All known fungal retrotransposons, as well as solo LTRs, are flanked by a 5-bp duplication of the genomic target site (3, 7, 10, 12, 13, 19).

Insertion and excision of retrotransposons or recombination between retrotransposons may affect the expression of genes in which they integrate or of genes neighboring the insertion site; this is well documented for *S. cerevisiae* (for reviews, see references 3, 6, and 12). Such events may cause phenotypic instabilities, especially under conditions which induce transposition of retrotransposons. Phenotypic instability is a common feature of dimorphic fungi and of many pathogenic fungi, but no direct involvement of retrotransposons in such processes has been demonstrated.

We have previously described mutants of the dimorphic ascomycetous fungus *Yarrowia lipolytica* harboring dominant mutations in the *GPR1* (glyoxylate pathway regulator) gene which inhibit utilization of two-carbon compounds like acetate and ethanol (17). Further studies have shown that these mutants exhibit a pleiotropic phenotype and that some of them are highly unstable after transfer onto media containing two-carbon compounds (23a). The high frequency of reversion suggested that a mobile element is involved.

It was the aim of this study to determine whether a repetitive sequence exists in the *GPR1* gene or in its surroundings. Here we describe the detection, isolation, and characterization of Ylt1, a repetitive retrovirus-like element located upstream of the *GPR1* gene in the dimorphic fungus *Y. lipolytica*. This repetitive element is bounded by an LTR, zeta. Cloning and sequencing of several genomic copies of zeta revealed features conserved within the LTRs of retroviruses and retrotransposons. We found that zeta also exists as a solo element divorced from the composite element.

MATERIALS AND METHODS

Strains, media, and cultivation conditions. *Y. lipolytica* B204-12C (*MATa met6-1 spo1-1*) and the dominant, acetate-nonutilizing mutant B204-12C-112 (*MATa GPR1-1 met6-1 spo1-1*), as well as five of its revertants, were used for isolation of genomic DNA.

Escherichia coli JA221 (*hsdR lacY leuB6 recA trpE5*) (8) was used for amplification of plasmid DNA.

Plasmids pYLD1, harboring a 5-kb fragment containing the *GPR1-1* allele, and pNS3, which contains the wild-type *GPR1* allele, will be described elsewhere (23b). Plasmids pYLH1 to pYLH11 were selected after colony hybridization from the gene library of strain B204-12C-112 during this work.

* Corresponding author. Mailing address: Department of Microbiology, Biozentrumy University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland. Phone: +41 61 267 21 32. Fax: +41 61 267 21 18. Electronic mail address: barth@urz.unibas.ch.

Complete medium (YPD) and synthetic minimal medium (MMT) with 2% glucose (MMT-G) or 0.4% sodium acetate (MMT-A) were prepared as previously described (1). Amino acids were supplied at a final concentration of 50 mg/liter in MMT. Cultivation of cells took place at 28°C. Cultivation of *E. coli* was done in LB medium (20) at 37°C.

Genetic techniques. Spontaneous revertants of mutant strain B204-12C-112 were selected by growth on MMT-A. Cells were grown in YPD for 20 h at 28°C, harvested, washed once with MMT medium without a carbon source, and plated onto MMT-G plates (10^{-6} dilution) for calculation of the number of living cells and onto MMT-A plates (undiluted). Plates were incubated at 28°C. Revertants were picked from the MMT-A plates after 4 days.

DNA manipulations. DNAs of *Y. lipolytica* B204-12C and B204-12C-112 were isolated as described by Cryer et al. (9). The gene library of strain B204-12C-112 was constructed by ligation of *Sau*3A fragments of the genomic DNA into the *Bam*HI site of plasmid pINA237 (23a).

Southern hybridization (24), colony hybridization, and nonradioactive detection of DNA fragments were performed with the Amersham ECL Kit. The blots were stripped as previously described (22) before reprobing.

DNA sequencing (23) was performed with the T7-Sequencing Kit of Pharmacia LKB Biotechnology with double-stranded DNA preparations as templates. The sequence was obtained from both strands after subcloning of fragments into pUCBM20 (Boehringer GmbH, Mannheim, Germany) by using M13 primers and synthetic oligonucleotides.

PCR analysis was done under standard conditions (22), with 10 ng of chromosomal strain B204-12C-112 DNA as the template. Oligonucleotides OH5 (5′-GGTGATGTTGATGA AGCA-3′) and LTR1 (5′-CTTTGTGCGTACCAGGGA-3′) were used as primers for amplification of a 4.2-kb fragment of Ylt1.

Nucleotide sequence accession number. The nucleotide sequence of LTR zeta has been submitted to the EMBL data base in Heidelberg, Germany (accession number, X74146).

RESULTS

Detection, cloning, and sequencing of the zeta element. We have previously cloned and sequenced the *GPR1* gene and its dominant allele, *GPR1-1* (23a), which harbors a mutation in the 5′-upstream region of the open reading frame (23b). The dominant *GPR1-1* allele reverted during incubation on two-carbon compounds at a frequency of approximately 2×10^{-5} revertants per cell. This suggested that a transposable element was involved. Therefore, Southern hybridizations with the genomic DNA of mutant B204-12C-112 with probes containing different portions of the *GPR1-1* fragment (Fig. 1A and B) were done. Hybridization with an *Eco*RI-*Bsa*AI fragment from the 5′-upstream region of the *GPR1-1* allele as the probe revealed a highly repetitive sequence in the genome (Fig. 1B). At least 100 copies per haploid genome were present. We call this repetitive element zeta. Conversely, the *Nco*I-*Hin*dIII fragment encompassing the *GPR1* gene hybridized to only a single genomic fragment (Fig. 1B).

To determine the borders of the zeta element, eight other fragments containing the repetitive sequence were isolated from the genomic DNA of strain B204-12C-112. This was done by colony hybridization of a gene library of *Y. lipolytica* B204-12C-112 in *E. coli* JA221 with the *Eco*RV-*Bsa*AI fragment from the 5′-upstream region of the *GPR1-1* allele (Fig. 1B) as the probe. Additionally, a DNA fragment containing the wild-type *GPR1* allele (plasmid pNS3 in Fig. 2) (23b) was also included. Sequencing of these fragments revealed a highly conserved sequence of 714 bp bounded by nucleotides 5′-TG . . . CA-3′ (Fig. 2), which were part of an inverted trinucleotide repeat (TGT/ACA). This is a feature conserved in the LTRs of retroviruses and retrotransposons.

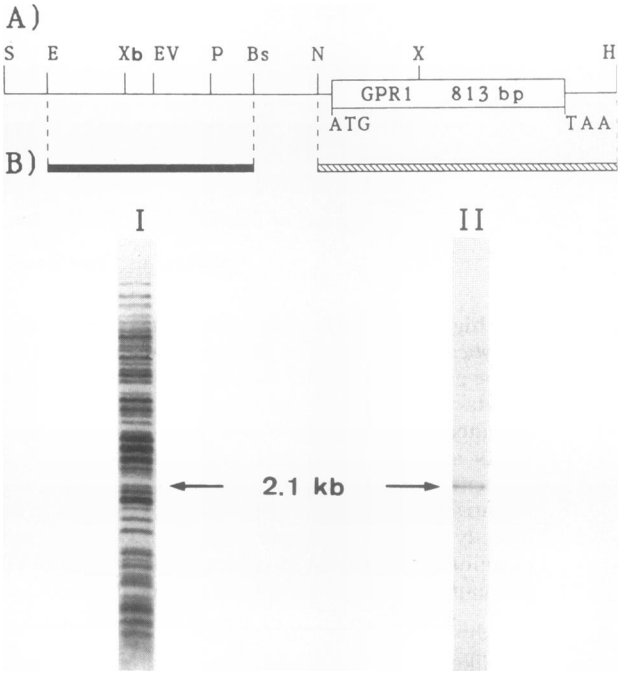

FIG. 1. The 2.1-kb genomic DNA fragment containing the *GPR1* gene together with the results of Southern blot hybridizations. (A) Restriction map of the 2.1-kb *Sal*I-*Hin*dIII fragment and the open reading frame of the *GPR1* gene. Restriction enzyme abbreviations: Bs, *Bsa*AI; E, *Eco*RI; EV, *Eco*RV; H, *Hin*dIII; N, *Nco*I; P, *Pst*I; S, *Sal*I; X, *Xho*I; Xb, *Xba*I. (B) *Eco*RI-*Bsa*AI and *Nco*I-*Hin*dIII fragments used as probes in Southern blot hybridizations with the genomic DNA of strain B204-12C-112 digested with *Sal*I and *Hin*dIII. Lanes: I, labelled fragments after hybridization with the *Eco*RI-*Bsa*AI fragment; II, labelled 2.1-kb fragment after hybridization with the *Nco*I-*Hin*dIII fragment.

Of the 10 repetitive sequences characterized, 5 (pYLH2, pYLH3, pYLH4, pYLH5, and pYLH11) are flanked by a 4-bp direct repeat (Fig. 2) which resembles integration sites of retroviruses and retrotransposons. Two such putative target site sequences (CATG and TATG), which are very similar, were detected. In two cases (pYLH2 and pYLH3 [Fig. 2]), the putative target site duplication differs at the first base (CATG/GATG), which may indicate that the first base is exchangeable or that a mutation had occurred since the time of integration, as is known for several solo LTRs in other organisms. In four cases (pYLD1, pNS3, pYLH1, and pYLH7), the 4-bp target site CATG sequence is present only at the 5′ end.

These data indicate that zeta elements closely resemble the LTRs of retroviruses and retrotransposons. LTRs have characteristic sequence features related to their mechanisms of transcription and transposition (3, 26). Consequently, the nucleotide sequences of the zeta elements (Fig. 3) were compared with those of the LTRs of other known fungal retrotransposons. This analysis revealed that putative signals for the start and termination of transcription are present in zeta (Fig. 3).

In some cases, the border between U3 and R within an LTR is formed by an inverted repeat (25). In the zeta element, a

pYLH4
cgagtaCATG**TGT**...708 bp...**ACA**CATGgggagatg

pYLH11
cagagcCATG**TGT**...708 bp...**ACA**CATGgagctac

pYLH5
gtcacgTATG**TGT**...708 bp...**ACA**TATGtagtgca

pYLH2
cctgggCATG**TGT**...708 bp...**ACA**GATGgatgaat

pYLH3
cctgggCATG**TGT**...708 bp...**ACA**GATGgatgaat

pYLD1
attccaCATG**TGT**...708 bp...**ACA**tcTGGTGGACG

pNS3
attccaCATG**TGT**...708 bp...**ACA**tcTGGTGGACG

pYLH1
gattgaCATG**TGT**...708 bp...**ACA**tcTGGTGGACG

pYLH7
cctagaCATG**TGT**...708 bp...**ACA**tcTGGTGGACG

pYLH9
ctcac**ATACAT**...543 bp...**CTCACA**tcTGGTGccca

FIG. 2. Borders of zeta elements written 5′ to 3′. Zeta itself is underlined. The plasmids containing these fragments are identified at the left. The trinucleotide inverted repeats at the ends of zeta are in capital letters. The duplicated 4-bp genomic target sites bounding zeta are also in capital letters but are not underlined. Plasmids pYLH2 and pYLH3 harbor duplicated regions from the genome which differ by exchanges of 3 bp in zeta. The zeta-containing fragments in plasmids pYLD1, pNS3, pYLH1, and pYLH7 have identical sequences at the 3′ side of zeta (also in capital letters).

```
  1 tgtaacactc gctctggaga gttagtcatc cgacagggta actctaatct
 51 cccaacacct tattaactct gcgtaactgt aactcttctt gccacgtcga
101 tcttactcaa ttttcctgct catcatctgc tggattgttg tctatcgtct
151 ggctctaata catttattgt ttattgccca aacaactttc attgcacgta
201 agtgaattgt tttataacag cgttcgccaa attgctgcgc catcgtcgtc
251 cggctgtcct accgttaggg gtagtgtgtc tcacactacc gaggttacta
301 gagttgggaa agcgatactg cctcggacac accacctggg tcttacgact
351 gcagagagaa tcggcgttac ctctctcaca aagcccttca gtaccgccgc
401 ctgtcgggaa ccgcgttcag gtggaacagg accacctccc ttgcacttct
451 tggTATAtca gtaTAGgctg atgtattcaT AGTGgggTTT ttcatAATAA
501 Atttactaac ggcaggcaac attcactcgg cttaaacgca aaacggaccg
551 tcttgatatc ttctgacgca ttgaccaccg agaaatagtg ttagttaccg
601 ggtgagttat tgttcttcta cacagcgacg cccatcgtct agagttgatg
651 tactaactca gatttcacta cctaccctat ccctggtacg cacaaagcac
701 tttatttttct caca
```

FIG. 3. Nucleotide sequence of zeta from 5′ to 3′. The trinucleotide inverted repeats at the ends are marked by arrowheads. The putative TATA box is in capital letters. The potential termination signals and the polyadenylation site are in capital letters and are underlined. An internal 37-bp degenerate inverted repeat and the 9-bp internal complete inverted repeat are marked by arrows.

37-bp degenerate inverted repeat exists upstream of the putative TATA box. Additionally, the middle of the degenerate repeat forms a perfect inverted repeat of 9 bp (Fig. 3). The function of this region is not known, but its structure and location suggest that it is the U3-R border.

Sequence comparisons of the cloned zeta elements have shown that the entire element is highly conserved. Only 2- to 12-bp differences were detected in the sequences of nine of these repeats (data not shown).

However, one repeat (pYLH9 [Fig. 2]) contained only 556 bp of the 714-bp conserved sequence of zeta. This repeat has lost the 5′ end and is bounded by direct repeats (5′-ACATA CAT . . . ACAT-3′) instead of the inverted repeat TGT/ACA found in the other zeta elements. Furthermore, this direct repeat is overlapped at the 3′ and the 5′ ends by a 7-bp direct repeat which includes target site sequences (Fig. 2). The function of these direct repeats is not known. It is very likely that this element is a truncated version of zeta.

It is known that some retrotransposons integrate in front of specific genes, and solo LTRs are found at these sites (for a review, see reference 3). The most extensively studied example is the Ty3 retrotransposon of *S. cerevisiae*, which integrates in front of tRNA genes in *S. cerevisiae* (4). However, no tRNA genes or other categories of specific genes were found in the areas surrounding the zeta elements in *Y. lipolytica*.

Zeta as part of composite element Ylt1. Comparison of the DNA sequences surrounding the zeta elements in 10 cloned fragments (Fig. 2) showed that the downstream regions of the zeta elements are identical in four cases (pNS3, pYLD1, pYLH1, and pYLH7). However, the upstream sequences are completely different, with the exception of pNS3 and pYLD1, which contain allelic regions. Several Southern hybridizations were done to determine whether or not these identical downstream regions of these zeta elements represent another repetitive sequence which forms, together with zeta, a composite retrotransposon-like element.

Analysis of the cloned *GPR1-1* allele showed that this gene is localized, together with a zeta element, on a 4-kb genomic *Hin*dIII fragment with a *Sal*I site in the middle (Fig. 4A). Therefore, genomic DNA was digested with *Hin*dIII and *Sal*I and then used in Southern hybridizations. The genomic digest was first probed with the *Eco*RV-*Bsa*AI fragment, which contains an internal part of zeta (Fig. 4B, lane II). This revealed many bands, the sizes of which depend on the distance to the next *Sal*I or *Hin*dIII site in the genome outside of zeta. The second hybridization of the same filter was done with the *Eco*RI fragment from the upstream region of zeta harboring the *Sal*I site (Fig. 4B, lane I), which should reveal only two bands of 1.9 kb (*Hin*dIII-*Sal*I fragment upstream of zeta and *GPR1* [Fig. 4A]) and 2.1 kb (*Sal*I-*Hin*dIII fragment containing zeta and *GPR1* [Fig. 4A]) if the upstream region is not repeated, but several fragments hybridized with this *Eco*RI probe. The 1.9-kb fragment gave a very strong signal, showing that many copies of this fragment are present in the genome. This demonstrates that the upstream region of zeta and zeta itself are parts of a composite element, which we called Ylt1.

The fragments which hybridized to the *Eco*RI probe thus contained the *Sal*I-*Eco*RI part of the Ylt1 composite element. The sizes of these fragments are variable, depending on the location of the next *Hin*dIII or *Sal*I site in the genome. Taking into account the fact that stronger bands contain more than one fragment, the number of these fragments (about 35) is the minimal number of copies of the composite element.

Analysis of the sequence at the border of zeta and the

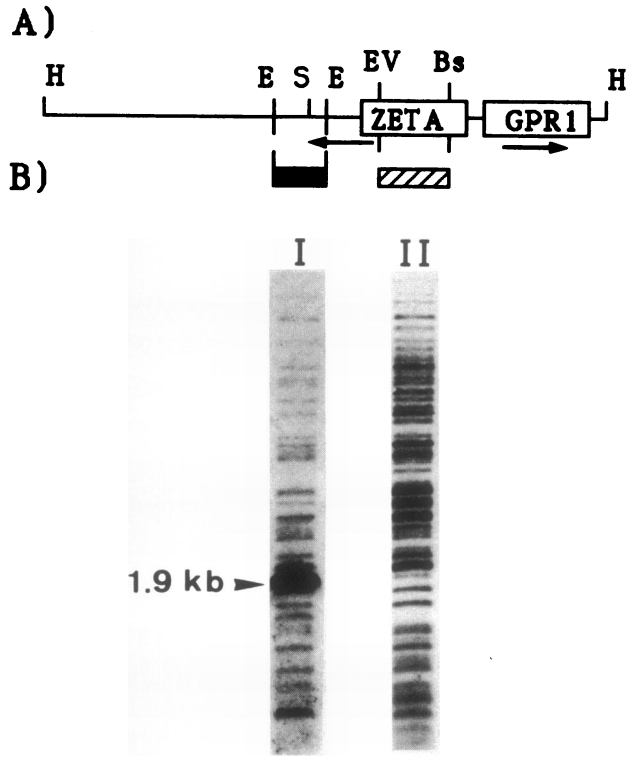

FIG. 4. The 4-kb DNA fragment containing the *GPR1* gene and zeta together with the results of Southern blot hybridizations. (A) Restriction map of the 4-kb *Hin*dIII fragment of genomic DNA. Abbreviations of restriction enzymes are defined in the legend to Fig. 1. The arrows indicate the putative directions of transcription. (B) Fragments used as probes in Southern blot hybridizations of genomic DNA digested with *Hin*dIII and *Sal*I. This digestion creates a 1.9-kb fragment and a 2.1-kb fragment which harbors zeta and the *GPR1* gene. Lane I, labelled fragments after hybridization with the *Eco*RI fragment from the region outside of zeta. The 1.9-kb fragment gave a strong hybridization signal, indicating that many copies are present in the genome. Lane II, labelled fragments after hybridization with an internal *Eco*RV-*Bsa*AI fragment of zeta as the probe.

internal part of Ylt1 revealed the sequence TGGTGGACG 2 bp from zeta in Ylt1 (Fig. 2). This motif is very similar to the primer-binding site (PBS) in Ty1, Ty2, and Ty3 (TGGTAG CGC/T) (3, 11). This suggests that transcription of the Ylt1 element adjacent to the *GPR1* gene proceeds in the direction opposite to that of transcription of the *GPR1* gene (Fig. 4A).

Size of composite element Ylt1. The question of whether composite element Ylt1 is bounded at both ends by the zeta repeat, like retroviruses and retrotransposons, arose. To determine the size of the complete Ylt1 element, isolates pYLH1 to pYLH11 were analyzed for the presence of zeta elements; however, none of them contained a second zeta. Therefore, another approach was used. After digestion of genomic DNA from strain B204-12C-112 with *Pst*I, for which a single site exists in zeta, a strongly hybridizing band of 4.8 kb was detected with an internal probe (*Eco*RI fragment [Fig. 4]) of Ylt1. *Pst*I fragments of this size were excised from the agarose gel and subsequently cloned into vector pUCBM20. One of the resulting plasmids (pYLH22), which gave a strong hybridization signal with the *Eco*RI fragment of Ylt1, was sequenced, but this insert contained a zeta element only at the same end as known before.

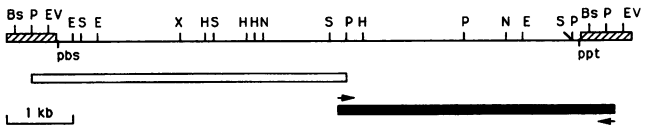

FIG. 5. Restriction map of retrotransposon Ylt1. For abbreviations of restriction enzymes, see the legend to Fig. 1. The boxes at both ends are the zeta elements bounding Ylt1. pbs, primer-binding site; ppt, polypurine tract. The empty box below the restriction map represents the 4.8-kb *Pst* fragment which was cloned as described in the text. The dark box represents the 4.2-kb fragment isolated by PCR from the genomic DNA. The arrows indicate the primers used in the PCR.

To find the second end of putative retrotransposon Ylt1, a PCR was done with an oligonucleotide priming toward the zeta-free end of the 4.8-kb fragment and another oligonucleotide priming in the opposite orientation in the zeta element. A fragment of about 4.2 kb which gave a strong hybridization signal with the zeta-free end of the 4.8-kb fragment was isolated. To ensure that the 4.2-kb fragment was part of a repetitive element, a 1.7-kb zeta-less *Pst*I fragment of the 4.2-kb fragment was hybridized to the genomic *Hin*dIII-*Sal*I digest of B204-12C-112. This Southern analysis also showed a pattern of several hybridizing bands (data not shown). This pattern was different from that which arose from hybridization with the *Eco*RI fragment of the 4.8-kb fragment, indicating that the 4.2-kb fragment was not identical to the previously characterized 4.8-kb fragment.

Sequencing of the ends of the 4.2-kb fragment showed that one end overlaps the zeta-free region of the 4.8-kb fragment, while the other end contains a zeta element. Immediately at the 5′ end of such 3′ LTRs of retrotransposons, a highly conserved polypurine tract acting as the plus-strand primer is present. In Ylt1, such a polypurine tract (5′-AGGGCAGG-3′) was also found directly at the 5′ border of zeta in the 4.2-kb fragment.

These data showed that the 4.2-kb fragment is, in fact, the second half of Ylt1. The resulting size of Ylt1 is therefore about 9.4 kb, and its restriction pattern is shown in Fig. 5.

Southern hybridizations of genomic DNA digested with *Cla*I, which does not cut within Ylt1, were performed to confirm that Ylt1 is a large repetitive element and to determine the approximate number of solo zetas (Fig. 6). If the calculated size of Ylt1 was correct, then no bands smaller than 9.4 kb should be detected after hybridization of the *Cla*I-cut DNA with an internal fragment of Ylt1. In fact, this hybridization resulted only in bands larger than 12 kb (Fig. 6A). Hybridization with an internal fragment of zeta resulted in 21 additional bands (Fig. 6B). Taking into account the facts that stronger bands contain more than one fragment and that fragments are difficult to count in the region above 10 kb, it was concluded that solo zeta elements have a copy number of at least 30.

Transposition of Ylt1. To determine whether Ylt1 can transpose, spontaneous revertants of mutant strain B204-12C-112 were analyzed by Southern hybridization. DNA preparations of the mutant and five of its revertants were digested with *Cla*I, which does not cut in Ylt1, and *Xho*I, which cuts once in Ylt1, to reduce the sizes of the fragments for better separation by electrophoresis. Hybridizations were done with a probe of the internal part of Ylt1 and an internal part of zeta (Fig. 7). Whereas the restriction patterns were unchanged in two revertants (R4 and R5), one band disappeared (Fig. 7A and B) and an additional band occurred after hybridization with the zeta probe in the samples of revertants R1 and R3 (Fig. 7B).

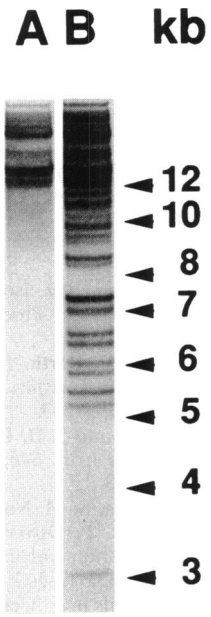

FIG. 6. Southern blot hybridizations of *Cla*I-digested genomic DNA of strain B204-12C-112 with two different probes of Ylt1. Lanes: A, hybridization with an internal *Eco*RI fragment of Ylt1 as the probe; B, hybridization of the same blot with an internal *Eco*RV-*Bsa*AI fragment of zeta as the probe. The first probe was removed before by washing as described in Materials and Methods.

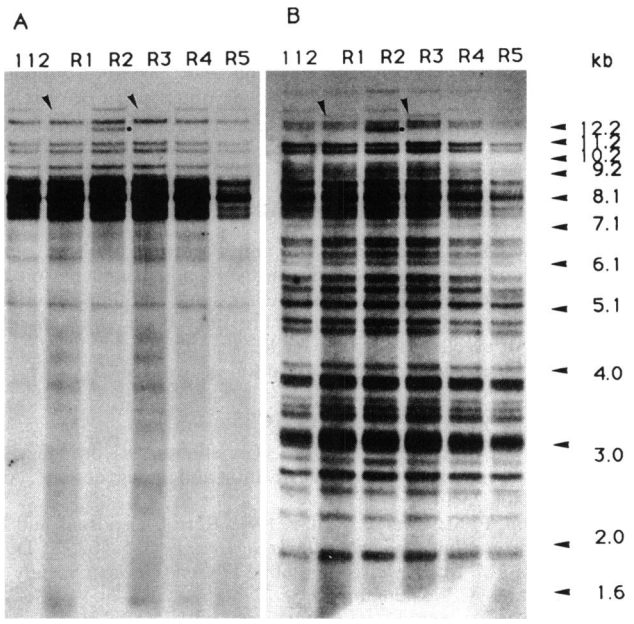

FIG. 7. Southern blot hybridizations of genomic DNAs of strain B204-12C-112 and five revertants. The DNA was digested with *Cla*I and *Xho*I before electrophoretic separation of the fragments. The same blot was hybridized with two different fragments of Ylt1. (A) Hybridization with an internal *Eco*RI fragment of Ylt1 as the probe. (B) Hybridization of the same blot with an internal *Eco*RV-*Bsa*AI fragment of zeta as the probe. The first probe was removed before by washing as described in Materials and Methods. The arrowheads at the tops of the lanes indicate the positions of lost (A) and new (B) fragments in the DNAs of revertants R1 and R3. The dots mark the newly arisen fragment in the DNA of revertant R2 (A and B).

This indicates that a recombination which resulted in a new solo zeta in these two revertants, R1 and R3, had occurred. In revertant R2, an additional band which hybridized with the internal part of Ylt1 and with zeta was detected (Fig. 7A and B). This suggests that a copy of Ylt1 had integrated into another site in the genome.

We conclude from these results that repetitive element Ylt1 is an active retrotransposon which can transpose in the genome. In addition, new copies of solo LTRs can also be created by recombination within the retrotransposon.

DISCUSSION

We have detected a new retrovirus-like element, Ylt1, during analysis of the genomic loci of two *GPR1* alleles in the dimorphic fungus *Y. lipolytica*. Ylt1 shares several features with other retrotransposons found in fungi and other eukaryotes. It is bounded by an LTR, zeta, which contains several sequence motifs commonly found in LTRs. Transcription of retroviruses and retrovirus-like elements starts in the 5′ LTR and terminates in the 3′ LTR. Therefore, transcription initiation signals like the TATA box and termination signals should be found in putative LTR sequences. A TATA-like sequence is present 458 bp from the 5′ end of zeta. This putative TATA box is closely followed by a sequence identical to that of the proposed eukaryotic termination signal (TAG . . . TAGTG . . . TTT) (30) and a polyadenylation signal (AATAAA) (Fig. 3).

Zeta itself is bounded by 5′-TG . . . CA-3′, as are most of the LTRs of retroviruses and retrotransposons. These dinucleotides are part of the inverted trinucleotide 5′-TGT . . . ACA-3′, as in the LTR of Tfl1 in *S. pombe* (18). Zeta is considerably longer than the longest hitherto known LTR from fungi (19) or from *D. melanogaster* (2) but comparable in length to LTRs from plants (15). Copies of solo zeta and Ylt1 are flanked by a direct repeat of the target site, as in all retrotransposons. However, in contrast to all other fungal target site duplications of 5 bp, only 4 bp are duplicated in *Y. lipolytica*, as are most target sites of retrotransposons in drosophila (2). The target sites analyzed at 14 borders of zeta differed only in the first base. This implies conservation of three bases of the target site (consensus sequence, C/G/TATG), but more data are necessary to confirm this conclusion.

Ylt1 therefore belongs to the group of high-copy-number fungal retrotransposons, together with Ty1 of *S. cerevisiae* and Tfl1 of *S. pombe*. Ylt1, with a size of 9.4 kb, is one of the largest retrotransposons known and is comparable to del1 of *Lilium henryi* and B104 and springer in drosophila (2; 15)

For DNA synthesis, retroviruses and retrotransposons require a PBS downstream of the 5′ LTR and a polypurine tract upstream of the 3′ LTR (11, 25). The PBS in Ylt1 starts 2 nucleotides away from the 3′ border of the 5′ zeta, as in Ty3 of *S. cerevisiae* (3) and in retroviruses (26). It contains the 5′-TGG-3′ sequence which is complementary to the highly conserved CCA sequence at the 3′ end of tRNA molecules. The 9-nucleotide PBS of Ylt1 is similar to the PBS found in Ty1, Ty2, and Ty3 of *S. cerevisiae* (for a review, see reference 3). The polypurine tract immediately upstream of the 3′ zeta in Ylt1 resembles the plus-strand PBS commonly found at this site in retrotransposons.

It is strongly suggested by the detection of new copies in spontaneous revertants of a dominant mutant that the Ylt1 element can transpose and recombine in the genome.

The data show that Ylt1 is an active retrotransposon bounded by LTR zeta. The sequence of the entire element, which is presently being determined, will finally show to which

subclass of retrotransposons Ylt1 belongs, but its length and the long LTRs indicate that Ylt1 is a new type of retrotransposon previously undetected in fungi. The possible effects of Ylt1 on expression of genes and its significance in the biology of the dimorphic fungus *Y. lipolytica* are under further investigation.

ACKNOWLEDGMENTS

We thank T. Bickle, P. Linder, and P. Philippsen for helpful discussion and suggestions during the work and comments on the manuscript. We thank also M. MacWilliams and E. Stoop for critical reading of the manuscript. Thanks are due to R. Doelz for the availability of excellent computer facilities.

This work was supported by the Canton of Basel-Stadt.

REFERENCES

1. **Barth, G., and H. Weber.** 1983. Genetic studies on the yeast *Saccharomycopsis lipolytica*. Inactivation and mutagenesis. Z. Allg. Mikrobiol. **23**:147–157.
2. **Bingham, P. M., and Z. Zachar.** 1989. Retrotransposons and FB transposon from *Drosophila melanogaster*, p. 485–502. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
3. **Boeke, J. D., and S. B. Sandmeyer.** 1991. Yeast transposable elements, p. 193–261. *In* J. R. Broach, J. R. Pringle, and E. W. Jones (ed.), The molecular biology of the yeast *Saccharomyces*, vol. 1. Genome dynamics, protein synthesis, and energetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
4. **Brodeur, G. M., S. B. Sandmeyer, and M. V. Olson.** 1983. Consistent association between sigma elements and tRNA genes in yeast. Proc. Natl. Acad. Sci. USA **80**:3292–3296.
5. **Cameron, J. R., E. Y. Loh, and R. W. Davis.** 1979. Evidence for transcription of dispersed repetitive DNA families in yeast. Cell **16**:736–751.
6. **Chaleff, D. T., and G. R. Fink.** 1980. Genetic events associated with an insertion mutation in yeast. Cell **21**:227–237.
7. **Chen, J.-Y., and W. A. Fonzi.** 1992. A temperature-regulated, retrotransposon-like element from *Candida albicans*. J. Bacteriol. **174**:5624–5632.
8. **Clarke, L., and J. Carbon.** 1976. A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E. coli* genome. Cell **9**:91–97.
9. **Cryer, D. R., R. Eccleshall, and J. Marmur.** 1975. Isolation of yeast DNA. Methods Cell Biol. **12**:39–44.
10. **Eibel, H., J. Gafner, A. Stotz, and P. Philippsen.** 1981. Characterization of the yeast mobile element Ty1. Cold Spring Harbor Symp. Quant. Biol. **45**:609–617.
11. **Errede, B., T. S. Cardillo, G. Wever, and F. Sherman.** 1981. Studies on transposable elements in yeast. I. ROAM mutations causing increased expression of yeast genes: their activation by signals directed toward conjugation functions and their formation by insertion of Ty1 repetitive elements. Cold Spring Harbor Symp. Quant. Biol. **45**:593–602.
12. **Farabaugh, P. J., and G. R. Fink.** 1980. Insertion of the eukaryotic transposable element Ty1 creates a 5 bp duplication. Nature (London) **286**:352–356.
13. **Gafner, J., and P. Philippsen.** 1980. The yeast transposon Ty1 generates duplications of target DNA on insertion. Nature (London) **286**:414–418.
14. **Genbauffe, F. S., G. E. Chrisholm, and T. G. Cooper.** 1984. Tau, sigma, and delta. A family of repeated elements in yeast. J. Biol. Chem. **259**:10518–10525.
15. **Grandbastien, M.-A.** 1992. Retroelements in higher plants. Trends Genet. **8**:103–108.
16. **Kinsey, J. A., and J. Helber.** 1989. Isolation of a transposable element from Neurospora crassa. Proc. Natl. Acad. Sci. USA **86**:1929–1933.
17. **Kujau, M., H. Weber, and G. Barth.** 1992. Characterization of mutants of the yeast *Yarrowia lipolytica* defective in acetyl-coenzyme A synthetase. Yeast **8**:193–203.
18. **Levin, H. L., D. C. Weaver, and J. D. Boeke.** 1990. Two related families of retrotransposons from *Schizosaccharomyces pombe*. Mol. Cell. Biol. **10**:6791–6798.
19. **McHale, M. T., I. N. Roberts, S. M. Noble, C. Beaumont, M. P. Whitehead, D. Seth, and R. P. Oliver.** 1992. CfT-I: an LTR-retrotransposon in Cladosporium fulvum, a fungal pathogen of tomato. Mol. Gen. Genet. **233**:337–347.
20. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 433. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
21. **Roeder, G. S., and G. R. Fink.** 1980. DNA rearrangements associated with a transposable element in yeast. Cell **21**:239–249.
22. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

23a. **Scheuber, T., and G. Barth.** Unpublished data.

23b. **Schmid-Berger, N., and G. Barth.** Unpublished data.

24. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
25. **Temin, H. M.** 1981. Structure, variation and synthesis of retrovirus long terminal repeat. Cell **27**:1–3.
26. **Varmus, H., and P. Brown.** 1989. Retroviruses, p. 53–108. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
27. **Voytas, D. F., and J. D. Boeke.** 1992. Yeast retrotransposon revealed. Nature (London) **358**:717. (Letter.)
28. **Weaver, D. C., G. V. Shpakovski, E. Caputo, H. L. Levin, and J. D. Boeke.** 1993. Sequence analysis of closely related retrotransposon families from fission yeast. Gene **131**:135–139.
29. **Xiong, Y., and T. H. Eickbush.** 1988. Similarity of reverse transcriptase-like sequences of viruses, transposable elements, and mitochondrial introns. Mol. Biol. Evol. **5**:675–690.
30. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription in yeast. Cell **28**:563–573.