

## Hypermutable of Homonucleotide Runs in Mismatch Repair and DNA Polymerase Proofreading Yeast Mutants

HIEP T. TRAN,<sup>1</sup> J. DIANNE KEEN,<sup>1</sup>† MAJA KRICKER,<sup>1</sup> MICHAEL A. RESNICK,<sup>1</sup>  
AND DMITRY A. GORDENIN<sup>1,2\*</sup>

*Genome Stability Section, Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709,<sup>1</sup> and Department of Genetics, St. Petersburg State University, St. Petersburg 199034, Russia<sup>2</sup>*

Received 27 December 1996/Returned for modification 13 February 1997/Accepted 25 February 1997

**Homonucleotide runs in coding sequences are hot spots for frameshift mutations and potential sources of genetic changes leading to cancer in humans having a mismatch repair defect. We examined frameshift mutations in homonucleotide runs of deoxyadenosines ranging from 4 to 14 bases at the same position in the *LYS2* gene of the yeast *Saccharomyces cerevisiae*. In the *msh2* mismatch repair mutant, runs of 9 to 14 deoxyadenosines are 1,700-fold to 51,000-fold, respectively, more mutable for single-nucleotide deletions than are runs of 4 deoxyadenosines. These frameshift mutations can account for up to 99% of all forward mutations inactivating the 4-kb *LYS2* gene. Based on results with single and double mutations of the *POL2* and *MSH2* genes, both DNA polymerase  $\epsilon$  proofreading and mismatch repair are efficient for short runs while only the mismatch repair system prevents frameshift mutations in runs of  $\geq 8$  nucleotides. Therefore, coding sequences containing long homonucleotide runs are likely to be at risk for mutational inactivation in cells lacking mismatch repair capability.**

Predisposition to certain types of cancers and occurrence of sporadic tumors are often correlated with the replication error (RER<sup>+</sup>) phenotype (15), which is identified by frequently occurring length variation of microsatellites (defined as multiply reiterated 1- to 4-nucleotide [nt] repeats [3]). The RER<sup>+</sup> phenotype in model *Saccharomyces cerevisiae* systems (53) and in human cells (reviewed in reference 56) is often due to defects in postreplication mismatch repair (MMR), which also results in increased mutation in coding DNA sequences (5, 6, 13, 14, 16, 24). Elevated rates of mutation in genes responsible for controlling cell proliferation (tumor suppressors) were proposed to be a source of malignant transformation in Mmr<sup>-</sup> cells (21, 25, 30, 33, 41).

The proportion of small frameshifts in short runs of two to six homonucleotides was increased in the forward mutation spectra observed in Mmr<sup>-</sup> prokaryotic and eukaryotic cells (4, 20, 32, 34, 46, 47). Homonucleotide runs in coding sequences in Mmr<sup>-</sup> cells appear to be an important source of mutations leading to carcinogenesis. In RER<sup>+</sup> tumor cell lines, the frequencies of frameshift mutations are high in three- to six-homonucleotide runs within the *APC* gene (21), which plays an important role in colorectal carcinogenesis (25). Most RER<sup>+</sup> cell lines derived from colon tumors contained frameshift mutations in the A<sub>10</sub> homonucleotide run within the coding region of the type II transforming growth factor  $\beta$  receptor (*RII*) gene. Malignancy is proposed to be due to mutated *RII* resulting in loss of growth control (33, 41, 61).

According to the replication slippage model of Streisinger and coworkers (54, 55), the incidence of frameshift mutations should increase with the length of the homonucleotide run. Mutation spectra obtained with bacteriophages (12, 42, 45),

bacteria (46, 47), and yeast (58, 59) as well as with in vitro replication (2, 26–28) support this proposal. The increase of frameshift mutation frequencies with increased size of a dinucleotide tract was demonstrated in M13 phage and in *Escherichia coli* plasmid systems (18, 29). The frameshift mutation rates for homonucleotide runs of different lengths and at various locations in the T4 phage *rII* and lysozyme genes were shown to increase exponentially as the length increased from 3 to 7 nt (42, 55). Since T4 DNA is not subject to MMR (12), these results along with mutation spectra suggest that homonucleotide runs are hypermutable in Mmr<sup>-</sup> cells.

Another means for preventing mutations during replication is DNA polymerase (DNA-Pol) proofreading. Proofreading defects in yeast appeared to result in a much weaker mutator phenotype than MMR defects for frameshifts in long stretches of GT dinucleotide repeats (53). This supported the suggestion of Streisinger and Owen (55) that only if the replication slippage loop is located very close to the 3' end of a nascent strand would it be efficiently detected by DNA-Pol proofreading. Bulged nucleotides resulting from replication slippage within a long run of simple repeats could locate at a distance from the 3' end of the growing DNA strand and thus escape from DNA-Pol proofreading. Using an in vitro system, Kroutil et al. (26) showed that as the length of a homonucleotide run was increased, the impact of proofreading in preventing frameshift mutations was decreased. Their data suggested that if this situation holds in vivo, MMR is the only system that prevents frameshift mutations in long runs of homonucleotides within coding sequences.

Based on the results summarized above, we propose that long homonucleotide runs within coding sequences are a major source of genetic change in MMR-deficient cells. To address this hypothesis we have examined (i) frameshift mutation rates in relation to the size of homonucleotide runs, (ii) the impact of MMR and proofreading defects on mutation rates, and (iii) the relative rates of mutation in runs versus the rest of a gene. We used the model eukaryote *S. cerevisiae* because of common features of MMR between this yeast and other organisms (36).

\* Corresponding author. Mailing address: Genome Stability Section, Laboratory of Molecular Genetics, NIEHS, 101 TW Alexander Dr., P.O. Box 12233, Research Triangle Park, NC 27709. Phone: (919) 541-5190. Fax: (919) 541-7593. E-mail: gordenin@niehs.nih.gov.

† Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

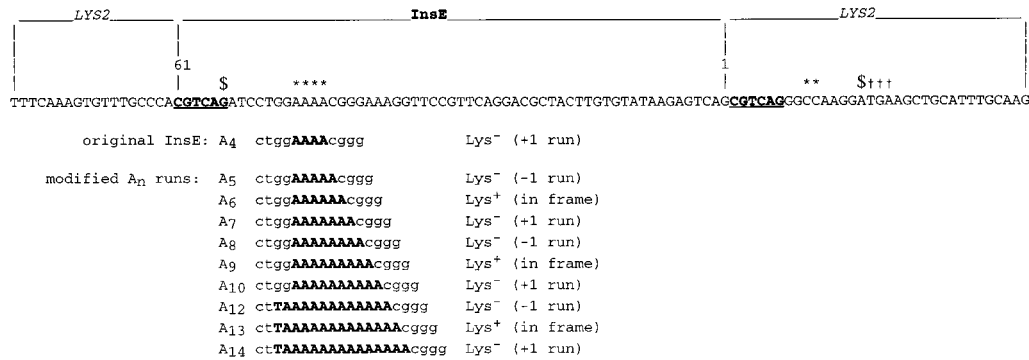


FIG. 1. Series of homonucleotide runs in *LYS2::InsE*. The *InsE* insert in the *LYS2* gene and its reversions have been described previously (58, 59). The *InsE* sequence is a remnant of an original Tn5 insert. The insert is flanked by 6-bp short direct repeats (underlined), one belonging to the insert and the other belonging to the *LYS2* gene. The insert causes a +1 shift of the *LYS2* reading frame, generating a TGA stop codon (\*\*\*), downstream from the insert. Reversions can occur via extended (>1-bp) complete or incomplete deletions and via 1-bp deletions that restore the original *LYS2* reading frame. The limits of the regions over which 1-bp deletions were detected are indicated (\$). The A<sub>4</sub> homonucleotide run hot spot of 1-bp deletions (\*\*\*\*) was increased up to A<sub>14</sub>. The initial A<sub>4</sub> sequence and modified sequences are shown in boldface capital letters, and the adjacent unchanged nucleotides are shown by lowercase letters.

To examine specifically the impact of run length on mutability, we developed a series of strains with homonucleotide runs of between 4 and 14 dAs (A<sub>4</sub> to A<sub>14</sub>) at the same position in the *LYS2* gene. Such runs are present in the coding sequences of genes from various organisms, including yeast and humans (see Discussion). Consistent with previous in vitro results (26), in vivo proofreading was only effective in removing mutations arising in short homonucleotide runs. In comparison with MMR, proofreading was relatively ineffective in dealing with mutations arising in runs of >7 nt. Thus, MMR is the only system preventing mutations in long homonucleotide runs.

There was an exponential increase in the rate of -1 and +1 frameshift mutations when the run was increased from A<sub>4</sub> to A<sub>10</sub> in an Mmr<sup>-</sup> (*msh2*) strain, and there was a 51,000-fold increase in mutation rate when the run was extended to A<sub>14</sub>. This contrasts with a 400-fold increase (for A<sub>14</sub> relative to A<sub>4</sub> runs) in the Mmr<sup>+</sup> strain. Furthermore, when the 4-kb *LYS2* coding region of the Mmr<sup>-</sup> (*msh2*) strain contained an A<sub>13</sub> run, nearly all forward mutations occurred in the run. Therefore, long homonucleotide runs in coding regions are an important class of at risk sequences in Mmr<sup>-</sup> cells.

#### MATERIALS AND METHODS

All yeast strains were isogenic with CG379 (37) (S1 in our collection) (*MATα ade5-1 his7-2 leu2-3,112 trp1-289 ura3-52*). The isogenic *pol2-4* mutant strain XAM8A (S3 in our collection) was obtained from A. Morrison (37); the *pol3-01* mutant (38) (S4 in our collection) was obtained from P. Shcherbakova (48). The *msh2*-null mutations were obtained by using the disrupting plasmid p203 (57), which replaces the 557-bp internal *Sall-Sall* fragment of the *MSH2* by the yeast *LEU2* gene. Changes increasing the A<sub>4</sub> homonucleotide run inside the *LYS2::InsE* to A<sub>10</sub> (Fig. 1) were generated in the plasmid p93 (58) by using the site-directed, double-stranded mutagenesis kit Chameleon (Stratagene). Double-stranded p93 was annealed with a mix of a mutation oligonucleotide [5'-CGTCCTGAACGGAAACCTTCCCG(T)<sub>n</sub>CCAGGATCTGACGTGGGC-3' (*n* is the number of bases in the homonucleotide dT run)] and the selection oligonucleotide 5'-GTAATTGCAAGTGGATATCTGAACAGTCCTAAAACGAG-3' (this replaces the unique *EcoRI* site by an *EcoRV* site in p93). Replacements of the wild-type *LYS2* gene were made by two-step gene replacement using *HpaI*-digested integrative versions (without the *Clal-Clal* *ARS-CEN* cassette) of the plasmids with poly(A) runs. A +2 revertant of the A<sub>10</sub> allele isolated in the wild-type strain also contained a change of two bases immediately upstream to the poly(A) run (dGG→dTA), creating an in-frame (Lys<sup>+</sup>) A<sub>13</sub> run. Mutant strains with A<sub>12</sub> and A<sub>14</sub> runs were obtained as -1-nt or +1-nt forward Lys<sup>-</sup> mutations from the strain with the A<sub>13</sub> allele (see Results). All poly(A) alleles in the plasmid and in the chromosome were verified by sequencing.

We established that A<sub>6</sub>, A<sub>9</sub>, and A<sub>13</sub> in-frame Lys<sup>+</sup> alleles also prevented growth on medium with  $\alpha$ -amino adipic acid (10, 11). Forward *lys2* mutants were selected in the strains containing A<sub>9</sub> and A<sub>13</sub> runs by  $\alpha$ -amino adipate selection. The rate of forward mutations of the *lys2::InsE-A13* allele in the *msh2* strain was determined under nonselective conditions by replica plating colonies grown on

yeast extract-peptone-dextrose onto synthetic medium lacking lysine. All genetic and molecular methods of mutation analyses used have been described previously (58, 59).

#### RESULTS

**Experimental system.** In order to study frameshift mutations in homonucleotide runs, we used the 61-bp *lys2::InsE* insert described previously (58, 59) (Fig. 1). This insert results in a +1 shift in the reading frame, leading to a Lys<sup>-</sup> phenotype. Deletions and insertions that restore the reading frame and Lys<sup>+</sup> phenotype can occur in a 71-bp region that includes the insert (Fig. 1). The longest homonucleotide run in *InsE* is an A<sub>4</sub> sequence which was previously shown to be a hot spot for +1 frameshift mutations (58, 59). We created a series of homonucleotide run alleles in *lys2::InsE* that resulted in either +1 frameshift (A<sub>4</sub>, A<sub>7</sub>, A<sub>10</sub>, A<sub>14</sub>), -1 frameshift (A<sub>5</sub>, A<sub>8</sub>, A<sub>12</sub>), or in-frame *lys2::InsE* (A<sub>6</sub>, A<sub>9</sub>, A<sub>13</sub>) alleles (Fig. 1; see also Materials and Methods). All runs are located either in the same context or in a context differing by the two adjacent bases (A<sub>12</sub>, A<sub>13</sub>, A<sub>14</sub>). To study mutations in homonucleotide runs, we selected Lys<sup>+</sup> revertants from +1 and -1 frameshift alleles and either selected or scored Lys<sup>-</sup> forward mutants from in-frame alleles.

**Mutations arising in homonucleotide runs of various sizes.** Reverse mutations of +1 and -1 frameshift alleles and forward mutations of in-frame alleles were selected, and mutations that occurred in homonucleotide runs of A<sub>n</sub> were identified as described in Materials and Methods (Table 1). Both reversions and forward mutations, when they occurred in A<sub>n</sub> runs, were always due to a frameshift in that sequence. Among mutations presented in Table 1, nearly all (225 of 227) reversions arising in +1 runs were due to deletions of a single nucleotide. The two exceptions, isolated among revertants of A<sub>10</sub> (one in a wild-type strain and one in a *pol3-01* strain), were insertions of 2 nt. Nearly 95% (206 of 218) of reversions in the -1 runs were due to addition of a single dA nucleotide. A few 2-nt deletions occurred: 6 of 84 reversion mutations of the A<sub>5</sub> run, 4 of 117 reversion mutations of the A<sub>8</sub> run, and 2 of 16 reversion mutations of the A<sub>12</sub> run. Since the system enabled us to select specific classes of reversion mutations, i.e., -1-nt reversions for the +1 run or +1-nt reversions for the -1 run, the mutabilities of these runs have been analyzed separately. All forward mutations that arose in the *LYS2* gene containing an A<sub>13</sub> in-frame run (see below) were single-nucleotide dele-

TABLE 1. Mutation rates of the *lys2::InsE* alleles with various sizes of dA homonucleotide runs

Run <sup>a</sup>	Wild type				<i>msh2</i>				<i>pol3-01</i>				<i>pol2-4</i>				<i>pol2-4 msh2</i>			
	Mut. rate (10 <sup>-9</sup> )	Ratio <sup>c</sup>	CI <sup>d</sup>	Mut. rate (10 <sup>-9</sup> )	Ratio	CI	Mut. rate (10 <sup>-9</sup> )	Ratio	CI	Mut. rate (10 <sup>-9</sup> )	Ratio	CI	Mut. rate (10 <sup>-9</sup> )	Ratio	CI	Mut. rate (10 <sup>-9</sup> )	Ratio	CI		
A <sub>4</sub> (-1)	5.4	4/55	3.9-6.2	46	20/30	32-57	140	4/37	116-181	6.5	7/20	3.6-10	1,660	26/29	1,270-2,960	ND	ND	ND		
A <sub>7</sub> (-1)	10	9/24	7-13	1,550	31/31	1,250-1,940	550	9/26	342-682	60	16/25	47-124	9,400	24/24	7,300-15,600	ND	ND	ND		
A <sub>10</sub> (-1)	47	9/9	41-93	314,000	10/10	143,000-599,000	3,520	20/20	2,400-4,120	133	9/9	83-170	133,000	11/11	90,800-181,000	ND	ND	ND		
A <sub>13</sub> (-1)	ND <sup>e</sup>	ND	ND	1,910,000	24/24	590,000-2,640,000	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND		
A <sub>14</sub> (-1)	164	10/10	117-237	1,600,000	8/8	1,370,000-2,140,000	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND		
A <sub>5</sub> (+1)	3.3	7/21	2.8-8.7	74	16/30	51-104	118	19/22	33-349	16	13/27	9.2-22	11,900	30/30	8,880-15,300	ND	ND	ND		
A <sub>8</sub> (+1)	16	17/27	14-19	3,440	34/34	2,640-3,780	3,010	17/18	2,440-4,970	75	20/25	50-103	8,000	29/29	5,620-20,400	ND	ND	ND		
A <sub>12</sub> (+1)	190	8/8	149-271	84,400	8/8	66,500-235,000	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND		

<sup>a</sup> Reversion rates were determined for -1 runs (A<sub>5</sub>, A<sub>8</sub>, and A<sub>12</sub> strains) and for +1 runs (A<sub>4</sub>, A<sub>7</sub>, A<sub>10</sub>, and A<sub>14</sub> strains). For the A<sub>3</sub> in-frame run, the forward mutation rate was determined (see text and Fig. 1). The predominant type of frameshift mutation in the run is given in parentheses (in Materials and Methods). (The -1-nt frameshifts as the predominant type of forward mutations in the A<sub>13</sub> run were determined only for the *msh2* background.) Mutation rates calculated within the runs are given in Table 2.

<sup>b</sup> Mut., mutation.

<sup>c</sup> The ratio of the number of revertants with the frameshift mutation in the run (N<sub>r</sub>) to the total number of revertants analyzed by sequencing (N<sub>t</sub>) is given. As described in references 58 and 59, reversion mutations that did not occur in the dA run were 1-nt frameshifts and extended (>1-nt) deletions (the complete spectrum of reversion mutations is available upon request).

<sup>d</sup> CI, 95% confidence interval for mutation rate (10<sup>-9</sup>).

<sup>e</sup> ND, not determined.

tions; therefore, these mutation results are included in the -1-nt reversion category.

**Length-dependent increase in mutations in wild type and Mmr<sup>-</sup> and proofreading mutants.** Increasing the homonucleotide run length greatly increased the incidence of -1-nt and +1-nt mutations for all strains tested (Tables 1 and 2). For the wild-type strain, the mutation rate for +1 and -1 reversions increased exponentially over the range 5 to 12 nt and 4 to 10 nt, respectively, with a maximal increase of reversion in the homonucleotide runs A<sub>4</sub> to A<sub>14</sub> of over 400-fold (Table 2 and Fig. 2). The proportion of reversions that were specifically due to deletions or additions in the homonucleotide runs among all frameshift mutations increased from only 7% (4 of 55) for the A<sub>4</sub> and 33% (7 of 21) for the A<sub>5</sub> run to 100% in the A<sub>10</sub>, A<sub>12</sub>, and A<sub>14</sub> runs in the wild-type strains.

Two factors likely to contribute to homonucleotide run stability are DNA-Pol proofreading and MMR. Defects in proofreading led to a relatively small increase in mutation, while the Mmr<sup>-</sup> (*msh2*) defect produced the largest mutator effect that has been observed for yeast or other eukaryotes. Elimination of proofreading exonuclease activity of either DNA-Pol δ (*pol3-01* mutation [40]) or DNA-Pol ε (*pol2-4* mutation [37]) increased frameshift rates for both -1 and +1 frameshift reversions in homonucleotide runs. The DNA-Pol δ defect led to increases of ~40- to ~300-fold over wild-type levels (Table 2) for the range of homonucleotide runs from A<sub>4</sub> to A<sub>10</sub>, indicating that DNA synthesis errors can be a significant source of mutations in the runs. The DNA-Pol ε proofreading defect had a smaller effect on mutation (Table 2).

The impact of run length on mutation rate was greatest in the *msh2* mutant strains (Tables 1 and 2). The reversion rate due to -1-nt frameshifts increased 51,000-fold as the run length was increased from A<sub>4</sub> to A<sub>14</sub>, compared to the ~400-fold increase in the wild-type strain. Over the range of A<sub>4</sub> to A<sub>10</sub> runs, the rates of reversion increased exponentially (Fig. 2). As the run was increased beyond 10 nt, the mutation rate of -1-nt frameshifts increased to a lesser extent and reached a plateau. The rate for the A<sub>13</sub> run was sixfold higher than that for the A<sub>10</sub> run, while there was a 200-fold increase when the A<sub>7</sub> run was expanded to A<sub>10</sub> in the *msh2* background (Table 2).

We examined the combined effects of deficiencies in MMR (*msh2*) and DNA-Pol ε proofreading on mutation in the homonucleotide runs. (Similar to observations of Morrison et al. [38], who used a *pms1* Mmr<sup>-</sup> mutation, an *msh2 pol3-01* DNA-Pol δ proofreading-deficient double mutant was not viable.) As shown in Table 2 and Fig. 2, the mutation rates in the A<sub>4</sub> and A<sub>5</sub> runs for the double mutant were at least 50-fold higher than those for either single mutant. As the run increased, however, there was a considerable reduction in the effect of the proofreading mutation so that the mutation rate for the A<sub>10</sub> run in the double mutant was approximately equal to that of the *msh2* single mutant (see the last column of Table 2).

**Forward mutations in the *LYS2* gene containing A<sub>9</sub> and A<sub>13</sub> homonucleotide runs.** The above results led us to propose that long homonucleotide runs are the primary source of forward mutations in an MMR mutant. Therefore, we examined forward mutations inactivating the Lys<sup>+</sup> alleles *LYS2::InsE-A<sub>9</sub>* and *LYS2::InsE-A<sub>13</sub>*. The *lys2* mutants were selected on medium containing L-α-amino adipic acid (11), and those occurring in the long homonucleotide runs were identified by sequencing of the PCR-amplified *InsE* regions (see Materials and Methods). (Rates of forward mutation in *lys2::InsE-A<sub>13</sub>* were determined by a nonselective technique [see Materials and Methods and Table 1]; rates of mutations on α-amino adipate medium were not determined because of residual growth.)

Forward *lys2* mutations located outside the homonucleotide

TABLE 2. Effects of DNA-Pol proofreading and MMR defects on rates of frameshift mutations in homonucleotide runs

Run <sup>a</sup>	w.t. <sup>b</sup>		<i>msh2</i>		<i>pol3-01</i>		<i>pol2-4</i>		<i>pol2-4 msh2</i>		
	Rate in run (10 <sup>-9</sup> ) <sup>c</sup>	Rate in run (10 <sup>-9</sup> )	Fold increase <sup>d</sup> over w.t.	Rate in run (10 <sup>-9</sup> )	Fold increase over w.t.	Rate in run (10 <sup>-9</sup> )	Fold increase over w.t.	Rate in run (10 <sup>-9</sup> )	Fold increase over:	w.t.	<i>pol2-4</i>
A <sub>4</sub> (-1)	0.4	31	78	15	38	2.3	5.8	1,448	3,720	647	48
A <sub>7</sub> (-1)	3.8	1,550	408	170	45	38	10	9,400	2,473	247	6.1
A <sub>10</sub> (-1)	47	314,000	6,681	3,520	75	133	2.8	133,000	2,830	1,000	0.4
A <sub>14</sub> (-1)	164	1,600,000	9,756	ND <sup>e</sup>	ND	ND	ND	ND	ND	ND	ND
A <sub>5</sub> (+1)	1.1	37	34	102	93	7.7	7	11,900	10,818	1,545	321
A <sub>8</sub> (+1)	10	3,440	344	2,843	284	60	6	8,000	800	133	2.3
A <sub>12</sub> (+1)	190	84,400	444	ND	ND	ND	ND	ND	ND	ND	ND

<sup>a</sup> The predominant type of frameshift mutation in the run is given in parentheses (in nucleotides).

<sup>b</sup> w.t., wild type.

<sup>c</sup> The mutation rate in the run ( $R_r$ ) was calculated as follows:  $R_r = R_t(N_r/N_t)$ , where  $R_t$  is the total rate of reversions. Values of  $R_t$  and  $N_r/N_t$  are taken from Table 1.

<sup>d</sup> Fold increase corresponds to the ratio of the rate for a given genotype to the rate in the wild-type or mutant strain with a poly(dA) run of the same length.

<sup>e</sup> ND, not determined.

run could have arisen by changes anywhere in the ~4-kb *LYS2* gene. In the wild-type background, only 5 of 24 *lys2* mutations in *LYS2::InsE-A<sub>13</sub>* and 2 of 24 mutations in *LYS2::InsE-A<sub>9</sub>* were due to frameshifts in the homonucleotide runs. (For *LYS2::InsE-A<sub>13</sub>*, four mutations were due to loss of a dA nucleotide and one mutation resulted from a dA nucleotide addition; for *LYS2::InsE-A<sub>9</sub>*, one mutation was due to an addition of a single dA and the other mutation was due to the addition of two dAs.) In contrast to these results, each of the 24 *lys2* mutants in the *msh2* *LYS2::InsE-A<sub>13</sub>* strain was due to a single-base deletion in the long homonucleotide run. Similarly, 19 of 24 *lys2* mutations in the *LYS2::InsE-A<sub>9</sub>* allele were due to loss of a single base in the run. Thus, although the entire coding sequence of *LYS2* is 4,179 bp, 100% (24 of 24) of the forward mutations occurred in a run of only 13 nt.

## DISCUSSION

We have developed a system in yeast that has enabled us to address the impact of MMR and DNA-Pol  $\epsilon$  proofreading on frameshift mutations occurring in homonucleotide runs. The results presented here have implications for mechanisms of mutation and genome stability in a variety of organisms, including humans, because of the similarities of mutation correction.

**Mutator effects of proofreading mutations.** Both DNA-Pol  $\epsilon$  and  $\delta$  replicate chromosomal DNA (1, 7) and participate in DNA repair (8), providing multiple possibilities for mutator effects. There is a substantial and approximately constant (except for the A<sub>8</sub> run) increase in rate for both +1 and -1 frameshifts in the proofreading mutants. As reported for other types of mutations in nonreiterated sequences (38, 39) and in GT dinucleotide repeats (53), the DNA-Pol  $\epsilon$  proofreading *pol2-4* mutant is a weaker mutator than the DNA-Pol  $\delta$  *pol3-01* mutant. In agreement with results obtained with dinucleotide repeats (53), the mutator effect of proofreading defects in long homonucleotide runs was much less than the effect of an MMR deficiency (Table 2; also see below).

**Hypermutability in homonucleotide runs caused by an MMR defect.** A mutation in the *msh2* gene dramatically increased mutation rates in homonucleotide runs (Table 2 and Fig. 2). By examining mutations in forward mutation systems that could detect both +1 and -1 frameshifts in A<sub>9</sub> and A<sub>13</sub>, we found that all mutations within those runs were -1-nt frameshifts. The strong preference for deletions is further demonstrated by the observation that the rate of -1 mutations in the A<sub>10</sub> run was higher than the rate of +1 mutations in the A<sub>12</sub> run within

the *msh2* strain. These results agree with the data of Petes and coworkers for multiple repeats of 4, 5, and 8 nt and contrast with the data for mutations in GT dinucleotide repeats, where 2-nt deletions exceeded additions not more than twofold (49, 52, 53).

The rate of frameshift mutation increased with run length; the rate of -1 frameshifts in A<sub>14</sub> was 51,000-fold higher than that in A<sub>4</sub>. The largest relative increase observed in the *msh2* background was found over the range A<sub>4</sub> to A<sub>10</sub> (Table 2 and Fig. 2). Overall, the mutability of homonucleotide runs reached the rates reported for runs of unstable GT<sub>14</sub> to GT<sub>16</sub> dinucleotide repeats (22, 52, 53).

We established that frameshifts in A<sub>9</sub> runs or longer homonucleotide runs could account for most of the mutations arising in the *LYS2* gene of the Mmr<sup>-</sup> strain. To assess their relative contributions to total mutations in the gene, we first estimated the rate of mutation in the rest of the gene. Based on

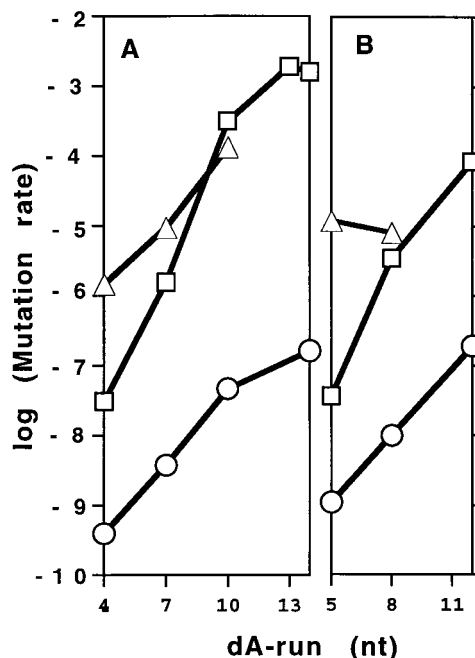


FIG. 2. Dependence of frameshift rates on length of homonucleotide run in the wild-type (○), *msh2* (□), and *msh2 pol2-4* (△) strains. Mutation rates are from Table 2. (A) -1 frameshifts; (B) +1 frameshifts.

interpolation over the range of  $A_4$  to  $A_{10}$  runs (Fig. 2), the rate of  $-1$ -nt deletions in an  $A_9$  run of the *msh2* strain would be expected to be  $\sim 5.3 \times 10^{-5}$  (the rate of  $+1$ -nt mutations would be much lower). Since 25% of the *lys2* forward mutations arising in the *lys2::InsE-A\_9* were not in the run, we estimate the rate of *lys2* mutation for the rest of the gene to be  $\sim 1.8 \times 10^{-5}$ , which is comparable to forward mutation rates for the *CAN1* gene in other *msh2* strains (22, 34). Thus, when MMR is inactive, the rates of 1-bp deletions in the  $A_9$ ,  $A_{10}$ , and  $A_{13}$  runs are 3-fold, 17-fold, and 108-fold, respectively, higher than the rate of all other mutations that inactivate the 4,179-nt *LYS2* coding sequence. We conclude that in  $Mmr^-$  yeast cells, long homonucleotide runs (of nine or more A's) are a major source of gene inactivation. Results from the laboratories of T. Petes (49) and S. Jinks-Robertson (19) also suggest that homonucleotide runs in  $Mmr^-$  yeast are hypermutable and that frameshift mutations in runs are a major component of the mutation spectrum.

There was an increase in the mutation rate in homonucleotide runs in the *msh2* strain relative to wild type for both  $-1$ -nt and  $+1$ -nt frameshifts (from  $\sim 100$ -fold for  $A_4$  to  $\sim 10,000$ -fold for  $A_{14}$  [Table 2]). Since mutations in homonucleotide runs can occur via both MMR-dependent and MMR-independent pathways (59), the above-mentioned relative increases are a minimal estimate of MMR efficiency. Possibly *msh2* has a smaller apparent effect on mutability of shorter runs because frameshift mutations in these runs occur more frequently via MMR-independent pathways. Another possibility is that the pMsh2 complex is directly involved in control of elongation fidelity and this function operates predominantly in long runs. The pMsh2 complex could (i) interact with frameshift intermediates (double-stranded DNA with a bulge formed by an extra nucleotide in either the template or nascent strand) and/or (ii) interact with the replication complex to assure that elongation occurs from properly paired nascent DNA.

**Interaction of DNA-Pol  $\epsilon$  proofreading and MMR defects.** There was a synergistic increase (much higher than additive) in  $-1$  and  $+1$  frameshifts in the short  $A_4$  and  $A_5$  runs when the *msh2* mutation and DNA-Pol  $\epsilon$  proofreading mutation *pol2-4* were combined. This is similar to the increased rate of mutation observed in a *pol2-4 pms1* ( $Mmr^-$ ) double mutant (39). The synergistic interaction may be due to two processes acting in series (38, 39) to remove errors arising during replication, i.e., proofreading followed by MMR of the remaining frameshift intermediates. The impact of proofreading in the double mutant (Table 2, last column) decreased considerably as the run length was increased from  $A_4$  to  $A_{10}$  for  $-1$  frameshifts and from  $A_5$  to  $A_8$  for  $+1$  frameshifts. These results suggest that in the *msh2* background many frameshift intermediates that appear in short ( $A_4$  and  $A_5$ ) runs can be corrected by the DNA-Pol  $\epsilon$  proofreading function.

The occurrence of frameshift intermediates leading to mutations increases dramatically with longer run lengths in the  $Mmr^-$  (*msh2*) strain (Table 2, second column). The decreased *pol2-4* effect in the *msh2* background indicates that nearly all frameshift intermediates occurring in the longer runs are inaccessible to proofreading (i.e., they escape), but they are correctable by MMR. Frameshifts in short runs are efficiently proofread in vitro, while intermediates generated in long homonucleotide runs escape proofreading (26). This hypothesis was proposed to explain the length-dependent increase of frameshift rates in homonucleotide runs. The present results provide in vivo evidence of this mechanism.

Based on similar rates of  $+1$ -nt frameshifts in  $A_5$  and  $A_8$  runs in the *msh2 pol2-4* double mutant (Table 2), the escape of

TABLE 3. Occurrence of poly(dA) and poly(dT) homonucleotide runs in yeast coding sequences<sup>a</sup>

Run length (nt)	CDS containing runs	
	No.	% of total CDS
8	990	15.9
9	303	4.9
10	93	1.5
11	45	0.70
12	26	0.40
13	10	0.16
14	9	0.14

<sup>a</sup> The occurrence of runs was determined using the database of all coding sequences (CDS) of the yeast genome. The total length of all coding sequences in the yeast genome is 8,720,211 bp; the number of coding sequences in the yeast genome is 6,218. Not shown in the table are longer runs found in yeast:  $A_{15}$  (one run),  $T_{15}$  (two runs),  $A_{16}$  (one run),  $T_{16}$  (one run),  $A_{17}$  (one run), and  $T_{17}$  (one run). Lists of coding sequences with runs and run locations are available from the authors upon request and from the Internet site <http://www.niehs.nih.gov/science/HTML/tablelink.html>. The database was obtained as the file yeast\_coding.fasta.Z downloaded from the pub/yeast/yeast\_GenBank directory at genome-ftp.stanford.edu, the guest ftp server of the Stanford DNA Sequence & Technology Center and the *Saccharomyces* Genome Database project (Internet address, [ftp://genome-ftp.stanford.edu/pub/yeast/yeast\\_GenBank/](ftp://genome-ftp.stanford.edu/pub/yeast/yeast_GenBank/)).

$+1$ -nt frameshift intermediates from DNA-Pol  $\epsilon$  proofreading could be the major mechanism underlying the increase of  $+1$ -nt frameshift rates in the *msh2* strain when the run was increased from  $A_5$  to  $A_8$ . On the other hand, there was a length-dependent increase for  $A_4$  to  $A_{10}$  in the *msh2 pol2-4* background. This could be due to the increased likelihood of  $-1$ -nt frameshift intermediates in the longer runs and also due to escape from DNA-Pol  $\delta$  proofreading. We could not determine the impact of the *msh2 pol3-01* double mutant on frameshift mutation, since haploids with two mutations were not viable. This question can be addressed if a diploid yeast strain carrying *msh2 pol3-01* appears viable similar to a diploid double mutant *pms1 pol3-01* (39).

**Implications of long homonucleotide runs on genetic instability in yeast and human cells.** Based on our measurements of genetic instability in the *msh2* mutant and the similarity of MMR in a variety of organisms, we suggest that genes containing long homonucleotide runs could account for many of the gene mutations which occur in  $Mmr^-$  cells. For example, as noted above, the *LYS2* gene with an  $A_9$ ,  $A_{10}$ , or  $A_{13}$  run is inactivated 3-, 17-, and 108-fold more, respectively, than the same coding sequence (which is  $\sim 4$  kb) without such runs.

Since coding sequences containing long homonucleotide runs could be a major target for genetic inactivation in  $Mmr^-$  cells, we determined their frequencies among all yeast genes. Approximately one-quarter of yeast genes contain dA or dT runs that are eight bases or longer, while 8% have runs that are nine bases or longer (Table 3). Therefore, yeast cells contain many genes that appear to be hypermutable targets in  $Mmr^-$  strains. This could explain the high rates of recessive lethal mutations that accumulate in diploid  $Mmr^-$  (*pms1* and *msh2*) yeast (43, 44, 62).

The  $Mmr^-$  phenotype in humans is often associated with malignant growth (see reference 56). Based on the above-mentioned results, homonucleotide runs within the coding sequences in such tumors could be mutation prone. In a small number of  $Mmr^-$  cells, there could be a high mutation rate in tumor suppressor coding sequences containing long runs (21, 33, 41, 51). In addition, the lack of MMR in cancer tissue could lead to inactivation of genes with long homonucleotide runs that are important for cancer progression and for secondary

effects of cancer. For example, the coding sequence of the structural gene for human parathyroid hormone-like protein (9, 50, 63) contains an A<sub>11</sub> run. Since this protein is associated with hypercalcemia of malignancy identified with many tumors (35), the controlling gene could be a frequent target for mutation in Mmr<sup>-</sup> tumors, possibly resulting in the hypercalcination syndrome. Based on our results with yeast, there are likely to be many human genes containing long homonucleotide runs that are at risk for inactivation in Mmr<sup>-</sup> cells. It will be necessary to establish directly in various cell lines the relationship between run length and likelihood of mutation in order to identify such genes.

In conclusion, we have demonstrated that homonucleotide runs existing in coding regions can be highly unstable. In MMR-deficient yeast cells, genes that have such runs are much more prone to inactivation than other genes. To further understand the relevance of homonucleotide runs in genome stability it is important to identify genetic factors and environmental agents that can destabilize homonucleotide runs. Factors affecting eukaryotic MMR could include replication components (60), nucleases (23), and genes controlling other repair processes, as well as environmental factors or conditions that suppress MMR. For example, in *E. coli* MMR gene products are depleted during starvation or in the stationary phase (17) and the mutation spectrum in these cells resembles the spectrum in Mmr<sup>-</sup> strains (31). Many of these issues can be conveniently addressed by utilizing the highly sensitive mutation system we have described.

#### ACKNOWLEDGMENTS

D.A.G. and M.A.R. contributed equally to this paper.

We are grateful to P. Shcherbakova and A. Morrison for yeast strains; to J. Taylor, M. Strand, W. Glaab, and L. Kroutil for critical comments on the manuscript; and to T. Petes and S. Jinks-Robertson for discussing the manuscript and communicating their prior to publication.

This work was supported by grant 96-04-48996 from the Russian Foundation for Fundamental Research (to D.A.G.), International Research Grant 75195-545401 from the Howard Hughes Medical Institute (to D.A.G. and M.A.R.), and Interagency Agreement DE-A105-94ER61940 from the DOE.

#### REFERENCES

- Araki, H., P. A. Ropp, A. L. Johnson, L. H. Johnston, A. Morrison, and A. Sugino. 1992. DNA polymerase II, the probable homolog of mammalian DNA polymerase epsilon, replicates chromosomal DNA in the yeast *Saccharomyces cerevisiae*. *EMBO J.* **11**:733-740.
- Bebenek, K., W. A. Beard, F. J. Casas, H. R. Kim, T. A. Darden, S. H. Wilson, and T. A. Kunkel. 1995. Reduced frameshift fidelity and processivity of HIV-1 reverse transcriptase mutants containing alanine substitutions in helix H of the thumb subdomain. *J. Biol. Chem.* **270**:19516-19523.
- Beckmann, J. S., and J. Weber. 1992. Survey of human and rat microsatellites. *Genomics* **12**:627-631.
- Bhattacharyya, N. P., A. Ganesh, G. Phear, B. Richards, A. Skandalis, and M. Meuth. 1995. Molecular analysis of mutations in mutator colorectal carcinoma cell lines. *Hum. Mol. Genet.* **4**:2057-2064.
- Bhattacharyya, N. P., A. Skandalis, A. Ganesh, J. Groden, and M. Meuth. 1994. Mutator phenotypes in human colorectal carcinoma cell lines. *Proc. Natl. Acad. Sci. USA* **91**:6319-6323.
- Branch, P., R. Hampson, and P. Karran. 1995. DNA mismatch binding defects, DNA damage tolerance, and mutator phenotypes in human colorectal carcinoma cell lines. *Cancer Res.* **55**:2304-2309.
- Budd, M., and J. L. Campbell. 1987. Temperature-sensitive mutations in the yeast DNA polymerase I gene. *Proc. Natl. Acad. Sci. USA* **84**:2838-2842.
- Budd, M. E., and J. L. Campbell. 1995. DNA polymerases required for repair of UV-induced damage in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **15**:2173-2179.
- Campos, R. V., L. Zhang, and D. J. Drucker. 1994. Differential expression of RNA transcripts encoding unique carboxy-terminal sequences of human parathyroid hormone-related peptide. *Mol. Endocrinol.* **8**:1656-1666.
- Chattoo, B. B., E. Palmer, B. Ono, and F. Sherman. 1979. Patterns of genetic and phenotypic suppression of *lys2* mutations in the yeast *Saccharomyces cerevisiae*. *Genetics* **93**:67-79.
- 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  $\alpha$ -amino adipate. *Genetics* **93**:51-65.
- Drake, J. W., and L. S. Ripley. 1994. Mutagenesis, p. 98-124. In J. D. Karam, J. W. Drake, K. N. Kreuzer, G. M. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, K. Carlson, and E. S. Miller (ed.), *Molecular biology of bacteriophage T4*. ASM Press, Washington, D.C.
- Dunlop, M. G. 1996. Mutator genes and mosaicism in colorectal cancer. *Curr. Opin. Genet. Dev.* **6**:75-80.
- Eshleman, J. R., E. Z. Lang, G. K. Bowerfind, R. Parsons, B. Vogelstein, J. K. Willson, M. L. Veigl, W. D. Sedwick, and S. D. Markowitz. 1995. Increased mutation rate at the *hprt* locus accompanies microsatellite instability in colon cancer. *Oncogene* **10**:33-37.
- Eshleman, J. R., and S. D. Markowitz. 1995. Microsatellite instability in inherited and sporadic neoplasms. *Curr. Opin. Oncol.* **7**:83-89.
- Eshleman, J. R., S. D. Markowitz, P. S. Donover, E. Z. Lang, J. D. Lutterbaugh, G. M. Li, M. Longley, P. Modrich, M. L. Veigl, and W. D. Sedwick. 1996. Diverse hypermutability of multiple expressed sequence motifs present in a cancer with microsatellite instability. *Oncogene* **12**:1425-1432.
- Feng, G., H.-C. T. Tsui, and M. E. Winkler. 1996. Depletion of the cellular amounts of the MutS and MutH methyl-directed mismatch repair proteins in stationary-phase *Escherichia coli* K-12 cells. *J. Bacteriol.* **178**:2388-2396.
- Freund, A. M., M. Bichara, and R. P. Fuchs. 1989. Z-DNA-forming sequences are spontaneous deletion hot spots. *Proc. Natl. Acad. Sci. USA* **86**:7465-7469.
- Greene, C. N., and S. Jinks-Robertson. 1997. Frameshift intermediates in homopolymer runs are removed efficiently by yeast mismatch repair proteins. *Mol. Cell. Biol.* **17**:2844-2850.
- Hess, P., G. Aquilina, E. Dogliotti, and M. Bignami. 1994. Spontaneous mutations at *aprt* locus in a mammalian cell line defective in mismatch recognition. *Somat. Cell Mol. Genet.* **20**:409-421.
- Huang, J., N. Papadopoulos, A. J. McKinley, S. M. Farrington, L. J. Curtis, A. H. Wyllie, S. Zheng, J. K. V. Willson, S. D. Markowitz, P. Morin, K. W. Kinzler, B. Vogelstein, and M. G. Dunlop. 1996. *APC* mutations in colorectal tumors with mismatch repair deficiency. *Proc. Natl. Acad. Sci. USA* **93**:9049-9054.
- Johnson, R. E., G. K. Kovvali, L. Prakash, and S. Prakash. 1996. Requirement of the yeast *MSH3* and *MSH6* genes for *MSH2*-dependent genomic stability. *J. Biol. Chem.* **271**:7285-7288.
- Johnson, R. E., G. K. Kovvali, L. Prakash, and S. Prakash. 1995. Requirement of the yeast *RTH1 5'* to *3'* exonuclease for the stability of simple repetitive DNA. *Science* **269**:238-240.
- Kat, A., W. G. Thilly, W. H. Fang, M. J. Longley, G. M. Li, and P. Modrich. 1993. An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair. *Proc. Natl. Acad. Sci. USA* **90**:6424-6428.
- Kinzler, K. W., and B. Vogelstein. 1996. Lessons from hereditary colorectal cancer. *Cell* **87**:159-170.
- Kroutil, L. C., K. Register, K. Bebenek, and T. A. Kunkel. 1996. Exonucleolytic proofreading during replication of repetitive DNA. *Biochemistry* **35**:1046-1053.
- Kunkel, T. A. 1986. Frameshift mutagenesis by eucaryotic DNA polymerases in vitro. *J. Biol. Chem.* **261**:13581-13587.
- Kunkel, T. A., S. S. Patel, and K. A. Johnson. 1994. Error-prone replication of repeated DNA sequences by T7 DNA polymerase in the absence of its processivity subunit. *Proc. Natl. Acad. Sci. USA* **91**:6830-6834.
- Levinson, G., and G. A. Gutman. 1987. High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in *Escherichia coli* K-12. *Nucleic Acids Res.* **15**:5323-5338.
- Loeb, L. A. 1991. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.* **51**:3075-3079.
- Longerich, S., A. M. Galloway, R. S. Harris, C. Wong, and S. M. Rosenberg. 1995. Adaptive mutation sequences reproduced by mismatch repair deficiency. *Proc. Natl. Acad. Sci. USA* **92**:12017-12020.
- Malkhosyan, S., A. McCarty, H. Sawai, and M. Perucho. 1996. Differences in the spectrum of spontaneous mutations in the *hprt* gene between tumor cells of the microsatellite mutator phenotype. *Mutat. Res.* **316**:249-259.
- Markowitz, S., J. Wang, L. Myeroff, R. Parsons, L. Sun, J. Lutterbaugh, R. S. Fan, E. Zborowska, K. W. Kinzler, B. Vogelstein, M. Brattain, and J. K. V. Willson. 1995. Inactivation of the type II TGF- $\beta$  receptor in colon cancer cells with microsatellite instability. *Science* **268**:1336-1338.
- Marsischky, G. T., N. Filosi, M. F. Kane, and R. Kolodner. 1996. Redundancy of *Saccharomyces cerevisiae MSH3* and *MSH6* in *MSH2*-dependent mismatch repair. *Genes Dev.* **10**:407-420.
- Martin, T. J., J. M. Moseley, and M. T. Gillespie. 1991. Parathyroid hormone-related protein: biochemistry and molecular biology. *Crit. Rev. Biochem. Mol. Biol.* **26**:377-395.
- Modrich, P., and R. Lahue. 1996. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **65**:101-133.
- Morrison, A., J. B. Bell, T. A. Kunkel, and A. Sugino. 1991. Eukaryotic DNA polymerase amino acid sequence required for *3'*→*5'* exonuclease activity. *Proc. Natl. Acad. Sci. USA* **88**:9473-9477.

38. Morrison, A., A. L. Johnston, L. H. Johnston, and A. Sugino. 1993. Pathway correcting DNA replication errors in *S. cerevisiae*. *EMBO J.* **12**:1467–1473.
39. Morrison, A., and A. Sugino. 1994. The 3'-5' exonucleases of both DNA polymerase  $\delta$  and  $\epsilon$  participate in correcting errors of DNA replication in *S. cerevisiae*. *Mol. Gen. Genet.* **242**:289–296.
40. Morrison, A., and A. Sugino. 1993. DNA polymerase II, the epsilon polymerase of *Saccharomyces cerevisiae*. *Prog. Nucleic Acid Res. Mol. Biol.* **46**:93–120.
41. Parsons, R., L. L. Myeroff, B. Liu, J. K. Willson, S. D. Markowitz, K. W. Kinzler, and B. Vogelstein. 1995. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. *Cancer Res.* **55**:5548–5550.
42. Pribnow, D., C. Sigurdson, L. Gold, B. S. Singer, C. Napoli, J. Brosius, T. J. Dull, and H. F. Noller. 1981. rII cistrons of bacteriophage T4: DNA sequence around the intercistronic divide and positions of genetic landmarks. *J. Mol. Biol.* **149**:337–376.
43. Prolla, T. A., D.-M. Christie, and R. M. Liskay. 1994. Dual requirement in yeast DNA mismatch repair for *MLH1* and *PMS1*, two homologs of the bacterial *mutL* gene. *Mol. Cell. Biol.* **14**:407–415.
44. Reenan, R. A. G., and R. D. Kolodner. 1992. Characterization of insertion mutations in the *Saccharomyces cerevisiae* *MSH1* and *MSH2* genes: evidence for separate mitochondrial and nuclear functions. *Genetics* **132**:975–985.
45. Ripley, L. S., A. Clark, and J. G. deBoer. 1986. Spectrum of spontaneous frameshift mutations: sequences of bacteriophage T4 *rII* gene frameshifts. *J. Mol. Biol.* **191**:601–613.
46. Schaaper, R. M., and R. L. Dunn. 1987. Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: the nature of *in vivo* DNA replication errors. *Proc. Natl. Acad. Sci. USA* **84**:6220–6224.
47. Schaaper, R. M., and R. L. Dunn. 1991. Spontaneous mutations in the *Escherichia coli* *lacI* gene. *Genetics* **129**:317–326.
48. Shcherbakova, P. V., and Y. I. Pavlov. 1996. 3'-5' exonucleases of DNA polymerases  $\epsilon$  and  $\delta$  correct base analog induced DNA replication errors on opposite DNA strands in *Saccharomyces cerevisiae*. *Genetics* **142**:717–726.
49. Sia, E. A., R. J. Kokoska, M. Dominska, P. Greenwell, and T. D. Petes. 1997. Microsatellite instability in yeast: dependence on repeat unit size and DNA mismatch repair genes. *Mol. Cell. Biol.* **17**:2851–2858.
50. Southby, J., L. M. Murphy, T. J. Martin, and M. T. Gillespie. 1996. Cell-specific and regulator-induced promoter usage and messenger ribonucleic acid splicing for parathyroid hormone-related protein. *Endocrinology* **137**:1349–1357.
51. Souza, R. F., R. Appel, J. Yin, S. Wang, K. N. Smolinski, J. N. Abraham, T. T. Zou, J. Q. Shi, J. Lei, J. Cottrell, K. Cymes, K. Biden, L. Simms, B. Leggett, P. M. Lynch, M. Fraizer, S. M. Powell, N. Harpaz, H. Sugimura, J. Young, and S. J. Meltzer. 1996. Microsatellite instability in the insulin-like growth factor II receptor gene in gastrointestinal tumors. *Nat. Genet.* **14**:255–257.
52. Strand, M., M. C. Earley, G. F. Crouse, and T. D. Petes. 1995. Mutations in the *MSH3* gene preferentially lead to deletions within tracts of simple repetitive DNA in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **92**:10418–10421.
53. Strand, M., T. A. Prolla, R. M. Liskay, and T. D. Petes. 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* **365**:274–276.
54. Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye. 1966. Frameshift mutations and the genetic code. *Cold Spring Harb. Symp. Quant. Biol.* **31**:77–84.
55. Streisinger, G., and J. Owen. 1985. Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. *Genetics* **109**:633–659.
56. Thomas, D. C., A. Umar, and T. A. Kunkel. 1996. Microsatellite instability and mismatch repair defects in cancer. *Mutat. Res.* **350**:201–205.
57. Tran, H., N. Degtyareva, D. Gordenin, and M. A. Resnick. 1997. Altered replication and inverted repeats induce mismatch repair-independent recombination between highly diverged DNAs in yeast. *Mol. Cell. Biol.* **17**:1027–1036.
58. Tran, H. T., N. P. Degtyareva, N. N. Koloteva, A. Sugino, H. Masumoto, D. A. Gordenin, and M. A. Resnick. 1995. Replication slippage between distant short repeats in *Saccharomyces cerevisiae* depends on the direction of replication and the *RAD50* and *RAD52* genes. *Mol. Cell. Biol.* **15**:5607–5617.
59. Tran, H. T., D. A. Gordenin, and M. A. Resnick. 1996. The prevention of repeat-associated deletions in *Saccharomyces cerevisiae* by mismatch repair depends on size and origin of deletions. *Genetics* **143**:1579–1587.
60. Umar, A., A. B. Buermeier, J. A. Simon, D. C. Thomas, A. B. Clark, M. R. Liskay, and T. A. Kunkel. 1996. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell* **87**:65–73.
61. Wang, J., L. Sun, L. Myeroff, X. Wang, L. E. Gentry, J. Yang, J. Liang, E. Zborowska, S. Markowitz, J. K. Willson, and M. G. Brattain. 1995. Demonstration that mutation of the type II transforming growth factor beta receptor inactivates its tumor suppressor activity in replication error-positive colon carcinoma cells. *J. Biol. Chem.* **270**:22044–22049.
62. Williamson, M. S., J. C. Game, and S. Fogel. 1985. Meiotic gene conversion mutants in *Saccharomyces cerevisiae*. I. Isolation and characterization of *pms1-1* and *pms1-2*. *Genetics* **110**:609–646.
63. Yasuda, T., D. Banville, G. N. Hendy, and D. Goltzman. 1989. Characterization of the human parathyroid hormone-like peptide gene. Functional and evolutionary aspects. *J. Biol. Chem.* **264**:7720–7725.