

# Transformation of *Saccharomyces cerevisiae* with Nonhomologous DNA: Illegitimate Integration of Transforming DNA into Yeast Chromosomes and In Vivo Ligation of Transforming DNA to Mitochondrial DNA Sequences

ROBERT H. SCHIESTL,† MARGARET DOMINSKA, AND THOMAS D. PETES\*

Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280

Received 4 November 1992/Returned for modification 22 December 1992/Accepted 29 January 1993

**When the yeast *Saccharomyces cerevisiae* was transformed with DNA that shares no homology to the genome, three classes of transformants were obtained. In the most common class, the DNA was inserted as the result of a reaction that appears to require base pairing between the target sequence and the terminal few base pairs of the transforming DNA fragment. In the second class, no such homology was detected, and the transforming DNA was integrated next to a CTT or GTT in the target; it is likely that these integration events were mediated by topoisomerase I. The final class involved the in vivo ligation of transforming DNA with nucleus-localized linear fragments of mitochondrial DNA.**

Recombination events involving little or no sequence homology (illegitimate recombination) have been observed in many types of organisms. In mammals, such events can result in mutations or structural rearrangements associated with tumors (27). In most eukaryotic organisms, DNA molecules introduced into the cell by transformation usually integrate into the genome quasi-randomly by illegitimate recombination events (33). In contrast, in the yeast *Saccharomyces cerevisiae*, transforming DNA is usually inserted into homologous loci (20). To determine whether yeast cells were capable of nonhomologous integration of transforming fragments, in previous experiments, we transformed a yeast strain that lacked *URA3* sequences with a *Bam*HI *URA3* fragment (34). Two unusual types of transformants were obtained. If the *Bam*HI fragments were transformed in the presence of the restriction enzyme *Bam*HI, most of the transformants contained an insertion of *URA3* in a *Bam*HI site in the yeast genome. These events (called restriction enzyme-mediated [REM] events) are a consequence of cleavage of genomic sequences by the exogenously added enzyme followed by incorporation of the transforming DNA; similar events have also been observed in a *Dictyostelium* sp. (24). In the absence of the enzyme, in three of three transformants examined, the *URA3* fragment inserted into a genomic GATC site. Since this sequence is identical to that found at the ends of *Bam*HI fragments, this result suggests that *S. cerevisiae* possesses a recombination system that is sensitive to extremely small amounts of homology.

Below, we report two new types of illegitimate integration events in *S. cerevisiae*. We analyzed the DNA sequences flanking the transforming DNA in 44 independent transformants. In 15 of these transformants, we also obtained the target sequence for the insertion. In addition to insertions resulting from recombination events involving very small

amounts of homology (4 bp or less), we observed integration events involving no sequence homology that may be catalyzed by topoisomerase I. Some of the transformants contained target site deletions or target site duplications, usually involving only a few base pairs. In addition, we found transformants resulting from the ligation of the transforming DNA to yeast mitochondrial DNA.

## MATERIALS AND METHODS

**Strains, plasmids, and media.** The yeast strain used for most of the transformation experiments was RSY12 (a *leu2-3,112 his3-11,15 ura3Δ::HIS3*). In this strain, the coding sequence of *URA3* has been deleted and replaced by *HIS3* sequences (34). The plasmid used in most of the transformation procedures was YIplac211 (19), which contains the *URA3* coding sequences inserted into the *Eco*O109 site of pUC19. Growth conditions and medium preparation were standard (35).

**Genetic and molecular techniques.** For experiments involving plasmid isolation and analysis, standard procedures were used (26). Small-scale plasmid preparations were done by the boiling method (21). Small-scale yeast DNA isolations were done by the method of Denis and Young (14). For Southern analysis of transformant DNA, both uncut DNA and DNA treated with *Bgl*II were analyzed; the hybridization probe used in these experiments was YIplac211.

Yeast transformations were done as described previously (34). In all experiments, about 20 μg of YIplac211 DNA was used per transformation. The DNA was treated with a restriction enzyme (either *Bam*HI, *Eco*RI, *Hind*III, *Pst*I, or *Asp* 718) before transformation. For some experiments, the enzyme (100 U) was removed by proteinase K treatment followed by phenol extraction (34); in other experiments, the enzyme was not removed. The frequency of transformation for YIplac211 (with or without restriction enzymes) was about two to five transformants per μg in all experiments.

From a number of transformants, we rescued the YIplac211 and junction sequences. DNA from the transformants was treated with *Bgl*II or *Cla*I and treated with DNA

\* Corresponding author.

† Present address: Department of Molecular and Cellular Toxicology, Harvard University School of Public Health, Boston, MA 02115.

ligase under conditions that promoted self-ligation. These DNA molecules were used to transform *Escherichia coli* DH5 $\alpha$  to ampicillin resistance. Plasmid DNA was isolated from *E. coli*, and this double-stranded DNA was sequenced by using Sequenase (23) and a Bio-Rad sequencing cell. The primers used for the sequencing were primers 1233 and 1212 from New England Biolabs.

Most targets for the insertions were determined by using the junction sequences in searches of the GenBank DNA data base with the Genetics Computer Group program FASTA. For one of the insertions (WT133), the sequence information obtained from the junctions was used to design oligonucleotide primers to allow polymerase chain reaction amplification of target DNA from the untransformed strain (34).

The stability of the Ura<sup>+</sup> transformed phenotype was determined by inoculating a colony isolated from medium lacking uracil to 3 ml of the rich growth medium YPD. After overnight growth at 32°C (about eight divisions of unselective growth), the culture was diluted and plated on either YPD (to determine the total number of cells) or medium containing 5-fluoro-uracil (to determine the number of Ura<sup>-</sup> cells [7]).

## RESULTS

In our previous study (34), we transformed the yeast strain RSY12, which lacked homology to the *URA3* gene, with a *Bam*HI fragment containing the *URA3* coding sequences. Since this fragment did not contain either a plasmid replication origin or a selectable drug resistance marker, the process of analyzing the sequences that flanked the insertion was rather tedious. Consequently, in the experiments reported here, we transformed RSY12 with a restriction fragment containing YIplac211 sequences (19). This plasmid is 3.8 kb in size and contains the *URA3* gene integrated into the *Eco*O109 site of pUC19. In the polylinker, the plasmid has a single site for the enzymes *Bam*HI, *Eco*RI, *Hind*III, *Asp*718, and *Pst*I. Three transformation experiments were done with *Bam*HI or *Eco*RI. In the first experiment, *Bam*HI-treated YIplac211 was used to transform either RSY12 in the absence of the *Bam*HI restriction enzyme. In the other two experiments, YIplac211 was treated with *Eco*RI before transformation. One of these experiments was done in the presence of *Eco*RI, and the other was done in the absence of the enzyme.

Following transformation, individual Ura<sup>+</sup> strains were purified and analyzed. Plasmids containing the insertions and flanking sequences were rescued by treating genomic DNA with *Bgl*II or *Cla*I (which do not cleave within YIplac211), allowing self-ligation of the resulting fragments, and transforming the ligated mixture in *E. coli*. Plasmids from the resulting ampicillin-resistant colonies were purified, and the junctions between plasmid DNA and genomic DNA were sequenced. To determine target sequences, the junction sequences were compared with existing GenBank sequences.

**Analysis of the integration of *Bam*HI-treated YIplac211 into RSY12.** The sequences flanking the YIplac211 DNA were analyzed in 19 transformants. Precise target sequences were obtained from seven transformants. Figure 1 shows the inferred structure of the junctions of plasmid YIplac211 and the genomic flanking DNA. End 1 is the end closest the *lacZ* gene, and end 2 is closest to the *lacI* gene. Also shown is potential base pairing between the ends of the plasmid and the target, as well as sites of cleavage necessary to explain

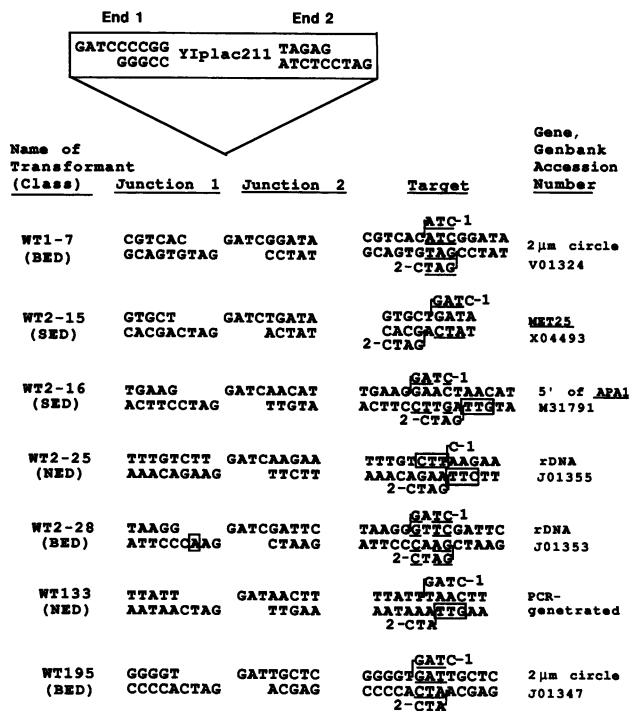


FIG. 1. Flanking sequences and targets of *Bam*HI-treated YIplac211 DNA transformed into RSY12. In this experiment, targets for seven genomic insertions were identified. Both ends of the *Bam*HI-cut YIplac211 had 5' protruding GATC sequences. Following rescue of the transforming DNA and flanking sequences, we sequenced each end by using oligonucleotide primers; although we analyzed about 200 to 250 bp from each junction, only 5 to 8 bp contiguous with the YIplac211 fragment are shown. Junction 1 and 2 sequences are contiguous with ends 1 and 2, respectively, of YIplac211. The double-stranded depiction of the junction represents sequences that are unambiguously derived from the target; the single-stranded portion represents sequences derived from the plasmid or (for end-directed events) derived from the plasmid and the target (as shown in Fig. 2). Above and below the target sequences, we show the inferred sequence of the 5' protruding ends derived from the plasmid; numbers adjacent to the end sequences indicate end 1 or end 2 of YIplac211. For some transformants, analysis of the junction sequences indicates that one or more bases were lost from the end of the transforming DNA. For example, in transformant WT1-7, one base was lost from end 1, but no bases were lost from end 2. The underlined regions in the target and plasmid ends represent homology between the plasmid ends and the target. The vertical lines drawn between bases in the target sequences represent the sites at which the targets were cleaved to allow integration of the transforming DNA. For example, the analysis of the junction sequences indicates that the terminal A base of the end 1 plasmid sequences was ligated to a C in the target. The boxed regions in the target represent CTT or GTT (consensus topoisomerase I binding sites) adjacent to the cleavage sites; the boxed base in junction 1 of WT2-28 represents an inferred mismatch repair event (see text). Classification as BED, SED, or NED indicates the number of ends that share homology to the target.

the sequence of the junctions. The assumptions involved in this analysis are the following: (i) if the same junction can be formed by deletions of bases either from the target or from the plasmid, bases are deleted from the ends of the plasmid, and (ii) base pairing between the target and the plasmid ends is relevant to the integration event if two or more contiguous bases (or three of four noncontiguous bases) are homologous

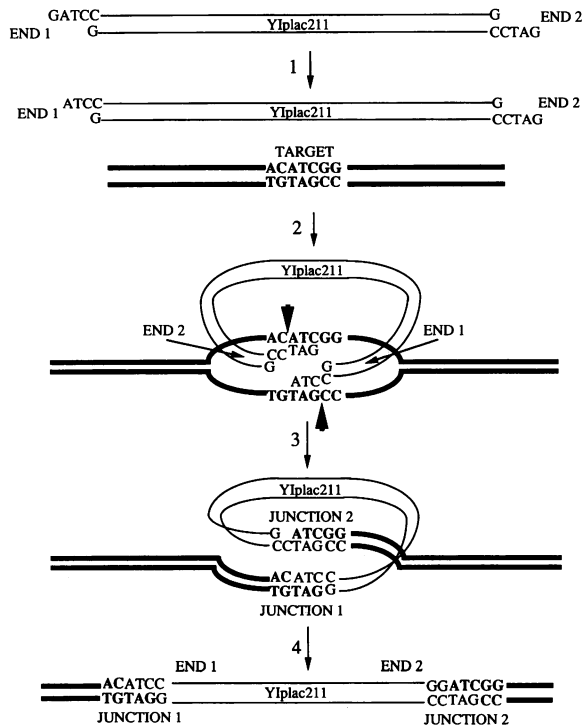


FIG. 2. Proposed mechanism of integration of transforming DNA when both ends share homology with the target. The target shown is that observed for WT1-7 (Fig. 1). One of the bases is lost from end 1. We suggest that both ends of the plasmid (indicated by thin lines) invade the homologous target site (indicated by thick lines), and base pairs are formed. In this example, only three of the four bases of the plasmid termini are paired with the chromosome. The DNA strands of the chromosome are then nicked (as indicated by thick arrows), and the plasmid DNA sequences are ligated to the chromosomal sequences, integrating the plasmid. At junction 2, a base (G) is inserted as the result of repair synthesis.

between the plasmid end and the target at the position of the nick in the target.

Of the seven integration events shown in Fig. 1, five exhibit homology between the ends of the plasmid and the target for either one (SED [single end directed]) or both (BED [both ends directed]) ends. This result confirms our previous conclusion (based on integration of the *URA3* fragment) that homology between the terminal base pairs of the integrating fragment and the target is important in many illegitimate integration events in *S. cerevisiae* (34). Unlike our previous observations, however, most end-directed events did not involve all four terminal base pairs of the plasmid. In our previous study, three of three integrants of the *Bam*HI *URA3* fragment were insertions into GATC targets. We do not know whether the lack of such insertions for YIplac211 represents an effect of the size of the plasmid, the sequences adjacent to the ends, or some other factor. It is also unclear whether the pairing of the ends of the plasmid to the target is important in determining the site of strand cleavage as shown in Fig. 2. An alternative explanation is that the DNA lesion occurs spontaneously at the target site, but the integration is promoted by pairing between the ends of the plasmid and the broken ends of the chromosome.

Although most of the integration events shown in Fig. 1 require only cut-paste interactions between the plasmid and the chromosome, to explain the junctions of WT2-28, we

need to postulate mismatch repair during the integration event. More specifically, if the A/A mismatch shown in Fig. 1 between end 1 and the target is corrected to insert a T in the end 1 strand, the observed junctions would be obtained. Since yeast cells efficiently correct mismatches in heteroduplexes (30), this assumption seems reasonable.

In two of the transformants (WT2-25 and WT133), the insertions were not end directed (NED). In both of these transformants, there were GTT or CTT sequences to the 5' side of the nick in one (WT133) or both (WT2-25) strands. Since CTT and GTT sequences are preferred sites for cleavage of DNA by topoisomerase I (3, 31), this observation suggested that NED integration events might be a consequence of topoisomerase I cleavage.

For all three BED transformants shown in Fig. 1, the predicted target site cleavage would result in ends with 5' overhangs of the same size as the region base paired between the ends and the target. In the other transformants, cleavages that would result in 5' overhangs, 3' overhangs, or blunt ends were observed. For NED events, the 5' overhangs result in target sequence deletions, and the 3' overhangs result in target site duplications. With a few exceptions to be discussed later, the amount of DNA deleted or duplicated from the targets usually involved only a few base pairs. Similarly, the ends of the plasmid either were intact or lost only a few base pairs.

In addition to the seven transformants shown in Fig. 1, we obtained target information for two other transformants, WT2-10 and WT2-11. The WT2-10 insertion occurred within a  $\delta$  repeated element. Since these repeats are quite diverged in DNA sequence (6), we could not identify a precise target for this insertion by a comparison with GenBank sequences. Southern analysis of DNA isolated from WT2-11 (described below) indicated that this transformant contained an autonomously replicating plasmid rather than an integrated plasmid. Sequence analysis of the WT2-11 junctions indicated that end 1 was contiguous with a region of the mitochondrial DNA located about 180 bp upstream of a *tRNA<sup>Cys</sup>* gene (43). Since the sequences of the end 2 junction were not contiguous with those of end 1, the YIplac211 sequences did not represent a simple insertion of the plasmid into a mitochondrial DNA target. Further analysis indicated that the WT2-11 transformant contained a plasmid consisting of YIplac211 sequences fused to two noncontiguous pieces of mitochondrial DNA (Fig. 3). We suggest that this plasmid reflects the *in vivo* ligation of the transforming DNA with mitochondrial DNA fragments located in the nucleus. The rationale for this conclusion and alternative possibilities will be reviewed in Discussion.

**Analysis of the integration of *Eco*RI-treated YIplac211 into strain RSY12 in the presence and absence of *Eco*RI.** Previously, we showed that when RSY12 cells were transformed with the *Bam*HI *URA3* fragment in the presence of *Bam*HI, the fragment inserted into genomic *Bam*HI sites (34). To test whether this method could be universally employed, we transformed *Eco*RI-treated YIplac211 into RSY12 in the presence of *Eco*RI. Sequences flanking the plasmid insertions were determined for 10 *Ura*<sup>+</sup> transformants, and four targets were obtained (Fig. 4). Since none of these four targets represent *Eco*RI sites, we assume that none of these events are REM events. One explanation of these results is that *Eco*RI, unlike *Bam*HI, does not enter the cells during transformation or, alternatively, does not function after it enters. It is also possible that REM events occur more efficiently with the smaller *URA3* fragment used in our first studies than with plasmid YIplac211.

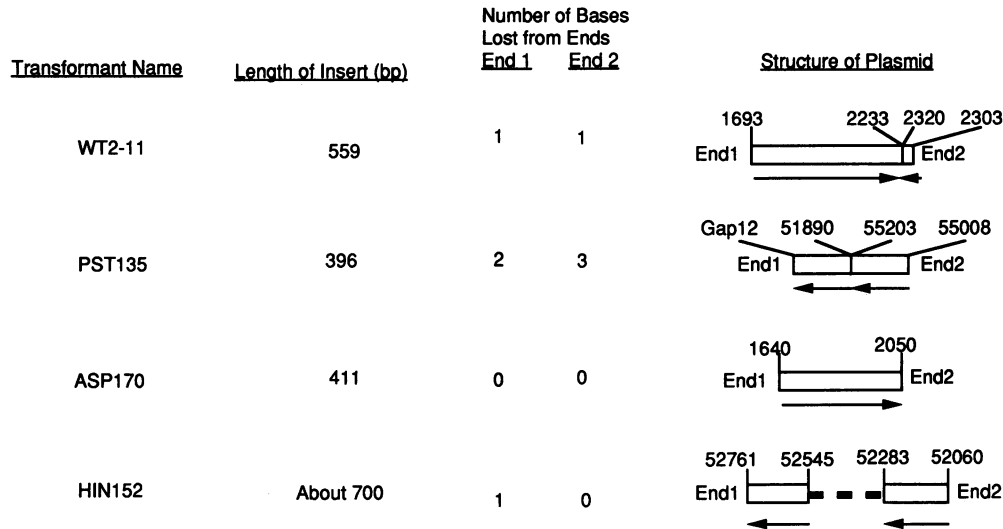


FIG. 3. Structures of plasmids containing mitochondrial DNA sequences. The mitochondrial DNA sequences are indicated by rectangles; more than one rectangle indicates noncontiguous mitochondrial DNA sequences. Arrows indicate the relative orientation of the sequences. Numbers above the rectangles signify the positions of the sequences relative to the mitochondrial map of Zamaroczy and Bernardi (43). The designation Gap12 for PST135 indicates that the junction has one breakpoint in a 336-bp region (located between 51844 and 51845) sequenced by Skelly and Clark-Walker (36) after the mapping study of Zamaroczy and Bernardi (43). The dotted lines shown for HIN152 indicate that the whole 700-bp insertion was not sequenced. The size and structure of this insertion, however, indicate that it is likely to represent a single continuous mitochondrial DNA fragment.

We also examined the integration of *EcoRI*-treated YIplac211 into the genome of RSY12 in the absence of the restriction enzyme (Fig. 4). The types of integration were similar in the presence and absence of *EcoRI*. In addition, similar types of integration events were detected in these experiments as were observed in the experiments involving the *BamHI* enzyme. End-directed events (SED or BED) were more common than events that did not involve shared homology (NED). Small (several-base-pair) deletions and duplications of target sequences were usually observed, although one insertion (ECO<sup>-</sup>197) involved neither loss nor duplication of the target. Three of the integration events (ECO<sup>+</sup>9-1, ECO<sup>-</sup>196, and ECO<sup>-</sup>197) occurred adjacent to a CTT sequence.

One of the integrants (ECO<sup>-</sup>128) represented an insertion of YIplac211 associated with a chromosomal rearrangement. End 2 of the plasmid was contiguous with sequences located 1,272 bp upstream of *YEF3*, the essential gene encoding elongation factor 3 (32), whereas end 1 was contiguous with sequences located 342 bp downstream of the initiating codon. Southern analysis indicated that the insertion of the YIplac211 sequences was associated with a duplication. Two different types of mechanisms could result in a duplication associated with the integration event. First, the two ends could invade different regions of a replication bubble (Fig. 5), as suggested by Roth and Wilson (33). Alternatively, in a mechanism similar to that proposed for repair of gapped plasmids (37), the two ends could invade different positions within a single DNA molecule, priming DNA synthesis and generating the duplication.

**Analysis of transformants obtained by treating YIplac211 with restriction enzymes other than *EcoRI* and *BamHI*: fusions between YIplac211 and mitochondrial DNA.** As described above, following *BamHI* treatment of YIplac211, we obtained a transformant in which the YIplac211 sequences were fused to mitochondrial DNA. To determine whether this fusion required *BamHI* ends, we examined RSY12-

derived transformants obtained by treating cells with YIplac211 DNA digested with another restriction enzyme (*PstI*, *HindIII*, or *Asp* 718). We examined by sequence analysis only those transformants in which the Ura<sup>+</sup> phenotype was unstable, suggesting that the *URA3* gene was plasmid borne. Three transformants (of approximately 80 examined) were obtained in which the YIplac211 sequences were fused to mitochondrial DNA at both junctions (Fig. 3). Each of these transformants is discussed below.

Transformant PST135 (resulting from transformation of RSY12 with *PstI*-treated YIplac211) contained a 396-bp insertion of mitochondrial DNA. End 1 was contiguous with a 200-bp fragment located between position 51890 in the Zamaroczy-Bernardi map (43) and a region in a gap in this map sequenced by Skelly and Clark-Walker (36). This fragment was connected (in direct orientation) with a 196-bp fragment extending from positions 55203 to 55008. This plasmid could have been formed by the ligation of YIplac211 to two small mitochondrial DNA fragments. Alternatively, the plasmid may reflect the fusion of YIplac211 with a single mitochondrial fragment. This fragment may have undergone a deletion prior to its entry into the nucleus or may have had a deletion after its fusion to YIplac211.

Transformant ASP170 was obtained by treating RSY12 with *Asp* 718-digested YIplac211. This plasmid contained a single 411-bp insertion of mitochondrial DNA derived from positions 1640 to 2050 on the Zamaroczy-Bernardi map. This plasmid was presumably formed by the in vivo ligation of the linear YIplac211 sequences to a single linear mitochondrial DNA fragment.

The last transformant (HIN152) was obtained by treating YIplac211 with *HindIII*. In this plasmid, the 200 bp contiguous with end 1 have good (88%) but not perfect homology to bases 52761 to 52545 on the Zamaroczy-Bernardi map. Similarly, the 230 bp contiguous with end 2 have 90% homology to bases 52060 to 52283. Since the size of the mitochondrial DNA insertion is about 700 bp (as determined

Name of Transformant (Class)	End 1		End 2		Gene, Genbank Accession Number
	Junction 1	Junction 2	Target	Target	
ECO <sup>+</sup> 9-1 (SED)	GGCTT CCGAATTAA	ATCTCAGCA AGAGTCGT	GGCTT ATCTCAGCA	ATCTCAGCA AGAGTCGT	rDNA J01355
ECO <sup>+</sup> 11-1 (NED)	TTGAG AACTCTTAA	AAGATACAA CTATGTT	TTGAG AAGATACAA	AAGATACAA CTATGTT	rDNA M35588
ECO <sup>-</sup> 11-2 (BED)	AAACT TTGATTAA	AATGTTAAA CAATTT	AAACT AATGTTAAA	AATGTTAAA CAATTT	ARS M37752
ECO <sup>+</sup> 16-1 (SED)	TGGGG ACCCCTTAA	AATGCCTTG ACGGAAC	TGGGG AATGCCTTG	AATGCCTTG ACGGAAC	rDNA M35588
ECO <sup>+</sup> 25-1 (BED)	ACAAAC TGTTTGTA	AACATTGCT GTAACGA	ACAAAC AACATTGCT	AACATTGCT GTAACGA	2 $\mu$ circle V01322
ECO <sup>-</sup> 128 (SED)	GGCGG CCGCCCTAA	AATTCAAAG GTTTC	GGCGG AATTCAAAG	AATTCAAAG GTTTC	YEF3 J0515
ECO <sup>-</sup> 196 (NED)	TACGT ATGCATTAA	AATTAAGGC TTCCG	TACGT AATTAAGGC	AATTAAGGC TTCCG	SHE30 gene X07673
ECO <sup>-</sup> 197 (SED)	GTCTT CAGAATTAA	AATTCATCT GTAGA	GTCTT AATTCATCT	AATTCATCT GTAGA	HOM2 gene X15649

FIG. 4. Flanking sequences and targets of *EcoRI*-treated YIplac211 DNA transformed into RSY12. The plasmid was transformed in the presence (ECO<sup>+</sup>) or absence (ECO<sup>-</sup>) of *EcoRI*. Since the transformants obtained in the presence of the enzyme did not have insertions in *EcoRI* sites, we assume that these events were not enzyme mediated. Notation is as in Fig. 1.

by gel analysis), although the entire insertion has not been sequenced, the structure of the plasmid is consistent with an in vivo ligation between YIplac211 and a linear mitochondrial DNA fragment containing the sequences located between positions 52060 and 52761. The difference between the mitochondrial DNA sequences in this plasmid and the other plasmids described above is the amount of sequence divergence. In the other plasmids, the homology levels are 98% or more, rather than the 88 to 90% observed in HIN152. The observed differences include single-base-pair changes as well as small deletions and additions. It is possible that these changes represent differences in sequence between the wild-type mitochondrial DNA in RSY12 compared with the wild-type mitochondrial DNA used in the Zamaroczy-Bernardi map. Alternatively, these changes may reflect mutational alterations that occurred in *S. cerevisiae*, perhaps associated with petite formation.

It should be emphasized that the structures of these plasmids (Fig. 3) are those expected from the ligation of linear DNA fragments rather than the structure expected if the YIplac211 sequences had inserted into a circular mitochondrial DNA fragment located in the nucleus. For an insertion event, one would expect the DNA sequences flanking the insertion to be a continuous region of the mitochondrial genome.

**Other physical and genetic characterizations of the transformants.** For 15 of the transformants with known targets, we performed Southern analysis and examined the stability of the Ura<sup>+</sup> phenotype (as described in Materials and Methods). By DNA sequence analysis, three transformants

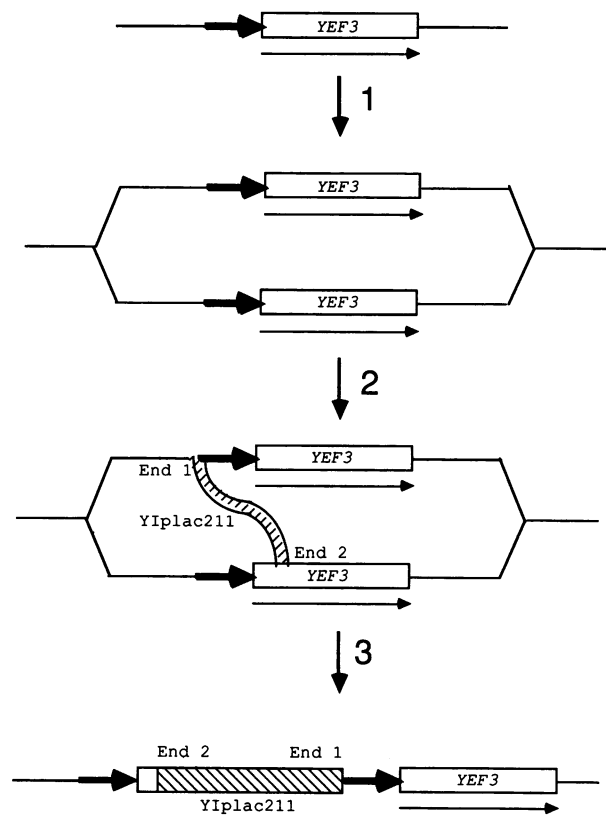


FIG. 5. Integration of insertion associated with a duplication (transformant ECO<sup>-</sup>128). This model is similar to that proposed by Roth and Wilson (33). The top portion represents the *YEF3* chromosomal sequences before the insertion. The thick horizontal arrow represents upstream sequences extending from -1272 to +1; the thin horizontal arrow represents the direction of transcription. In step 1, the chromosomal DNA is shown in the process of DNA replication. In step 2, unequal nonhomologous recombination occurs between the ends of YIplac211 (shown as a thick cross-hatched line) and different sides of the replication bubble. End 1 invades at position -1272, and end 2 invades at position +342. In step 3, one of the expected products of the exchange is a chromosome with an insertion of YIplac211 flanked by duplicated chromosomal sequences. A cell containing the reciprocal product containing the deletion (not shown) would probably not be viable, since *YEF3* is an essential gene (32).

(WT2-15, WT2-16, and ECO<sup>-</sup>11-2) had insertions in or near single-copy nuclear genes, five (WT2-25, WT2-28, ECO<sup>+</sup>9-1, ECO<sup>+</sup>11-1, and ECO<sup>+</sup>16-1) had insertions in rDNA, two (WT1-7 and ECO<sup>+</sup>25-1) had insertions in the 2 $\mu$  plasmid, and four (WT2-11, PST135, ASP170, and HIN152) had associations with mitochondrial DNA. By Southern analysis, all strains with genomic insertions in either single-copy or repeated sequences had a single band of hybridization with the intensity expected for a single-copy insertion, and those with insertions in the 2 $\mu$  plasmid or associated with mitochondrial DNA had bands of hybridization with the intensity expected for high-copy-number plasmids (data not shown). As expected, transformants with insertions into single-copy genomic sequences had a stable Ura<sup>+</sup> phenotype (0 to 0.5% Ura<sup>-</sup> cells after eight generations of nonselective growth), whereas transformants associated with plasmid were quite unstable (3 to 20% Ura<sup>-</sup> cells). Most transformants with insertions in rDNA had intermediate levels of

instability (0.5 to 3% Ura<sup>-</sup> cells); the loss of the Ura<sup>+</sup> phenotype in these transformants presumably reflects unequal sister strand recombination within the rDNA (38).

## DISCUSSION

In both prokaryotes and eukaryotes, recombination events that do not involve extensive DNA sequence homology (illegitimate recombination events) have been observed (16, 27). Such events result in chromosomal deletions, duplications, and translocations. Some of these alterations are associated with very small (<10-bp) regions of sequence homology between the interacting DNA molecules, whereas other alterations appear to be completely homology independent. There appear to be several mechanisms involved in this type of recombination, including the action of break-and-join enzymes (for example, topoisomerases and site-specific recombinases), DNA polymerase slippage, and random DNA breakage (16).

In many types of eukaryotes, transforming DNA is integrated by illegitimate recombination (illegitimate integration). Such events have been reported in *Schizosaccharomyces pombe* (42), *Neurospora crassa* (2, 11), *Coprinus cinereus* (4), and a number of other fungi. In mammalian cells, illegitimate integration events represent the predominant form of integration (33). To our knowledge, there have been only four integrants into the mammalian genome in which the both junctions of the insert with genomic DNA and the target have been sequenced. In three of these integrants, no gross rearrangements of target sequences were observed; two integrants were associated with small (9- and 12-bp) target deletions (28), and no alteration in the target was observed in the third (18). The fourth integrant was associated with a 5-kb duplication (41). In most of these integrants, small numbers of bases were deleted from the ends of the transforming DNA. In one integrant, there was 6-bp homology between one end of the transforming DNA and the target (28). In studies of end-joining reactions of linear simian virus 40 genomes transformed into mammalian cells, small amounts (1 to 5 bp) of homology are often observed at the junctions of the rejoined molecules (33).

We can make several generalizations concerning illegitimate integration in *S. cerevisiae*. First, most (11 of 15) integrants involve base pairing between one or both of the ends of the transforming DNA with the chromosomal target. Second, in the class of integrant that does not involve base pair interactions between the target and the plasmid, there appears to be a preference for insertion next to CTT or GTT in the target. Third, little DNA (0 to 4 bp) is lost from the ends of the transforming molecules, and no DNA is lost from the double-stranded portion of the vector. Fourth, most integrants are not associated with large deletions or duplications. Although small (1- to 4-bp) deletions and duplications are common, only one integrant with larger changes (1,600-bp duplication) was observed.

As described above, four transformants (WT2-25, WT133, TOP2-10, ECO<sup>+</sup>11-1, and ECO<sup>-</sup>196) were observed in which neither end had homology to the target. Of these four transformants, three had a GTT or CTT sequence adjacent to one (WT133 and ECO<sup>-</sup>196) or both (WT2-25) junctions. Thus, for these transformants, four of the eight junctions were next to a CTT or GTT. The probability that a single junction will have a CTT or GTT is about  $2 \times (1/4)^3$ , or 1/32. Thus, the chance of one or more junctions with a CTT or GTT in a sample of eight is  $1 - (31/32)^8$ , or 0.22. Using the binomial expansion, we calculate that the probability of two

or more such junctions is 0.02. Thus, the observation of four such junctions is highly statistically significant.

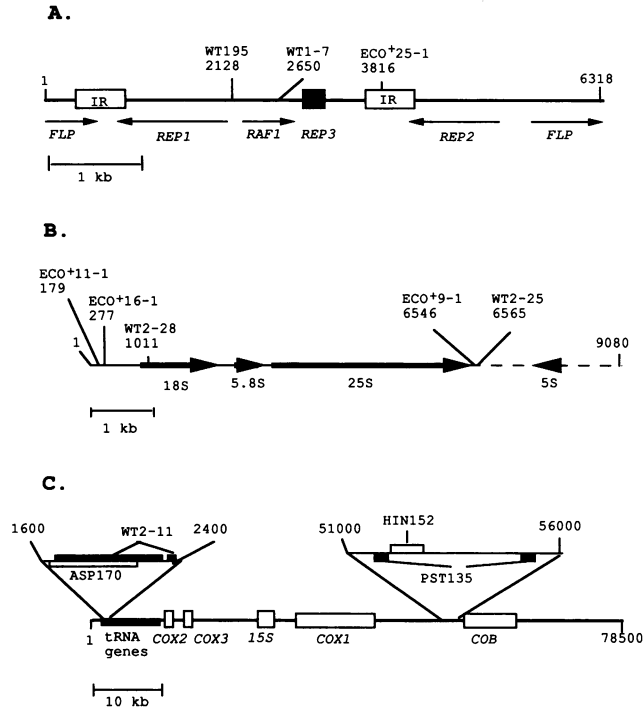
The biological significance of this observation is that CTT and GTT are the preferred sites for cleavage by topoisomerase I (3, 31). Although the cleavage specificity of yeast topoisomerase I has not been determined, since the cleavage specificity of the enzyme isolated from diverse sources is very similar, we assume that the yeast enzyme shares the same specificity. We suggest, therefore, that although most of the illegitimate integration events are the result of base pairing between the terminal sequences of the transforming DNA and the target, the remaining integration events may be catalyzed by topoisomerase I cleavage of the target. The integration event could be the result of a topoisomerase I-catalyzed break, followed by ligation of the transforming DNA to the broken chromosome. For example, a single-stranded break caused by topoisomerase I action near a nick on the opposite strand would cause a double-stranded break that could be ligated to the transforming DNA. Double-strand breaks caused by this type of mechanism could represent one class of lesion for the initiation of spontaneous mitotic recombination in yeast. Alternatively, the topoisomerase I enzyme may directly ligate the 3' T of the GTT or CTT to the 5' end of the transforming DNA. One of the NED transformants did not have a GTT or CTT at the junction. These transformants may represent either events catalyzed by another topoisomerase (such as the *HPR1* [1] or *TOP3* [40] enzyme) or events reflecting another mechanism of breakage.

It has been previously suggested that topoisomerase I may be involved in catalyzing illegitimate recombination events in both mammalian cells (12) and bacteria (16). In mammalian cells, GTT or CTT sequences have been observed at the junctions of insertions of transforming DNA (41) and at the junctions formed by excision of simian virus 40 from the chromosome (10). In addition, in one of two nonhomologous integration events examined in *N. crassa*, there was a CTT at the junction (2).

Our studies cannot rule out the possibility that in both yeast and mammalian cells, there is a preference for insertion next to GTT or CTT for reasons other than the action of topoisomerase I. In preliminary studies, we have examined insertions of transforming DNA in a *top1* strain (data not shown). Of 10 integrants obtained, only one NED transformant was observed; this transformant did not represent an insertion next to GTT or CTT. Convincing evidence for (or against) the involvement of topoisomerase I in NED integration events would require the analysis of many more NED events from both *TOP1* and *top1* strains.

Of the 15 insertions associated with a previously characterized gene, 8 were inserted into repeated DNA: 5 into rDNA and 3 into the 2 $\mu$ m plasmid. Since 44 transformants were sequenced, 11% of the insertions were in the rDNA and 7% were in the 2 $\mu$ m plasmid. Since 7 to 14% of the genome is rDNA (29) and about 2 to 4% of the genome is 2 $\mu$ m plasmid (8), there does not appear to be any striking preference for insertion into either of these targets. As shown in Fig. 6A, the position of the insertions within the 2 $\mu$ m plasmid appears random. For the insertion in the rDNA (Fig. 6B), the insertions appear more clustered, although no two insertions were identical. There were also insertions into or nearby a number of single-copy genes, including *MET25*, *APAI*, *YEF3*, *SNR30*, and *HOM2*. These procedures, therefore, should be a useful method of insertional mutagenesis.

Approximately 10% of the transformants observed in our experiments contained plasmids with mitochondrial inser-



**FIG. 6.** Target sites for insertions in the 2µm plasmid, rRNA genes, and mitochondrial DNA. (A) Insertions in the 2µm plasmid. Arrows indicate the positions and orientations of various transcripts, the dark rectangle represents the position of the structural element *REP3*, and the light rectangles represent the locations of inverted repeats (IR) (9). The number under the name of each transformant indicates the position of the insertion within form A of the plasmid; position 1 represents the *EcoRI* site within the small unique region. (B) Insertions within the rRNA gene. Five such insertions were mapped. Arrows indicate the direction of transcription. Thick lines represent the sequences encoding the mature transcripts, thin lines represent transcribed precursor RNA, and the dashed line indicates untranscribed spacer. The number below the name of each transformant indicates the position within the rRNA gene with the +1 position representing the first base in the 37S precursor rRNA. (C) Regions of the yeast mitochondrial genome that were ligated in vivo to YIplac211. Plasmids derived from four transformants contained mitochondrial sequences. In the lower portion, we indicate various coding regions in mitochondrial DNA. In the upper portion, we show the locations of the fragments within the genome; the numbers are those used in the map of Zamaroczy and Bernardi (43). For transformants ASP170 and HIN152, single fragments (indicated by white rectangles) were ligated to YIplac211. For transformants WT2-11 and PST135, two noncontiguous fragments (indicated by black rectangles) were ligated to YIplac211.

tions. In these plasmids, either one piece or several noncontiguous pieces of mitochondrial DNA were connected to the YIplac211 sequences. To be maintained as a plasmid in *S. cerevisiae*, at least one fragment must have autonomously replicating sequence (ARS) activity. This condition is not particularly restrictive, since many mitochondrial DNA fragments have ARS activity (5, 13, 22); it should be noted, however, that there appears to be some nonrandomness in the recovery of particular portions of the mitochondrial chromosome (Fig. 6C). We believe that these plasmids are likely to be localized in the nucleus, rather than the mitochondria, because mitochondrial *URA3* genes do not lead to a *Ura*<sup>+</sup> phenotype (39). Previously, Thorsness and Fox (39) showed that a plasmid containing *URA3* sequences associ-

ated with mitochondrial DNA escaped from the mitochondria to the nucleus at a rate of about  $2 \times 10^{-5}$  per cell per generation. We suggest that the plasmids observed in our experiments represent the in vivo ligation of YIplac211 sequences to one or more fragments of escaped mitochondrial DNA. The deletions and inversions observed in several of the plasmids may reflect the random ligation of several mitochondrial fragments. A second possibility is that the mitochondrial DNA in the plasmid was derived from a petite genome in which a rearrangement had occurred. The rearranged fragment was then transported into the nucleus and ligated to YIplac211. It has been shown that petite genomes frequently contain deletions and inversions, although the mechanisms involved in these rearrangements are not yet understood (15).

This relatively high fraction of transformants with mitochondrial DNA observed in our experiments suggests that the frequency of transfer of mitochondrial DNA into the nucleus may be even higher than that observed by Thorsness and Fox (39). The efficiency of nonhomologous transformation is about 1% of that observed for homologous transformation with linear DNA fragments (34). If the efficiency of ligation of plasmid sequences to mitochondrial DNA is 100%, then the frequency of nonhomologous events involving mitochondrial DNA relative to the frequency of homologous transformation is about  $10^{-3}$  to  $10^{-4}$ . If we assume that the fraction of cells that contain nuclear-located mitochondrial DNA is not different in untransformed cells, approximately 1 in 1,000 to 1 in 10,000 cells may contain this species of DNA. The difference between this frequency and that observed by Thorsness and Fox ( $2 \times 10^{-5}$ ) may be explained by the observation that most of the mitochondrial DNA fragments are small. In the experiment by Thorsness and Fox, the detection of nucleus-localized mitochondrial DNA required the entry of an intact *URA3* gene. Alternatively, the transformation procedure that we are using may facilitate the entry of mitochondrial DNA into the nucleus.

The two observations of mitochondrial DNA in the nucleus and end-directed integration of transforming fragments suggest that entry of mitochondrial DNA into the nuclear genome may be an important feature of genome evolution in *S. cerevisiae*. Two such events have been clearly documented. In one strain, a 289-bp mitochondrial DNA fragment containing the *var1* gene and several noncontiguous other portions of the mitochondrial genome was inserted within the nuclear genome (17). In another strain, a 227-bp region derived from the fourth intron of cytochrome *b* was inserted into telomeric sequences (25). If such insertion events were very frequent, one might expect some spontaneous mutations of nuclear genes to be a consequence of insertions of mitochondrial DNA. To our knowledge, no such mutations have been reported. However, since most nonhomologous integration events appear end directed, one would expect most mitochondrial insertions would be within the AT-rich DNA that separates coding sequences rather than into the coding sequences themselves. It is conceivable that the AT-rich spacers may have originated as a consequence of the insertion of mitochondrial DNA. In addition, since the mitochondrial DNA is rich in ARS elements, the nonhomologous insertion of mitochondrial DNA may represent a mechanism for the introduction of DNA replication origins into the yeast chromosome.

In summary, in studies of illegitimate integration events in *S. cerevisiae*, we have obtained three types of transformants. The most common type involves the insertion of the transforming DNA into the target as the result of an inter-

action between a small number of bases on the ends of the transforming DNA with bases within the target. The second class of transformant involves an insertion of transforming DNA next to a CTT or GTT sequence in the target; we suggest that these events are mediated by topoisomerase I. The third class of transformant involves the *in vivo* ligation of the transforming DNA with nucleus-localized linear mitochondrial DNA fragments.

#### ACKNOWLEDGMENTS

We thank Nick Carls and Arturo Saavedra for excellent technical assistance, J. Huberman for providing a copy of the rDNA sequence, and M. Goebel for help with DNA sequence analysis. We thank members of the Petes and Schiestl laboratories for comments on the manuscript.

This research was supported by NIH grant GM47256 to T.D.P. and by funds from the Department of Molecular and Cellular Toxicology and ACS grant CN-83 to R.H.S.

#### REFERENCES

- Aguilera, A., and H. L. Klein. 1990. *HPRI*, a novel yeast gene that prevents intrachromosomal excision recombination shows carboxy-terminal homology to the *Saccharomyces cerevisiae TOPI* gene. *Mol. Cell. Biol.* **10**:1439-1451.
- Asch, D. K., G. Frederick, J. N. Kinsey, and D. D. Perkins. 1992. Analysis of junction sequences resulting from integration at nonhomologous loci in *Neurospora crassa*. *Genetics* **130**:737-748.
- Been, M. D., R. R. Burgess, and J. J. Champoux. 1984. Nucleotide sequence preference at rat liver and wheat germ type I DNA topoisomerase breakage sites in duplex SV40 DNA. *Nucleic Acids Res.* **12**:3097-3114.
- Binninger, D. M., C. Skryznia, P. Pukkila, and L. A. Casselton. 1987. DNA-mediated transformation of the basidiomycete *Coprinus cinereus*. *EMBO J.* **6**:835-840.
- Blanc, H. 1984. Two modules from the hypersuppressive rho<sup>-</sup> mitochondrial DNA are required for plasmid replication in yeast. *Gene* **30**:47-61.
- Boeke, J. D. 1989. Transposable elements in *Saccharomyces cerevisiae*, p. 335-374. *In* D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- Boeke, J. D., F. Lacroute, and G. R. Fink. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.* **197**:345-346.
- Broach, J. R. 1981. The yeast plasmid 2 $\mu$  circle, p. 445-470. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), *The molecular biology of the yeast Saccharomyces*, vol. I. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Broach, J. R., and F. C. Volkert. 1991. Circular DNA plasmids of yeasts, p. 297-331. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), *The molecular biology of the yeast Saccharomyces*, vol. I. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Bullock, P., J. J. Champoux, and M. Botchan. 1985. Association of crossover points with topoisomerase I cleavage sites: a model for nonhomologous recombination. *Science* **230**:954-958.
- Case, M. E. 1986. Genetical and molecular analyses of QA-2 transformants in *Neurospora crassa*. *Genetics* **113**:569-587.
- Champoux, J. J., and P. A. Bullock. 1988. Possible role for the eucaryotic type I topoisomerase in illegitimate recombination, p. 655-666. *In* R. Kucherlapati and G. R. Smith (ed.), *Genetic recombination*. American Society for Microbiology, Washington, D.C.
- Delouya, D., and F. G. Nobrega. 1991. Mapping of the ARS-like activity and transcription initiation sites in the non-canonical yeast mitochondrial *ori6* region. *Yeast* **7**:51-60.
- Denis, C. L., and E. T. Young. 1983. Isolation and characterization of the positive regulatory gene *ADR1* from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **3**:360-370.
- Dujon, B. 1981. Mitochondrial genetics and functions, p. 505-635. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), *The molecular biology of the yeast Saccharomyces*, vol. I. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Ehrlich, S. D. 1989. Illegitimate recombination in bacteria, p. 799-832. *In* D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- Farrelly, F., and R. A. Butow. 1983. Rearranged mitochondrial genes in the yeast nuclear genome. *Nature (London)* **301**:296-301.
- Gahlmann, R., and W. Doerfler. 1983. Integration of viral DNA into the genome of the adenovirus type 2-transformed hamster cell line HE5 without loss or alteration of cellular nucleotides. *Nucleic Acids Res.* **11**:7347-7361.
- Gietz, R. D., and A. Sugino. 1988. New yeast-*Escherichia coli* shuttle vectors constructed with *in vitro*-mutagenized yeast genes lacking six-base-pair restriction sites. *Gene* **74**:527-534.
- Hinnen, A., J. B. Hicks, and G. R. Fink. 1978. Transformation of yeast. *Proc. Natl. Acad. Sci. USA* **75**:1929-1933.
- Holmes, D. S., and M. Quigley. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**:193-197.
- Hyman, B. C., J. H. Cramer, and R. H. Rownd. 1982. Properties of a *Saccharomyces cerevisiae* mtDNA segment conferring high-frequency yeast transformation. *Proc. Natl. Acad. Sci. USA* **79**:1578-1582.
- Kraft, R., J. Tardiff, K. S. Krauter, and L. A. Leinwand. 1988. Using mini-prep plasmid DNA for sequencing double stranded templates with Sequenase. *BioTechniques* **6**:544-545.
- Kuspa, A., and W. F. Loomis. 1992. Tagging developmental genes in *Dictyostelium* by restriction enzyme-mediated integration of plasmid DNA. *Proc. Natl. Acad. Sci. USA* **89**:8803-8807.
- Louis, E. J., and J. E. Haber. 1991. Evolutionarily recent transfer of a group I mitochondrial intron to telomere regions in *Saccharomyces cerevisiae*. *Curr. Genet.* **20**:411-415.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Meuth, M. M. 1989. Illegitimate recombination in mammalian cells, p. 833-860. *In* D. E. Berg and M. M. Howe (ed.), *Mobile DNA*. American Society for Microbiology, Washington, D.C.
- Murnane, J. P., M. J. Yezzi, and B. R. Young. 1990. Recombination events during integration of transfected DNA into normal human cells. *Nucleic Acids Res.* **18**:2733-2738.
- Olson, M. V. 1991. Genome structure and organization in *Saccharomyces cerevisiae*, p. 1-39. *In* J. R. Broach, E. W. Jones, and J. R. Pringle (ed.), *The molecular and cellular biology of the yeast Saccharomyces*, vol. I. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Petes, T. D., R. E. Malone, and L. S. Symington. 1991. Recombination in yeast, p. 407-521. *In* J. R. Broach, E. W. Jones, and J. R. Pringle (ed.), *The molecular and cellular biology of the yeast Saccharomyces*, vol. I. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Porter, S. E., and J. J. Champoux. 1989. Mapping *in vivo* topoisomerase I sites on simian virus 40 DNA: asymmetric distribution of sites on replicating molecules. *Mol. Cell. Biol.* **9**:541-550.
- Qin, S., A. Xie, M. Bonato, and C. McLaughlin. 1990. Sequence analysis of the translation elongation factor 3 from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **265**:1903-1912.
- Roth, D., and J. Wilson. 1988. Illegitimate recombination in mammalian cells, p. 621-653. *In* R. Kucherlapati and G. R. Smith (ed.), *Genetic recombination*. American Society for Microbiology, Washington, D.C.
- Schiestl, R. H., and T. D. Petes. 1991. Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **88**:7585-7589.
- Sherman, F., G. R. Fink, and J. B. Hicks. 1986. *Methods in yeast genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Skelly, P. J., and G. D. Clark-Walker. 1990. Conversion at large



- intergenic regions of mitochondrial DNA in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **10**:1530–1537.
37. Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl. 1983. The double strand break repair model for recombination. *Cell* **33**:25–35.
  38. Szostak, J. W., and R. Wu. 1980. Unequal crossing-over in the ribosomal DNA of *Saccharomyces cerevisiae*. *Nature (London)* **284**:426–430.
  39. Thorsness, P. E., and T. D. Fox. 1990. Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae*. *Nature (London)* **346**:376–379.
  40. Wallis, J. W., G. Chrebet, G. Brodsky, M. Rolfe, and R. Rothstein. 1989. A hyper-recombination mutation in *S. cerevisiae* identifies a novel eukaryotic topoisomerase. *Cell* **58**:409–419.
  41. Wilkie, T. M., and R. D. Palmiter. 1987. Analysis of the integrant in MyK-103 transgenic mice in which males fail to transmit the integrant. *Mol. Cell. Biol.* **7**:1646–1655.
  42. Wright, A. P. H., K. Moundrell, and S. Shall. 1986. Transformation in *Schizosaccharomyces pombe* by non-homologous, unstable integration of plasmids in the genome. *Curr. Genet.* **10**:107–113.
  43. Zamaroczy, M., and G. Bernardi. 1986. The primary structure of the mitochondrial genome of *Saccharomyces cerevisiae*—a review. *Gene* **47**:155–177.