Theor Appl Genet (2002) 105:423–430
DOI 10.1007/s00122-002-0955-8

M. Horwath · W. Kramer · R. Kunze

Structure and expression of the *Zea mays mutS*-homologs *Mus1* and *Mus2*

Received: 17 October 2001 / Accepted: 7 March 2002 / Published online: 19 June 2002

Abstract DNA mismatch repair proteins play an important role in maintaining the integrity of the genetic information during replication and homologous recombination. The MutS-homologous (MSH) and MutL-homologous (MLH) proteins are highly conserved among all prokaryotes and eukaryotes. We have isolated two *mutS* homologous genes from *Zea mays*, named *Mus1* and *Mus2*. Phylogenetic analysis identifies Mus1 as a member of the MSH2 protein family. Mus2 is an ortholog of the *Arabidopsis thaliana* MSH7 protein and belongs to a subgroup of MSH proteins that is possibly plant-specific. *Mus1* and *Mus2* are expressed at very low levels. *Mus1* is located on chromosome 7L near locus b32B, and *mus2* maps on chromosome 3S.

Keywords *Zea mays* · DNA mismatch repair · MutS · MSH protein

Introduction

In all organisms, DNA repair processes are crucial to maintain the integrity of the genome through generations. Many animals minimize the number of cell divisions, and thus the risk to accumulate mutations, from one to the next generation by means of the germline. Plants have no germline, and the gametophytes develop from sporophytic tissue. Thus, mutations in sporophytic cells can be transmitted to the progenitors. Moreover, as plants are sessile organisms and cannot actively avoid UV-light exposure, they receive much higher radiation doses than animals. In spite of this, it was suggested that the evolutionary mutation rate in plants appears not to be much higher than in animals (Klekowski 1997). This is especially surprising, as it has recently been observed that in *Arabidopsis thaliana* the somatic point mutation frequency is more than 100-fold higher than is estimated for other eukaryotes, and that it dramatically increases beyond that following UV-C irradiation (Kovalchuk et al. 2000). To-date, however, the plant DNA repair systems that have to cope with DNA damage are largely unknown (reviewed in Britt 1999), and only recently were the first studies about this subject published.

The DNA mismatch repair (MMR) apparatus has an important function in stabilizing the genome, and the protein components of this system are highly conserved in all prokaryotic and eukaryotic systems. The MMR system is responsible for the post-replicative repair of mismatches and small single-stranded DNA loops that result from DNA replication errors, and were not corrected by the proof reading activity of the DNA polymerase. Single-stranded loops frequently result from DNA polymerase slippage during the replication of microsatellite sequences (simple sequence repeats) (Kunkel 1992). In addition, the MMR system is critically involved in preventing recombination between homoeologous (non-identical) DNA sequences.

In *Escherichia coli* the MMR core system consists of the proteins MutH, MutL and MutS. MutS recognizes and binds to mispaired nucleotides or small single-stranded loops as a homodimer. Subsequently, in a coordinated cascade of reactions involving MutL, MutH, exonuclease, DNA polymerase III and other proteins, the newly synthesized DNA strand encompassing the mismatched nucleotide(s) is removed and re-synthesized

Communicated by F. Salamini

M. Horwath · R. Kunze (✉)
Institute of Genetics, University of Munich,
Maria-Ward-Strasse 1a, 80638 München, Germany
e-mail: reinhard.kunze@uni-koeln.de
Tel.: +49-221-470 6596, Fax: +49-221-470 5039

W. Kramer
Department of Molecular Genetics and Preparative Molecular Biology, Institute for Microbiology and Genetics, Georg-August University, Göttingen, Grisebachstrasse 8, D-37077 Göttingen, Germany

Present addresses:
M. Horwath, Icon Genetics AG, Blumenstrasse 16, 85354 Freising, Germany

R. Kunze, Botanical Institute, University of Cologne, Gyrhofstrasse 15, 50931 Köln, Germany

by the DNA polymerase III holoenzyme (reviewed in Cox 1997).

The MMR proteins play a second major role in selecting against mispaired recombination intermediates, thus preventing recombination between homoeologous sequences. It is supposed that MutS and MutL inhibit RecA catalyzed strand transfer between diverged DNAs and block branch migration (Worth et al. 1994). Normally, recombination between *E. coli* and its near relative *Salmonella thyphimurium*, having 16% DNA sequence divergence, is very rare. However, it is common when the recipient is *mutS* (Rayssiguier et al. 1989).

The eukaryotic MMR system is more complex than the bacterial one. *Saccharomyces cerevisiae* has six different MutS homologs (MSH1–MSH6) and four MutL homologs (MLH1–MLH3 and PMS1) which have different functions. MSH1 is involved in repair and stabilization of the mitochondrial genome (Reenan and Kolodner 1992a, b). MSH2 forms heterodimers with MSH3 and MSH6, that recognize and bind to mismatches and short insertion/deletion mispairs in the nucleus (Alani et al. 1995; Marsischky et al. 1996). *Msh2* mutants exhibit a pleiotropic phenotype, most prominently an elevated mutation rate and instability of microsatellites. *Msh3/Msh6* double mutants show the same phenotype. A defect in *Msh3* alone leads to microsatellite instability, whereas a mutation in *Msh6* causes an increase in point mutations and instability of microsatellites with single-nucleotide repeats (Alani 1996; Iaccarino et al. 1996; Harfe and Jinks-Robertson 1999). Two more *msh* genes have been identified whose gene products are not involved in MMR during replication. MSH4 and MSH5 proteins are active during meiosis and take part in the regulation of chromosome pairing and meiotic recombination (Ross-Macdonald and Roeder 1994; Hollingsworth et al. 1995; de Vries et al. 1999).

Similarly to MutS, eukaryotic MSH and MLH proteins are involved in regulating recombination. In yeast and in mouse cells, mutations in MSH2 or MLH1 effect a strong increase in general and homoeologous recombination (de Wind et al. 1995; Selva et al. 1995; Baker et al. 1996; Edelmann et al. 1996; Datta et al. 1997; Hunter and Borts 1997; Chen and Jinks-Robertson 1998, 1999). Moreover, it is now well-established that in mouse and humans defects in the *Msh2* or *Mlh1* loci correlate with a predisposition to certain kinds of tumors, for example heritable colorectal cancer, HNPCC (Fishel et al. 1993; Leach et al. 1993; Bronner et al. 1994).

The investigation of the MMR system in plants has only recently begun. MutS and MutL homologs have been identified and partially characterized in *A. thaliana* (Culligan and Hays 1997; Adé et al. 1999; Jean et al. 1999; Culligan and Hays 2000; Culligan et al. 2000; Adé et al. 2001), and partial cDNAs of putative MMR genes were isolated from *Triticum aestivum* (Korzun et al. 1999), rice, *Brassica napus* and soybean. Here we report on the identification, transcription and gene structure of two *MutS* homologous genes from maize, termed *Mus1* and *Mus2*.

Materials and methods

Polymerase chain reactions

PCR reactions were carried out similarly as described in Laging et al. (2001). Briefly, reactions were performed in a 50-µl total volume containing template DNA (cDNA: approximately 20 ng; genomic DNA: approximately 0.5 µg), 0.4 µM of each primer, 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1% DMSO and 2 u of *Taq* DNA polymerase (AGS, Heidelberg). Cycling conditions were: 2 min denaturation at 95 °C followed by 25 cycles (30 s at 94 °C, 30 s annealing at the primer-specific temperature, 1 min at 72 °C per 1 kb of amplified DNA fragment), and a terminal extension step of 5 min at 72 °C.

RNA analyses

Fifteen to 20 etiolated 12-day old maize seedlings were frozen in liquid nitrogen and ground to a fine powder. Poly(A) RNA was prepared as described in Pawlowski et al. (1994) with minor modifications. For RNA blotting experiments 3–8 µg of poly(A) RNA were electrophoresed, capillary blotted and hybridized essentially as described (Kunze et al. 1987). As probes, PCR fragments were amplified on maize seedling cDNA using the *Mus1*-specific oligonucleotides zmMutS-1 (5′-ATGCGAGACAAGCGCAGGGTTTC-ATC-3′) and zmMutS-2 (5′-CTCGAACTGCAGCAGAATCAAGT-CGCAT-3′), or the *Mus2* -specific oligonucleotides zmMus2-A (5′-TGGGTTCAGTCCAAGATGATGACTCGTCT-3′) and zmMus2-H (5′-ATACCCAAGTGTTGTAGGGCCTAGTCCA-3′), and were ^{32}P-labeled by random priming.

DNA analyses

DNA was isolated from etiolated maize seedlings in parallel to RNA extraction. After LiCl-precipitation of the RNA, DNA was precipitated from the supernatant with 0.6 vol of isopropanol, washed in 70% ethanol and subsequently purified by CsCl density gradient centrifugation.

For DNA gel-blot experiments 4 µg of DNA was digested with restriction enzymes, electrophoresed, capillary blotted onto Hybond-N^+ nylon membranes (Amersham) and hybridized in Roti Hybri Quick hybridization buffer (Roth) as recommended by the manufacturer. Hybridization probes were prepared as described for the RNA gel blots.

Reverse transcription – PCR

Five micrograms of maize seedling poly(A) RNA was reverse transcribed using Superscript reverse transcriptase (Life Technologies). First-strand cDNA synthesis was primed with the degenerate oligonucleotide oMSH-D (5′-RWARTGNGTNRCRAA-3′). PCR was performed on first-strand cDNA using a combination of the degenerate oligonucleotide MSH-A (5′-GCTCTAGACNGGNCCNAAYATGG-G-3′) and one of the primers MSH-C1 to MSH-C8 (C1: 5′-CAC-GGTACCNCGNCCYAAYTCRTC-3′, C2: 5′-CACGGTACCYCTN-CCYAAYTCRTC-3′, C3: 5′-CACGGTACCNCGNCCNAGYTCR-TC-3′, C4: 5′-CACGGTACCYCTNCCNAGYTCRTC-3′, C5: 5′-CACGGTACCNCGNCCDATYTCRTC-3′, C6: 5′-CACGGTACC-YCTNCCDATYTCRTC-3′, C7: 5′-CACGGTACCNCGNCCNAC-YTCRTC-3′, C8: 5′-CACGGTACCYCTNCCNACYTCRTC-3′; R = G/A; Y = T/C; W = T/A; D = G/A/T; N = G/A/T/C). The amplification cycle was: 2 min at 95 °C, 5 min at 85 °C – during this period 1 u of *Tfl* polymerase (Epicentre Technologies) was added to the reaction mixture, 35 cycles (20 s at 95 °C, 30 s at 46 °C, 45 s at 72 °C), followed by a final step of 5 min at 72 °C. The resulting PCR products were fractionated on a 1.3% agarose gel, the expected approximately 270-bp fragments were excised and eluted, digested with *Kpn*I and *Xba*I and ligated into *Kpn*I–*Xba*I-restricted M13mp18. Three cDNA clones, named S-39, S-110 and S-410, were identified that show homology to *S. cerevisiae* MSH2 and MSH6 genes.

5′-RACE

The 5′-end of the *Mus1* transcript was identified by rapid amplification of the 5′ cDNA end (5′-RACE). Three micrograms of maize seedling poly(A)-RNA were reverse transcribed using reverse transcriptase (Promega) as recommended by the manufacturer. First-strand cDNA synthesis was primed with the oligonucleotide zmMSH-3 (5′-GCACCTAATGCAACAAGAGCTG-3′). The first-strand cDNA was 3′-terminally extended with a poly(dG) tail by treatment with Terminal desoxynucleotidyl Transferase (MBI Fermentas).

PCR was performed on the dG-tailed first-strand cDNA using High Fidelity Polymerase (Roche) and the oligonucleotides zmMSH-2 (5′-GGTAGTATGTCCTTGCGATAAACG-3′) and RKo51-C11 (5′-AGCCATGGCATCTTGACTCGGATCCCCCCC-CCCCC-3′). The PCR parameters were: 2 min at 95 °C, five cycles (30 s at 95 °C, 1 min at 45 °C, 2 min at 72 °C), 20 cycles (30 s at 95 °C, 1 min at 54 °C, 2 min at 72 °C), and 5 min at 72 °C. The resulting PCR products were agarose gel-purified, reamplified using primers zmMSH-2 and RKo51 (5′-AGCCAT-GGCATCTTGACTCGGATC-3′), digested with *Bam*HI and ligated into *Bam*HI-linearized pBluescript KSI f(+).

Screening of bacteriophage λ libraries

Mus1 and *Mus2* cDNAs were isolated from a maize strain CI31A seedling cDNA library in λ-phage NM1149 (a gift from Dr. Monika Frey, Technical University Munich). A *Mus2* 3′-end cDNA clone was isolated from a seedling cDNA library in λZAP-Express (Stratagene) from a maize line carrying the *wx-m7::Ac* mutation (McClintock 1951). *Mus1* and *Mus2* genomic clones were isolated from a maize strain CI31A library in λ-phage EMBL4 (a gift from Dr. Monika Frey, Technical University Munich). The cDNA libraries were screened by hybridization with ^{32}P-labeled RT-PCR probes. For screening of the genomic library DIG-11-dUTP-labeled probes were generated and detected using the DIG-HighPrime and DIG-Luminescense Detection Kits (Roche). Phage λ-DNA was prepared using the Lambda Mini Kit (Qiagen, Hilden). cDNA inserts from λZAP-phages were subcloned by in vivo excision.

DNA Sequencing

From the *Mus1* and *Mus2* cDNA and genomic clones, overlapping deletion subclones were generated with a Double-stranded Nested Deletion Kit (Pharmacia). Plasmids were sequenced using the Big-Dye Terminator Cycle Sequencing Kit and Ampli*Taq* polymerase (ABI). The reactions were analyzed with an ABI Prism 377 DNA automatic sequencer.

Results

Isolation of *mus1* and *mus2* cDNAs and expression analysis

Based on the highly conserved C-terminal domain of MSH proteins that carries four putative nucleotide binding sites, degenerate oligonucleotides were designed (Fig. 1A). Primer oMSH-D is derived from the fourth putative nucleotide binding site of MSH proteins and was used to prime reverse transcription of poly(A) RNA from etiolated maize seedlings. The cDNA was used as a

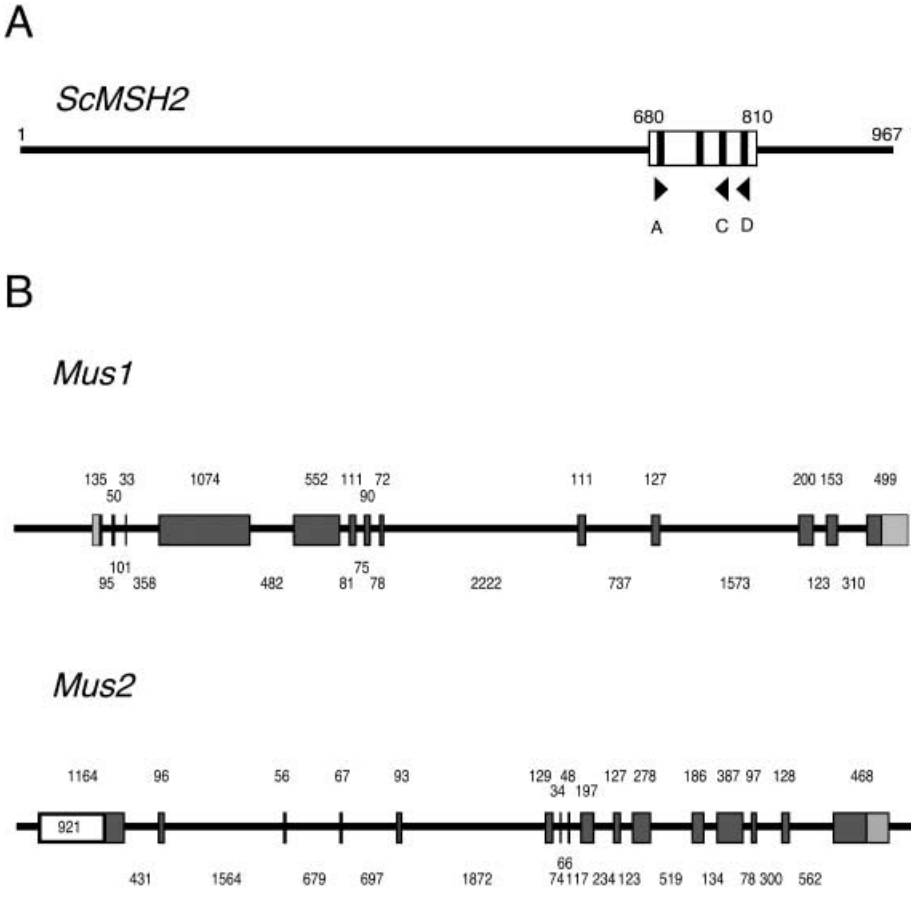

Fig. 1 **A** Schematic of the 967 amino-acid yeast MSH2 protein. The conserved C-terminal region of the MSH proteins is shown as a *bar* with the four nucleotide binding sites as *vertical black boxes*. D: degenerate cDNA synthesis primer oMSH-D; A and C: degenerate PCR primers oMSH-A and oMSH-C1 to oMSH-C8. **B** Exon-intron structure of *Mus1* and *Mus2*. Exons are depicted by *boxes*, their lengths are indicated above. The lengths of the intervening sequences is shown below. *Lighter shading* of the terminal exons indicates the 5′- and 3′-untranslated sequences. The transcription start site of *Mus2* has not been experimentally determined, and the putative start of the open reading frame was deduced from the genomic DNA sequence. The predicted 5′-end of the ORF is shown as an *open box*

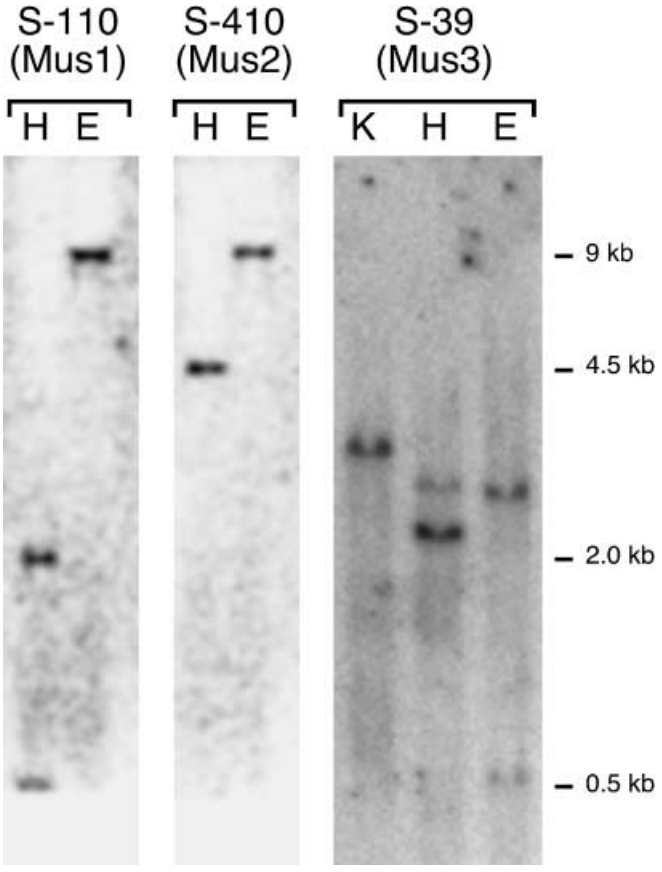

Fig. 2 DNA gel-blot analysis of *Mus* genes. Four micrograms of genomic maize DNA were digested with *Kpn*I (*K*), *Hin*dIII (*H*) or *Eco*RI (*E*), electrophoresed, blotted and hybridized with ^{32}P-labeled cDNA fragments of the *Mus1* (S-110), *Mus2* (S-410), or *Mus3* (S-39) genes. The approximate sizes of the *Mus1* and *Mus2* bands are indicated

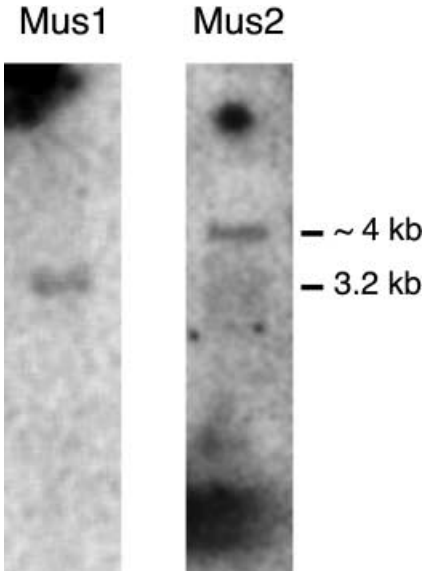

Fig. 3 Northern-blot analysis of *Mus1* and *Mus2* transcription. Four (*Mus1*) and eight micrograms of poly(A)RNA (*Mus2*) from etiolated maize seedlings were size fractionated by denaturating formaldehyde gel electrophoresis, blotted and hybridized with ^{32}P-labeled *Mus1* (S-110) and *Mus2* (S-410) cDNA fragments. The approximate sizes of the *Mus1* and *Mus2* mRNAs are indicated

template for PCR amplification with the forward primer oMSH-A and one of the reverse primers oMSH-C1 to oMSH-C8. As the distance between the four nucleotide-binding sites is conserved in all MutS homologs, PCR products of the expected size were purified, cloned and sequenced. Four clones contained products with similarity to MSH proteins.

In gel-blot hybridization experiments only three of these cDNA clones hybridized to genomic DNA from maize. Each of the three cDNA clones S-39, S-110 and S-410 hybridized to distinct single-copy sequences (Fig. 2). The deduced amino-acid sequences of S-39 and S-410 show homology to MSH6 proteins, whereas the third fragment, S-110, is more closely related to MSH2 proteins.

A maize seedling cDNA library was screened with the three cDNA fragments. Hybridizing clones were isolated for S-110 and S-410; however, no positive clones were identified for S-39 among 2×10^6 plaques. After subcloning, sequencing of the positive clones confirmed their classification as MSH2- and MSH6-homologous genes. As in maize the designation ‘msh’ is already in use (for ‘male sterility homologue’); we termed the MSH2-like gene *Mus1* and the MSH6-like gene *Mus2*. The S-39 cDNA clone corresponds to *Mus3*, a second, distinct MSH6-like gene (Genbank accession AF227632).

The *Mus1* cDNA clone contains a poly(A) tail, but is incomplete at the 5′ end. The 5′ end of the transcript was identified by 5′-RACE. The full-length *Mus1* cDNA is 3.2 kb in length and contains a 2,826 nucleotide open reading frame, putatively coding for a 942 amino-acid maize MSH2-homologous protein.

The original *Mus2* cDNA clone S-410 is incomplete at both ends. By screening of a cDNA library using the 3′-end of S-410 as a probe, a cDNA clone containing the *Mus2* 3′-end including the poly(A) tail was isolated. The 5′-end of the gene could not be isolated as a cDNA clone and 5′-RACE was not successful.

The expression of *Mus1* and *Mus2* was investigated by RNA gel-blot analyses. As in total RNA the messages could not be detected, the blots were repeated with 5 to 8 μg of poly(A)RNA from young maize seedlings. Here, the messenger RNAs of both genes were detected as faint bands (Fig. 3). Consistent with the low abundance of *Mus1* and *Mus2* cDNA clones in the library, the weak signals indicate that both genes are expressed at very low levels. The size of the *Mus1* mRNA is approximately 3.2 kb, which agrees with the overall length of the cDNA. Using an RNA-ladder as a marker, the *Mus2* mRNA is estimated to be approximately 4 kb in size. Accordingly, at least 1 kb of the 5′-end of the transcript is missing in the cDNA clones.

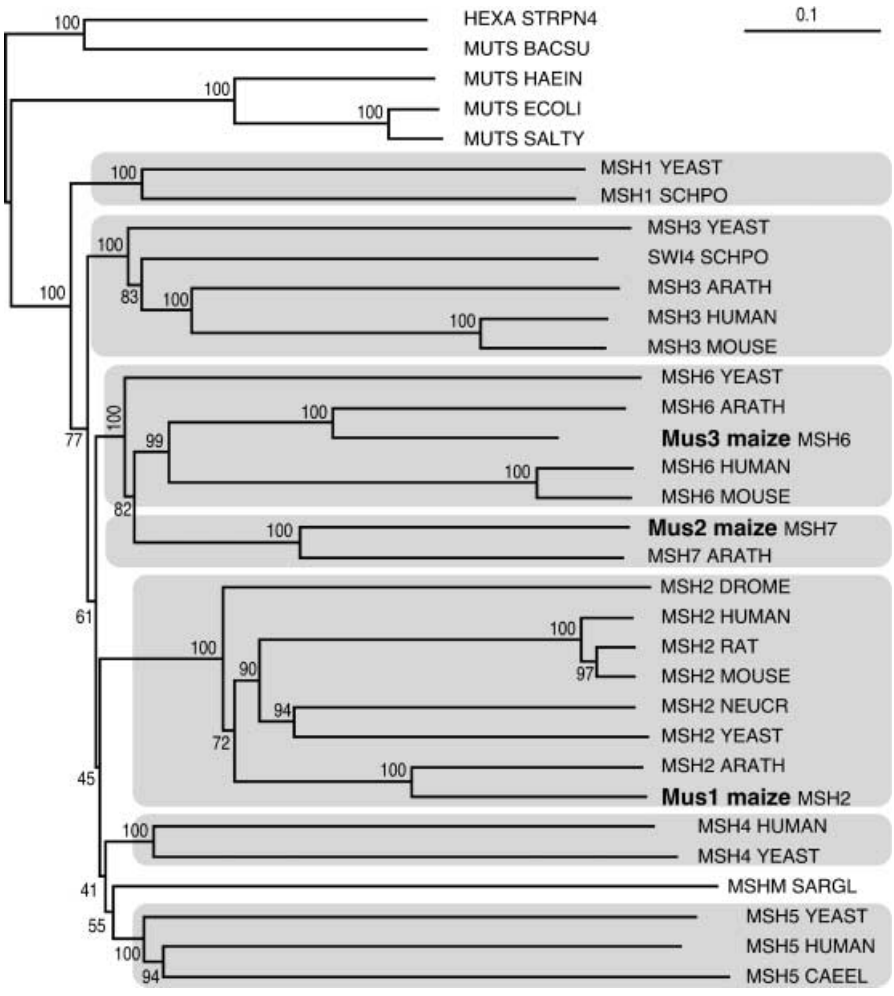

Fig. 4 Phylogenetic analysis of a subset of MutS-homologous proteins. The full-length protein sequences were aligned by ClustalX (1.81) using the Gonnet 250 protein weight matrix (Thompson et al. 1997). The phylogram was calculated by the Neighbor Joining method and drawn using TreeView 1.6.5 (Page 1996), defining the MutS proteins from the gram-positive bacteria *B. subtilis* and *S. pneumoniae* as an outgroup. The *numbers* above or below each branch indicate the frequency (%) of appearance of the branch in 1,000 bootstrap replicas. The six MSH protein families are highlighted by *shading*. The SWISS-PROT (or Genbank) protein entry names, source organism names, and accession numbers in the order of their appearance in the tree are: HEXA_STRPN, *Streptococcus pneumoniae*, P10564; MUTS_BACSU, *Bacillus subtilis*, P49849; MUTS_HAEIN, *Haemophilus influenzae*, P44834; MUTS_ECOLI, *Escherichia coli*, P23909; MUTS_SALTY, *Salmonella typhimurium*, P10339; MSH1_YEAST, *S. cerevisiae*, P25846; MSH1_SCHPO, *S. pombe*, O13921; MSH3_YEAST, *S. cerevisiae*, P25336; SWI4_SCHPO, *Schizosaccharomyces pombe*, P26359; MSH3_ARATH, *Arabidopsis thaliana*, O65607; MSH3_HUMAN, *H. sapiens*, P20585; MSH3_MOUSE, *Mus musculus*, P13705; MSH6_YEAST, *S. cerevisiae*, Q03834; MSH6_ARATH, *A. thaliana*, O04716; Mus3_MAIZE_MSH6, *Z. mays*, AAF35250 (Genbank); MSH6_HUMAN, *H. sapiens*, P52701; MSH6_MOUSE, *M. musculus*, P54276; Mus2_MAIZE_MSH7, *Zea mays*, CAB42555 (Genbank); MSH7_ARATH, *A. thaliana*, Q9SMV7; MSH2_DROME, *Drosophila melanogaster*, P43248; MSH2_HUMAN, *H. sapiens*, P43246; MSH2_RAT, *Rattus norvegicus*, P54275; MSH2_MOUSE, *M. musculus*, P43247; MSH2_NEUCR, *Neurospora crassa*, O13396; MSH2_YEAST, *S. cerevisiae*, P25847; MSH2_ARATH,*A. thaliana*, O24617; Mus1_MAIZE_MSH2, *Z. mays*, Q9XGC9; MSH4_HUMAN, *Homo sapiens*, O15457; MSH4_YEAST, *Saccharomyces cerevisiae*, P40965; MSHM_SARGL, mitochondrion *Sarcophyton glaucum* (a possible case of gene transfer from the nucleus to the mitochondrial genome), O63852; MSH5_YEAST, *S. cerevisiae*, Q12175; MSH5_HUMAN, *H. sapiens*, O43196; MSH5_CAEEL, *Caenorhabditis elegans*, Q19272

Gene structures and chromosomal location of *Mus1* and *Mus2*

A maize genomic λ library was probed with the *Mus1* and *Mus2* cDNAs. Two overlapping clones containing the complete *Mus1* gene were isolated. To determine the gene structure the clones were fully sequenced (Fig. 1B). From the transcription start to the polyadenylation site the *Mus1* gene is 9.4 kb in length and consists of 13 exons. The first three exons are remarkably short (55 bp, 50 bp and 33 bp), the longest exon covers 1,074 bp. The 12 intervening sequences range from 75 bp to 2,222 bp in length. No sequence deviations between cDNA and genomic clones in the exons were observed. The analysis of the 3-kb genomic DNA upstream sequence of *Mus1* revealed that the promoter has no apparent TATA-box.

For *Mus2* also two overlapping genomic λ clones were isolated. The sequence analysis of the cDNA and genomic

clones, and the alignment of the putative coding region with MSH6 proteins from *A. thaliana*, yeast and mammals, suggests that the protein coding region extends 921 bp beyond the 5′-end of the *Mus2* cDNA (Fig. 1B). The predicted full-length *Mus2* gene extends over approximately 11.3 kb, consists of 16 exons (15 introns) varying between 34 bp and 1,161 bp (65 bp and 1,875 bp) in length, and encodes a 1,184 amino-acid protein. The *Mus1* and *Mus2* sequences were deposited in GenBank under accession numbers AJ238785, AJ238786 and AJ238787.

The chromosomal location of *Mus1* and *Mus2* was determined by using a collection of 48 T232/CM37 recombinant inbred lines (Burr et al. 1988; Burr and Burr 1991). A *Mus1* specific probe detects a *Hin*dIII-RFLP in digested genomic DNA of the 48 inbred lines (data not shown). *Mus1* cosegregates with the locus *b32B* on chromosome 7L that codes for a ribosome inactivation protein. Using a *mus2* specific probe that detects a *Bam*HI-RFLP, *Mus2* was mapped on chromosome 3S. Accordingly, the *Mus1* and *Mus2* mismatch repair genes are not clustered.

Relationship of Mus1 and Mus2 to other MSH proteins

To determine the relationship of Mus1 and Mus2 to other MutS-homologous proteins, alignments of MutS-related proteins from prokaryotes, yeast, animals and plants were performed. A phylogenetic tree was calculated by the neighbor-joining method, using the MutS-homologs of the gram-positive bacteria *Bacillus subtilis* and *Streptococcus pneumoniae* as an outgroup (Fig. 4).

With an overall amino-acid sequence identity of 65% Mus1 is most similar to MSH2 from *A. thaliana*. It shares 41% and 40% amino-acid sequence identity with the MSH2 proteins from mouse and human, respectively. Thus, the similarity of Mus1 to MSH2 orthologs from other organisms is significantly higher than to the paralog Mus2, which has only a 23% amino-acid sequence identity with Mus1 (not counting the unaligned, protruding N-terminus of Mus2). Mus2 is 51% identical to AtMSH7 (Swissprot acc. no. Q9SMV7) and 27% identical to AtMSH6 (acc. no. O04716).

Discussion

The high degree of sequence conservation in the C-terminal part of all known MutS-homologous proteins allowed us to amplify cDNA fragments from MutS-homologous *Zea mays* genes by RT-PCR using degenerate oligonucleotide primers. The respective genes were termed *Mus1* and *Mus2*. Mus1 is the maize ortholog of the *Arabidopsis*, yeast and mammalian MSH2 proteins. Mus2 belongs to the MSH7 family, whose only other known member is the *A. thaliana* MSH7 protein. Both genes are expressed at very low levels, as has also been observed for the *Arabidopsis*, yeast and mammalian MSH genes (Kramer et al. 1996; Marra et al. 1998; Adé et al. 1999).

The 942 amino-acid Mus1 protein is most similar to *A. thaliana* MSH2. It shares 65% identical amino acids with the 937-aa AtMSH2 over the whole sequence and 83% sequence identity in the conserved, approximately 120 amino-acid C-terminal region that contains the four putative nucleotide binding sites. The two plant MSH2 orthologs AtMSH2 and ZmMus1 are much closer related to each other than to non-plant MSH2 proteins as for example human MSH2, which shares 40% and 63% conserved residues with ZmMus1 over the whole length and the conserved C-terminus, respectively. The similarity of Mus1 to its orthologs in plants and non-plants is significantly higher than to its paralog Mus2 (23% overall identity and 40% in the C-terminus).

The pronounced conservation of the MSH2 proteins throughout all eukaryotes indicates that these proteins fulfill very similar tasks in all organisms. In yeast and mammals MSH2 forms heterodimers with MSH3 and MSH6, which specifically recognize and bind to different DNA lesions (Strand et al. 1995; Iaccarino et al. 1998). The mismatch-bound MSH heterodimers interact with a MutL heterodimer and supposedly also with the PCNA component of the replication complex (for a review see Jiricny 1998). Recently it was shown that the *Arabidopsis* AtMSH2 forms heterodimers with AtMSH6 and AtMSH7, that differ in binding affinities to certain DNA mismatches (Culligan and Hays 2000), indicating that the two proteins have distinct functions.

In yeast and animal genomes only a single *Msh6* gene has been identified. Mus2 and AtMSH7 (that has also been termed AtMSH6-2; Adé et al. 1999) are more closely related to the MSH6 proteins than to any other MSH protein family. However, although the primary structures of the members of the MSH6 protein family are in-general less conserved than those of the MSH2 proteins, Mus2 and AtMSH7 form a distinct clade between the *S. cerevisiae* MSH6 and the MSH6 proteins from higher eukaryotes. With sizes of 1,185 and 1,110 amino acids the Mus2 and AtMSH7 proteins are smaller than the MSH6 proteins (e.g. AtMSH6 – 1,362 aa, *S. cerevisiae* MSH6 – 1,242 aa, HmMSH6 – 1,360 aa). Thus, Mus2 belongs to a branch of MSH6-related proteins that appears to occur exclusively in plants and is more similar to the mammalian MSH6 proteins than to yeast MSH6 (Adé et al. 1999; Culligan and Hays 2000) (Fig. 4). From rice an EST of a gene was isolated, which presumably also is a AtMSH7 homolog (Genbank acc. no. CAA10614).

A partial cDNA sequence of another maize MSH homolog, termed Mus3 (Genbank acc. no. AAF35250), has been deposited in Genbank (A. Franklin, unpublished). The encoded 629 amino-acid sequence is 62% identical to AtMSH6 and 36% identical to yeast MSH6. Thus, *ZmMus3* presumably encodes the authentic maize MSH6 ortholog (Fig. 4).

Acknowledgments We thank Berthold Fartmann for support with the RT-PCRs, Monika Frey (Technische Universität München, Germany) for the gift of libraries, help with the recombinant inbred mapping, and helpful discussions. This work was supported by Deutsche Forschungsgemeinschaft through a Heisenberg fellowship to R.K.

References

AdéJ, Belzile F, Philippe H, Doutriaux M-P (1999) Four mismatch repair paralogues coexist in *Arabidopsis thaliana*: AtMSH2, AtMSH3, AtMSH6-1 and AtMSH6-2. Mol Gen Genet 262: 239–249

AdéJ, Haffani Y, Belzile FJ (2001) Functional analysis of the *Arabidopsis thaliana* mismatch repair gene MSH2. Genome 44:651–657

Alani E (1996) The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. Mol Cell Biol 16:5604–5615

Alani E, Chi N-W, Kolodner R (1995) The *Saccharomyces cerevisiae* Msh2 protein specifically binds to duplex oligonucleotides containing mismatched DNA base pairs and insertions. Genes Dev 9:234–247

Baker SM, Plug AW, Prolla TA, Bronner CE, Harris A, Yao X, Christie D-M, Monell C, Arnheim N, Bradley A, Ashley T, Liskay RM (1996) Involvement of mouse *Mlh1* in DNA mismatch repair and meiotic crossing over. Nature Genet 13:336–342

Britt AB (1999) Molecular genetics of DNA repair in higher plants. Trends Plant Sci 4:20–25

Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, Kane M, Earabino C, Lipford J, Lindblom A, Tannergard P, Bollag RJ, Godwin AR, Ward DC, Nordenskjold M, Fishel R, Kolodner R, Liskay RM (1994) Mutation in the DNA mismatch repair gene homologue *hmlh1* is associated with hereditary nonpolyposis colon cancer. Nature 368: 258–261

Burr B, Burr FA (1991) Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. Trends Genet 7 No. 2:55–60

Burr B, Burr FA, Thompson KH, Albertson MC, Stuber CW (1988) Gene mapping with recombinant inbreds in maize. Genetics 118:519–526

Chen W, Jinks-Robertson S (1998) Mismatch repair proteins regulate heteroduplex formation during mitotic recombination in yeast. Mol Cell Biol 18:6252–6537

Chen W, Jinks-Robertson S (1999) The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. Genetics 151: 1299–1313

Cox EC (1997) Perspectives. Anecdotal, historical and critical commentaries on genetics. MutS, proofreading and cancer. Genetics 146:443–446

Culligan KM, Hays JB (1997) DNA mismatch repair in plants. An *Arabidopsis thaliana* gene that predicts a protein belonging to the *msh2* subfamily of eukaryotic *mutS* homologs. Plant Physiol 115:833–839

Culligan KM, Hays JB (2000) Arabidopsis MutS homologs – AtMSH2, AtMSH3, AtMSH6, and a novel AtMSH7 – form three distinct protein heterodimers with different specificities for mismatched DNA. Plant Cell 12:991–1002

Culligan KM, Meyer-Gauen G, Lyons-Weiler J, Hays JB (2000) Evolutionary origin, diversification and specialization of eukaryotic MutS homolog mismatch repair proteins. Nucleic Acids Res 28:463–71

Datta A, Hendrix M, Lipsitch M, Jinks-Robertson S (1997) Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. Proc Natl Acad Sci USA 94:9757–9762

Edelmann W, Cohen PE, Kane M, Lau K, Morrow B, Bennett S, Umar A, Kunkel T, Cattoretti G, Chaganti R, Pollard JW, Kolodner RD, Kucherlapati R (1996) Meiotic pachytene arrest in *mlh1*-deficient mice. Cell 85:1125–1134

Fishel R, Lescoe MK, Rao MRS, Copeland NG, Jenkins NA, Garber J, Kane M, Kolodner R (1993) The human mutator gene homolog *msh2* and its association with hereditary nonpolyposis colon cancer. Cell 75:1027–1038

Harfe BD, Jinks-Robertson S (1999) Removal of frameshift intermediates by mismatch repair proteins in *Saccharomyces cerevisiae*. Mol Cell Biol 19:4766–4773

Hollingsworth NM, Ponte L, Halsey C (1995) MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. Genes Dev 9:1728–1739

Hunter N, Borts RH (1997) Mlh1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis. Genes Dev 11:1573–1582

Iaccarino I, Marra G, Palombo F, Jiricny J (1998) hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSa. EMBO J 17: 2677–2686

Iaccarino I, Palombo F, Drummond J, Totty NF, Hsuan JJ, Modrich P, Jiricny J (1996) MSH6, a *Saccharomyces cerevisiae* protein that binds to mismatches as a heterodimer with MSH2. Curr Biol 6 No. 4:484–486

Jean M, Pelletier J, Hilpert M, Belzile F, Kunze R (1999) Isolation and characterization of AtMLH1, a MutL homologue from *Arabidopsis thaliana*. Mol Gen Genet 262:633–42

Jiricny J (1998) Replication errors: cha(lle)nging the genome. EMBO J 17:6427–6436

Klekowski Jr. EJ (1997) Somatic mutation theory of clonality. In: de Kroon H, van Groenendael J (eds) The ecology and evolution of clonal plants. Backbuys Publishers, Leyden, The Netherlands, pp 1–15

Korzun V, Borner A, Siebert R, Malyshev S, Hilpert M, Kunze R, Puchta H (1999) Chromosomal location and genetic mapping of the mismatch repair gene homologs *MSH2*, *MSH3* and *MSH6* in rye and wheat. Genome 42:1255–7

Kovalchuk I, Kovalchuk O, Hohn B (2000) Genome-wide variation of the somatic mutation frequency in transgenic plants. EMBO J 19:4431–8.

Kramer W, Fartmann B, Ringbeck EC (1996) Transcription of *mutS* and *mutL*-homologous genes in *Saccharomyces cerevisiae* during the cell cycle. Mol Cell Biol 252:275–283

Kunkel TA (1992) DNA replication fidelity. J Biol Chem 267: 18251–18254

Kunze R, Stochaj U, Laufs J, Starlinger P (1987) Transcription of transposable element *Activator* (*Ac*) of *Zea mays* L. EMBO J 6:1555–1563

Laging M, Fartmann B, Kramer W (2001) Isolation of segments of homologous genes with only one conserved amino-acid region via PCR. Nucleic Acids Res 29:e8

Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, Pelltomäki P, Sistonen P, Aaltonen LA, Nyström-Lahti M, Guan X-Y, Zhang J, Meltzer PS, Yu J-W, Kao F-T, Chen DJ, Cerosaletti KM, Fournier REK, Todd S, Lewis T, Leach RJ, Naylor SL, Weissenbach J, Mecklin J-P, Järvinen H, Petersen GM, Hamilton SR, Green J, Jass J, Watson P, Lynch HT, Trent JM, de la Chapelle A, Kinzler KW, Vogelstein B (1993) Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. Cell 75:1215–1225

Marra G, Iaccarino I, Lettieri T, Roscilli G, Delmastro P, Jiricny J (1998) Mismatch repair deficiency associated with overexpression of the *msh3* gene. Proc Natl Acad Sci USA 95:8568–8573

Marsischky GT, Filosi N, Kane MF, Kolodner R (1996) Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev 10:407–420

McClintock B (1951) Chromosome organization and genic expression. Cold Spring Harbor Symp Quant Biol 16:13–47

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–8.

Pawlowski K, Kunze R, de Vries S, Bisseling T (1994) Isolation of total, poly(A) and polysomal RNA from plant tissues. In: Gelvin SB, Schilperoort RA (eds) Plant molecular biology manual. Kluwer Academic Publishers, Dordrecht, pp D5/1–13

Rayssiguier C, Thaler DS, Radman M (1989) The barrier to recombination between *E. coli* and *S. thyphimurium* is disrupted in mismatch-repair mutants. Nature 342:396–401

Reenan RAG, Kolodner RD (1992a) Characterization of insertion mutations in the *Saccharomyces cerevisiae msh1* and *msh2* genes: evidence for separate mitochondrial and nuclear functions. Genetics 132:975–985

Reenan RAG, Kolodner RD (1992b) Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. Genetics 132:963–973

Ross-Macdonald P, Roeder GS (1994) Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. Cell 79:1069–1080

Selva EM, New L, Crouse GF, Lahue RS (1995) Mismatch correction acts as a barrier to homoeologous recombination in *Saccharomyces cerevisiae*. Genetics 139:1175–1188

Strand M, Earley MC, Crouse GF, Petes TD (1995) Mutations in the *msh3* gene preferentially lead to deletions within tracts of simple repetitive DNA in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 92:10418–10421

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–82.

Vries SS de, Baart EB, Dekker M, Siezen A, de Rooij DG, de Boer P, te Riele H (1999) Mouse MutS-like protein Msh5 is required for proper chromosome synapsis in male and female meiosis. Genes Dev 13:523–531

Wind N de, Dekker N, Berns A, Radman M, te Riele H (1995) Inactivation of the mouse *msh2* gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination and predisposition to cancer. Cell 82:321–330

Worth L, Clark S, Radman M, Modrich P (1994) Mismatch repair proteins MutS and MutL inhibit RecA-catalyzed strand transfer between diverged DNAs. Proc Natl Acad Sci USA 91:3 238–3241