Human Molecular Genetics, 2001, Vol. 10, No. 18 **1889–1900**

Functional analysis of human MLH1 and MSH2 missense variants and hybrid human–yeast MLH1 proteins in *Saccharomyces cerevisiae*

Aaron R. Ellison, Joan Lofing and Grant A. Bitter*

BitTech, Inc., Westlake Village, CA 91361, USA

Received April 13, 2001; Revised and Accepted June 24, 2001

Hereditary non-polyposis colorectal cancer (HNPCC) is an autosomal dominant inherited disease caused by defects in the process of DNA mismatch repair (MMR), and mutations in the *hMLH1* or *hMSH2* genes are responsible for the majority of HNPCC. In addition to clear loss-of-function mutations conferred by nonsense or frameshift alterations in the coding sequence or by splice variants, genetic screening has revealed a large number of missense codons with less obvious functional consequences. The ability to discriminate between a loss-of-function mutation and a silent polymorphism is important for genetic testing for inherited diseases like HNPCC where the opportunity exists for early diagnosis and preventive intervention. In this study, quantitative *in vivo* DNA MMR assays in the yeast *Saccharomyces cerevisiae* were performed to determine the functional significance of amino acid replacements observed in the human population. Missense codons previously observed in human genes were introduced at the homologous residue in the yeast *MLH1* or *MSH2* genes. This study also demonstrated feasibility of constructing genes that encode functional hybrid human–yeast MLH1 proteins. Three classes of missense codons were found: (i) complete loss of function, i.e. mutations; (ii) variants indistinguishable from wild-type protein, i.e. silent polymorphisms; and (iii) functional variants which support MMR at reduced efficiency, i.e. efficiency polymorphisms. There was a good correlation between the functional results in yeast and available human clinical data regarding penetrance of the missense codon. The results reported here raise the intriguing possibility that differences in the efficiency of DNA MMR exist between individuals in the human population due to common polymorphisms.

INTRODUCTION

Colorectal cancer (CRC) is one of the most common cancers, affecting 3–5% of the population in the western world by the age of 70. Worldwide, each year approximately one million individuals are diagnosed with CRC and 500 000 individuals die from CRC (1). Epidemiological studies have shown that 5–15% of CRCs have a genetic basis (reviewed in refs 2 and 3). The most common form of familial CRC is hereditary non-polyposis colorectal cancer (HNPCC), which manifests with a high rate of mortality in the absence of early detection and treatment (reviewed in refs 2,4 and 5). Diagnosis of HNPCC in a family is based on kindred analysis using the Amsterdam criteria (6) which require: (i) three or more family members to have had histologically verified CRC, with one being a first-degree relative of the other two; (ii) CRC in at least two generations; and (iii) at least one individual diagnosed with CRC before the age of 50. At the molecular level, HNPCC is associated with mutations in genes encoding proteins involved in DNA mismatch repair (MMR).

The process of DNA MMR corrects non-native DNA structures that form, primarily, during DNA replication. These aberrant structures include incorrectly paired bases resulting from nucleotide misincorporation by DNA polymerases as well as insertion/deletion loops in DNA which form, for example, as a result of microsatellite instability. The amino acid sequences of MMR protein functional domains are conserved from *Escherichia coli* to humans, and the eukaryotic MMR proteins are named based on their homology to the *E.coli* MutS and MutL proteins. Mechanistic studies of MMR in yeast and human cells have elucidated similar processes and functional protein complexes (reviewed in refs 7–9). The MutSα complex is a heterodimer of MSH2 and MSH6 while MutSβ is a heterodimer of MSH2 and MSH3. MutSα recognizes base:base mismatches as well as single base insertion/deletion mispairs. MutSβ also recognizes single base insertion/deletion mispairs but is primarily responsible for recognition of larger insertion/deletion mispairs. Heterodimers of the MutL homologs bind to the MutSα or MutSβ DNA mismatch complex to effect repair. The yeast MLH1-PMS1 heterodimer (MLH1-PMS2 in humans) binds both MutSα and MutSβ while the yeast MLH1-MLH3 complex (MLH1-PMS1 in humans) binds MutSβ (reviewed in ref. 9). In addition to DNA MMR, MMR proteins act as a molecular switch (10,11) and appear to be involved in the induction of apoptosis in response to DNA damage (12–14).

HNPCC has been shown to be caused by mutations in the *hMLH1*, *hMSH2*, *hPMS2* or *hMSH6* genes. To date, more than

*To whom correspondence should be addressed at: PO Box 1499, Agoura Hills, CA 91376-1499, USA. Tel: +1 805 495 7515; Fax: +1 805 495 1866; Email: gab@bittech.net

240 mutations have been described, and the vast majority occur in either *hMLH1* (60%) or *hMSH2* (35%). It is probable that the majority of HNPCC is associated with mutations in either *hMLH1* or *hMSH2* since inactivation of either of these genes results in impaired replication of a broad spectrum of mismatches (single base:base mismatches and both small and large insertion/deletion loops). Furthermore, CRC patients that do not fit the strict Amsterdam criteria ('suspected HNPCC') have been shown to possess germline mutations in MMR genes (15,16). It has also been reported that defects in MMR are involved in sporadic cancers, and can be caused by somatic *hMLH1* gene silencing due to hypermethylation (17).

In the genetic analyses of HNPCC kindreds, >25% of the gene alterations observed are minor variants such as amino acid replacements or small in-frame deletions. These sequence variants, furthermore, are scattered throughout the gene coding region. If an observed amino acid replacement can be shown to segregate with disease in the affected family, it suggests, but does not prove, that the amino acid replacement is a pathogenic mutation. Frequently, however, small family size or unavailability of clinical samples has precluded attempts to correlate the amino acid replacement with pathogenic effect. As genetic analyses of HNPCC kindreds have continued, an increasing number of minor variants have been documented (http://www.nfdht.nl). To date, missense codons resulting in 53 different amino acid replacements have been described in *hMLH1* while 35 have been reported in *hMSH2*. It is now generally acknowledged (3,8,18) that accurate and effective genetic testing for HNPCC will require determination of the functional significance of these minor variants, since the utility of genetic tests is severely compromised if there is any ambiguity in the results.

The process of MMR is conserved from yeast to humans, as are the amino acid sequences of the protein functional domains. We previously described a standardized and quantitative *in vivo* assay of MMR in the yeast *Saccharomyces cerevisiae* which is useful for assessing the functional significance of missense codons observed in human MMR genes (19). In the current study, the effect of 20 amino acid replacements on MLH1p and MSH2p function were evaluated using a reporter gene which measures the stability of an in-frame $(GT)_{16}$ tract. Codon changes previously identified in human genes were introduced by site-directed mutagenesis at the homologous codon in the yeast gene and tested for *in vivo* function in *S.cerevisiae*. We also demonstrated feasibility of constructing genes which encode hybrid human–yeast MLH1 proteins that are functional in MMR *in vivo*. These hybrid genes allow functional assays of variant proteins containing human amino acid replacements at residues that are not conserved in yeast and/or where the equivalent residue in yeast is uncertain. In addition to identification of silent polymorphisms and mutations conferring complete loss of MMR function, we found that certain codon changes in *hMLH1* and *hMSH2* give rise to proteins with a reduced efficiency of MMR. Some of these amino acid replacements occurred as germline mutations in individuals who developed CRC but whose families did not satisfy the criteria of HNPCC. These observations raise the intriguing possibility that differences in the efficiency of DNA MMR exist between individuals in the population due to common polymorphisms, and that such polymorphisms may predispose to early onset cancer.

RESULTS

Function of the MMR protein variants was determined using a standardized MMR assay that measures stability of a $(GT)_{16}$ tract *in vivo* (19). For these experiments, we constructed haploid yeast strains that have complete deletions of the chromosomal coding region of *MLH1* (strain YBT24) or *MSH2* (strain YBT25). Utilizing the reporter gene $(GT)_{16}$-*URA3* (pSH91; Materials and Methods), both YBT24 and YBT25 exhibit a mutation frequency (Materials and Methods) >100-fold greater than an MMR wild-type strain. As demonstrated below, the mutation frequency can be reduced to approximately that observed in the wild-type strain by complementing the chromosomal null mutations with a single copy plasmid-expressed (Materials and Methods) wild-type yeast *MLH1* gene in YBT24 or *MSH2* gene in YBT25. The mean mutant frequency in plasmid-complemented chromosomal null mutants was used as the basis for comparing activity of the variant MMR proteins described below. MMR defect is defined as the mutant frequency conferred by the variant protein in the chromosomal null mutant divided by the mutant frequency observed in the chromosomal null mutant with a plasmid-expressed wild-type yeast gene. All *MLH1* or *MSH2* variants were analyzed in the same host containing the same reporter gene and were expressed from the same expression vector as the wild-type yeast gene. For all mutagenized genes, at least three independent clones were tested for MMR function with identical results. The DNA sequence on both strands of the mutagenized gene was confirmed for one mutant, and this clone was assayed in replicate cultures ($n \geq 4$) for determination of the mutation frequencies reported below.

MLH1 orthologs from various organisms, including yeast and humans, were aligned to identify homologous amino acid residues in the conserved regions of the proteins (Fig. 1). These alignments, together with previously identified human missense codons, were used to target yeast residues for alterations that would mimic variants observed in the human population (Table 1). Strain YBT24 ('Mutator', Fig. 2), exhibited a mutation frequency of 1.8×10^{-3}, a level 160-fold higher than that exhibited by the wild-type parental strain YBT5-1. When the wild-type *MLH1* gene was expressed from a plasmid in YBT24 ('Complemented', Fig. 2) the mutation frequency was reduced >65-fold to nearly the wild-type levels. Cells expressing MLH1p with the amino acid replacements A41F, G64R, I65N, E99K, I104R, T114R, Q552L and R672P exhibited mutation frequencies of $1.7–3.6 \times 10^{-3}$ (Fig. 2). These mutation frequencies represent MMR defects of 62–130 (i.e. mutation frequencies 62–130-fold higher than 'Complemented') and are not significantly different from that exhibited by the *mlh1*Δ deletion strain YBT24 which lacked a complementing plasmid ('Mutator'). These results demonstrate that amino acid replacements A41F, G64R, I65N, E99K, I104R, T114R, Q552L and R672P result in complete loss of MLH1p-mediated MMR function. These missense codons are therefore mutations.

Strain YBT24 expressing the A41S, V216I, I326V or A694T variants exhibited mutation frequencies of $1.0–3.2 \times 10^{-5}$, which were not significantly different from the mutation frequency observed when strain YBT24 expressed the wild-type yeast *MLH1* gene ('Complemented', Fig. 2). The A41S and I326V variants are conservative amino acid replacements, and these codon changes convert the wild-type yeast residue to the amino acid in the corresponding position of the wild-type

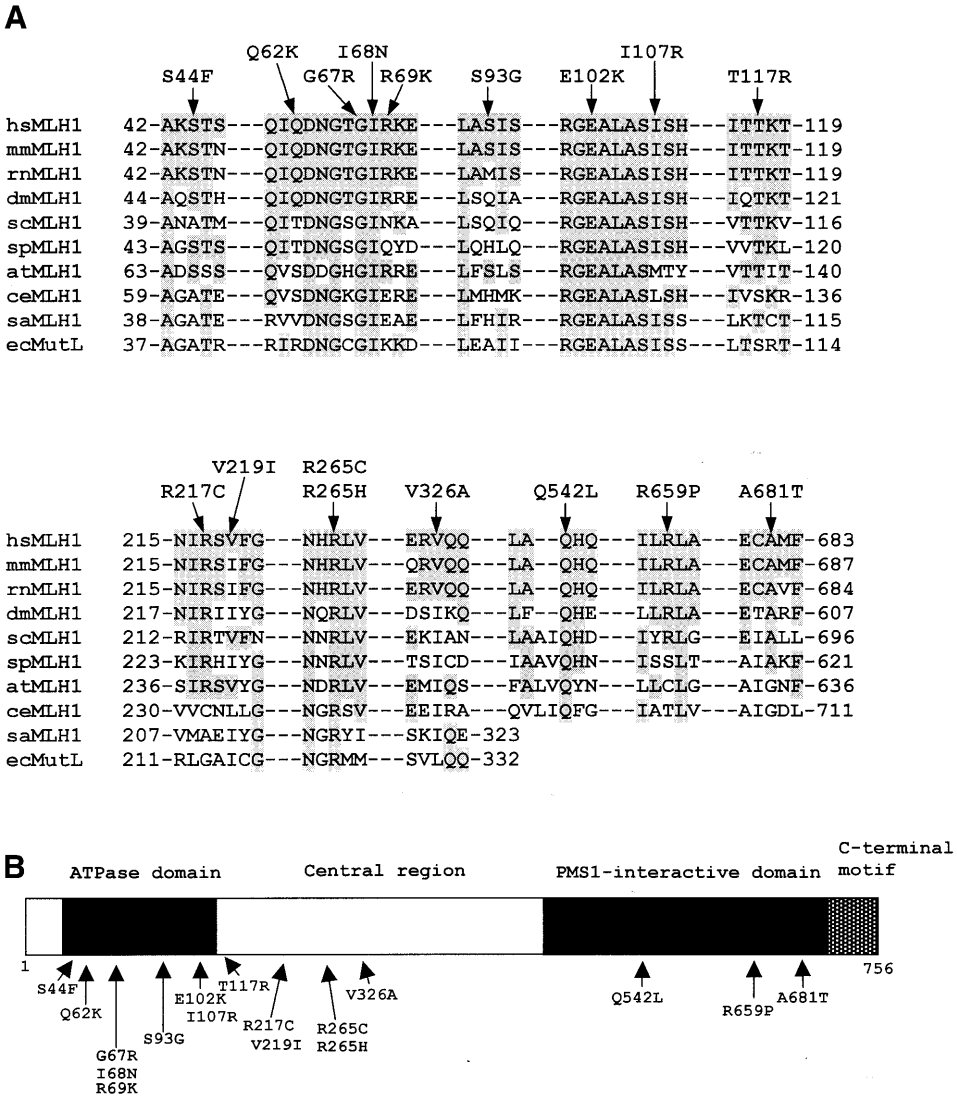

Figure 1. (**A**) Alignment of MLH1 polypeptides. Amino acid sequences from human (hs, *Homo sapiens*), mouse (mm, *Mus musculus*), rat (rn, *Rattus norvegicus*), fruit fly (dm, *Drosophila melanogaster*), yeast (sc, *Saccharomyces cerevisiae* and sp, *Schizosaccharomyces pombe*), plant (at, *Arabidopsis thaliana*), flatworm (ce, *Caenorhabditis elegans*), and bacteria (sa, *Staphylococcus aureus* and ec, *Escherichia coli*) were aligned using BLAST algorithms (48). Shaded amino acid residues are identical to those in the human protein. The human *hMLH1* missense mutations examined in this study are noted above the appropriate amino acid residue in the human sequence. (**B**) Representation of the missense alteration relative to the full length MLH1 protein and the defined functional domains (40,49).

human protein. The data in Figure 2 demonstrate that the A41S, V216I, I326V and A694T amino acid replacements do not alter MMR function, and therefore represent silent polymorphisms (see Discussion).

Four of the codon changes in *MLH1* encode proteins which appear to support intermediate efficiencies of DNA MMR. The R214C, R265C, R265H and I326A amino acid changes exhibit MMR defects of 2.3, 13.8, 4.0 and 1.6, respectively, which are intermediate between that of the wild-type *MLH1* complemented mutant (MMR defect = 1.0) and the *mlh1*Δ null mutant (MMR defect = 65). The mutation frequencies measured with the R214C, R265C, R265H and I326A amino acid replacements are significantly less than the *mlh1*Δ null mutant lacking a complementing plasmid ($P < 0.0005$). The mutation frequencies of the R214C, R265C and R265H variants are clearly different from the strain complemented with the wild-type *MLH1* gene with *P*-values of 0.0005, 0.0001 and 0.0001, respectively. The 1.6-fold MMR defect of the I326A variant also appears to be significant ($P = 0.0279$) compared with the strain expressing wild-type *MLH1*. These results demonstrate that the R214C, R265C, R265H and I326A missense variants function in DNA MMR at a reduced efficiency, and we refer to this class of variants as efficiency polymorphisms.

Previous work from this laboratory demonstrated that a P640L alteration in yeast MSH2p was a loss-of-function MMR mutation while a H658Y alteration had no effect on MMR (19). In the current study, an additional amino acid replacement, G317D, in yeast MSH2p was evaluated. Mutation frequencies of strain YBT25 expressing the G317D allele were determined and compared (Fig. 3) with strain YBT25 without a complementing gene ('Mutator') and YBT25 complemented with the wild-type *MSH2* gene ('Complemented'). The strain expressing the G317D variant exhibited a low mutation frequency (2.3×10^{-5}) that was significantly less than the *msh2*Δ null mutant,

Table 1. Human *MLH1* and *MSH2* variants examined in this study and oligonucleotides used for site-directed mutatgenesis

Variation in human MMR gene		Equivalent substitution in yeast	Oligonucleotide sequence (sense strand)[a] (5′→3′)	Restriction site alteration[b]
MLH1	S44F	A41F	CCATCGATGCGAACTTTACAATGATTGATATTC	– *Bsa*MI
	–	A41S	CCATCGATGCGAATTCTACAATGATTGATATTC	+ *Eco*RI
	Q62K	–	GAAGTTGATTCAGATTAAAGACAATGGCACCG	– *Bst*XI
	G67R	G64R	GATAACGGATCGCGAATTAATAAAGCAGACCTGCC	+ *Nru*I
	I68N	I65N	GATAACGGATCTGGAAATAATAAAGCAGAC	– *Vsp*I
	R69K	–	CAATGGCACCGGTATCAAGAAAGAAGATCTGG	+ *Age*I
	S93G	–	CTTTGAGGATTTAGCCGGTATTTCTACCTATG	+ *Bsr*FI
	E102K	E99K	GGATTCCGAGGAAAGGCTTTAGCCAGTATCTC	– *Hin*dIII
	I107R	I104R	GAAGCTTTAGCAAGTAGATCTCATGTGGCAAGAG	+ *Bgl*II
	T117R	T114R	GAGTCACAGTAACGCGTAAAGTTAAAGAAGAC	+ *Mlu*I
	R217C	R214C	CCAGGATAGGATTTGTACAGTGTTCAATAAATC	+ *Bsr*GI
	V219I	V216I	GATAGGATTAGGACAATATTCAATAAATCTGTG	+ *Ssp*I
	R265C	R265C	TTTTTTTCATTAATAATTGTCTAGTGACATGTG	– *Spe*I
	R265H	R265H	TTTTTTTCATTAATAATCACCTAGTGACATGTG	– *Spe*I
	V326A	I326A	GAGATCATAGAGAAAGCGGCCAATCAATTGC	+ *Eae*I
	–	I326V	GAGATCATAGAGAAAGTCGCGAATCAATTGC	+ *Nru*I
	Q542L	Q552L	GATTAGCCGCTATTCTTCATGACTTAAAGC	+ *Bsp*HI
	R659P	R672P	CCATTTTTTATATATCCCTTAGGTAAAGAAGTTG	+ *Bsu*36I
	A681T	A694T	GTATTTTAAGAGAGATTACATTGCTCTATATACCTG	+ *Bsr*DI
MSH2	G322D	G317D	CAAAATCCATTCGATAGCAACAATTTAGC	None[c]

[a]Both sense and antisense oligonucleotides were employed for site-directed mutations in yeast and human MMR genes as described in Materials and Methods.
[b]Restriction site alterations are silent, except for the indicated amino acid substitutions. +, restriction site addition; –, restriction site loss.
[c]Alteration screened by DNA sequencing.

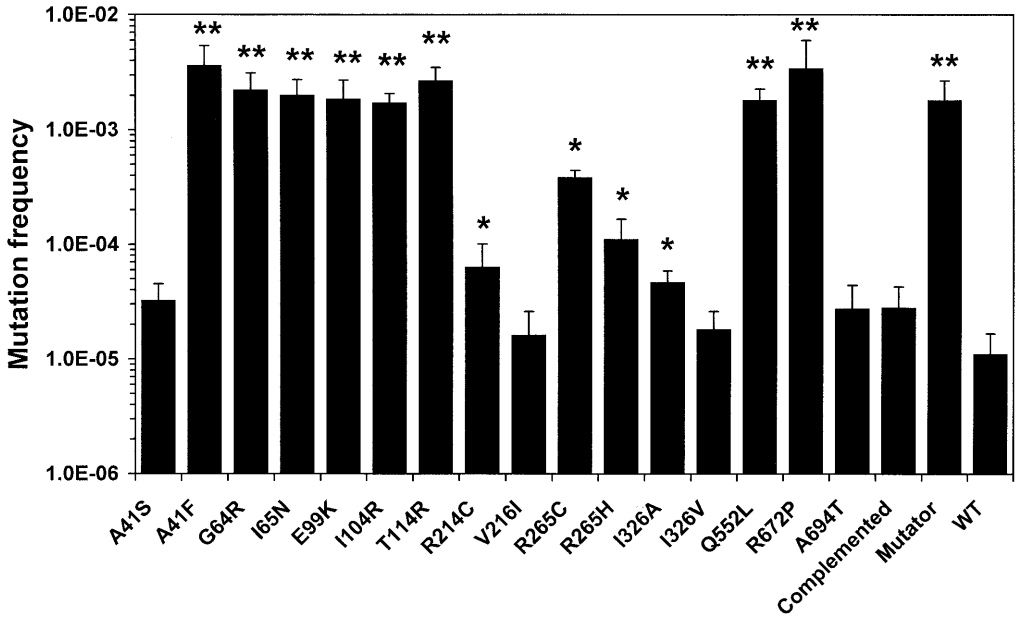

Figure 2. Mutation frequencies conferred by missense codons in the yeast *MLH1* gene. Strain YBT24 containing pSH91 was transformed with pMLH1 (Complemented) or the indicated mutant form of pMLH1. Mutator refers to YBT24 containing pSH91 but without a pMLH1 plasmid, while WT refers to strain YBT5-1 containing pSH91 but without a pMLH1 plasmid. Mutation frequencies were determined as described in Materials and Methods. The mutant frequencies are presented as the mean ± SD of at least four replicate cultures of a single yeast clone that expresses the indicated *MLH1* gene alteration. **, a mutation frequency which is not significantly different from the mutator strain. *, a mutation frequency which is intermediate between the mutator strain and complemented strain. Mean mutation frequencies are: WT, 1.1×10^{-5}; Mutator, 1.8×10^{-3}; Complemented, 2.8×10^{-5}; A41S, 3.2×10^{-5}; A41F, 3.58×10^{-3}; G64R, 2.15×10^{-3}; I65N, 1.97×10^{-3}; E99K, 1.85×10^{-3}; I104R, 1.65×10^{-3}; T114R, 2.65×10^{-3}; R214C, 6.3×10^{-5}; V216I, 1.6×10^{-5}; R265C, 3.77×10^{-4}; R265H, 1.09×10^{-4}; I326A, 4.6×10^{-5}; I326V, 1.8×10^{-5}; Q552L, 1.77×10^{-3}; R672P, 3.42×10^{-3}; A694T, 1.7×10^{-5}.

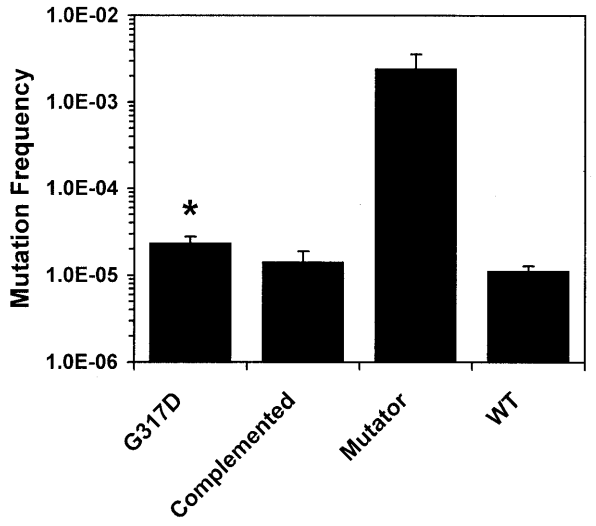

Figure 3. Mutation frequencies conferred by missense codons in the yeast *MSH2* gene. Strain YBT25 containing pSH91 was transformed with pMETc/MSH2 (Complemented) or the indicated mutagenized form of pMETc/MSH2. Mutator refers to YBT25 containing pSH91 but without a pMetc/MSH2 plasmid, while WT refers to strain YBT5-1 containing pSH91 but without a pMetc/MSH2 plasmid. Mutation frequencies were determined as described in Materials and Methods. The mutant frequencies are presented as the mean ± SD of six replicate cultures of a single yeast clone that expresses the indicated MSH2 gene alteration. Mean mutation frequencies are: WT, 1.0×10^{-5}; Mutator, 2.39×10^{-3}; Complemented, 1.4×10^{-5}; G317D, 2.3×10^{-5}.

demonstrating MMR activity of the protein variant. Cells expressing G317D, however, exhibit an MMR defect of 1.7, and this difference from YBT25 complemented with wild-type *MSH2* appears significant ($P = 0.0073$). These results indicate that the G317D replacement encodes a functional protein of slightly reduced efficiency in MMR.

A number of the amino acid residues in human MLH1 which have been linked to the development of HNPCC are not conserved in the yeast protein. As shown above (for the human S44F and V326A variants), alterations at non-conserved residues can be evaluated by replacing the corresponding yeast codon with that found in the native human gene. If the change is a silent polymorphism, the effect of the variant human codon can subsequently be analyzed in yeast. The utility of this approach, however, may be limited. In some cases the wild-type human amino acid may not functionally replace the yeast amino acid, or there may be instances in which the homology between the human and yeast genes is too weak to unambiguously assign the corresponding yeast codon. An alternative approach would be to construct gene fusions that encode hybrid human–yeast proteins which retain MMR function *in vivo*.

Seven gene fusions were constructed (Materials and Methods) by replacing a portion of the yeast *MLH1* gene with the homologous coding sequence from the human gene (Fig. 4A). The gene fusions were expressed from the same parental expression vector (Materials and Methods) in the *mlh1*Δ-null strain YBT24 and quantitative MMR assays were carried out to evaluate function of the hybrid human–yeast protein (Fig. 4B). Complementation of the *mlh1*Δ-null strain was not observed with two hybrids, MLH1_h(498–756) and MLH1_h(498–584), which contain portions of the human C-terminal domain (Fig. 4B). In contrast, five hybrids that contained portions of the human N-terminal domain complemented strain YBT24, yielding substantial reductions in mutation frequency (Fig. 4B). The functional hybrid proteins spanned amino acids 1–177 and conferred mutation frequencies that tended to correlate with the length of the human portion of the hybrid protein. Whereas the *mlh1*Δ-null strain exhibited an MMR defect of 74.5, strains expressing hybrids MLH1_h(1–177), MLH1_h(1–86), MLH1_h(41–130) and MLH1_h(41–86) exhibited MMR defects of 39.6, 9.9, 8.5 and 4.8, respectively. The most efficient hybrid was MLH1_h(77–134), containing a 57 amino acid segment from the human coding sequence, which conferred an efficiency of DNA MMR which is within a factor of 2 of the wild-type native yeast protein.

The ability to generate functional hybrid human–yeast proteins allowed evaluation of human amino acid replacements at residues that are not conserved in yeast. Three replacements (Q62K, R69K, S93G) that appear in one or more of the MLH1 mutation databases were engineered by site-directed mutagenesis into an appropriate gene fusion. These genes were expressed in strain YBT24 and mutation frequencies compared with that obtained with the gene encoding the original hybrid human–yeast protein (Fig. 5). The Q62K and R69K alterations conferred mutation frequencies which were significantly lower than the mutator strain without a complementing plasmid. However, these mutation frequencies were significantly ($P < 0.05$) greater (1.7- and 1.9-fold, respectively) than those obtained with the MLH1_h(41–86) hybrid, demonstrating these codon changes to be efficiency polymorphisms. The S93G alteration did not appear to affect function in MMR since the mutation frequency was not significantly different ($P = 0.298$) from that observed with the MLH1_h(77–134) hybrid and, accordingly, S93G appears to be a silent polymorphism.

DISCUSSION

In this study, the functional significance in MMR of 20 different amino acid replacements in MLH1p and MSH2p was determined. Codon changes were engineered into the yeast gene based on previous observation of the same amino acid replacement in the human protein and a potential implication of the missense variant in cancer development. Quantitative *in vivo* assays of DNA MMR allowed classification of the amino acid replacements as either silent polymorphsims, loss-of-function mutations or efficiency polymorphisms. The quantitative *in vivo* assay of MMR utilized in this study measured microsatellite instability of a $(GT)_{16}$ tract (19). Some of our results have been confirmed by investigators using different reporter genes to assess MMR activity *in vivo* (Table 2). Functional results identical to ours but obtained using different reporter genes are consistent with the role of MLH1p and MSH2p in repairing a broad spectrum of DNA mismatches (9). Moreover, functional analyses of sequence variants using an *in vitro* protein:protein interaction assay (20) and a dominant yeast mutator model (21) were largely complementary to our results. It should be pointed out, however, that MMR proteins also appear to be involved in other biological activities, including induction of apoptosis in response to DNA damage (12–14). Therefore, the formal possibility exists that sequence variants that have no detectable effect on MMR may have clinical significance due to effects of the variation on other biological activities of the proteins.

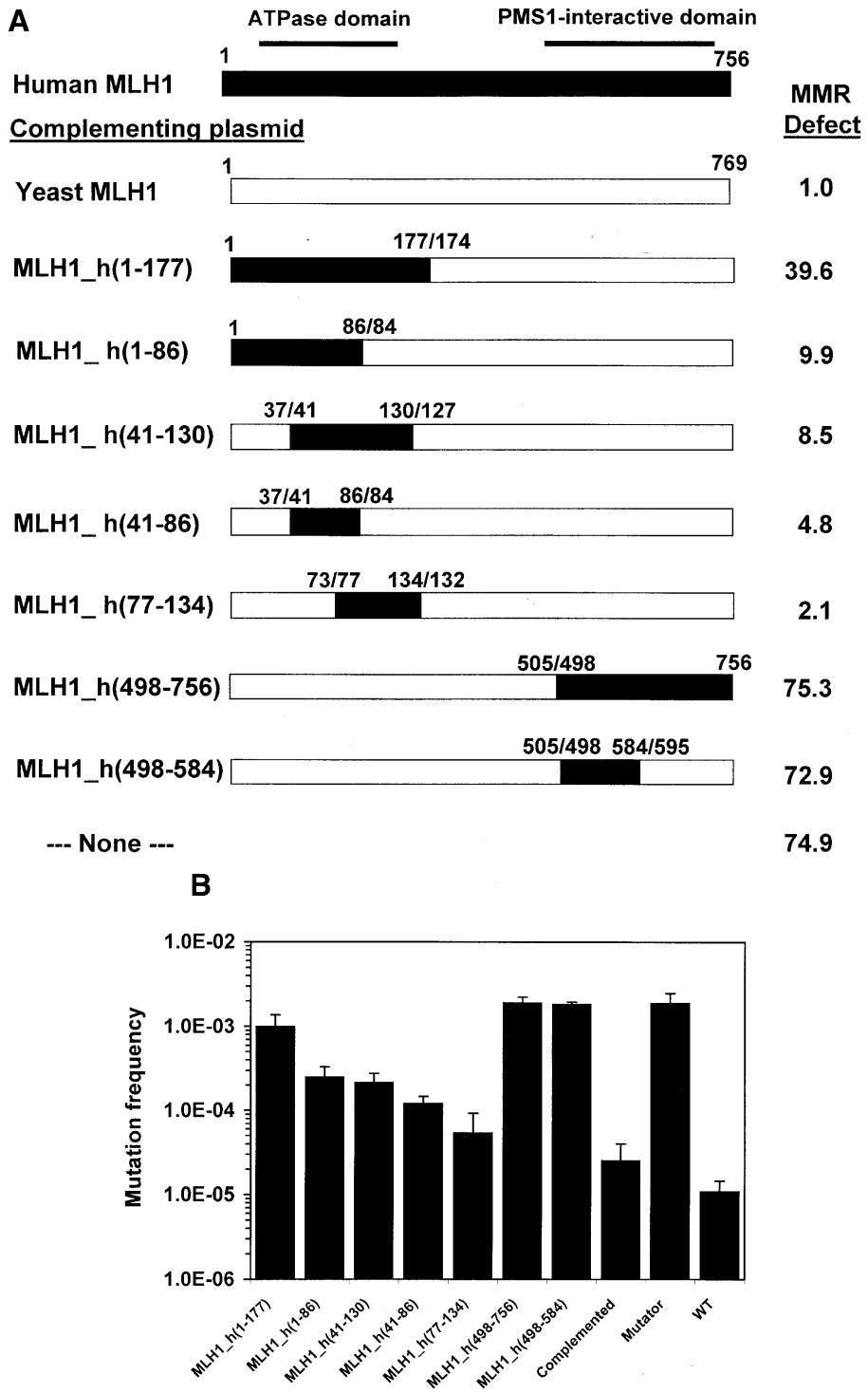

Figure 4. Structure and function of hybrid human–yeast MLH1 proteins. (**A**) Schematic representation of the seven hybrid MLH1 proteins in comparison to full-length native human and yeast MLH1. Portions of the hybrid protein representing human sequences are represented by solid bars. Numbers above each bar indicate the amino acid residue of the human portion of each gene. For hybrid genes where the fusion is within the protein coding region, the number of the flanking yeast residue is also indicated. The MMR defect (normalized to the strain expressing wild-type *MLH1*; Materials and Methods) is listed to the right of each protein. (**B**) Mutation frequencies of yeast strains expressing native and hybrid MLH1 proteins. Each hybrid gene was expressed in YBT24 containing pSH91 and assayed for MMR activity as described in Materials and Methods. The mutant frequencies are presented as the mean ± SD of at least four replicate cultures of a single yeast clone that expresses the indicated hybrid protein. Mean mutation frequencies are: WT, 1.1×10^{-5}; Mutator, 1.88×10^{-3}; Complemented, 2.5×10^{-5}; MLH1_h(1–177), 9.95×10^{-4}; MLH1_h(1–86), 2.49×10^{-4}; MLH1_h(41–130), 2.14×10^{-4}; MLH1_h(41–86), 1.20×10^{-4}; MLH1_h(77–134), 5.4×10^{-5}; MLH1_h(498–756), 1.89×10^{-3} ; MLH1_h(498–584), 1.8×10^{-3}.

There is very good agreement between the results of our functional genetic tests (Table 2) and available human clinical data concerning the amino acid replacement (reviewed in ref. 5; see also http://www.nfdht.nl and references therein).

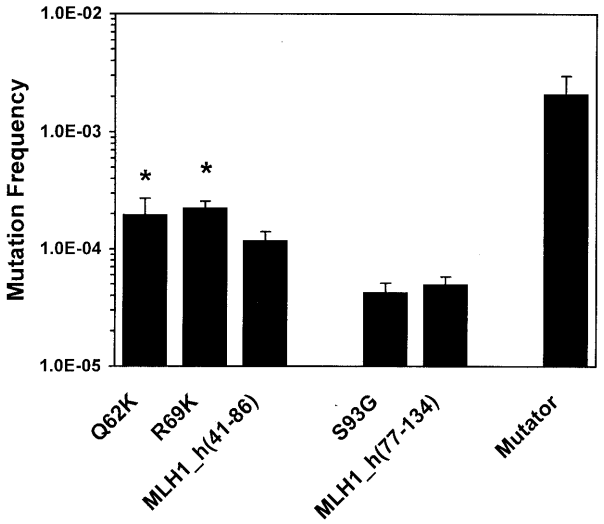

Figure 5. Mutation frequencies conferred by missense codons in human–yeast hybrid MLH1 genes. MLH1 Q62K and R69K alterations were made in pMLH1_h(41–86). MLH1 S93G was made in pMLH1_h(77–134). The resulting constructs were introduced into YBT24 containing pSH91 (Mutator) and mutation frequencies determined as described in the Materials and Methods section. The mutator strain complemented with the parental pMLH1_h(41–86) and pMLH1_h(77–134) constructs served as controls. *, a mutation frequency which is intermediate between the mutator strain and the strain complemented with the human–yeast hybrid MLH1. Mutation frequencies are from a representative experiment and are presented as the mean ± SD of five replicate cultures of a single yeast clone that expresses the indicated missense alteration. Mean mutation frequencies are: Mutator, 2.09×10^{-3}; Q62K, 1.95×10^{-4}; R69K, 2.22×10^{-4}; MLH1_h(41−86), 1.17×10^{-4}; S93G, 4.2×10^{-5}; MLH1_h(77−134), 4.9×10^{-5}.

Human variants corresponding to the yeast *MLH1* A41F, I65N and I104R replacements were identified in HNPCC kindreds (families that satisfied the Amsterdam criteria) (6) and shown to segregate with disease in these families (22–27). When these codon changes were introduced to the yeast gene, the variant protein was non-functional in MMR (mutant defect equivalent to the *mlh1*Δ-null strain without a complementing plasmid), demonstrating these amino acid replacements to be loss-of-function mutations. Human variants corresponding to the yeast *MLH1* G64R, T114R, Q552L and R672P replacements were identified in HNPCC kindreds (23,26,28–30) but no data was available regarding segregation with disease. These amino acid replacements were shown here to be loss-of-function mutations. The human variant corresponding to the yeast *MLH1* E99K alteration was entered as an unpublished observation in the ICG-HNPCC mutation database without clinical information. This alteration is a loss-of-function mutation (Results).

The GenBank entry for human MLH1 has a V at amino acid position 219 while both I and L have been reported as common polymorphisms at this position with a population incidence which ranges from 31 to 83% in different geographic regions (27,30,31). The high incidence in the population and lack of linkage to disease make it likely these alterations are silent polymorphisms. However, until this time no data on the *in vivo* function of these variant proteins was available. The native yeast gene contains V at this position and we demonstrated that a yeast MLH1 protein with a V216I alteration retained full MMR function. Thus, an amino acid alteration suspected to be a polymorphism in humans was confirmed as a silent polymorphism in the functional studies reported here.

Four different amino acid substitutions (R214C, R265C, R265H, I326A) resulted in MLH1 proteins that are functional, but at reduced efficiency. That is, the efficiency of DNA MMR in strains expressing these variants is intermediate between the *mlh1*Δ-null mutant and the null mutant complemented with the wild-type yeast gene. The R265H variant was observed in an HNPCC family that satisfied the Amsterdam criteria and cosegregated with disease (32). However, this substitution occurred in the same allele, which contained a frameshift and may have segregated with, but been unrelated to, cancer progression. The V326A allele in humans was identified in HNPCC kindreds (23,33). The yeast gene encodes an I at this position, and substitution of a V is a silent polymporphism (Results). Replacement of this amino acid with an A, however, results in a protein with reduced activity. The R265C variant was reported as a pathogenic mutation in the ICG-HNPCC database, citing unpublished data. The R214C replacement was reported in two 'suspected' HNPCC individuals from separate families that did not satisfy the Amsterdam criteria (16,34). The data reported here demonstrate that certain amino acid replacements result in partial inactivation of DNA MMR, and such decreased efficiencies of MMR can be associated with early onset colon cancer. We refer to these substitutions as efficiency polymorphisms.

The G322D alteration in *hMSH2* was initially identified in a family that satisfied the Amsterdam criteria (29). A number of subsequent studies, however, identified this change in 1–6% of both sporadic cancers and unaffected controls (31,35) suggesting that it might be a common polymorphism which cosegregated with another mutation in the HNPCC family. In our experiments, the yeast G317D allele was functional since the mutation frequency was reduced >100-fold when the variant was expressed in the *msh2*Δ-null mutant, although the variant exhibited a slightly reduced efficiency of MMR compared with the wild-type yeast MSH2p. In a previous study (36), the yeast G317D allele partially complemented an *msh2*Δ-null mutant when expressed at high levels from a *GAL10* promoter, but did not provide any complementation when expressed from the native *MSH2* promoter. Cumulatively, these results indicate that the G317D allele is an efficiency polymorphism and that *in vivo* function of this variant may be sensitive to the levels of expression.

The finding that the yeast protein with the A694T replacement has full MMR activity was unexpected. This missense codon was found in three individuals from separate HNPCC kindreds, and was reported to segregate with disease in these families (37). There are a number of possibilities why a clinical association with disease did not correlate with our functional data. First, there is the possibility that we did not target the correct codon in the yeast gene. In this 26 amino acid region of yeast MLH1p, the alanine represented only one of three amino acids which is perfectly conserved in the human protein. The computer-generated alignment, therefore, may be insufficient to unequivocally assign corresponding amino acids in this region. Utilization of functional hybrids incorporating human segments in this region of MLH1p will overcome the limitations of computer-generated alignments. A second possibility is that the A694T variant retains the ability to repair mismatches in a (GT) microsatellite but is ineffective in repairing other types of mismatches. However, it was recently reported, using the *lys2::InsE-A*$_{14}$ reporter gene, that the A694T MLH1p variant is functional in

Table 2. Comparison of functional and clinical data concerning MMR gene variants

Variation in human MMR gene		Equivalent subsitution in yeast	Patient satisfies Amsterdam-criteria for HNPCC	Mutation segregates with disease	Functional classification[a] (results)	Other MMR functional data
MLH1	S44F	A41F	Yes	Yes	Mutant	Mutant (40)
	–	A41S	–	–	Polymorphism	Polymorphism (40)
	Q62K	–	Yes	?	Δ efficiency	–
	G67R	G64R	Yes	No	Mutant	Mutant (40,50,52)
	I68N	I65N	Yes	?	Mutant	Mutant (40,50)
	R69K	–	No	?	Δ efficiency	–
	S93G	–	No	?	Polymorphism	–
	E102K	E99K	?	?	Mutant	–
	I107R	I104R	Yes	?	Mutant	–
	T117R	T114R	Yes	No	Mutant	Mutant (40)
	R217C	R214C	No	?	Δ efficiency	–
	I219V	V216I	No	No	Polymorphism	–
	R265C	R265C	?	?	Δ efficiency	–
	R265H	R265H	Yes	Yes	Δ efficiency	–
	V326A	I326A	Yes	No	Δ efficiency	–
	–	I326V	–	–	Polymorphism	–
	Q542L	Q552L	Yes	No	Mutant	–
	R659P	R672P	Yes	No	Mutant	–
	A681T	A694T	Yes	Yes	Polymorphism	Polymorphism (38)
MSH2	G322D	G317D	Yes/No	?	Δ efficiency	Δ efficiency (51)
	P622L	P640L	Yes	Yes	Mutant	Mutant (51)
	H639Y	H658Y	No	No	Polymorphism	–

[a]Mutant, the gene conferred a mutation frequency that was not significantly different from the chromosomal null mutant lacking a complementing plasmid; Polymorphism, the gene conferred a mutant frequency that was not different from the the null mutant strain that was complemented with the wild-type yeast gene; Δ efficiency, the gene confered a mutant frequency that is intermediate between the null mutant and the null mutant complemented with the wild-type yeast gene (i.e. efficiency polymorphism).
?, no information available.

MMR, is capable of interacting with PMS1p, and does not affect the mutator phenotype observed by overexpression of wild-type MLH1p (38). Thirdly, the A694T replacement may affect a function of MLH1 in humans that is not involved in MMR. Fourthly, it is possible that the sequence alteration in humans functions as a mutation by disrupting the normal gene expression process. For example, it has recently been demonstrated that some missense codons, as well as translationally silent point mutations, can cause exon skipping during the process of splicing (39). Finally, the clinical data may be incomplete or misleading. In two reports of this variant (http://www.nfdht.nl), the data is unpublished and thus impossible to evaluate. In the study of Froggatt *et al.* (37), the codon was present in all five affected relatives of an HNPCC family but not in 53 normal controls. However, these investigators did not sequence the entire *hMLH1* cDNA. Thus, this codon change could be a silent polymorphism with the actual mutation in this family due to a sequence alteration elsewhere in the gene.

In the experiments reported here, all variant genes were analyzed in the same yeast strain, utilizing the same reporter gene and identical gene expression conditions. The only difference in expression of the variant MMR genes was the altered codon. We did not evaluate steady-state protein levels, and thus did not determine whether some alterations affect protein stability. The *MLH1* efficiency polymorphisms (R214C, R265C, R265H, I326A) were also tested in the quantitative MMR assay at temperatures of 25 and 35°C. Both the mutation frequencies and MMR defects were identical to those observed at 30°C (A.R.Ellison and G.A.Bitter, unpublished data), suggesting that the decreased efficiency of these variants is not due to decreased protein stability. Regardless of whether the amino acid replacement directly affects the process of MMR, or indirectly affects MMR through altered stability of the protein, the consequence for the cell is an increased mutation rate.

Our results demonstrate the feasibility of expressing hybrid human–yeast MMR proteins that are functional *in vivo*. Two hybrids that incorporated large regions from the C-terminal end of human MLH1 were not active in MMR. The C-terminal portion of yeast MLH1p has been shown to interact with PMS1p (40). It is possible that the regions of the human protein (greater than 86 amino acids; Results) that were incorporated into the hybrids do not functionally interact with yeast PMS1p and/or MLH3p. Based on the results with the N-terminal hybrids, it is expected that utilization of smaller human coding sequences in the C-terminus will result in functional hybrids. Human–yeast hybrid proteins containing N-terminal regions of human MLH1 were functional and were used to determine the effect of missense replacements directly in the human coding sequence. In general, the efficiency of the hybrid protein in

DNA MMR was inversely correlated with the length of the human segment. The most efficient hybrid human–yeast MLH1p was functional in MMR at an efficiency within a factor of two of the native yeast protein (Fig. 4). The potential for making a series of hybrids to examine a wider range of human amino acid variants is now established.

Using hybrid human–yeast MLH1 proteins, we have been able to investigate the functional consequences of codon changes observed in clinical studies at residues not conserved in the yeast gene and which, until now, have lacked biochemical characterization. One of these changes (S93G) was a silent polymorphism and two (Q62K, R69K) were efficiency polymorphisms. Our results concerning Q62K and R69K are further examples of missense codons that may partially disrupt MMR and, if present in the germline, predispose an individual to colon cancer yet have incomplete penetrance in a family. The fact that S93G is an apparent silent polymorphism indicates that missense codons that become evident upon gene sequencing should be interpreted with caution until a more complete analysis, including extensive kindred analysis and/or functional data, is available.

A surprising finding in this study was that certain missense codons do not inactivate MMR, but result in lower efficiency of DNA MMR. Some of these amino acid replacements had a weak clinical association with cancer development. The substitutions were observed as germline mutations in individuals that developed CRC but whose families did not satisfy the criteria of HNPCC. These observations raise the intriguing possibility that differences in the efficiency of DNA MMR exist between individuals in the population due to common polymorphisms. Certain polymorphisms could exhibit weak penetrance yet still predispose to cancer development. If true, this would predict specific, genetically determined differences in susceptibility to cancer development. The systems described in this report will be useful for determining whether such genetic variation in the population exists and is associated with cancer development. Elucidation of such relationships may facilitate the future implementation of appropriate preventive strategies.

Genetic testing of individuals at a high risk of developing certain hereditary conditions is a powerful emerging strategy for the prevention of disease. However, a persistent ambiguity will arise regarding the functional significance of missense codons identified by gene sequencing. If sufficient biochemical, clinical and population data are lacking, then it becomes impossible to state with confidence if a sequence variation is pathogenic or simply a natural variation in the human population. Systems for assessing the *in vivo* significance of amino acid replacements will increase the effectiveness of genetic testing programs.

MATERIALS AND METHODS

MLH1 and MSH2 mutations and polymorphisms

Human *hMLH1* and *hMSH2* missense codons that we examined in functional assays were previously reported in public databases maintained on-line by the ICG-HNPCC (http://www.nfdhtl.nl), Human Gene Mutation Database (http://www.uwcm.ac.uk) and SWISS-PROT (http://www.expasy.ch). The databases are partially overlapping, and several other missense codons were reported in publications and review articles as noted in the text.

Bacterial strains, growth conditions and plasmid expression vectors

Eschericia coli strains JM109, DH5α and XL1-Blue (Stratagene, La Jolla, CA) were used for construction and amplification of plasmids. Unless otherwise described, standard bacterial growth conditions and gene cloning methods were employed (41). Yeast centromeric expression vector pMETc contains a HIS3 selectable marker and a multicloning site positioned between the MET25 promoter and CYC1 terminator (p413MET25) (42). Plasmid pSH91 is a yeast expression vector carrying the *URA3* coding sequence containing an in-frame $(GT)_{16}$ tract (43).

Yeast *MLH1* and *MSH2* expression vectors

Expression of the yeast *MSH2* gene from plasmid pMETc was previously demonstrated to complement *msh2* chromosomal mutations (19). For expression of MLH1p, the *MLH1* gene coding region plus 1.5 kb of 5′ flanking DNA was amplified by PCR from *S.cerevisiae* strain 288C genomic DNA (Promega, Madison, WI) with primers 5′-GCG CGA GCT CCC TCT AGA GAG TTA GAT GAG CC-3′ (B496-1) and 5′-GCG CCT CGA GGG TAT TAC AGC CAA AAC G-3′ (T941-6). The primers introduced a *Sac*I restriction site 5′ to the *MLH1* promoter region and an *Xho*I restriction site at 3′ to the *MLH1* coding sequence. The 3.9 kb PCR product was restricted with *Sac*I and *Xho*I and cloned between the unique *Sac*I and *Xho*I sites of pMETc. This construction deleted the MET25 promoter and placed the expression of *MLH1* coding sequences under control of the *MLH1* promoter.

Construction of hybrid human–yeast MLH1 genes

Genes encoding hybrid human–yeast MLH1 proteins were constructed by replacing portions of the yeast *MLH1* gene with the homologous human coding sequence from the *hMLH1* gene. The hybrid constructions were designed to maintain the open reading frame and precisely substitute conserved regions between the human and yeast proteins. DNA sequencing was performed to verify correct clones. DNA sequences of the cloned human fragments were found to exactly match the GenBank report (accession no. U07418).

Hybrid MLH1-h(1–177) was constructed by overlap extension PCR using previously described strategies (44). A 591 bp fragment containing the yeast *MLH1* 5′ regulatory region (nucleotides –586 to +6 relative to A of translation initiator codon) was amplified from *S.cerevisiae* 288C genomic DNA using primers 5′-GAC GAC ACA ATG CCA TAT AGG-3′ (D650-1) and 5′-AGA CAT TGC TTA TTG ATA GG-3′ (D650-2). Codons 3–177 of human MLH1 were amplified from a *hMLH1* cDNA clone (ATCC no. 217884 ; the first two yeast and human amino acids are identical) using primers 5′-CCT ATC AAT AAG CAA TGT CTT TCG TGG CAG GGG TTA TTC G-3′ (D650-3) and 5′-CGT ATC GCC CGA CAA CAT CCA AAA TTT TCC CAT ATT CTT C-3′ (D650-4). A 1.8 kb C-terminal portion of yeast *MLH1*, including codons 174–769, was amplified with primers 5′-GAT GTT GTC GGG CGA TAC G-3′ (D173-5) and T941-6. Approximately equimolar ratios of the three fragments were combined and amplified with the D650-1 and T941-6 primers. The resulting 2.9 kb overlap extension product was gel purified, digested with *Afl*II

and *Xho*I and cloned into plasmid pMLH1, replacing the native *MLH1* gene.

Hybrid MLH1_h(1–86) was constructed using a two-piece overlap extension reaction. A fragment of the human *hMLH1* cDNA containing codons 3–86 was amplified by PCR using primers D650-3 and 5′-CAG TTT AGA CGT CGT GAA CCT TTC ACA TAC-3′ (E466-1). The 255 bp PCR product was diluted and mixed with an approximately equimolar amount of the 591 bp yeast *MLH1* 5′ regulatory region fragment (described above) and overlap extension PCR was carried out using primers D650-1 and E466-1. The 0.85 kb overlap extension product was digested with *Afl*II and *Aat*II and cloned into plasmid pMLH1, replacing the native yeast *MLH1* segment.

Hybrid gene MLH1_h(41–130) was constructed by cloning a 285 bp fragment of the human *hMLH1* cDNA containing codons 41–130 between the *Cla*I and *Nde*I sites in pMLH1 and replacing codons 38–126 of the native yeast *MLH1* gene. The human fragment was amplified using primers 5′-GAA CTG TAT CGA TGC AAA ATC CAC-3′ (E179-1) and 5′-GCC ATA TGA TGC TCT GTA TGC ACA CTT TC-3′ (D173-2), which introduce *Cla*I and *Nde*I sites, respectively. The PCR product was digested with *Cla*I and *Nde*I to allow in-frame cloning into the yeast *MLH1* gene which had been subcloned as a *Sac*I–*Xho*I fragment into pBluescript II (Stratagene). The hybrid MLH1 gene was subsequently recloned into pMLH1, replacing the native yeast gene.

Hybrid gene MLH1_h(41–86) was constructed by PCR amplification of a 140 bp fragment containing codons 41–86 of the *hMLH1* cDNA and direct cloning into the yeast *MLH1* gene between the *Cla*I and *Aat*II sites. The human segment was amplified with primers 5′-GAA CTG TAT CGA TGC AAA ATC CAC-3′ (E179-1) and E466-1, which introduce *Cla*I and *Aat*II sites, respectively. The PCR product was digested with *Cla*I and *Aat*II to allow in-frame cloning into the yeast *MLH1* gene in expression vector pMLH1.

Hybrid gene MLH1_h(77–134) was constructed in a two-piece overlap extension reaction. A 170 bp fragment of the *hMLH1* gene containing codons 77–134 was PCR amplified using primers 5′-GTT CAC GAC GTC TAA ACT GCA GTC C-3′ (E466-2) and 5′-GGG GCT TTC CAA CAT TTT TCC ATC TGA GTA AC-3′ (E466-4), which introduces an *Aat*II site at the 5′ end. A 1.9 kb fragment of yeast *MLH1* containing codons 132–769 was amplified by PCR using primers 5′-GGA AAA ATG TTG GAA AGC CCC-3′ (E466-3) and T941-6. The two fragments were gel purified, diluted, mixed in equimolar amounts and amplified using primers E466-2 and T941-6. The overlap extension product was digested with *Aat*II and *Xho*I to allow in-frame cloning into pMLH1, replacing the native yeast *MLH1* gene.

Hybrid gene MLH1_h(498–756) was constructed by direct cloning of a 829 bp fragment containing codons 498–756 of human *hMLH1* into the yeast gene. The human fragment was amplified by PCR using primers 5′-GAC CTA AGG AGA GAA GGA TCA TTA ACC TC-3′ (D173-3) and 5′-CAC CTC GAG AAA GAA GAA CAC ATC C-3′ (D173-4). These introduce a *Bsu*36I site at the 5′ end and an *Xho*I site at the 3′ end and allow in-frame cloning into the yeast *MLH1* gene in pMLH1 as a *Bsu*36I–*Xho*I fragment, replacing codons 506–769 of the yeast gene.

Hybrid gene MLH1_h(498–584) was constructed using a two-piece overlap extension reaction. A 290 bp fragment of *hMLH1* containing codons 498–584 was PCR amplified with primers D173-3 and 5′-GGA GAT TAT ACA ATA CGA TGT CAA ATA GCG GTG CTG GCT C-3′ (E179-6). A 560 bp fragment of yeast *MLH1*, containing codons 595–769, was amplified with primers 5′-ATC GTA TTG TAT AAT CTC C-3′ (E179-2) and T941-6. Approximately equimolar amounts of each fragment were mixed and subjected to overlap extension PCR using primers D173-3 and T941-6. The ~800 bp product was digested with *Bsu*36I and *Xho*I for replacement of the equivalent fragment in pMLH1.

Site-directed mutagenesis

Mutations were introduced into MMR genes using the QuikChange Site-Directed Mutagenesis kit (Stratagene) following the manufacturer's instructions. The protocol employs multiple rounds of synthesis with *Pfu* DNA polymerase using plasmid DNA as template, but no amplification of the *in vitro* product. Templates for the mutagenesis reaction were as follows: pMLH1 for yMLH1 variants A41F, A41S, G64R, I65N, E99K, I104R, T114R, R214C, V216I, R265C, R265H, I326A, I326V, Q552L, R672P and A694T; pMetc/MSH2 for yMSH2 variant G317D; pMLH1_h(41–86) for *hMLH1* variants Q62K and R69K; and pMLH1_h(77–134) for *hMLH1* variant S93G. Sense and antisense oligonucleotide primers were PAGE-purified and, to facilitate screening for mutant clones, included a silent restriction site change in addition to the desired missense alteration (Table 1). For all mutations, at least three independent mutant clones were tested for function in yeast with identical results. At least one clone that contained the appropriate restriction site alteration was sequenced on both strands over the region of interest to confirm the mutation and verify the native sequence over ≥100 bp on either side of the introduced mutation. The data in Results are derived from replicate cultures of a single mutant clone that had been confirmed by DNA sequence analysis.

PCR

Routine PCR was carried out with ~3–15 ng plasmid DNA or 400 ng genomic DNA using 2.5 U *Taq* DNA polymerase and 0.1 μM forward and reverse primers. When high fidelity PCR was required, such as for the production of DNA fragments for generating hybrid human–yeast genes, *Pfu* DNA polymerase (Stratagene) and PCR conditions recommended by the manufacturer were used.

DNA sequencing

DNA sequencing was performed at commercial sequencing facilities using ABI BigDye Terminator chemistry and ABI automated DNA sequencers (models 377 and 3700).

Yeast strains, growth conditions and transformations.

All yeast strains used in this study are derivatives of *S.cerevisae* YPH500 which has the genotype *MAT*α *ade2-101 his3-Δ200 leu2-Δ1 lys2-801 trp1-Δ63 ura3-5* (45). Strain YBT5-1 was described previously (19) and has the genotype *MAT*α *ade2-101 his3-Δ200 lys2-801 trp1-Δ63 ura3-52*.

Strain YBT24 (*MAT*α *ade2-101 his3-Δ200 leu2-Δ1 lys2-801 trp1-Δ63 ura3-52 mlh1Δ::LEU2*) contains a deletion of the entire *MLH1* coding region and was generated by chromosomal

targeting using a DNA fragment constructed by overlap extension PCR procedures (44). Briefly, nucleotides –140 to +6 of yeast *MLH1* (relative to the A of the translation initiator codon at +1) and nucleotides 2299 (termination codon at 2306) to 2684 were PCR amplified from *S.cerevisiae* S288C genomic DNA. The yeast *LEU2* gene coding region plus 440 bp of 5′ flanking and 40 bp of 3′ flanking DNA was also PCR amplified from *S.cerevisiae* S288C genomic DNA. The 5′ end of the *LEU2* 5′ primer was homologous to the *MLH1* upstream 3′ primer while the 5′ end of the *MLH1* downstream 5′ primer was homologous to the *LEU2* 3′ primer. Approximately equimolar amounts of each PCR product were mixed and subjected to overlap extension PCR using the outermost *MLH1* 5′ and 3′ primers. The resulting 2.2 kb *MLH1 5′-LEU2–MLH1-3′* fusion was transformed into *S.cerevisiae* strain YPH500 and leucine prototrophs were selected. Genomic DNA was isolated from one clone, strain YBT24, which exhibited a mutator phenotype using the pSH91 reporter gene and confirmed by PCR analysis to have the entire *MLH1* coding region deleted and replaced by the *LEU2* gene (data not shown). Strain YBT25 (*MATα ade2-101 his3-Δ200 leu2-Δ1 lys2-801 trp1-Δ63 ura3-52 msh2Δ::LEU2*) contains a deletion of *MSH2* from codon 2 through the termination codon. Generation, selection and confirmation of YBT25 was similar to construction of YBT24 except that a *MSH2 5′-LEU2–MSH2-3′* overlap extension product was used for gene targeting.

The MMR reporter plasmid pSH91 (43) was introduced into YBT5-1, YBT24 and YBT25, selecting for tryptophan prototrophs. The strains were maintained in SD medium supplemented with adenine, histidine and lysine. The additional selection for uracil prototrophy maintains the cultures with 100% of the pSH91 containing an in-frame (GT) tract. YBT24 and YBT25 containing pSH91 were also transformed with pMLH1 and pMetc/MSH2, respectively, and maintained on SD medium supplemented with adenine and lysine. Transformations were carried out by the polyethylene glycol-lithium acetate method (46). Yeast strains were stored at –80°C in 15% glycerol.

In vivo MMR assay

The standardized *in vivo* assay for DNA MMR has been described in detail elsewhere (19). The assay is based on instability of 33 bp $(GT)_{16}$ tract which is inserted in-frame in the 5′ end of the yeast *URA3* gene coding region in plasmid pSH91. The $(GT)_{16}$ microsatellite is unstable during DNA replication and, if insertion/deletion loops are not repaired by MMR, *ura3* mutants form due to frameshift mutations. Selection on plates containing 5-fluoroorotic acid (FOA) is used to quantitate *ura3* mutants (47). Briefly, yeast strains were grown for 24 h in SD media (0.67% yeast nitrogen base without amino acids, 2% dextrose) supplemented with adenine and lysine. An additional supplement of histidine was included for strains that did not carry MLH1 or MSH2 expression vectors. An equivalent number of cells (~1×10^6) from each saturated culture was subcultured (1:200) into fresh media (20 ml) containing the same supplements as above plus uracil, and grown for an additional 24 h. The presence of uracil allows growth of any newly formed *ura3* mutants in this culture. After overnight growth to saturation, aliquots (25 and 100 µl) of each culture were plated on SD plates containing adenine, lysine, uracil and 1 mg/ml FOA to determine the concentration (cells/ml) of *ura3* mutants. The concentration (cells/ml) of total viable plasmid-containing cells was determined by serial dilution and plating on SD plates supplemented with adenine, lysine and uracil. Additional supplementation with histidine was included if the strain did not carry MLH1 or MSH2 expression vectors. Colonies were counted 2–3 days later. Mutant frequencies were calculated by dividing the concentration of FOA-resistant colonies by the concentration of viable cells. At least four independent cultures of each yeast strain carrying variant MMR genes were assayed. The MMR defect is defined as the ratio of the mutation frequency in the test strain divided by that observed in the same strain complemented with the wild-type yeast gene.

Statistical analysis

Differences in mutant frequencies between yeast strains were assessed by ANOVA/multiple *t*-tests (Statview 4.5). *P*-values of <0.05 were considered statistically significant.

ACKNOWLEDGEMENTS

We thank Nevin Sabet for technical assistance in the early stages of this project. This work was supported by a research grant from the National Institutes of Health to G.A.B. (R43 CA-81965).

REFERENCES

1. Vasen, H.F.A. (1994) What is hereditary nonpolyposis colorectal cancer (HNPCC). *Anticancer Res.*, **14**, 1613–1616.
2. Papadopoulos, N. and Lindblom, A. (1997) Molecular basis of HNPCC: mutations of MMR genes. *Hum. Mutat.*, **10**, 89–99.
3. Peltomaki, P., Vasen, H. and The International Collaborative Group on HNPCC (1997) Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. *Gastroenterology*, **113**, 1146–1158.
4. Kinzler, K.W. and Vogelstein, B. (1996) Lessons from hereditary colorectal cancer. *Cell*, **87**, 159–170.
5. Peltomaki, P. and de la Chapelle, A. (1997) Mutations predisposing to hereditary nonpolyposis colorectal cancer. *Adv. Cancer Res.*, **71**, 93–119.
6. Vasen, H.F., Mecklin, J.P., Khan, P.M. and Lynch, H.T. (1991) The International Collaborative Group on Hereditary Non-Polypolposis Colorectal Cancer (ICG-HNPCC). *Dis. Colon Rectum*, **34**, 424–425.
7. Fishel, R. and Wilson, T. (1997) *MutS* homologs in mammalian cells. *Curr. Opin. Genet. Dev.*, **7**, 105–113.
8. Jiricny, J. and Nystrom-Lahti, M. (2000) Mismatch repair defects in cancer. *Curr. Opin. Genet. Dev.*, **10**, 157–161.
9. Kolodner, R.D. and Marisischky, G.T. (1999) Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.*, **9**, 89–96.
10. Gradia, S., Acharya, S. and Fishel, R. (1997) The human mismatch recognition complex hMSH2-hMSH6 functions as novel molecular switch. *Cell*, **91**, 995–1005.
11. Gradia, S., Acharya, S. and Fishel, R. (2000) The role of mismatched nucleotides in activating the hMH2-hMSH6 molecular switch. *J. Biol. Chem.*, **275**, 3922–3930.
12. Hickman, M.J. and Samson, L.D. (1999) Role of DNA mismatch repair and p53 signaling in induction of apoptosis by alkylating agents. *Proc. Natl Acad. Sci. USA*, **96**, 10764–10769.
13. Wu, J., Gu, L., Wang, H., Geacintov, N.E. and Li, G.-M. (1999) Mismatch repair processing of carcinogen-DNA adducts triggers apoptosis. *Mol. Cell. Biol.*, **19**, 8292–8301.
14. Gong, J.G., Costanzo, A., Yang, H.-Q., Melino, G., Kaelin, W.G., Levrero, M. and Wang, J.Y.J. (1999) The tyrosine kinase c-Abl regulates p73 in apoptotic response to cisplatin-induced DNA damage. *Nature*, **399**, 806–809.

15. Yuan, Y., Han, H.-J., Zheng, S. and Park, J.-G. (1998) Germline mutations of hMLH1 and hMSH2 genes in patients with suspected hereditary non-polyposis colorectal cancer and sporadic early-onset colorectal cancer. *Dis. Colon Rectum*, **41**, 434–440.
16. Han, H.-J., Yuan, Y., Ku, J.-L., Oh, J.H., Won, Y.J., Kang, K.J., Kim, K.Y., Kim, S., Kim, C.Y., Kim, J.-P. *et al.* (1996) Germline mutations of *hMLH1* and *hMSH2* genes in Korean hereditary nonpolyposis colorectal cancer. *J. Natl Cancer Inst.*, **88**, 1317–1319.
17. Herman, J.G., Umar, A., Polyak, K., Graff, J.R., Ahuja, N., Issa, J.P., Markowitz, S., Willson, J.K.V., Hamilton, S.R., Kinzler, K.W. *et al.* (1998) Incidence and functional consequences of *hMLH1* promoter hypermethylation in colorectal carcinoma. *Proc. Natl Acad. Sci. USA*, **95**, 6870–6875.
18. Kolodner, R.D. (2000) Guarding against mutations. *Nature*, **407**, 687–689.
19. Polaczek, P., Putzke, A.P., Leong, K. and Bitter, G.A. (1998) Functional genetic tests of DNA mismatch repair protein activity in *Saccharomyces cerevisiae*. *Gene*, **213**, 159–167.
20. Guerette, S., Acharya, S. and Fishel, R. (1999) The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer. *J. Biol. Chem.*, **274**, 6336–6341.
21. Shimodaira, H., Filosi, N., Shibata, H., Suzuki, T., Radice, P., Kanamaru, R., Friend, S.H., Kolodner, R.D. and Ishioka, C. (1998) Functional analysis of human *MLH1* mutations in *Saccharomyces cerevisiae*. *Nat. Genet.*, **19**, 384–389.
22. Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A. *et al.* (1994) Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature*, **368**, 258–261.
23. Buerstedde, J.-M., Alday, P., Torhorst, J., Weber, W., Muller, H. and Scott, R. (1995) Detection of new mutations in six out of ten Swiss HNPCC families by genomic sequencing of the hMSH2 and hMLH1 genes. *J. Med. Genet.*, **32**, 909–912.
24. Hackman, P., Tannergard, P., Osei-Mensa, S., Chen, J., Kane, M.F., Kolodner, R., Lambert, B., Hellgren, D. and Lindblom, A. (1997) A human compound heterozygote for two *MLH1* missense mutations. *Nat. Genet.*, **17**, 135–136.
25. Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M. *et al.* (1993) Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. *Cell*, **75**, 1215–1225.
26. Nystrom-Lahti, M., Wu, Y., Moisio, A.-L., Hofstra, R.M.W., Osinga, J., Mecklin, J.-P., Jarvinen, H.J., Leisti, J., Buys, C., de la Chapelle, A. and Peltomaki, P. (1996) DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. *Hum. Mol. Genet.*, **5**, 763–769.
27. Tannergard, P., Lipford, J.R., Koldner, R., Frodin, J.E., Nordenskjold, M. and Lindblom, A. (1995) Missense mutations in the *MLH1*gene in Swedish HNPCC families. *Cancer Res.*, **55**, 6092–6096.
28. Hutter, P., Couturier, A., Scott, R.J., Alday, P., Delozier-Blanchet, C., Cachat, F., Antonarakis, S.E., Joris, F., Gaudin, M., D'Amato, L. and Buerstedde, J.-M. (1996) Complex genetic predisposition to cancer in an extended HNPCC family with an ancestral hMLH1 mutation. *J. Med. Genet.*, **33**, 636–640.
29. Maliaka, Y.K., Chudina, A.P., Belev, N.F., Alday, P., Bochkov, N.P. and Buerstedde, J.-M. (1996) CpG dinucleotides in the hMSH2 and hMLH1 genes are hotspots for HNPCC mutations. *Hum. Genet.*, **97**, 251–255.
30. Moslein, G., Tester, D.T., Lindor, N.M., Honchel, R., Cunningham, J.M., French, A.J., Halling, K.C., Schwab, M., Goretski, P. and Thibodeau, S.N. (1996) Microsatellite instability and mutation analysis of *hMSH2* and *hMLH1* in patients with sporadic, familial and hereditary colorectal cancer. *Hum. Mol. Genet.*, **9**, 1245–1252.
31. Liu, B., Nicolaides, N.C., Markowitz, S., Willson, J.K.V., Parsons, R.E., Jen, J., Papadopoulos, N., Peltomaki, P., de la Chapelle, A., Hamilton, S.R., Kinzler, K.W. and Vogelstein, B. (1995) Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat. Genet.*, **9**, 48–55.
32. Viel, A., Genuardi, M., Capozzi, E., Leonardi, F., Bellacosa, A., Paravatou-Petsotas, M., Pomponi, M.G., Fornasarig, M., Percesepe, A., Roncucci, L. *et al.* (1997) Characterization of *MSH2* and *MLH1* mutations in Italian families with hereditary nonpolyposis colorectal cancer. *Genes Chromosomes Cancer*, **18**, 8–18.
33. Liu, B., Nicolaides, N.C., Markowitz, S., Parsons, R.E., Papadopoulous, N., Nicolaides, N.C., Lynch, H.T., Watson, P., Jass, J.R., Dunlop, M. *et al.* (1996) Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nat. Med.*, **2**, 169–174.
34. Miyaki, M., Konishi, M., Muraoka, M., Kikuchi-Yanoshita, R., Tanaka, K., Iwama, T., Mori, T., Koike, M., Ushio, K., Chiba, M., Nomizu, S. and Utsunomiya, J. (1995) Germ line mutations of *hMSH2* and *hMLH1* genes in Japanese families with hereditary nonpolyposis colorectal cancer (HNPCC): usefulness of DNA analysis for screening and diagnosis of HNPCC patients. *J. Mol. Med.*, **73**, 515–520.
35. Tomlinson, I.P.M., Beck, N.E., Homfray, T., Harocopos, C.J. and Bodmer, W.F. (1997) Germline HNPCC gene variants have little influence on the risk for sporadic colorectal cancer. *J. Med. Genet.*, **34**, 39–42.
36. Drotschmann, K., Clark, A.B., Tran, H.T., Resnick, M.A., Gordenin, D.A. and Kunkel, T.A. (1999) Mutator phenotypes of yeast strains heterozygous for mutations in the MSH2 gene. *Proc. Natl Acad. Sci. USA*, **96**, 2970–2975.
37. Froggatt, N.J., Brasset, C., Koch, D.J., Evans, D.G., Hodgson, S.V., Popnder, B.A. and Maher, E.R. (1996) Mutation screening of *MSH2* and *MLH1*mRNA in hereditary non-polyposis colon cancer syndrome. *J. Med. Genet.*, **33**, 726–730.
38. Shcherbakova, P.V., Hall, M.C., Lewis, M.S., Bennett, S.E., Martin, K.J., Bushel, P.R., Afshari, C.A. and Kunkel, T.A. (2001) Inactivation of DNA mismatch repair by increased expression of yeast *MLH1*. *Mol. Cell. Biol.*, **21**, 940–951.
39. Liu, H.-H., Cartegni, L., Zhang, M.Q. and Krainer, A.R. (2001) A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA1* and other genes. *Nat. Genet.*, **27**, 55–58.
40. Pang, Q., Prolla, T.A. and Liskay, R.M. (1997) Functional domains of the *Saccharomyces cerevisiae* Mlh1p and Pms1p DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations. *Mol. Cell. Biol.*, **17**, 4465–4473.
41. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
42. Mumberg, D., Muller, R. and Funk, M. (1994) Regulatable promoters of *Saccharomyces cerevisiae*: comparison of transcriptional activity and their use for heterologous expression. *Mol. Gen. Genet.*, **22**, 5767–5768.
43. Strand, M., Prolla, T.A., Liskay, R.M. and Petes, T.D. (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature*, **365**, 274–276.
44. Bitter, G.A. (1998) Function of hybrid human-yeast cyclin dependent kinases in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.*, **260**, 120–130.
45. Sikorski, R.S. and Hieter, P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*, **122**, 19–27.
46. Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.*, **153**, 163–168.
47. Boeke, J.D., Lacroute, F. and Fink, G. (1984) A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.*, **197**, 345–346.
48. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
49. Tran, P.T. and Liskay, M.R. (2000) Functional studies on the candidate ATPase domains of *Saccharomyces cerevisiae* MutL. *Mol. Cell. Biol.*, **20**, 6390–6398.
50. Shcherbakova, P.V. and Kunkel, T. (1999) Mutator phenotypes conferred by *MLH1* overexpression and by heterozygosity for *mlh1* mutations. *Mol. Cell. Biol.*, **19**, 3177–3183.
51. Drotschmann, K., Clark, A.B. and Kunkel, T.A. (1999) Mutator phenotypes of common polymorphisms and missense mutations in MSH2. *Curr. Biol.*, **9**, 907–910.
52. Jeyaprakash, A., Das Gupta, R. and Kolodner, R. (1996) *Saccharomyces cerevisiae pms2* mutations are alleles of *MLH1*, and *pms2-2* corresponds to a hereditary nonpolyposis colorectal carcinoma-causing missense mutation. *Mol. Cell. Biol.*, **16**, 3008–3011.