Instability of Simple Sequence DNA in *Saccharomyces cerevisiae*

SAMUEL T. HENDERSON[1] AND THOMAS D. PETES[2]*

Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, Illinois 60637,[1] and Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280[2]

All eukaryotic genomes thus far examined contain simple sequence repeats. A particularly common simple sequence in many organisms (including humans) consists of tracts of alternating GT residues on one strand. Allelic poly(GT) tracts are often of different lengths in different individuals, indicating that they are likely to be unstable. We examined the instability of poly(GT) and poly(G) tracts in the yeast *Saccharomyces cerevisiae*. We found that these tracts were dramatically unstable, altering length at a minimal rate of 10^{-4} events per division. Most of the changes involved one or two repeat unit additions or deletions, although one alteration involved an interaction with the yeast telomeres.

The eukaryotic genome contains many regions (typically, 25 to 60 bp in length) in which a single base or a pair of bases is repeated multiple times. The most abundant types of these tracts consist of alternating GT, alternating GA, or poly(A) sequences (reviewed in reference 47). The poly(GT) tracts are particularly common, being found in about 10^5 copies per human genome (12); these tracts have the potential for forming regions of Z-DNA (reviewed in reference 33). The yeast *Saccharomyces cerevisiae*, like higher eukaryotes, has poly(GT) tracts, approximately 50 to 100 per genome (50). In addition to "pure" poly(GT) tracts, tracts of poly(G_{1-3}T) several hundred base pairs in length are located at and near the yeast telomeres (39, 49).

It is unclear whether poly(GT) tracts and other simple repetitive sequences have a functional role in the cell or represent selfish DNA. One argument for a functional role is the evidence that poly(GT) tracts are nonrandomly distributed in the *Drosophila* genome, poly(GT) sites being associated with regions of transcriptional activity (24). Other workers have found evidence that these tracts stimulate (3) or repress (27, 38) transcription and stimulate recombination (41, 44, 48). Although these results suggest that poly(GT) tracts may be functionally important, they do not exclude the possibility that these tracts simply represent some byproduct of DNA metabolism in eukaryotic cells.

Regardless of their functional role, simple repetitive sequences represent a potential source of two types of genetic instability. First, since the tracts are repeated many times in the genome, recombination events between different tracts could generate deletions, duplications, inversions, and translocations. Second, since the tracts are internally repetitive, their size could grow or shrink as the result of recombination or DNA polymerase slippage events. In prokaryotes, runs of the same base are hotspots for deletion mutations (8, 43) (reviewed in reference 34). In *Escherichia coli*, plasmid-borne poly(GT) tracts are unstable, altering size at a frequency of about 10^{-4} events per division (11, 14, 22). Although no direct measurement of the stability of poly(GT) tracts has been made in eukaryotes, the lengths of poly(GT) tracts at specific loci are known to be highly polymorphic within a species (15, 51), and these tracts have been used for various DNA fingerprinting studies (46).

Below, we describe two assays for tract stability in *S.*

cerevisiae. We find that the poly(GT) tracts in *S. cerevisiae* are unstable enough to account for the polymorphisms observed in higher eukaryotes.

MATERIALS AND METHODS

Media and growth conditions. Media for yeast growth were prepared as described by Sherman et al. (40). 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) plates for *S. cerevisiae* were prepared as described by Rose and Botstein (35). Medium containing 5-fluoroorotic acid (5-FOA) was used to select for *ura3* mutants (4), and medium with DL-aminoadipate was used to select for strains containing *lys2* mutations (6). Yeast strains were grown at 32°C.

E. coli strains were grown at 37°C in LB broth or in X-Gal medium (26). *E. coli* HB101, DH5α, and DK1 were used for cloning and plasmid rescues and were handled by following standard molecular techniques (20, 26).

Plasmid constructions. (i) pSH31, pSH36, pSH40, and pJA4. These plasmids were derived from plasmid pHT259 (provided by H. Tu, University of Chicago), which has the *LEU2* promoter and the first 12 codons of the yeast *LEU2* protein fused to the eighth codon of the *E. coli* β-galactosidase gene (Fig. 1). We annealed various oligonucleotides containing simple repetitive DNAs into the *Bam*HI site near the beginning of the β-galactosidase gene (Table 1) as follows: (i) 200 ng of each pair of complementary oligonucleotides was suspended in 25 μl of 1 mM Tris-HCl (pH 7.5)–100 mM NaCl; (ii) the mixture was heated to 100°C for 4 min, cooled to 50°C, and incubated 1 h at 50°C; and (iii) the sample was reheated to 75°C for 4 min, cooled to 50°C for 30 min, and then slowly cooled to room temperature. These double-stranded oligonucleotides were treated with T4 polynucleotide kinase (26). For construction of the plasmids pSH31, pSH36, and pJA4, pHT259 was digested with *Bam*HI, the oligonucleotides were added with a molar excess of 100:1, and the mixture was treated at 12°C with T4 DNA ligase. *E. coli* transformations were done with strain DH5α. Since the *LEU2* promoter functions in *S. cerevisiae* and in *E. coli*, transformants were screened for the loss of β-galactosidase activity on X-Gal plates, and the junctions were then confirmed by DNA sequencing.

The oligonucleotides used to construct plasmid pSH40 had cohesive ends compatible with *Xho*I, not *Bam*HI. Consequently, we inserted a 10-bp *Xho*I linker (New England Biolabs) into a filled-in *Bam*HI site in pHT259, creating

---

* Corresponding author.

FIG. 1. Plasmids used to assay frameshifting in tracts of simple repetitive DNA sequences. Frameshifting is monitored by examining either β-galactosidase expression (pHT259) or by analyzing *URA3* expression (pSH44). Plasmids pSH31, pSH36, pJA4, and pSH40 were derived from pHT259 by insertions of oligonucleotides (described in Table 1) into the *Bam*HI site. Plasmid pSH44 was derived from pHT259 by deleting two *Pvu*II fragments containing most of the β-galactosidase gene and inserting a fusion containing a small portion of *HIS4* and most of *URA3*.

plasmid pSH35. This plasmid was the immediate progenitor of pSH40.

(ii) **pSH37.** Plasmid pSH31.1 is a derivative of pSH31 generated by an in vivo frameshift within the poly(GT) tract; the resulting tract is 27 bp in length. The plasmid contains a single *Bam*HI site at one end of the tract. Plasmid pSH37 was made by filling in this *Bam*HI site and religating the blunt ends. The net result of these manipulations is a +1 frameshift in the β-galactosidase gene.

(iii) **pSH44.** Plasmid pSH31.4 is a derivative of pSH31 that has a 33-bp poly(GT) tract as the result of an in vivo frameshift. This plasmid was digested with *Pvu*II, deleting

TABLE 1. Insertions used in the construction of various plasmids

Plasmid	Sequence of insertion	Total no. of base pairs of insertion; length of simple sequence	Reading frame	Fusion marker
pSH31	GATCGGTCGACATGTGTGTGTGTGTGTGTGTGTGTGTGTACTCGAG CAGCTGTACACAGACACACACACACACACACACACACATGAGCTCCTAG	47; 29	−1	β-Gal[a]
pSH36	GATCCTCGAGTACACACACACACACACACACACACACACATGTCGAC GAGCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTACAGCTGCTAG	47; 29	−1	β-Gal
pJA4	GATCGAGCTCGGGGGGGGGGGGGGGGGGGGGCTCGAG CTCGAGCCCCCCCCCCCCCCCCCCCGAGCTCCTAG	37; 20	+1	β-Gal
pSH40	GATCCCCTCGACTGCAGAGTGGGTGTGGTGTGGGTGTGTGGGTGTGTGTGGGTGTGTGGGAAGCTGTCGAGGG GGGAGCTGACGTCTCACCCACCACACCCACCACACACCCACCACACCCTTCGACAGCTCCCCTAG	82; 51	+1	β-Gal
pSH44	GATCGGTCGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTACTCGAG CAGCTGTACACAGACACACACACACACACACACACACACACACATGAGCTCCTAG	51; 33	In frame	*URA3*
pSH60	GATCGGTCGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTACTCGAG CAGCTGTACACAGACACACACACACACACACACACACACACACATGAGCTCCTAG	51; 33	In frame	β-Gal

[a] β-Gal, β-galactosidase.

most of the *lacZ* gene but leaving the −40 primer site (Fig. 1). Into this site was inserted a 1.5-kb *Hin*dIII fragment (cohesive ends filled in with Klenow fragment) derived from pNKY48 (2) that contains a wild-type *URA3* gene fused to a small region of the *HIS4* gene. Thus, plasmid pSH44 encodes an unusual, but functional, *URA3* gene product.

(iv) **pSH60.** Plasmid pSH31.4 (described above) was treated with *Eco*RI and *Hin*dIII to generate a 3.0-kb fragment containing a *LEU2-lacZ* fusion with an in-frame 33-bp poly(GT) sequence. This fragment was gel purified, treated with the Klenow fragment of DNA polymerase to fill in the cohesive ends, and ligated into the filled-in *Xho*I site of plasmid pDP6 (9). Plasmid pDP6 contains the *LYS2* gene in a pUC vector, and the *Xho*I site is within the *LYS2* gene.

Yeast strain constructions. (i) SH31, SH36, SH40, SH44, and JA4. Strain PD3 (α *trp1 arg4-17 tyr7-1 ade6 ura3 his4-sal*; provided by P. Detloff, University of North Carolina) was transformed with plasmids pSH31, pSH36, pSH40, pSH44, and pJA4 to generate strains SH31, SH36, SH40, SH44, and JA4, respectively. Transformants were selected on minimal plates containing all required amino acids except tryptophan.

(ii) **SH52.** Strain SH52 was derived from strain MW13 (provided by M. White, University of North Carolina), which is isogenic with PD3 except for the genotype at the *HIS4* and *LEU2* loci; MW13 contains *HIS4* and *leu2-bst*. MW13 was transformed with a *Bam*HI fragment containing the *RAD52* gene disrupted by the insertion of *LEU2*; this fragment was derived from the plasmid pSM20 (provided by D. Schild, University of California, Berkeley). The one-step transplacement (37) was confirmed by Southern analysis and sensitivity of the transformants to methylmethane sulfonate.

(iii) **SH70.** The purpose of this strain is to allow monitoring of the rate of frameshifts for a poly(GT) tract inserted into the chromosome at the *LYS2* locus. Plasmid pSH60 was treated with *Nco*I and used to transform PD3. The resulting *Nco*I fragment contains a *lys2* gene with an insertion of the β-galactosidase fusion. Lys⁻ transformants were selected on plates containing α-aminoadipate. The resulting colonies were spotted to a plate containing X-Gal and screened for the appearance of blue colonies. This strain was shown by Southern analysis to contain an insertion at the *LYS2* locus.

(iv) **SH75.** Strain SH75 was derived from SH70 as the result of an in vivo frameshift of the poly(GT) tract at the *LYS2* locus. The tract length in SH75 is 29 bp (determined by sequence analysis of a polymerase chain reaction [PCR] product derived from the chromosome), 4 bp smaller than that in SH70.

Plasmid rescue. Most of the sequence analysis was done on plasmids rescued from *S. cerevisiae* and transformed into *E. coli*. The method used was that described by Hoffman and Winston (13). Plasmids containing the β-galactosidase gene were transformed into strains DH5α and DK1; plasmids containing the *URA3* gene were transformed into HB101.

DNA sequencing. Double-stranded sequencing templates were prepared by the method of Kraft et al. (16). Sequencing was done with a commercially available −40 primer (5′ GTTTTCCCAGTCACGAC) and the US Biochemical Sequenase kit by following the manufacturer's directions.

To determine the size of the chromosomal poly(GT) tracts, we performed asymmetric PCR amplification of a 125-bp DNA fragment containing the tract, using the procedure of Kreitman and Landweber (17). Yeast DNA was isolated by the spheroplast method (40). The primers flanking the poly(GT) tract had *Eco*RI sites on their 5′ ends. One primer (Leu2RI) contained sequences homologous to the 5′

end of the *LEU2* gene (5′ GAATTCCCTAAGAAGATC GTCG), and the second (−40RI) was homologous to the −40-bp primer described above. For the reaction, 1 μg of yeast DNA was used as a template with 150 ng of the Leu2RI primer and 5 ng of the −40 RI primer. The single-stranded product was sequenced by using the Sequenase (US Biochemical) kit and the −40 primer.

Frameshift rate determination. Two different methods were used to calculate the rates of frameshifting in poly(GT) tracts. For strains in which the frequencies of frameshifts were monitored by the production of β-galactosidase, cultures grown under selective conditions were diluted in rich growth medium (yeast extract-peptone-dextrose [YPD]) to a concentration of 100 cells per ml. Aliquots (0.1 ml) of this diluted culture were used to seed the wells of a 96-well microtiter dish, which was incubated at 32°C for 22 to 26 h. The contents of each well were harvested by centrifugation, resuspended in 1 ml of sterile water, and plated on X-Gal plates; these plates lacked leucine and threonine (to derepress the *LEU2* promoter) and tryptophan (to force retention of the plasmid). Plates were incubated for 10 days to allow the complete expression of the β-galactosidase phenotype to develop and then were scored for the appearance of blue colonies. From the proportion of cultures with no blue colonies, we calculated the rate of frameshifts by using the method of Luria and Delbruck (25), as modified by Lea and Coulson (21).

For strains in which the poly(GT) tract was inserted upstream of the *URA3* gene, a different method was used. In this strain, the poly(GT) tract was in frame, resulting in a Ura⁺ phenotype. Frameshifts were detected as derivatives that were Ura⁻ (5-FOAʳ). The strains were grown overnight in medium lacking tryptophan, diluted to 100 cells per ml in YPD medium, and inoculated into the wells of a microtiter dish. Cells were allowed to grow to approximately 10⁵ cells per well. A small aliquot from each well was removed to determine the total number of cells, and the remainder of the culture was plated on solid medium lacking tryptophan, threonine, and leucine but containing 5-FOA. The plates were incubated at 32°C for 1 week, and the number of 5-FOAʳ colonies was counted. Since all of the plates had some 5-FOAʳ colonies, the method of the median was used to derive a rate (21).

Rate measurements obtained by both procedures described above are quite sensitive to minor variations of technique. In particular, different batches of media appear to affect the rates by about a factor of 3.

RESULTS

Description of the experimental system. In order to examine the stability of simple repetitive tracts in *S. cerevisiae*, we used two systems for detecting frameshift mutations, both related to the system developed by Levinson and Gutman (22). The first system (Fig. 1) involved a plasmid in which the yeast *LEU2* promoter (and the first 12 codons of the *LEU2* gene) was fused to a polylinker/β-galactosidase gene; the *LEU2* promoter functions in both *S. cerevisiae* and *E. coli*. We inserted various simple repetitive sequences into the *Bam*HI site of the polylinker (Table 1). Since the insertions used were not a multiple of 3 in length, *S. cerevisiae* and *E. coli* strains containing the plasmids failed to produce functional β-galactosidase and therefore formed white colonies on X-Gal plates. We anticipated that alterations in the length of the simple repetitive tracts to produce an in-frame inser-

TABLE 2. Rates of frameshifts of simple repetitive DNA sequences inserted upstream of β-galactosidase

Strain	Tract (nontran-scribed strand)	Location of tract[a]	No. of cultures	Avg. no. of cells/culture	No. of cultures with no detectable frameshifts[b]	Rate[c] (10^{-4})	Avg. rate (10^{-4})
SH31	$(GT)_{14}G$	P	29	1,027	26	1.9 ± 0.6	1.4
			47	969	43	1.0 ± 0.4	
			20	2,210	15	1.3 ± 0.6	
SH36	$(CA)_{14}C$	P	9	2,500	6	1.6 ± 0.2	1.5
			24	2,128	17	1.6 ± 0.6	
			24	2,617	17	1.3 ± 0.5	
JA4	$(G)_{20}$	P	35	2,185	19	2.9 ± 0.1	2.9
SH40	$(C_{1-3}A)$	P	34	2,347	34	<0.1	<0.1
			45	2,000	45	<0.1	
SH75	$(GT)_{14}G$	C	15	9,360	12	0.2 ± 0.1	0.3
			15	9,790	12	0.2 ± 0.1	
			17	11,600	10	0.4 ± 0.2	

[a] P, plasmid; C, chromosome.

[b] Since all tracts were out-of-frame insertions in the β-galactosidase gene, frameshifts were detected by cells that formed blue colonies on X-Gal plates.

[c] Calculated by the Luria-Delbruck method as modified by Lea and Coulson (21).

tion could be detected as blue yeast colonies on X-Gal plates.

The second system involved a fusion of the same *LEU2* promoter, a small portion of the β-galactosidase gene, a small portion of *HIS4*, and a functional *URA3* gene. Since the plasmid containing this construction (pSH44; Fig. 1) had an in-frame insertion of poly(GT) sequences (Table 1), yeast cells containing the plasmid were Ura$^+$. We anticipated that alterations in the tract producing an out-of-frame insertion would result in Ura$^-$ cells; such cells can be selected as resistant to 5-FOA (4). As described below, both systems gave comparable results.

Plasmid-borne poly(GT) and poly(G) tracts are unstable in S. cerevisiae. The plasmids pSH31 and pSH36 contain poly(GT) tracts 29 bp in length inserted into the vector pHT259 (Fig. 1 and Table 1). The only difference between the plasmids is the orientation of the tracts: pSH31 contains a poly(GT) tract in the nontranscribed strand, whereas pSH36 contains poly(CA) in the nontranscribed strand. Since the total length of the insertions for both plasmids is 47 bp, the *lacZ* gene is translated in the wrong reading frame (−1), and the yeast strains containing these plasmids (SH31 containing pSH31 an pSH36 containing pSH36) form white colonies on X-Gal plates. When many cells of strain SH31 or SH36 were plated on X-Gal plates, a small number of blue colonies was observed. Sequence analysis (described below) indicated that all of these colonies represent alterations in the size of the poly(GT) tract. Thus, the rate of formation of blue colonies is a direct measurement of tract instability. Using a Luria-Dulbruck fluctuation test, we found that the rates of tract instability were very similar for the different orientations of the poly(GT) tract, 1.4×10^{-4} events per division for strain SH31 and 1.5×10^{-4} events per division for strain SH36 (Table 2).

The β-galactosidase assay requires that the poly(GT) tract alter in length to restore reading frame. To determine whether a similar rate would be observed with a less restrictive system, we analyzed a strain (SH44) containing a plasmid (pSH44) with an in-frame insertion of poly(GT) upstream of the *URA3* gene. As described above, alterations in tract length out of frame can be detected by looking for 5-FOA^r cells. We found that the stability of the poly(GT) tract was similar (0.7 × 10^{-4} events per division; see Table 4) to that detected previously. On the 5-FOA plates, the strains with +2 additions grew somewhat better than those

with −2 deletions. This result is likely to reflect translational frameshifting on the poly(GU) tract, as described below. Production of a small amount of *URA3* gene product would be expected to reduce resistance to 5-FOA. The slight reduction in the frequency of frameshifts in SH44 compared with the other strains is probably explained by the observation that the copy number of centromere-containing plasmids is a mixture of cells with one and two plasmids (36). Since the 5-FOA^r phenotype is recessive, cells with two plasmids could not readily acquire this phenotype. Another factor that would reduce the instability estimate for SH44 is that there is a short phenotypic lag in the expression of 5-FOA resistance (5).

The high rate of tract instability is not specific to poly(GT) tracts. We examined a strain (JA4) containing a plasmid (pJA4) in which a poly(GT) tract 20 bp in length was inserted into pHT259. We found that the rate of instability, 2.9 × 10^{-4} events per division, was slightly elevated relative to that of the poly(GT) tracts.

Chromosomal poly(GT) tracts are also unstable. The experiments described above concern plasmid-borne simple repetitive sequences. To determine whether simple repetitive tracts in the chromosome were similarly unstable, we constructed a strain (SH75) in which a poly(GT) tract and associated β-galactosidase sequences were integrated into chromosome II at the *LYS2* locus. The tract was the same length and in the same orientation as in SH31. The observed rate of instability (0.3 × 10^{-4} events per division) was about fivefold less than that seen for the plasmid. Part of this difference (as explained above) is likely to reflect the observation that some of the cells containing the plasmid have more than one copy.

Telomeric repeats are more stable than poly(GT) or poly(G). In *S. cerevisiae*, the sequence poly(G$_{1-3}$T) is found at and near the telomeres (39, 49). To examine the stability of these sequences, we inserted a 51-bp telomeric tract (Table 1) in pHT259 and monitored tract stability as described above. In two experiments (total of 79 independent cultures, each with about 2,000 colonies), no blue colonies were observed. By the Fisher exact test, this frequency is significantly less than that observed for strains containing poly(GT) tracts ($P < 0.01$) or the strain containing the poly(G) tract ($P < 0.01$). Blue colonies were observed in experiments in which very large numbers of cells were streaked. We estimate from these studies that the rate of
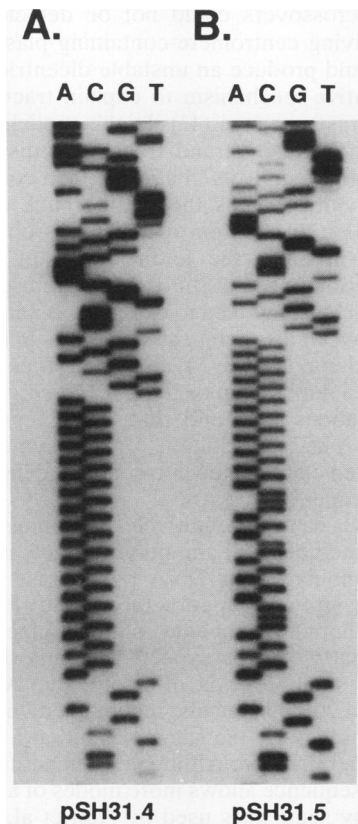
FIG. 2. DNA sequence analysis of two strains containing frameshifts in a poly(GT) tract. Both strains were blue derivatives of strain SH31. (A) Sequence of pSH31.4, containing a poly(GT) tract of 33 bp; the original tract was 29 bp in length. (B) Sequence of pSH31.5, containing a poly(GT) tract with a deletion of a GT pair and an insertion of 15 bp of telomeric [poly(G$_{1-3}$T)] sequences.

instability for the telomeric sequences is less than 10^{-5} (Table 2).

Alterations in size in the poly(GT) and poly(G) tracts usually involve deletions or additions of one or two repeats. To characterize frameshift events in plasmid-borne poly(GT) and poly(G) tracts, we "rescued" into *E. coli* the plasmids from yeast cells that were blue on X-Gal plates or that were capable of growth on 5-FOA plates. As expected, all sequenced plasmids (71 of 71) had changes in the length of the repetitive tracts (Fig. 2 and Table 3).

TABLE 3. Sequence changes in poly(GT) tracts
with frameshift mutations

Strain	Tract (nontran-scribed strand)	No. of base pairs changed[a]					
		−4	−2	+2	+4	Others	Total
SH31	(GT)$_{14}$G	1	17	2	3	2[b]	25
SH36	(CA)$_{14}$C	0	13	0	1	0	14
SH44	(GT)$_{16}$G	0	3	9	0	0	12
SH52	(GT)$_{16}$G	0	2	7	0	1[c]	10
SH75	(GT)$_{14}$G	0	8	0	1	1[d]	10

[a] +, addition; −, deletion.
[b] Includes one tract with a 14-bp deletion and one with a deletion of 2 bp and an addition of 15 bp of telomeric sequence.
[c] Includes one tract with a 14-bp deletion.
[d] Includes one tract with a 10-bp addition of poly(GT).

In the plasmids containing poly(GT), all events except one involved additions or deletions of GT repeats, with the most frequent events being 2-bp deletions (30 events) or 4-bp additions (4 events) (Table 3). Since the tracts in SH31 and SH36 were in the −1 reading frame, the minimal alterations to restore the correct reading frame involving GT repeats are deletions of one repeat or insertions of two repeats. These classes account for 87% of the total events.

Several of the derivatives of SH31 had unanticipated changes in tract length. Three tracts were in the +1 reading frame: two as the result of a 2-bp insertion, and one as the result of a 4-bp deletion. The yeast colonies containing plasmids with these tracts were pale blue on X-Gal plates, rather than the dark blue color of colonies containing the in-frame insertions. The two most likely explanations of this result are (i) that these plasmids contained a second frameshift elsewhere in the β-galactosidase gene that restored some enzymatic activity, and (ii) that slipping back by 1 bp occurs on poly(GU) tracts during translation, resulting in the production of some functional β-galactosidase gene product. Two arguments suggest that the second explanation is correct. First, when we constructed a plasmid with a poly(GT) tract in the +1 reading frame (pSH37), colonies containing this plasmid were pale blue on X-Gal medium. Second, translational frameshifting occurs at high frequency on poly(GU) tracts in prokaryotes (reviewed in reference 52). The mRNA transcribed from the pSH31 plasmid would contain a poly(GU) tract. As expected from this hypothesis, no pale blue colonies were observed that were derived from SH36, which has a poly(CA) tract in the mRNA.

All of the frameshifts derived from SH31 and SH36 involved changes in the number of GT repeats except one. One plasmid derived from SH31 had a deletion of one GT repeat and an addition of 15 bp of poly(G$_{1-3}$T) (telomeric sequences; Fig. 2). Possible mechanisms for this addition will be discussed below.

Strain SH75 contained the same poly(GT) sequence as SH31, except the tract was inserted into the chromosome. We used the PCR technique to determine the sequences of 10 independent frameshifts. The spectrum of events was similar to that observed with SH31 (Table 3). Thus, the instability of poly(GT) tracts on the plasmid and in the chromosome is qualitatively and quantitatively similar.

The poly(GT) tract in strain SH44 is in the correct reading frame. All of the altered tracts had insertions (nine derivatives) or deletions (three derivatives) of 2 bp. This result reinforces the conclusion that the tract sizes tend to change in the minimal number of repeat increments. A similar conclusion can also be drawn from analysis of strain JA4. This strain contains a plasmid with a poly(G) tract inserted into the β-galactosidase gene in the +1 reading frame. Ten of ten blue derivatives had a single base pair deletion.

We did a limited sequence analysis of frameshifts derived from strain SH40, which contained a plasmid with telomeric repeats [51 bp of poly(G$_{1-3}$T)] inserted upstream of the β-galactosidase gene. Of five blue derivatives examined, two had no frameshift within the poly(G$_{1-3}$T) tract and presumably represent insertions or deletions elsewhere in the gene. Two of the derivatives had deletions of 16 bp, and one had an insertion of 11 bp (Fig. 3).

The rate of frameshifting in the poly(GT) tract is *RAD52* independent. Most mitotic recombination events in *S. cerevisiae* require the *RAD52* gene product (reviewed in reference 30). To determine whether the stability of poly(GT) tracts was affected by *RAD52*, we constructed a yeast strain (SH52) that was isogenic with SH44 [which contained a
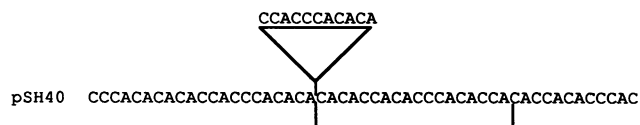
CCACCCACACA

pSH40   CCCACACACACACCACCCACACACACACCACACCCACACCACACCACACCCAC

FIG. 3. Insertions and deletions associated with a tract of telomeric [poly(G$_{1-3}$T)] sequences. The long sequence represents the original sequence in plasmid pSH40. The bases bracketed below this sequence represent those deleted in two of the frameshifted derivatives. The insertion observed in one frameshifted derivative is indicated above this sequence. Two other derivatives had no alteration in the telomeric tract.

poly(GT) tract upstream of the *URA3* gene], except that SH52 was homozygous for a null mutation of *RAD52*. The rate of tract stability was the same in the two strains (Table 4). In addition, we sequenced 10 plasmids containing a frameshift derived from SH52. The types of alterations were very similar to those seen in SH44. Thus, the *RAD52* gene product does not affect tract stability quantitatively or qualitatively.

DISCUSSION

In this paper, we show the following: (i) poly(GT) and poly(G) tracts alter in length at a rate of about 10^{-4} per division, whereas poly(G$_{1-3}$T) tracts are at least 10-fold more stable; (ii) the rate of instability of the poly(GT) tract is independent of the orientation of the tract in the plasmid; (iii) tracts can alter in length by either additions or deletions, and in general, the number of bases added or deleted is one or two repeat units for the poly(GT) tract; (iv) poly(GT) sequences are capable of interacting with telomeric repeats; and (v) the stability of the poly(GT) tracts is unaffected by the *RAD52* gene product. Since we will discuss these results in the contexts of the mechanism of tract instability, we will review some of the proposed mechanisms.

For tandemly repeated genes (or tandemly repeated bases), two mechanisms have been proposed to explain alterations in the number of repeated units: unequal recombination and DNA polymerase slippage (reviewed in references 29 and 34). In the first mechanism (Fig. 4A), two tandem arrays (possibly located on sister strands of a partially replicated molecule) pair in a misaligned configuration. Crossing-over between the misaligned arrays results in a deletion from one array and a duplication in the sister strand. Unequal sister-strand crossovers are quite common within the rRNA gene tandem array of *S. cerevisiae* (28, 45) but have not been reported to occur within a single tract of tandemly repeated bases. A mechanism related to unequal crossing-over is unequal gene conversion, in which the number of repeats is altered as the result of a nonreciprocal recombination event (10). It should be noted that unequal

sister-strand crossovers could not be detected in genetic systems involving centromere-containing plasmids, since a crossover would produce an unstable dicentric plasmid.

The alternative mechanism to explain tract instability is DNA polymerase slippage (43). In this model, during DNA replication, the primer strand transiently dissociates from the template and reanneals in a misaligned configuration. If the misalignment involves the folding back of the primer strand (Fig. 4B), an addition of repeats is observed; if the misalignment involves the folding back of the template strand, a deletion is seen. Slipped-strand mispairing during replication has been invoked to explain spontaneous deletion formation between direct repeat sequences in both prokaryotes (reviewed in reference 34) and eukaryotes (7). Das et al. (7) found that a duplication of 19 bp in *S. cerevisiae* reverted at a rate of about 10^{-5} and that the reversion rate was unaffected by *rad52*. In addition, DNA polymerase slippage in vitro has been clearly shown to generate deletions in tracts of repeated sequences (18).

Although we cannot definitively determine whether the instability of the poly(GT) and poly(G) tracts is the result of unequal recombination or DNA polymerase slippage, the evidence more strongly supports the second alternative. One argument is that the frequency of the frameshifts is high relative to what would be expected for mitotic recombination between small regions of homology. Ahn et al. (1) reported that a 26-bp sequence recombined with a region of sequence homology on the same plasmid at a rate of about 10^{-10} events per division. Although the repetitious nature of the poly(GT) sequence allows more modes of alignment than the single-copy sequences used by Ahn et al. (1), it seems unlikely that these extra modes of alignment contribute 6 orders of magnitude to the observed rate.

The high rate of tract instability observed with the poly(GT) tracts appears sufficient to explain the polymorphisms in tract length seen by other investigators (15, 51). Since we observed a similar rate with the poly(G) tract, this high rate does not appear to be a sequence-specific phenomenon. Since the telomeric sequences do not represent a pure simple repetitive DNA sequence, the relatively high stability of the telomeric sequences presumably reflects the decreased stability of misaligned configurations relative to that of perfectly aligned configurations during replication and/or recombination. It should be mentioned that the replication and/or recombination properties of the telomeric sequences located at the tip of the chromosome are likely to be different from the same sequences embedded within the chromosome.

It has been proposed that, in eukaryotes, different DNA polymerases replicate the continuous and discontinuous arms of the replication fork (42). Thus, in one orientation of the tract, the strand containing the poly(GT) sequence is replicated by one polymerase and the complementary poly(CA) sequence is replicated by the other; changing the

TABLE 4. Rates of frameshifts in poly(GT) tracts inserted upstream of *URA3*: effect of the *rad52* mutation

Strain	RAD52 genotype	Tract	No. of cultures	Avg. no. of cells/culture (10^4)	Median no. of 5-FOA^{r}cells/culture[a]	Rate[b] (10^{-5})	Avg. rate (10^{-5})
SH44	*RAD52*	(GT)$_{16}$G	15	10.6	26	7.4 ± 1.1	7.4
			17	12.5	32	7.4 ± 1.0	
SH52	*rad52*	(GT)$_{16}$G	14	8.0	20	7.9 ± 1.2	8.0
			19	7.8	19	8.0 ± 1.1	

[a] Since the poly(GT) insertions were in frame upstream of *URA3*, frameshifts were detected by determining the number of Ura$^-$ cells, using plates containing 5-FOA.

[b] Calculated by using the method of the median (Lea and Coulson [21]).
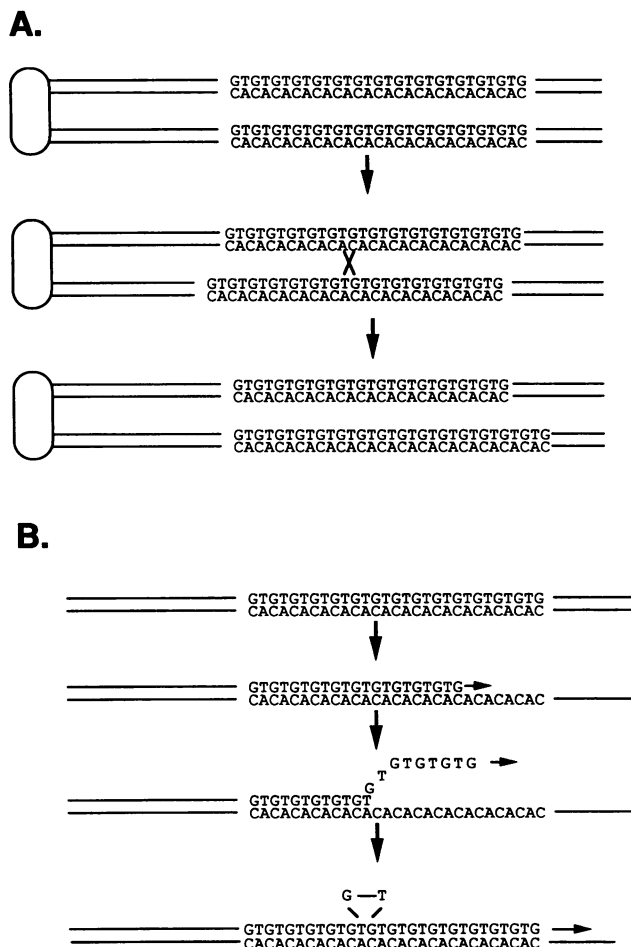
A.

B.

FIG. 4. Mechanisms of altering the length of poly(GT) tracts. (A) Unequal sister-strand crossing-over. Two poly(GT) tracts of 29 bp located on sister chromatids pair in a misaligned configuration, displaced by one GT pair. Crossing-over occurs between the mis-aligned tracts, resulting in one tract with 27 bp and one with 31 bp. Unequal sister-strand gene conversion events could also occur in which one strand gains (or loses) poly(GT) repeats, without a reciprocal alteration in the other strand. (B) DNA polymerase slippage. DNA replication occurs across a poly(GT) tract. The primer strand (indicated by an arrow) transiently dissociates from the template. When it reassociates, a loop of one GT pair is formed. When replication continues, the poly(GT) tract on the primer strand will now contain 31 instead of 29 bp. If the primer strand reassociated in a configuration in which the template strand formed a loop, the same mechanism could also result in a deletion.

orientation of the tract will change the type of DNA polymerase that replicates the individual strands of the tract. If the rate of frameshifts is a function of the type of DNA polymerase and the DNA sequence, then the orientation of the tract would be expected to affect the rate of frameshifts. For example, if DNA polymerase α had a rate of frameshifts of 10^{-4} on poly(GT) tracts and 10^{-3} on poly(CA) tracts, whereas DNA polymerase δ had a rate of frameshifts of 10^{-4} on both types of tracts, then the rate of frameshifts would be about 10-fold higher in one orientation of the tract than in the other. Since we found the same rate of tract instability regardless of the orientation, either the DNA polymerases do not have different sequence-specific rates of slippage or the tract instability is not a consequence of slippage.

Weston-Hafer and Berg (53) found that deletion formation between very small repeats on a plasmid was independent of the orientation of the repeats relative to the DNA replication origin.

DNA sequence analysis indicated that the alterations in length for the poly(GT) and poly(G) tracts were usually in units of one or two repeats, although more extensive changes were also detected. Additions, as well as deletions, were observed. Levinson and Gutman (23) previously suggested that the observed high frequency of repetitive DNA sequences seen in eukaryotes could be explained if DNA polymerase slippage produced expansions of the tract more frequently than deletions. Our results with strains SH44 and SH52 are consistent with the possibility of a bias of this sort. This conclusion, however, is not definitive, because the -2 frameshifts ($+1$ reading frame) grow more slowly on 5-FOA plates than the $+2$ frameshifts and therefore might be less likely to be chosen for DNA sequencing. It should also be pointed out that the frequency of frameshifts, as well as the direction of frameshifts, may be affected by the DNA sequences flanking the tracts.

In studies of the rate of frameshifts of poly(GT) tracts in M13 phage, Levinson and Gutman (22) found a 3:1 bias in favor of deletions. In a study of poly(GT)$_{12}$ tracts cloned in a plasmid, Freund et al. (11) found roughly comparable frequencies of $+2$, -4, $+8$, and -10 alterations. Longer poly(GT) tracts tended to have large deletions. In vitro studies involving the use of yeast DNA polymerase I (19) indicate that most frameshifts observed in regions of 4 T's or 5 C's are deletions of a single base pair, although a low frequency of base additions was also observed. It is possible that frameshifts on single base pair tracts tend to be deletions, whereas frameshifts on tracts of alternating base pairs represent both additions and deletions.

We found one poly(GT) tract that had undergone a frameshift as the result of incorporating telomeric sequences. One interpretation of this tract is that it represents a gene conversion event between the poly(GT) tract on the plasmid and telomeric sequences on the chromosome. A high rate of recombination between related, but nonidentical, telomeric sequences has been previously reported (31). Alternatively, the plasmid containing the tract could have broken open within the poly(GT) sequence, telomeric sequences could have been added to the ends, and the ends could have been religated. Whatever the mechanism, this result demonstrates that poly(GT) tracts are able to interact with telomeric sequences in *S. cerevisiae*.

In *S. cerevisiae*, most types of recombination, both cross-overs and gene conversion events, are reduced in a *rad52* genetic background (reviewed in reference 30); one exception to this generalization is that unequal sister-strand recombination in the rRNA genes does not require the *RAD52* gene product (32, 54). Recombination between related but heterologous telomeric sequences on the ends of a linear yeast plasmid are also *RAD52* independent (31). Since we find no effect of the *RAD52* gene product on the stability of the poly(GT) tracts, either the alterations in tract size reflect DNA polymerase slippage rather than recombination or the tract size is altered by a *RAD52*-independent mode of recombination.

In summary, we found that poly(GT) and poly(G) tracts change in size at a minimal rate of 10^{-4} events per division. Although the high rate and the *RAD52* independence of the events suggest that the tract instability reflects slippage of DNA polymerase, we cannot rule out a recombinational role in the instability. The relative importance of these two

processes may be clarified by a study of mutations that affect the process and by in vitro studies of DNA polymerase slippage on poly(GT) tracts.

ACKNOWLEDGMENTS

REFERENCES

1. Ahn, B.-Y., K. J. Dornfeld, T. J. Fagrelius, and D. M. Livingston. 1988. Effect of limited homology on gene conversion in a Saccharomyces cerevisiae plasmid recombination system. Mol. Cell. Biol. 8:2442–2448.
2. Alani, E., and N. Kleckner. 1987. A new type of fusion analysis applicable to many organisms: protein fusions to the URA3 gene of yeast. Genetics 117:5–12.
3. Berg, D. T., J. D. Walls, A. E. Reifel-Miller, and B. W. Grinnell. 1989. E1A-induced enhancer activity of the poly(dG-dT) · poly(dA-dC) element (GT element) and interactions with a GT-specific nuclear factor. Mol. Cell. Biol. 9:5248–5253.
4. Boeke, J. D., F. Lacroute, and G. Fink. 1984. A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoroorotic acid resistance. Mol. Gen. Genet. 197:345–346.
5. Boeke, J. D., J. Truehart, G. Natsoulis, and G. Fink. 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. 154:164–175.
6. Chattoo, B. B., F. Sherman, D. A. Azubalis, T. A. Fjellstedt, D. Mehnert, and M. Ogur. 1979. Selection of lys2 mutants of the yeast Saccharomyces cerevisiae by the utilization of alpha-aminoadipate. Genetics 93:51–65.
7. Das, G., S. Consaul, and F. Sherman. 1988. A highly revertible cyc1 mutant of yeast contains a small tandem duplication. Genetics 120:57–62.
8. Farabaugh, P. J., and J. H. Miller. 1978. Genetic studies of the lac repressor: on the molecular nature of spontaneous hotspots in the lacI gene of Escherichia coli. J. Mol. Biol. 126:847–863.
9. Fleig, U. N., R. D. Pridmore, and P. Philippsen. 1986. Construction of LYS2 cartridges for use in genetic manipulations of Saccharomyces cerevisiae. Gene 46:237–245.
10. Fogel, S., J. W. Welch, and E. J. Louis. 1984. Meiotic gene conversion mediates gene amplification in yeast. Cold Spring Harbor Symp. Quant. Biol. 49:55–65.
11. Freund, A., M. Bichara, and R. P. Fuchs. 1989. Z-DNA-forming sequences are spontaneous deletion hot spots. Proc. Natl. Acad. Sci. USA 86:7465–7470.
12. Hamada, H., M. G. Petrino, and T. Kakunaga. 1982. A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA 79:6465–6469.
13. Hoffman, C. S., and F. Winston. 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of Escherichia coli. Gene 57:267–272.
14. Jaworski, A., J. A. Blaho, J. E. Larson, M. Shimizu, and R. D. Wells. 1989. Tetracycline promoter mutations decrease non-B DNA structural transitions, negative linking differences and deletions in recombinant plasmids in Escherichia coli. J. Mol. Biol. 207:513–526.
15. Kashi, Y., Y. Tikochinsky, E. Genislav, F. Iraqi, A. Nave, J. S. Beckman, Y. Gruenbaum, and M. Soller. 1990. Large restriction fragments containing poly-TG are highly polymorphic in a variety of vertebrates. Nucleic Acids Res. 18:1129–1132.
16. 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.

17. Kreitman, M., and L. F. Landweber. 1989. A strategy for producing single-stranded DNA in the polymerase chain reaction. Gene Anal. Tech. 6:84–88.
18. Kunkel, T. A. 1986. Frameshift mutagenesis by eucaryotic DNA polymerase in vitro. J. Biol. Chem. 261:13581–13587.
19. Kunkel, T. A., R. K. Hamatake, J. Motto-Fox, M. P. Fitzgerald, and A. Sugino. 1989. Fidelity of DNA polymerase I and the DNA polymerase I-DNA primase complex from Saccharomyces cerevisiae. Mol. Cell. Biol. 9:4447–4458.
20. Kurnit, D. M. 1989. Escherichia coli recA deletion strains that are highly competent for transformation and for in vivo phage packaging. Gene 82:313–315.
21. Lea, D. E., and C. A. Coulson. 1949. The distribution of the number of mutants in bacterial populations. J. Genet. 49:264–285.
22. Levinson, G., and G. A. Gutman. 1987. High frequency of short frameshifts in poly-CA/GT tandem repeats borne by bacteriophage M13 in Escherichia coli K-12. Nucleic Acid Res. 15:5323–5338.
23. Levinson, G., and G. Gutman. 1987. Slipped strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4:203–221.
24. Lowenhaupt, K., A. Rich, and M. L. Pardue. 1989. Nonrandom distribution of long mono- and dinucleotide repeats in Drosophila chromosomes: correlations with dosage compensation, heterochromatin, and recombination. Mol. Cell. Biol. 9:1173–1182.
25. Luria, S. E., and M. Delbruck. 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics 28:491–511.
26. Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. Naylor, L. H., and E. M. Clark. 1990. d(TG)n-d(CA)n sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. Nucleic Acid Res. 18:1595–1601.
28. Petes, T. D. 1980. Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. Cell 19:765–774.
29. Petes, T. D., and C. W. Hill. 1988. Recombination between repeated genes in microorganisms. Annu. Rev. Genet. 22:147–168.
30. Petes, T. D., R. E. Malone, and L. S. Symington. 1991. Recombination in yeast, p. 407–521. In J. Broach, E. Jones, and J. Pringle (ed.), The molecular and cellular biology of the yeast Saccharomyces: genome dynamics, protein synthesis, and energetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. Pluta, A. F., and V. A. Zakian. 1989. Recombination occurs during telomere formation in yeast. Nature (London) 337:429–433.
32. Prakash, L., and T. Taillon-Miller. 1981. Effects of the RAD52 gene on sister chromatid recombination in Saccharomyces cerevisiae. Curr. Genet. 3:247–250.
33. Rich, A., A. Nordheim, and H.-J. Wang. 1984. The chemistry and biology of left-handed Z-DNA. Annu. Rev. Biochem. 53:791–846.
34. Ripley, L. S. 1990. Frameshift mutations: determinants of specificity. Annu. Rev. Genet. 24:189–213.
35. Rose, M., and D. Botstein. 1983. Construction and use of gene fusions to lacZ (β-galactosidase) that are expressed in yeast. Methods Enzymol. 101:167–180.
36. Rose, M., and J. Broach. 1991. Cloning genes by complementation in yeast. Methods Enzymol. 194:195–230.
37. Rothstein, R. J. 1983. One-step gene disruptions in yeast. Methods Enzymol. 101:202–211.
38. Santoro, C., F. Costanzo, and G. Ciliberto. 1984. Inhibition of eukaryotic tRNA transcription by potential Z-DNA sequences. EMBO J. 3:1553–1559.
39. Shampay, J., J. W. Szostak, and E. H. Blackburn. 1984. DNA sequences of telomeres maintained in yeast. Nature (London) 310:154–157.
40. Sherman, F., G. R. Fink, and J. B. Hicks. 1983. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. Slightom, J. L., A. E. Blechl, and O. Smithies. 1980. Human fetal

globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell **21:**627–638.

42. **Stillman, B.** 1988. Initiation of eukaryotic DNA replication *in vitro*. Bioessays **9:**56–60.

43. **Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye.** 1966. Frameshift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. **31:**77–84.

44. **Stringer, J. R.** 1985. Recombination between poly[d(GT) · d(CA)] sequences in simian virus 40-infected cultured cells. Mol. Cell. Biol. **5:**1247–1259.

45. **Szostak, J. W., and R. Wu.** 1980. Unequal crossing-over in the ribosomal DNA of *Saccharomyces cerevisiae*. Nature (London) **284:**426–430.

46. **Tautz, D.** 1990. Genomic fingerprinting goes simple. Bioessays **12:**44–46.

47. **Tautz, D., M. Trick, and G. Dover.** 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature (London) **322:**652–656.

48. **Treco, D., and N. Arnheim.** 1986. The evolutionarily conserved repetitive sequence d(GT · AC)ₙ promotes reciprocal exchange

and generates unusual recombination tetrads during yeast meiosis. Mol. Cell. Biol. **6:**3934–3947.

49. **Walmsley, R. M., C. S. M. Chan, B.-K. Tye, and T. D. Petes.** 1984. Unusual DNA sequences associated with the ends of yeast chromosomes. Nature (London) **310:**157–160.

50. **Walmsley, R. M., S. W. Szostak, and T. D. Petes.** 1983. Is there left-handed DNA at the ends of yeast chromosomes? Nature (London) **302:**84–86.

51. **Weber, J. L., and P. E. May.** 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. **44:**388–396.

52. **Weiss, R. B., D. M. Dunn, J. F. Atkins, and R. F. Gesteland.** 1987. Slippery runs, shifty stops, backward steps, and forward hops: -2, -1, $+1$, $+2$, $+5$, and $+6$ ribosomal frameshifting. Cold Spring Harbor Symp. Quant. Biol. **52:**687–693.

53. **Weston-Hafer, K., and D. E. Berg.** 1991. Deletions in plasmid pBR322: replication slippage involving leading and lagging strands. Genetics **127:**649–655.

54. **Zamb, T. J., and T. D. Petes.** 1981. Unequal sister strand recombination within yeast ribosomal DNA does not require the *RAD52* gene product. Curr. Genet. **3:**125–132.