Frequent Germline Mutations and Somatic Repeat Instability in DNA Mismatch-Repair-Deficient *Caenorhabditis elegans*

Marcel Tijsterman, Joris Pothof and Ronald H. A. Plasterk[1]

Hubrecht Laboratory, Centre for Biomedical Genetics, 3584 CT, Utrecht, The Netherlands

Manuscript received January 14, 2002
Accepted for publication March 9, 2002

ABSTRACT

Mismatch-repair-deficient mutants were initially recognized as mutation-prone derivatives of bacteria, and later mismatch repair deficiency was found to predispose humans to colon cancers (HNPCC). We generated mismatch-repair-deficient *Caenorhabditis elegans* by deleting the *msh-6* gene and analyzed the fidelity of transmission of genetic information to subsequent generations. *msh-6*-defective animals show an elevated level of spontaneous mutants in both the male and female germline; also repeated DNA tracts are unstable. To monitor DNA repeat instability in somatic tissue, we developed a sensitive system, making use of heat-shock promoter-driven *lacZ* transgenes, but with a repeat that puts this reporter gene out of frame. In genetic *msh-6*-deficient animals *lacZ*$^+$ patches are observed as a result of somatic repeat instability. RNA interference by feeding wild-type animals dsRNA homologous to *msh-2* or *msh-6* also resulted in somatic DNA instability, as well as in germline mutagenesis, indicating that one can use *C. elegans* as a model system to discover genes involved in maintaining DNA stability by large-scale RNAi screens.

DNA mismatch repair (MMR) mutants were originally found in screens directed at the identification of bacterial mutants that had a mutator phenotype and thus had elevated levels of spontaneous mutants in their progeny. Subsequent genetic as well as biochemical studies identified the MMR machinery as an enzymatic complex that could identify DNA mismatches resulting from single-nucleotide substitutions or small insertions/deletions, recognize the parental from the newly synthesized strand, excise the new strand around the lesion, and initiate repair to close the gap (reviewed in Kolodner 1996; Modrich and Lahue 1996; Jiricny 1998).

One of the greatest success stories of model organism genetics came when a human syndrome of cancer predisposition, hereditary nonpolyposis colon cancer (HNPCC), was found to result from a defect in human homologs of genes encoding components in the bacterial MMR machinery (Fishel *et al.* 1993; Leach *et al.* 1993; Bronner *et al.* 1994; Liu *et al.* 1994; Nicolaides *et al.* 1994; Papadopoulos *et al.* 1994). The fact that these cancers are characterized by an increased instability of simple DNA repeats provided the first clue that a replication-associated repair mechanism was involved (Peinado *et al.* 1992; Aaltonen *et al.* 1993; Ionov *et al.* 1993; Peltomaki *et al.* 1993). The notion that MMR defects are associated with human cancer provides strong support for the hypothesis that a so-called mutator phenotype, here as a result of elevated levels of unrepaired somatic DNA mismatches, can promote tumorigenesis (Loeb 1991). This model has been further supported by mouse knockouts of the MMR genes *Msh2*, *Msh6*, *Pms2*, or *Mlh1* that show enhanced cancer frequencies and repeat instability (de Wind *et al.* 1995; Reitmair *et al.* 1995; Baker *et al.* 1996; Edelmann *et al.* 1996, 1997; Narayanan *et al.* 1997; Prolla *et al.* 1998).

Also in humans that do not contain germline mutations in DNA MMR genes, tumors are often found that display repeat instabilities. Upon analysis these are often defective in known components of the MMR machinery; either they carry mutations within the genes themselves or the expression of these MMR genes is epigenetically downregulated as a result of hypermethylation (Buermeyer *et al.* 1999 and references therein). However, not all sporadic human tumors with repeat instability show a defect in the known DNA MMR genes. In addition, in numerous HNPCC cases no germline mutations are found in the known MMR genes (Peltomaki and de la Chapelle 1997). This suggests that there are additional genes in humans whose loss results in this specific type of genetic instability. Indeed, recently, two other genes, *i.e.*, *MLH3* and *EXO1*, that are involved in the correction of mismatched bases (Szankasi and Smith 1995; Flores-Rozas and Kolodner 1998) have been found to be mutated in the germline of patients suffering from HNPCC (Lipkin *et al.* 2001; Wu *et al.* 2001a,b).

Relatively little is known about effects of DNA mismatch repair mutations upon germline mutagenesis in animals. Hypermutability of simple sequence repeats in human sperm can be measured (Bacon *et al.* 2001), and sperm of patients with HNPCC have been analyzed for aneuploidy frequencies (Martin *et al.* 2000). However, the spontaneous levels of mutagenesis in common

[1] *Corresponding author:* Hubrecht Laboratory, Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands. E-mail: plasterk@niob.knaw.nl

laboratory animals such as the mouse are too low to measure with a degree of precision. Here, we analyzed the effect of MMR on transmission of the genetic material to subsequent generations by studying MMR-defective *Caenorhabditis elegans* strains and found that the male as well as the female germline is protected against spontaneous mutagenesis. This was monitored by (i) the segregation of different visible mutant phenotypes, (ii) an enhanced level of mutations in essential genes in a balanced segment of the genome, and (iii) an enhanced level of mutations in a monitor gene (*unc-93*). In addition, we followed the fate of mono- and dinucleotide repeats in the *C. elegans* genome over many generations in wild-type as well as in MMR-deficient animals.

Tumorigenesis as a result of somatic DNA changes is not easily scored in mutator mutants, due to the fact that *C. elegans* lives only 2 weeks and has only 959 somatic cells. Therefore, we constructed transgenic animals that contained hundreds of tandemly repeated copies of an out-of-frame *lacZ* gene, reasoning that replication errors during development that would bring only one copy back in frame could be visualized by staining of such tissues using 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). To enhance the chances of a frameshift we cloned mono- or dinucleotide repeats between the initiation ATG and the β-galactosidase-encoding *lacZ* gene. Such constructs were found to be stable in wild-type animals (*i.e.*, no or almost no β-galactosidase expression observed), while clear positive patches were seen in virtually every individual animal in a *msh-6* genetic background. Identical phenotypes were observed after RNAi of *msh-6* and of *msh-2*. RNAi is the experimental silencing of the expression of a given gene by administration, in this case feeding, of double-stranded RNA corresponding to that gene (Fire *et al.* 1998). Thus feeding on dsRNA homologous to a MMR gene is sufficient to trigger DNA instability in somatic cells. This suggests that we can use *C. elegans* as a model system to discover genes involved in maintaining DNA stability by large-scale RNAi screens. These genes may be homologs of the yet unknown non-*msh* genes protecting DNA against mutations.

MATERIALS AND METHODS

Strains and maintenance: General methods for culturing *C. elegans* strains were as described in Brenner (1974). Strains used in this study were CB1500 [*unc-93(e1500)*], MT765 [*unc-93(e1500 n224)*], and BC1958 [*dpy-18(e364)/eT1* III; *unc-46(e177)/eT1* V]. A deletion mutant of *msh-6: pk2504* was isolated from a chemical deletion library as described (Jansen *et al.* 1997).

Spontaneous mutation frequency: Growing cultures of *msh-6* strains segregate a plethora of visible mutants indicative of a mutator phenotype. From the brood of four *msh-6* hermaphrodites, 300 progeny animals were picked that had a wild-type appearance. These worms were grown individually and the progeny were inspected for Mendelian segregation of visible phenotypes. Plates were screened a second time 2 days after food deprivation; this allows the scoring of an embryonic lethal phenotype, here interpreted as the abundant presence of dead eggs on the culture dish.

To determine whether *msh-6* animals have a high incidence of male (him) phenotype, the broods of three to five animals of genotype *msh-6* or wild type were inspected for the presence of a male: *msh-6*, 1/1209 (0.08%); wild type, 1/1059 (0.09%). The genetic recombination frequency was analyzed by determining the genetic distance between the visible marker *unc-32* and *dpy-18* on LGIII in an *msh-6* and wild-type genetic background. For animals of genotype *msh-6; unc-32 dpy-18/ + +*, the brood consisted of 412 wild type, 20 Unc, 21 Dpy, and 112 Unc Dpy, resulting in a recombination frequency of 0.075 (map distance 7.5 cM). In a mismatch-proficient genetic background the frequency was 0.072 (map distance 7.2 cM): A total of 527 wild type, 26 Unc, 24 Dpy, and 140 Unc Dpy segregated from animals of genotype *unc-32 dpy-18/+ +*.

The mutator phenotype of *msh-6 C. elegans* was quantified using the reciprocal translocation *eT1(III;V)* as a balancer, as described by Rosenbluth *et al.* (1983). First, *msh-6* males were crossed with hermaphrodites that were homozygous for the translocated *eT1* chromosomes (this genotype results in a visible phenotype because the translocation disrupts the *unc-36* locus). F_1 males were subsequently crossed with hermaphrodites of genotype *dpy-18; unc-46* (to mark the nontranslocated chromosomes) and cross-progeny of genotype *msh-6/+* I; *dpy-18/eT1* III; *unc-46/eT1* V were selected. Next generation animals homozygous for *msh-6* and segregating both Dpy-18 Unc-46 and Unc-36 animals were used as starting strains in the following experimental setup: Phenotypically wild-type progeny of hermaphrodites of the above-described genotype were picked onto individual plates and scored for segregation of the Dpy-18 Unc-46 phenotype. The frequency of recessive lethal mutations induced in the balanced area of the genome is reflected by the percentage of animals that fail to segregate this phenotype: A lethal mutation in the crossover-suppressed region of the canonical chromosomes prevents embryos homozygous for these chromosomes from developing into adult Dpy Unc worms. Clonal lines that were positive in this screen were grown further and confirmed as carrying a lethal mutation inside one of the crossover-suppressed regions if showing no such Dpy Unc markers in at least 250 offspring.

For determining the germline frequency in male sperm of *msh-6* animals, males of genotype *msh-6* I; *dpy-18/eT1* III; *unc-46/eT1* V were crossed to hermaphrodites of genotype *eT1(III/V)*. Phenotypically wild-type progeny were analyzed for segregation of the marked chromosomes as described above. The germline mutation frequency of hermaphrodite oocytes was determined by analyzing the phenotypically wild-type cross-progeny of *dpy-18/eT1; unc-46/eT1* males crossed to *msh-6; dpy-18; unc-46* hermaphrodites. In both crossing schemes, the *msh-6* deficient animals that were used to start the analysis were homozygous for more than one generation. Therefore, to prevent scoring mutations that occurred in earlier generations (that result in so-called "Jackpots") >30 cross-progeny animals were tested from a single hermaphrodite.

RNAi of *msh-6* and *msh-2* was done by injecting hermaphrodites of strain BC1958 with cognate dsRNA and subsequent analysis of the mutator phenotype in the phenotypically wild-type F_1. Thus the F_2 was inspected for segregation of the Dpy Unc phenotype. In addition, RNAi was measured by culturing BC1958 animals on *msh-2* or *msh-6* dsRNA-producing bacteria (described below).

Mutation spectrum of *msh-6* worms: Phenotypic reversion of the uncoordinated "rubber-band," egg-laying-defective phenotype conferred by *unc-93(e1500)* was used to determine the nature of mutations that occurred in a *msh-6* genetic background. Cultures started with single hermaphrodites of geno-

type *msh-6 unc-93(e1500)* were inspected regularly for revertants that were recognized by their wild-type movement and normal egg-laying behavior. Intragenic reversion events [mutations in at least four other loci can suppress the *unc-93(e1500)*-associated phenotype] were identified by the failure of these alleles to complement *unc-93(e1500n224)*. Subsequently, the coding region of the *unc-93* locus was sequenced.

Microsatellite repeat instability in *msh-6* worms: From a single hermaphrodite (*msh-6* and Bristol N2), 55 progeny were picked to start lines that were maintained by transferring several L4 animals every 3–4 days to new plates. After 10 generations DNA was isolated from cultures started with a single animal (due to the mutator phenotype of *msh-6*, mutations will accumulate and often a sterile phenotype is observed when individual animals are cloned out). From these cultures different genomic loci were analyzed by sequencing PCR products. Primers used are (5′-3′) as follows: R03C1_A, cggcaaaca atttttccg; R03C_C, acggaggtgttcacggag; F59A3_A, cgtttgaaggat gatgtc; F59A3_C, gatgctcgatgacttcgg; C41D7_A, gattctcaagtc cacccg; C41D7_C, gacccgttctcctactcc; M03F4_A, cgaaatggatc tgagtggg; M03F4_C, atatcccatgatgacccc; C24A3_A, gagtgcgc ttgaagagactg; C24A3_C, cggaactcggagagagatag; Y54G11A_A, ggatcttggctcctggaacg; and Y54G11A_C, cattgagtgatactcggccg.

Detection of somatic repeat instability: To allow detection of somatic repeat instability we created several constructs that contained stretches of either mono- or dinucleotide repeats between the start of translation and the *lacZ* open reading frame (ORF), under the control of a heat-shock promoter.

In brief, vector L2681 (kindly provided by A. Fire), which has a green fluorescent protein (GFP)/LacZ fusion under the control of a heat-shock promoter, was digested with *Bam*HI and allowed to close to create pRP1820; this cloning step removes two upstream ATG sequences without affecting essential promoter sequences. Then, the original starting codon was removed by site-directed mutagenesis to create pRP1821. This construct was subsequently used as a recipient for insertion of DNA fragments containing different types of repeats: Partially complementing oligonucleotides were annealed and inserted into a *Kpn*I site near the beginning of the fusion protein-encoded sequences. All constructs had a similar molecular architecture: heat-shock promoter-(*Kpn*I)-ATG-(A or CA)n-GFP/*lacZ* ORF (sequences and cloning details available upon request). The different types of repeat used in this study were pRP1822, $(A)_{16}$; pRP1823, $(A)_{17}$; pRP1840, $(A)_{15}$; pRP1841, $(CA)_{15}$; pRP1842, $(CA)_{14}$; and pRP1843, $(CA)_{13}$. pRP1822 and pRP1842 contain an in-frame *lacZ* construct encoding functional β-galactosidase.

All constructs were injected separately (together with pRF4 containing the dominant marker *rol-6*) into the canonical *C. elegans* strain Bristol N2 to establish transgenic lines (Mello *et al.* 1991). The transgenic array containing pRP1822 was integrated by γ-irradiation and used for further detailed analysis of somatic reversion events.

To identify expression of β-galactosidase, nematodes were fixed and stained with X-gal.

RNAi: *By injection:* PCR fragments of *msh-6* [Y47G6A, nucleotides (nt) 22458–23143] and *msh-2* (H26D21, nt 467–946, GenBank accession no. AF106587) coding sequences were cloned into vector pCCM114 (kind gift of Craig Mello) that contains oppositely oriented T7 promoters. Plasmid DNA was isolated, linearized, and used as template to synthesize dsRNA *in vitro* with T7 RNA polymerase (Boehringer Mannheim, Indianapolis) according to the manufacturer's conditions. Hermaphrodites were injected with 500 ng/μl dsRNA.

By feeding: msh-6 and *msh-2* DNA segments (identical sequences as described above) were cloned into the "feeding vector" L4440 and subsequently transformed to HT115 bacterial cells that were used for RNAi by feeding using the protocol described by Ahringer and co-workers (Fraser *et al.* 2000).

RESULTS

Mutator phenotype in mismatch-repair-defective *C. elegans*: We screened the genome sequence of *C. elegans* for homologs of bacterial and human DNA MMR genes and found *msh-2* and *msh-6* (H26D21.2 and Y47G6A.11, respectively; homologous to prokaryotic mutS) and *mlh-1* and *pms-2* (T28A8.7 and H12C20.2; homologous to prokaryotic mutL). Surprisingly, an ortholog of *msh-3* was not detected. We then knocked out the *msh-6* gene, using a mutant library approach previously developed in our laboratory (Jansen *et al.* 1997). Figure 1 shows the genomic organization of the *C. elegans msh-6* gene, the human and *Saccharomyces cerevisiae* homologs aligned to *msh-6* of *C. elegans*, and the deletion mutant that was used in this study.

Homozygous *msh-6* mutants are viable, and the first indication of a mutator phenotype was the frequent occurrence of readily recognizable mutants (such as *dpy* and *unc*) among the progeny. Since *C. elegans* lines are maintained as self-fertilizing hermaphrodites, spontaneous new mutations can homozygose in self-progeny, so that recessive mutations are easily observed. At least 20 phenotypic mutations could be found in the brood ($n = 300$) of phenotypically wild-type hermaphrodites. In the parental strain such a level of spontaneous mutations is not seen (0 mutant phenotypes in the brood of 300 cloned individuals). Apart from the increased mutation frequency in the *msh-6* mutant, no other phenotypes that are indicative of specific defects in genome stability were noted: X-chromosomal nondisjunction is not affected by the *msh-6* deletion, indicated by the absence of a high incidence of male (him) phenotype. Also, no effect was observed on genetic recombination: The genetic distance between visible markers is similar in wild-type and *msh-6* animals (see materials and methods for details).

Increased germline mutagenesis in *msh-6* defective *C. elegans* in both the male and female germline: To validate and quantify this mutator effect, we measured the frequency of recessive lethal mutations induced within a sizable region of the *C. elegans* genome (Rosenbluth *et al.* 1983; see materials and methods for details). In a wild-type strain we detect spontaneous mutations in this region below a frequency of 10^{-3}, which is in line with the numbers reported in the literature for wild-type *C. elegans*, *i.e.*, 6×10^{-4} (Rosenbluth *et al.* 1983). In *msh-6* mutants this level is 1.6×10^{-2} (Figure 2), at least 28 times elevated. These mutations could theoretically arise from mutations that occur uniquely in the sperm or in the oocytes of the hermaphrodite parent. In humans, mutation frequencies determined for microsatellite loci revealed a clear paternal excess; males mutate about five times as often as females (reviewed in Ellegren 2000). To test whether the MMR machinery protects the male and female germlines equally, we performed experiments that scored for

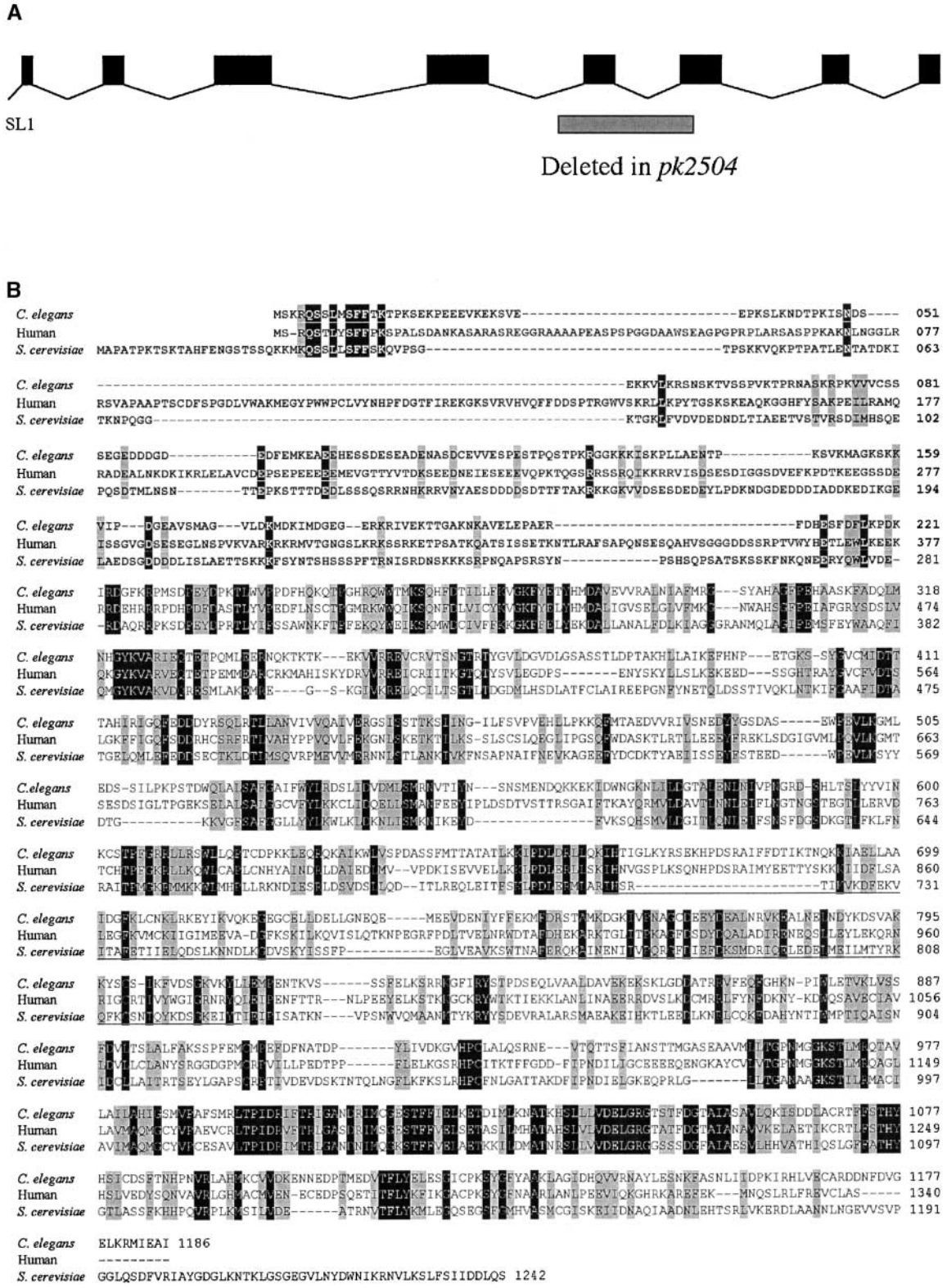

FIGURE 1.—The *C. elegans msh-6* gene. (A) Structure of the *C. elegans msh-6* gene deduced from genomic sequences and cDNA generated by reverse transcription-PCR from Bristol N2 RNA. The genomic region that is deleted in *pk2504* (nt 24180–25956 of Y47G6A, GenBank accession no. AC024791) and takes out exon 5 and part of exon 6 is indicated. (B) Alignment of the amino acid sequence of *C. elegans*, human, and *S. cerevisiae* MSH-6 using the CLUSTALW algorithm. Black background indicates amino acid identity and gray background indicates conserved amino acid substitutions. The amino acids deleted in *pk2504* are underlined. Possible alternative splicing of exon 4 onto exon 7 predicts an out-of-frame product.

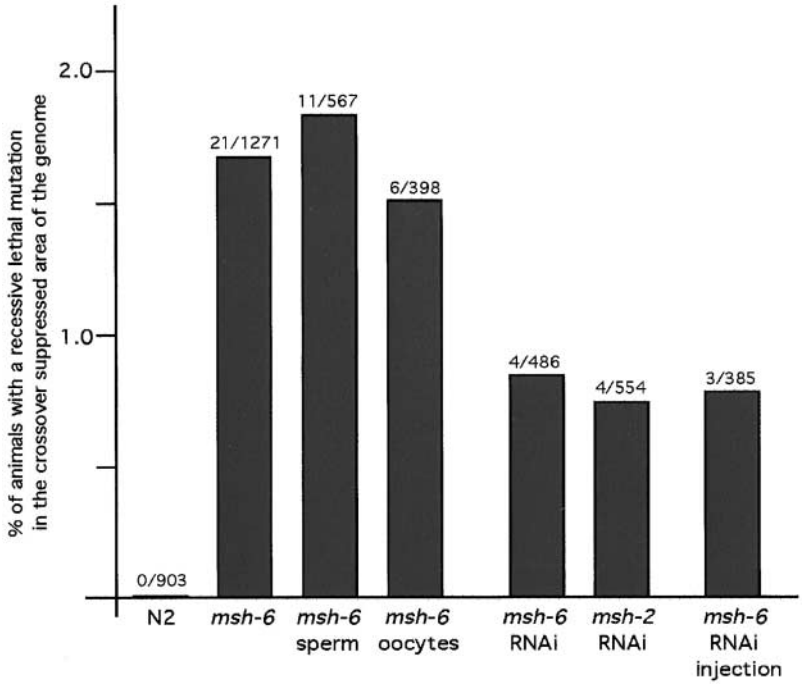

FIGURE 2.—Mismatch repair proteins MSH-6 and MSH-2 protect the *C. elegans* germline from spontaneous mutagenesis. The experimental setup that is used to measure the level of spontaneous mutagenesis is described in MATERIALS AND METHODS. This assay determines the absolute number of loss-of-function mutations in essential genes in a region that covers ~7% of the *C. elegans* genome (estimated number of target genes, ~300). The *y*-axis reflects the percentage of animals that acquire such a lethal mutation within one generation.

spontaneous mutants in progeny from crosses between males and hermaphrodites, in which either one of the parents was mutant and the other wild type for *msh-6* (see MATERIALS AND METHODS for details). As shown in Figure 2, both the oocytes of the hermaphroditic mother and the sperm from male fathers show a similar increase in the level of spontaneous mutagenesis in the *msh-6* mutant. We conclude two things: The frequency of original DNA replication errors is probably comparable in sperm and oocytes, and the level of protection by the MMR machinery is also similar. (Note that we did not test the sperm produced by hermaphrodites, since mutations there cannot be distinguished from those that arose in the oocytes of the same animal.)

As a second measure of mutation frequencies we scored for loss-of-function mutations in the *unc-93* monitor gene. Animals homozygous for the *unc-93(e1500)* mutation are uncoordinated and absolutely egg-laying defective, while complete loss of the *unc-93* gene has no visible phenotype, and thus loss-of-function mutants of the *unc-93(e1500)* gene can be scored by recognizing normally moving animals among contracted ones (GREENWALD and HORVITZ 1980). Therefore, this gene has been previously used to assay mutagenesis levels (DE STASIO *et al.* 1997). We found that the levels of mutations in *unc-93(e1500)* increase at least 30-fold in *msh-6* mutants compared to wild type (data not shown). The advantage of using the *unc-93* monitor gene is that once obtained, these mutants can also be identified at the molecular level by direct sequencing of the relatively small genomic *unc-93* gene. It is known that loss of four other genes (*sup-9*, *sup-10*, *sup-11*, and *sup-18*) also reverts the *unc-93(e1500)* phenotype, so we first sorted out the mutations that mapped to *unc-93* and sequenced only those (from 64 revertants that were independently isolated, 18 mapped to the *unc-93* locus). The nature of the mutations is shown in Table 1: Mostly we find frameshifts in short monomeric runs and single-base substitutions, which is similar to the spectrum seen in *S. cerevisiae MSH6*Δ strains (MARSISCHKY *et al.* 1996; GREENE and JINKS-ROBERTSON 1997) and in human cell lines defective for *hMSH6* (BHATTACHARYYA *et al.* 1994).

Microsatellite repeat instability in *msh-6*-defective *C. elegans*: As discussed in the Introduction, microsatellite instability is a hallmark of tumors derived from HNPCC patients [replication errors in repeats (RER^+)]. To see if *C. elegans* defective for *msh-6* display microsatellite instability, we started 55 parallel lines by cloning the progeny of one *msh-6* hermaphrodite. These lines were maintained for 10 generations, and then we picked one animal per line and sequenced various genomic loci containing microsatellites. As shown in Table 2, mono-, di-, and trinucleotide repeats become highly unstable in the absence of functional MSH-6, while no instability was observed in wild-type animals.

Visualization of somatic repeat instability in *C. elegans*: Having observed these frequent repeat length changes in the germline of *msh-6* mutants we wondered if these could also be observed in somatic cells. With a life span of only 2 weeks, and most somatic cells being only a few cell divisions removed from the zygote, one may not expect too many mutations. Therefore we devised a sensitive system for scoring repeat length instability. Inspired by the work of Petes and co-workers (HENDERSON and PETES 1992; STRAND *et al.* 1993), we cloned a repeat into a reporter gene in such a way that the repeat was

TABLE 1

Types of mutations that occur in a *msh-6* genetic background, using the *unc-93* gene as a mutational target

Type of mutation	Mutation	*unc-93* sequence and position	aa change
+1 frameshift	Insertion A	(221) TCGAGAA(A)TATTC (232)	
+1 frameshift	Insertion A	(232) CGAAAAA(A)CTTCG (243)	
+1 frameshift	Insertion A	(255) GCAAAAA(A)TTTGG (266)	
+1 frameshift	Insertion A	(255) GCAAAAA(A)TTTGG (266)	
+1 frameshift	Insertion A	(374) CCAAAAAA(A)GAAG (285)	
−1 frameshift	Deletion T	(359) AAGAGTTTTTCGAGG (373)	
Single-base-pair substitution	G → A	(789) ATTTAAC**G**GACTCCA (803)	Gly → Arg
	G → A	(1155) ACACTGC**G**GACAAGT (1169)	Gly → Arg
	G → A	(1551) TCTAGTT**G**GAGTTTA (1565)	Gly → Arg
	G → A	(1650) TTCCCTA**G**TCTTCGG (1664)	Val → Ile
	A → G	(1611) CTTTGTG**A**TGGCCTG (1625)	Met → Val
	A → C	(1492) AATATAA**A**GTTCATG (1506)	Lys → Thr
	G → C	(1707) CGGAGCA**G**TAGTGAA (1720)	Val → Leu
	T → G	(1578) CGTCGGA**T**GTGGCCT (1592)	Cys → Gly
	T → G	agg**t**ttcagAAAAAT (1439)	3′ splice site
Complex	—	(71) GTAGATC (77) to GTACGCATC	
Complex	—	(527) ATTTTTGC (534) to ACTTTTTC	Cys → Phe

aa, amino acid.

between the ATG initiation triplet and the domain of the gene encoding the enzymatic activity and would keep the latter out of frame. Unrepaired replication errors in the repeat will bring the gene in frame, which can be visualized using the reporter. To enhance the chances of finding such events, we took advantage of the fact that transgenes in *C. elegans* are usually tandem repeats of hundreds of copies of the injected DNA; one would hope that a frameshift in only one of those copies could be scored. Initial attempts to use GFP for this purpose failed. We then switched to using *lacZ*, which was present as a fusion protein in the same construct. A disadvantage of this reporter is that the animal needs to be impregnated with the reagent X-gal, which kills the animal. An advantage is that *lacZ* staining is more sensitive, especially because one can prolong the staining to increase the detection level.

Transgenic animals with the *lacZ* gene cloned "in frame" under the control of a heat-shock promoter show, as expected, a broad β-galactosidase expression pattern (Figure 3a). In contrast, in transgenic animals that have the *lacZ* gene out of frame, downstream of a repeat sequence, virtually no staining is observed after heat shock (Figure 3b). However, in *msh-6* mutant background, as shown in Figure 3c, almost every worm shows one or more blue patches. We conclude that these arise from repeat instability and restoration of the *lacZ* reading frame in different cell lineages. Unfortunately the fixed and stained worms have not allowed us to recognize specific sublineages, but we could see blue patches of multiple tissues.

To check the role of the repeat in this *msh-6*-dependent frameshift, we generated transgenic animals that contained identical constructs without the repeat and

TABLE 2

Microsatellite alterations in wild-type (N2) and MMR-deficient (*msh-6*) *C. elegans*

		msh-6						N2			
Repeat	Cosmid:	R03C1	F59A3	C41D7	M03F4	Y54G11A	C24A3	F59A3	M03F4	Y54G11A	C24A3
Repeat		$(A)_{15}$	$(A)_{15}$	$(CA)_{18}$	$(CA)_{18}$	$(GAG)_{15}$	$(CTT)_{13}$	$(A)_{15}$	$(CA)_{18}$	$(GAG)_{15}$	$(CTT)_{13}$
−1		2	3	7	5	1	2	0	0	0	0
0		38	42	32	34	43	44	44	44	48	47
+1		0	0	2	6	0	0	0	0	0	0
Total		40	45	41	45	44	46	44	44	48	47

Fifty-five subcultures of wild-type and *msh-6* were started and maintained for 10 generations by transferring three progeny animals (L4 stage) to new plates each generation. After 10 generations, several cultures were established from each line starting with a single animal. From 1 culture per line, DNA was isolated and several endogenous DNA repeats were amplified and analyzed by sequencing. The location on the *C. elegans* genome is indicated by the name of the corresponding cosmid. The primers that are used are listed in MATERIALS AND METHODS.

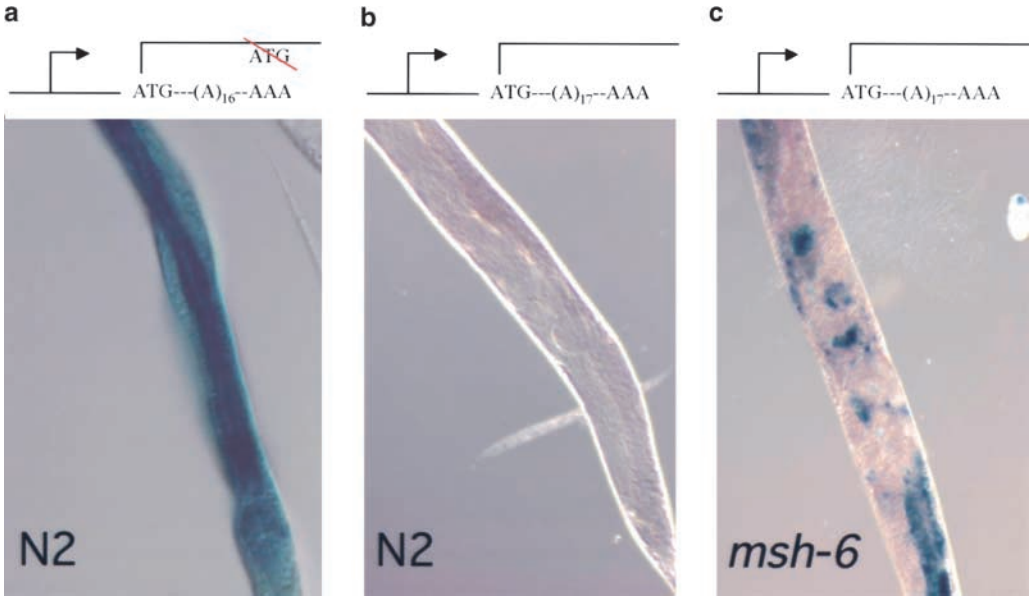

FIGURE 3.—Genetic instability in MMR-defective somatic cells. A schematic representation of the constructs that are used to measure somatic repeat instability is depicted above the images of the nematodes. (a) Transgenic *C. elegans* that carry multiple in-frame copies of heat-shock-driven *lacZ*. (b) MMR-proficient transgenic *C. elegans* (N2) that carry multiple copies of a *lacZ*-containing construct in which a repeat sequence is cloned immediately downstream of the ATG that puts the downstream-positioned β-galactosidase ORF out of frame. (c) The identical transgenic array crossed into an *msh-6* genetic background.

saw that the level of somatic cells that stain blue was not different in wild-type and *msh-6*-deficient animals (data not shown). This suggests that the increased staining that we observed for *msh-6* cells results from replication errors that occur at the repeat position, leading to one in-frame *lacZ* copy in that cell and its descendants.

Destabilizing the germline by feeding *msh-2* and *msh-6* dsRNA: RNA interference is the silencing of gene expression by administration of dsRNA that corresponds to exonic sequences of that gene (FIRE *et al.* 1998). One of the striking features of RNAi in *C. elegans* is that the dsRNA can be administered by feeding *Escherichia coli* that contain a plasmid that transcribes both strands, which combines to form dsRNA (TIMMONS and FIRE 1998). We fed worms on *E. coli* that produced dsRNA for *msh-6* and measured spontaneous mutation frequencies by scoring for mutants in the progeny. The results are shown in Figure 2: As for the genetic knockout, RNAi of *msh-6* resulted in an increased spontaneous mutation frequency. The administration of the *msh-6* dsRNA via another route, *i.e.*, direct injection of purified dsRNA into the gonad of the animal, had the identical outcome. We then also performed RNAi for another known MMR component, the *msh-2* gene, and found a comparable increase in mutation frequency (Figure 2).

Destabilizing the genetic contents of somatic cells by *msh-2* and *msh-6* RNAi: Combining the *lacZ* reporter with *msh-2* and *msh-6* RNAi, we fed dsRNA to worms and scored for restoration of the *lacZ* reading frame in somatic cells, and, as shown in Figure 4a, the effect is the same as that of the genetic null: Almost every animal has patches of β-galactosidase expression. The fact that we can phenocopy the genetic null mutation using RNAi allows us to analyze whether the staining was dependent on cell division as a marker for DNA replication. This prediction should be met if the patched phenotype is indeed the result of frameshift errors in somatic cells during development. We placed animals at different developmental stages on RNAi food and measured the level of staining after the animals developed to adulthood. We reasoned that because young larvae undergo more cell division cycles than mature animals, they are more prone to frameshift errors. Indeed, we found a reduction in the number of patches when the animals were exposed at later developmental stages (Figure 4b), consistent with a dependence on cell division.

We also tested a number of constructs that differed in the type or length of the repeat present between the ATG and the β-galactosidase coding regions (Figure 4c). Whereas MMR-proficient *C. elegans* transgenics failed to display significant somatic frameshift errors, RNAi of *msh-2* or *msh-6* resulted in the patched phenotype for all four transgenes tested, further demonstrating the requirement for MMR to stabilize repeated sequences in the *C. elegans* genome.

DISCUSSION

Most research on MMR function *in vivo* has focused either on unicellular organisms (because in those one can easily monitor mutator effects in large numbers of progeny) or on somatic cells or tissue culture cells of higher animals. The number of progeny animals that need to be inspected to recognize spontaneous mutants (that are not induced by chemicals or radiation) is prohibitively large. *C. elegans*, however, is a multicellular organism that can be studied in large numbers, allowing a systematic assessment of MMR in both somatic and germline tissues. Another advantage of this genetic system is that animals can be maintained as self-fertilizing hermaphrodites (spontaneous new mutations can homozygose in self-progeny) so that recessive mutations that lead to a phenotype will become visible in the next generation.

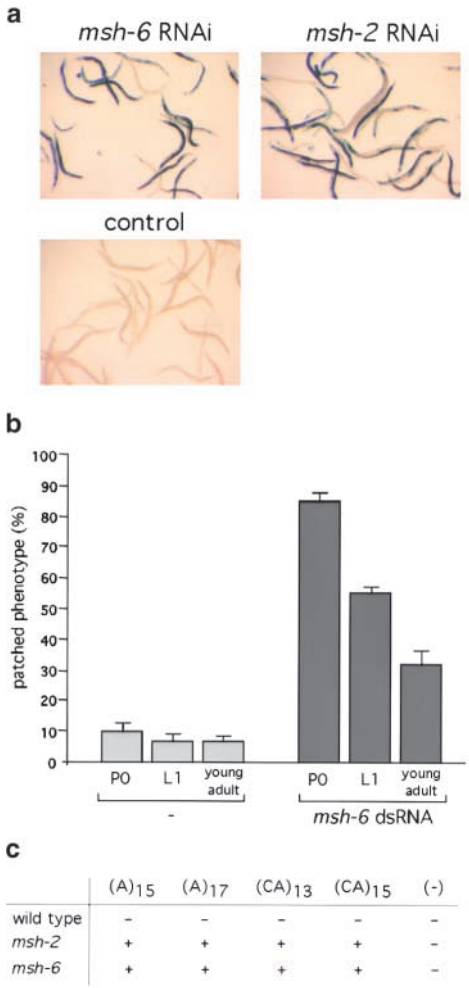

FIGURE 4.—Somatic repeat instability induced by *msh-2* and *msh-6* RNAi. (a) Animals are fed *E. coli* that produce dsRNA homologous to the *C. elegans* MMR genes *msh-2* and *msh-6* or a non-MMR control gene, *unc-22*. (b) Nematodes synchronized at different developmental stages (L1 larvae or L4 young adults) were exposed to RNAi by feeding, cultured to adulthood, and analyzed for somatic repeat instability. P0 indicates that progeny animals were analyzed from cultures in which the mother was transferred to the RNAi plates. The *y*-axis indicates the percentage of animals that display blue patches. (c) The stability of different types of repeats was assayed for their dependence on the MMR genes *msh-2* and *msh-6*. The absence of blue patches is indicated by −, while a frequent occurrence of blue-patched animals is indicated by +.

We show that loss of MMR in *C. elegans* results in a drastic increase in the level of spontaneous mutagenesis in both the male and female germlines. This severely affects the integrity of the hereditary material, thus threatening the existence of subsequent generations. We can extrapolate the total number of mutations that occur in MMR-deficient *C. elegans*, considering that the target region we assayed to quantify the mutation frequency covers ~7% of the total diploid genome. Because a high number of target genes are analyzed simultaneously, a possible sequence bias is ruled out. Considering that loss of essential genes is only part of the analysis (roughly estimated to account for 10% of all genes), we calculate the total number of loss-of-function mutations that occur in MMR-deficient animals to be approximately one to two per animal per generation. This high number predicts that growth of the animals (by self-fertilization, thus preventing mating and the exchange of genetic material) leads to rapid loss in fecundity. Indeed, when 50 individual *msh-6* lines were grown separately (by transferring single worms each generation), 57% became sterile before the tenth generation. Recently and independently, such reduced survival after serial passage was also shown—with elevated levels of microsatellite instability and increased mutation frequency—in a strain in which the *C. elegans msh-2* gene is disrupted by a transposable element (DEGTYAREVA *et al.* 2002).

One could in principle culture such strains (in population) for multiple generations and achieve quite significant accumulated levels of mutations, while maintaining selection pressure for viability. Possibly a strain like this may be of use in experiments aiming at experimental quantitative genetics, where genetic adaptation to specific environmental challenges can be studied more efficiently than in a wild-type isolate, because the rate of evolution is enhanced.

We found that inactivation of *msh-2* or *msh-6* results in a similar increase in the frequency of spontaneous mutagenesis in both the germline and somatic tissue. This situation differs from yeast and mammals where, in addition to *msh-6*, loss of *msh-3* is required to phenocopy the *msh-2* knockout to result in a complete MMR defect. In these systems two distinct protein complexes (both containing MSH2) with different substrate specificity function in damage recognition: MSH2/MSH6 acts primarily on single-base mispairs whereas the MSH2/MSH3 complex has a higher affinity for insertion/deletion loops (STRAND *et al.* 1993; JOHNSON *et al.* 1996; MARSISCHKY *et al.* 1996; SIA *et al.* 1997; UMAR *et al.* 1998). The phenotypic similarity in *msh-2*- and *msh-6*-defective *C. elegans*, together with the notion that the sequenced *C. elegans* genome does not contain an ortholog of the *msh-3* gene, strongly suggest that this species has only one damage recognition complex: Inactivating one partner of this MSH2/MSH6 heterodimer knocks out MMR. At present we do not know whether MSH-6 compensates for the absence of MSH-3. In other words, is the substrate specificity of *C. elegans* MSH2/MSH-6 greater than that in other animal systems? Alternatively, *C. elegans* just fails to repair larger mispairs that require the action of a specific MSH2/MSH3 complex. We found that MSH-6 is required to suppress tract expansions/contractions of endogenous mono-, di-, and trinucleotide microsatellite repeats, suggesting a role for this protein in the repair of insertion/deletion loops of at least 3 nt.

As stated, a significant fraction of human tumors is apparently caused by somatic mutations in genes that

affect genome stability, but not in all cases are these mutations in genes of the known MMR system. There seems to be no direct way to identify these genes, while they may be highly relevant as causative agents of human cancers. We here describe a *C. elegans* system that mimics the somatic repeat stability in human cancers, in that there is also an *msh-2*- and *msh-6*-dependent repeat length change, but here it can be recognized by an intense blue staining. Since feeding dsRNA homologous to MMR genes can also induce this effect, we now have a system to test any *C. elegans* gene for its role in repressing repeat length changes. Recently genome-wide libraries of dsRNAs of *C. elegans* have been described (Fraser *et al.* 2000; Gonczy *et al.* 2000), and we are currently testing all genes in this animal's genome for their mutator effect. If additional classes of mutator genes exist, possibly not at all related to MMR, but perhaps to replication factors, chromatin proteins that protect the genome, or totally novel protection systems, they can now be discovered. Eventually, possible human homologs can be tested for their role in human cancer etiology.

We thank Rik Korswagen, Edwin Cuppen, and Rene Ketting for critically reading the manuscript; members of the Plasterk Lab for discussions; Andy Fire and Julie Ahringer for reagents and protocols; and the *Caenorhabditis* Genetic Stock Centre for providing some of the strains used in this study.

LITERATURE CITED

Aaltonen, L. A., P. Peltomaki, F. S. Leach, P. Sistonen, L. Pylkkanen *et al.*, 1993 Clues to the pathogenesis of familial colorectal cancer. Science **260:** 812–816.

Bacon, A. L., M. G. Dunlop and S. M. Farrington, 2001 Hypermutability at a poly(A/T) tract in the human germline. Nucleic Acids Res. **29:** 4405–4413.

Baker, S. M., A. W. Plug, T. A. Prolla, C. E. Bronner, A. C. Harris *et al.*, 1996 Involvement of mouse *Mlh1* in DNA mismatch repair and meiotic crossing over. Nat. Genet. **13:** 336–342.

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.

Brenner, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics **77:** 71–94.

Bronner, C. E., S. M. Baker, P. T. Morrison, G. Warren, L. G. Smith *et al.*, 1994 Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature **368:** 258–261.

Buermeyer, A. B., S. M. Deschenes, S. M. Baker and R. M. Liskay, 1999 Mammalian DNA mismatch repair. Annu. Rev. Genet. **33:** 533–564.

Degtyareva, N. P., P. Greenwell, E. R. Hofmann, M. O. Hengartner, L. Zhang *et al.*, 2002 *Caenorhabditis elegans* DNA mismatch repair gene *msh-2* is required for microsatellite stability and maintenance of genome integrity. Proc. Natl. Acad. Sci. USA **99:** 2158–2163.

de Stasio, E., C. Lephoto, L. Azuma, C. Holst, D. Stanislaus *et al.*, 1997 Characterization of revertants of *unc-93*(*e1500*) in *Caenorhabditis elegans* induced by N-ethyl-N-nitrosourea. Genetics **147:** 597–608.

de Wind, N., M. Dekker, A. Berns, M. Radman and H. te Riele, 1995 Inactivation of the mouse *Msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. Cell **82:** 321–330.

Edelmann, W., P. E. Cohen, M. Kane, K. Lau, B. Morrow *et al.*, 1996 Meiotic pachytene arrest in MLH1-deficient mice. Cell **85:** 1125–1134.

Edelmann, W., K. Yang, A. Umar, J. Heyer, K. Lau *et al.*, 1997 Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. Cell **91:** 467–477.

Ellegren, H., 2000 Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet. **16:** 551–558.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver *et al.*, 1998 Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature **391:** 806–811.

Fishel, R., M. K. Lescoe, R. M. Rao, N. G. Copel, N. A. Jenkins *et al.*, 1993 The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell **75:** 1027–1038.

Flores-Rozas, H., and R. D. Kolodner, 1998 The *Saccharomyces cerevisiae MLH3* gene functions in *MSH3*-dependent suppression of frameshift mutations. Proc. Natl. Acad. Sci. USA **95:** 12404–12409.

Fraser, A. G., R. S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann *et al.*, 2000 Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. Nature **408:** 325–330.

Gonczy, P., G. Echeverri, K. Oegema, A. Coulson, S. J. Jones *et al.*, 2000 Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature **408:** 331–336.

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.

Greenwald, I. S., and H. R. Horvitz, 1980 *unc-93*(*e1500*): a behavioral mutant of *Caenorhabditis elegans* that defines a gene with a wild-type null phenotype. Genetics **96:** 147–164.

Henderson, S. T., and T. D. Petes, 1992 Instability of simple sequence DNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **12:** 2749–2757.

Ionov, Y., M. A. Peinado, S. Malkhosyan, D. Shibata and M. Perucho, 1993 Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature **363:** 558–561.

Jansen, G., E. Hazendonk, K. L. Thijssen and R. H. Plasterk, 1997 Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*. Nat. Genet. **17:** 119–121.

Jiricny, J., 1998 Replication errors: challenging the genome. EMBO J. **17:** 6427–6436.

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.

Kolodner, R., 1996 Biochemistry and genetics of eukaryotic mismatch repair. Genes Dev. **10:** 1433–1442.

Leach, F. S., N. C. Nicolaides, N. Papadopoulos, B. Liu, B. J. Jen *et al.*, 1993 Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell **75:** 1215–1225.

Lipkin, S. M., V. Wang, D. L. Stoler, G. R. Anderson, I. Kirsch *et al.*, 2001 Germline and somatic mutation analyses in the DNA mismatch repair gene MLH3: evidence for somatic mutation in colorectal cancers. Hum. Mutat. **17:** 389–396.

Liu, B., R. E. Parsons, S. R. Hamilton, G. M. Petersen, H. T. Lynch *et al.*, 1994 hMSH2 mutations in hereditary nonpolyposis colorectal cancer kindreds. Cancer Res. **54:** 4590–4594.

Loeb, L. A., 1991 Mutator phenotype may be required for multistage carcinogenesis. Cancer Res. **51:** 3075–3079.

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, R. H., J. Green, E. Ko, L. Barclay and A. W. Rademaker, 2000 Analysis of aneuploidy frequencies in sperm from patients with hereditary nonpolyposis colon cancer and an hMSH2 mutation. Am. J. Hum. Genet. **66:** 1149–1152.

Mello, C. C., J. M. Kramer, D. Stinchcomb and V. Ambros, 1991 Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. **10:** 3959–3970.

Modrich, P., and R. Lahue, 1996 Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu. Rev. Biochem. **65:** 101–133.

Narayanan, L., J. A. Fritzell, S. M. Baker, R. M. Liskay and P. M. Glazer, 1997 Elevated levels of mutation in multiple tissues of mice deficient in the DNA mismatch repair gene *Pms2*. Proc. Natl. Acad. Sci. USA **94:** 3122–3127.

Nicolaides, N. C., N. Papadopoulos, B. Liu, Y. F. Wei, K. C. Carter *et al.*, 1994 Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature **371:** 75–80.

Papadopoulos, N., N. C. Nicolaides, Y. F. Wei, S. M. Ruben, K. C. Carter *et al.*, 1994 Mutation of a mutL homolog in hereditary colon cancer. Science **263:** 1625–1629.

Peinado, M. A., S. Malkhosyan, A. Velazquez and M. Perucho, 1992 Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. Proc. Natl. Acad. Sci. USA **89:** 10065–10069.

Peltomaki, P., and A. de la Chapelle, 1997 Mutations predisposing to hereditary nonpolyposis colorectal cancer. Adv. Cancer Res. **71:** 93–119.

Peltomaki, P., R. A. Lothe, L. A. Aaltonen, L. Pylkkanen, M. Nystrom-Lahti *et al.*, 1993 Microsatellite instability is associated with tumors that characterize the hereditary non-polyposis colorectal carcinoma syndrome. Cancer Res. **53:** 5853–5855.

Prolla, T. A., S. M. Baker, A. C. Harris, J. L. Tsao, X. Yao *et al.*, 1998 Tumour susceptibility and spontaneous mutation in mice deficient in *Mlh1*, *Pms1* and *Pms2* DNA mismatch repair. Nat. Genet. **18:** 276–279.

Reitmair, A. H., R. Schmits, A. Ewel, B. Bapat, M. Redston *et al.*, 1995 MSH2 deficient mice are viable and susceptible to lymphoid tumours. Nat. Genet. **11:** 64–70.

Rosenbluth, R. E., C. Cuddeford and D. L. Baillie, 1983 A rapid eukaryotic mutagen test system using the reciprocal translocation, *eT1*(III;V). Mutat. Res. **110:** 39–48.

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.

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.

Szankasi, P., and G. R. Smith, 1995 A role for exonuclease I from S. pombe in mutation avoidance and mismatch correction. Science **267:** 1166–1169.

Timmons, L., and A. Fire, 1998 Specific interference by ingested dsRNA. Nature **395:** 854.

Umar, A., J. I. Risinger, W. E. Glaab, K. R. Tindall, J. C. Barrett *et al.*, 1998 Functional overlap in mismatch repair by human MSH3 and MSH6. Genetics **148:** 1637–1646.

Wu, Y., M. J. Berends, R. H. Sijmons, R. G. Mensink, E. Verlind *et al.*, 2001a A role for MLH3 in hereditary nonpolyposis colorectal cancer. Nat. Genet. **29:** 137–138.

Wu, Y., M. J. Berends, J. G. Post, R. G. Mensink, E. Verlind *et al.*, 2001b Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. Gastroenterology **120:** 1580–1587.

Communicating editor: B. J. Meyer