Microbiology (1996), 142, 2021–2029

Printed in Great Britain

Bacillus subtilis mutS mutL operon: identification, nucleotide sequence and mutagenesis

Francesca Ginetti,[1] Marta Perego,[1,2]† Alessandra M. Albertini[1] and Alessandro Galizzi[1]

Author for correspondence: Alessandro Galizzi. Tel: +39 382 505548. e-mail: albertini@ipvgen.unipv.it

[1] Dipartimento di Genetica e Microbiologia, Università degli Studi di Pavia, 207 via Abbiategrasso, 27100 Pavia, Italy

[2] Dipartimento Farmaceutico, Università degli Studi di Parma, Italy

The *Bacillus subtilis mutS* and *mutL* genes, involved in the DNA mismatch repair system, have been cloned and characterized. From sequence analysis the two genes appear to be organized in a single operon, located immediately downstream of the *cotE* gene (approximately 150° on the genetic map). The deduced MutS protein is 49% identical to HexA and MutL is 46% identical to HexB of *Streptococcus pneumoniae*. Deletion of both *mutS* and *mutL* resulted in an increase in the frequency of spontaneous mutations and abolished the marker effect observed in transformation. The expression of the *mut* operon was studied with the use of a *mutSL–lacZ* transcriptional fusion. An increase in expression was observed during late exponential growth.

Keywords: *Bacillus subtilis*, *mut* operon, DNA repair, mutagenesis

INTRODUCTION

The genetic integrity of bacterial genomes is guaranteed by, among other systems, the mismatch repair machinery, acting on mispaired regions derived from replication errors or recombination intermediates. Biosynthetic mistakes generated during chromosomal replication are eliminated by the *mutSHL* excision repair system (Modrich, 1994). In *Escherichia coli* and *Salmonella typhimurium*, mutation is avoided by discrimination between the newly synthesized DNA strand and the template DNA strand, the latter being tagged by *dam*-dependent methylation (Modrich, 1991). A second source of base mispairing is represented by the generation of heteroduplex DNA ensuing from genetic recombination. The dedicated mismatch repair system has been thoroughly described in *Streptococcus pneumoniae* and depends on the products of the genes *hexA* and *hexB* (Claverys & Lacks, 1986). In this case the discrimination between the endogenous strand and the incoming one, seems to be nick-directed and targeted to the incoming strand of the heteroduplex by the presence of strand discontinuities (Prudhomme *et al.*, 1991).

The Mut and Hex systems both involve long excision tracts and even if the initial event is different (the presence of hemimethylated sites in the case of Mut, the availability of single strand ends in the case of Hex), the two mechanisms are very similar and functionally equivalent. In fact the Mut repair pathway has been shown to be involved in maintaining the natural barrier to interspecific recombination between *E. coli* and *Sal. typhimurium* (Rayssiguier *et al.*, 1989). The mismatch recognition functions of MutS and MutL of *E. coli* also act to prevent homologous recombination between duplicated, diverged sequences (Petit *et al.*, 1991) and a possible role of these two repair proteins in RecA-catalysed strand exchange has been suggested by *in vitro* experiments (Worth *et al.*, 1994). Cloning and sequencing of the genes involved in the Hex and Mut systems of *E. coli*, *Sal. typhimurium*, *Strep. pneumoniae* and *Azotobacter vinelandii*, confirmed the functional similarities. The MutS and HexA proteins are highly similar, and so are MutL and HexB (Connoly & Winkler, 1992; Haber *et al.*, 1988; Le *et al.*, 1993; Mankovich *et al.*, 1989; Prudhomme *et al.*, 1989, 1991; Priebe *et al.*, 1988; Schlensong & Boeck, 1991). Negative complementation conferred by the expression in *E. coli* of the *Strep. pneumoniae* HexA function confirmed their common evolutionary origin (Prudhomme *et al.*, 1991).

Functions similar to the bacterial components of the mismatch repair machinery, have been identified in yeast

† **Present address:** The Scripps Research Institute, Division of Cell Biology, La Jolla CA92037, USA.

Abbreviations: i.s., identity score; PY, Penassay Broth; SM, Schaeffer's sporulation medium; Ap, ampicillin; Cm, chloramphenicol; Em, erythromycin; Rf, rifampicin.

The GenBank accession number for the nucleotide sequence reported in this paper is U27343.

Table 1. *B. subtilis* strains used in this study

Strain	Genotype*	Source or reference†
PB19	*Prototroph*	SB 19, E. W. Nester
PB168	*trpC2*	
PB1679	*ilvA1 metB5 nrdA*	E3113, D. Karamata
PB1701	*trpC2 metB4*	GSY228, C. Anagnostopoulos
PB1814	*dnaF69 glnA100*	A. Galizzi
PB1831	*trpC2 pheA1*	JH642, J. Hoch
PB1841	*trpC2 cotE*::pLZ100-HP *cat*	PB168 tf‡ with pLZ100-HP
PB1843	*recA260 metB5 SPβ(−) trpC2 xin-1 erm*	BGSC1A746
PB1856	*trpC2 pheA1 ΔmutSL::cat*	This work
PB1868	*trpC2 ilvC1*	GSY111, C. Anagnostopoulos
PB1869	*trpC2 ilvC1 ΔmutSL::cat*	PB1868 tf‡ with PB1856
PB1870	*trpC2metB*4 *ΔmutSL::cat*	PB1701 tf‡ with PB1856
PB5179	*trpC2 mutL*::pFG2792 *erm*	PB168 tf‡ with pFG2792
PB5180	*trpC2 cotE*::pLZ100-HP *cat* *mutL*::pFG2792 *erm*	PB1841 tf‡ with PB5179
PB5207	*trpC2 pheA1 mutSL::lacZ cat*	PB1831 tf‡ with pFG2783S
PB5211	*recA260 metB5 SPβ(−) trpC2 xin-1 erm* *mutSL::lacZ cat*	PB1843 tf‡ with pFG2783S

* *cat* and *erm* are the determinants for Cm^R and Em^R, respectively.

† BGSC, Bacillus Genetic Stock Center, Ohio, USA; E. W. Nester, Stanford, USA; D. Karamata, Lausanne, Switzerland; C. Anagnostopoulos, Gif-sur-Yvette, France; A. Galizzi, this laboratory; J. Hoch, La Jolla, USA.

‡ Constructed by transformation with either plasmid or chromosomal DNA.

Table 2. Plasmids used in this study

Plasmid	Description (source or reference)
Vectors	
pJH101	Ap Tc Cm; integrative vector (Ferrari *et al.*, 1993)
pJM103	Ap Cm; integrative vector (Perego, 1993)
pJM105A	Ap Cm; integrative vector (Perego, 1993)
pJM109	Em; integrative vector (Perego, 1993)
pDIA5304	Ap Cm; integrative vector, used for plasmid walking (Glaser *et al.*, 1993)
pDH32	Ap Cm; vector used to construct *lacZ* transcriptional fusions and recombine into the chromosome at *amyE* (Shimotsu & Henner, 1986)
Other plasmids	
pLZ100	Clone of *cotE* (2 kb *Hin*dIII fragment) in pBR322 (gift of L. Zheng, Cambridge, USA)
pLZ100-HP	0·6 kb *Hin*dIII–*Hpa*II fragment from pLZ100, upstream of *cotE*, cloned into pJH101
pFG2752	5·3 kb *Bgl*II chromosomal fragment cloned into pJM103
pFG2764A	0·5 kb *Hin*dIII–*Bgl*II fragment from pFG2752 cloned into pDIA5304
pFG2767	2·5 kb *Hin*dIII–*Sac*I fragment in pDIA5304, obtained by *Sac*I digestion and ligation of chromosomal DNA from transformants with pFG2764A
pFG2769	1·5 kb *Bgl*II fragment from pFG2767 cloned into pJM103
pFG2771	3·7 kb *Bgl*II–*Sph*I fragment in pJM103, obtained by *Sph*I digestion and ligation of chromosomal DNA from transformants with pFG2769
pFG2783S	333 bp *Eco*RI–*Bam*HI fragment obtained by PCR, cloned in pDH32, containing the *mutSL* promoter
pFG2791	pJM105A derivative, containing a 687 *Kpn*I–*Cla*I fragment (*mutS* start) and a 419 bp *Pst*I–*Bam*HI fragment (*mutL* end) at the sites of the *cat* cassette; used to make *ΔmutSL::cat*
pFG2792	pJM109 derivative, containing the *Pst*I–*Bam*HI fragment from pFG2791, corresponding to the end of *mutL*; used for genetic mapping

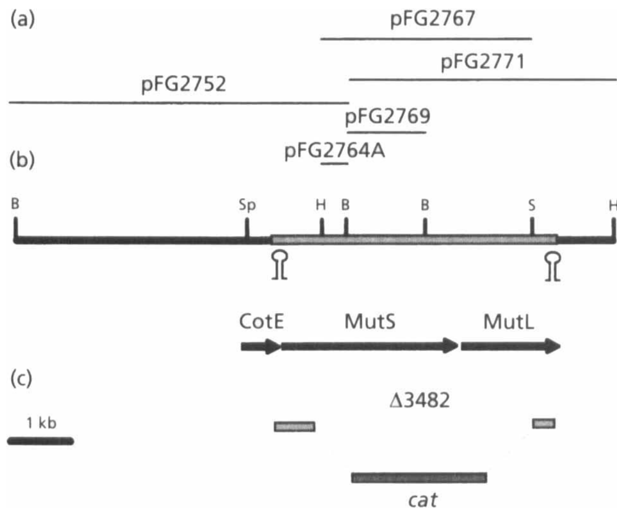

Fig. 1. Map of the *mutSL* operon. (a) Bars represent the inserts in the indicated plasmids. (b) Restriction endonuclease cleavage sites are represented as follows: B, *Bgl*II; H, *Hin*dIII; S, *Sac*I; Sp, *Sph*I. Only relevant sites are reported. The lightly shaded region indicates the sequence described in this paper. The position of two putative terminators is indicated by hairpins. The ORFs are represented by arrows; the direction of transcription is toward the terminus of DNA replication. (c) Δ3482 indicates the extent of the deletion obtained by double cross-over gene replacement obtained upon transformation of the linearized plasmid pFG2791. *cat* is the Cm^R determinant cassette.

(Kramer *et al.*, 1989; Reenan & Kolodner, 1992) and in human cells (Fishel *et al.*, 1993; Leach *et al.*, 1993; Nicolaides *et al.*, 1994; Papadopoulos *et al.*, 1994). In yeast, evidence has been obtained that MSH2 or PMS1 functions (homologues of MutS and MutL) can act as a system of control of symmetric heteroduplex length during recombination (Alani *et al.*, 1994). The human MutS homologue (hMSH2) and the three human MutL homologues (hMLH1, hPMS1 and hPMS2) are involved in a DNA surveillance defect in cancer cells: loss of mismatch repair proficiency is responsible for the common cancer predisposition syndrome known as HNPCC (hereditary nonpolyposis colorectal cancer) (Bronner *et al.*, 1994; Leach *et al.*, 1993; Nicolaides *et al.*, 1994; Papadopoulos *et al.*, 1994).

In this work we report the cloning, sequence determination and functional characterization by targeted mutagenesis of the *Bacillus subtilis mutS* and *mutL* genes. The two genes are organized in an operon: we present preliminary data on the regulation of expression, in particular during the transition to stationary phase.

METHODS

Media. *B. subtilis* was grown in Schaeffer's sporulation medium (SM) (Schaeffer *et al.*, 1965) or Penassay Broth (PY; Difco). Antibiotics were used at the following concentrations: ampicillin (Ap), 100 μg ml^{-1}; chloramphenicol (Cm), 5 μg ml^{-1}; erythromycin (Em), 1 μg ml^{-1}. Transformation of *B. subtilis* and PBS1-mediated transduction were performed according to Hoch *et al.* (1967). Non-temperature sensitive (ts^+) recombinants were selected on nutrient agar plates incubated at 47 °C.

Strains, plasmids and PCR. Standard *E. coli* strains were used for cloning (Sambrook *et al.*, 1989). The *B. subtilis* strains used are listed in Table 1, and the plasmids in Table 2. To evaluate the expression of the *mutSL* operon, we constructed a *mutS–lacZ* transcriptional fusion that contained 330 bp from the end of *cotE* to codon 50 of *mutS*. The fragment was obtained by PCR performed with chromosomal DNA derived from strain PB168 (*trpC2*). The forward primer (FG1001) was a 26-mer with the sequence 5′ GGAATTCCGGAGTTTTAGTGGGAGA 3′ (the underlined bases correspond to the end of *cotE*). The 5′ part of the primer was designed to contain an *Eco*RI recognition site. The backward primer (FG1002) was a 26-mer derived from *mutS* and with a *Bam*HI recognition site at its 5(end. The sequence of FG1002 was 5′ CGGATCC-GCACCGCCGTCTCTGCTCG 3′ (the underlined bases correspond to the sequence derived from *mutS*). PCR was run for 30 cycles with the following parameters: denaturation, 3 min at 93 °C for the first cycle and 1 min at 92 °C thereafter; annealing, 1 min at 51 °C; extension, 15 s at 72 °C. The PCR product was purified by PAGE and electroelution, then digested with *Eco*RI and *Bam*HI and cloned into pDH32 (Shimotsu & Henner, 1986) upstream of *lacZ*. The plasmid with the fusion was designated pFG2783S (Table 2). It was linearized and used to transform strains PB1831 and PB1843 selecting for Cm^R. Double cross-over events that resulted in the introduction of the *mutS–lacZ* fusion at the *amyE* locus caused an Amy^- phenotype.

A strain with a deletion of *mutS* and *mutL* was constructed by gene replacement by a double cross-over and insertion of a Cm^R determinant. To this purpose a 687 bp fragment containing the *mutS* promoter and the first 572 bp of the coding sequence was produced by PCR on chromosomal DNA. The forward primer (FG1003) was a 26-mer with the sequence 5′ GGGGTACC ACAGTACCCCCAAGTCTC 3′ (the underlined bases correspond to the *mutS* sequence, preceded 5′ by a *Kpn*I recognition site). The backward primer (FG1004), was a 26-mer with the sequence 5′ CCATCGATACCGCACCGCTCTCTCAG 3′, with a *Cla*I recognition site at its 5′ end. PCR was run for 35 cycles with the following parameters; denaturation, 3 min at 93 °C for the first cycle and 1 min at 92 °C thereafter; annealing, 1 min at 53 °C; extension, 45 s at 72 °C. The PCR product was purified as above and cloned into pJM105A restricted with *Kpn*I and *Cla*I. The downstream fragment of 418 bp, corresponding to the end of *mutL*, was also obtained by PCR with the following primers: FG1005, 5′ AACTGCAGCTGACGTT CCACTACTCC 3′, with a *Pst*I restriction site at its 5′ end, and FG1006, 5′ CGGGATCCTACCACCCCCGCTACATC3′, with a *Bam*HI site at its 5′ end. PCR was run for 30 cycles with the following parameters; denaturation, 3 min at 93 °C for the first cycle and 1 min at 95 °C thereafter; annealing, 1 min at 52 °C; extension, 20 s at 72 °C. The purified PCR product was digested with *Pst*I and *Bam*HI and cloned in the pJM105A derivative containing the upstream fragment. The plasmid containing the two fragments on either side of *cat*, pFG2791, was linearized and transformed into strain PB1831 selecting for Cm^R.

***β*-Galactosidase assays.** All assays were performed on triplicate samples and the values were averaged. *β*-Galactosidase specific activity is expressed as Miller units (mg $protein)^{-1}$ calculated as described previously (Albertini & Galizzi, 1990).

DNA sequencing. T7 DNA polymerase (Pharmacia LKB) was used to sequence double-stranded plasmid DNA. The sequence of both strands was determined using custom-made primers. Oligonucleotides were synthesized by the phosphoramidite method by means of a Cyclon Plus DNA synthesizer (Milligen Biosearch).

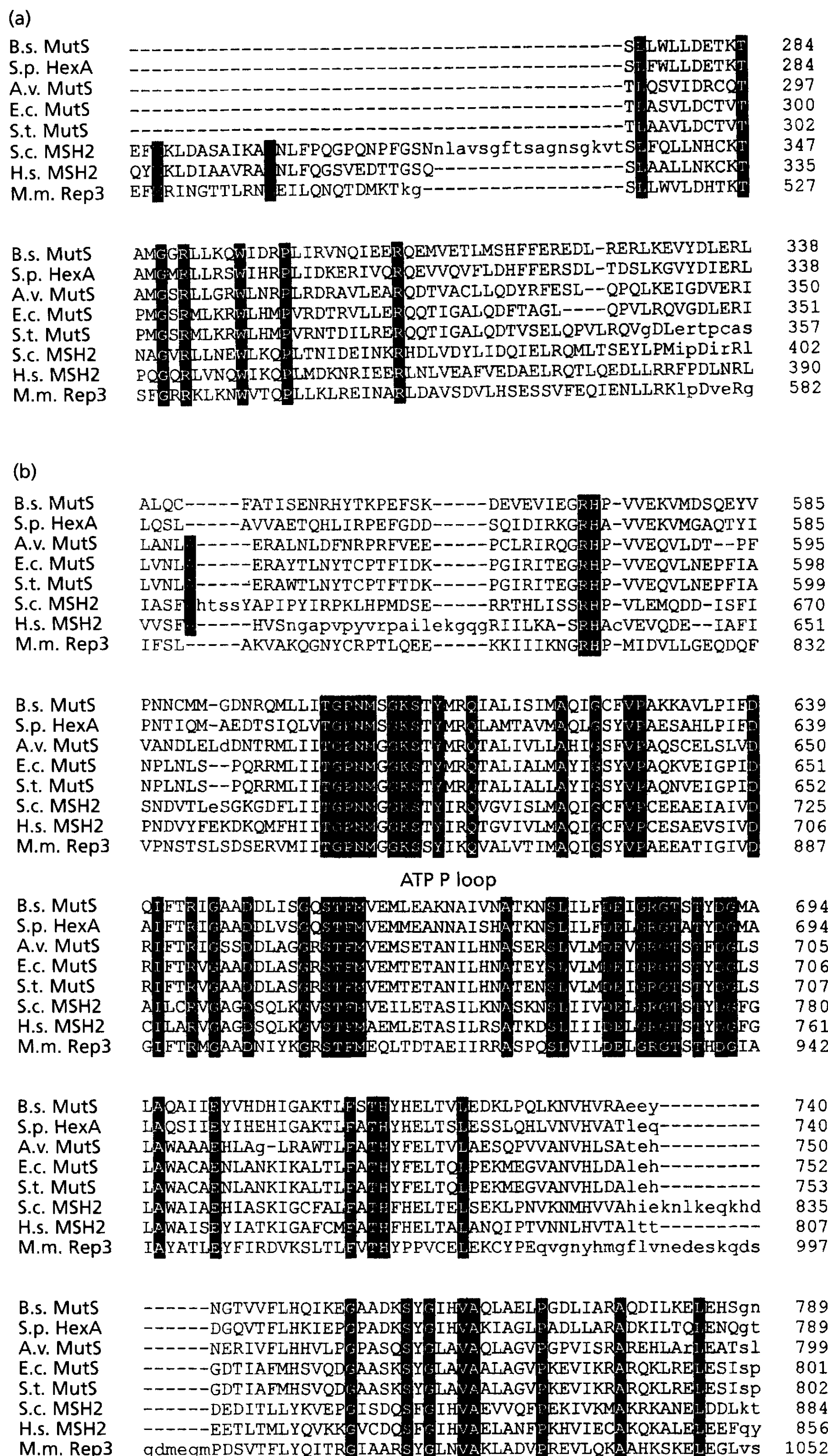

Fig. 2. Multiple alignment of the amino acid sequence of MutS proteins. Shaded boxes indicate universally conserved amino acids. Two stretches (a and b) of significant similarity are shown. The conserved signature of the ATP binding site (P loop) is indicated in (b). Species abbreviations used and the DNA or protein sequence accession numbers are: B.s., *B. subtilis* (U27343); S.p., *Strep. pneumoniae* (P10564); A.v., *A. vinelandii* (P27345); E.c., *E. coli* (P23909); S.t., *Sal. typhimurium* (P10399); S.c., *Saccharomyces cerevisiae* (P25847); H.s., *Homo sapiens* (U03911); M.m., *Mus musculus* (P13705).

RESULTS AND DISCUSSION

Cloning the *B. subtilis mutS* and *mutL* genes

As part of an international project to sequence the *B. subtilis* genome, we cloned three overlapping fragments in the *cotE* region at 150° on the chromosome. The first fragment (present in plasmid pFG2752) was obtained by cloning *Bgl*II-cut chromosomal DNA in the integrative vector pJM103 (Perego, 1993) and hybridizing with a 2 kb *Hin*dIII probe obtained from the *cotE*-carrying plasmid pLZ100 (Fig. 1). A pFG2752 subclone was constructed in the integrative plasmid pDIA5304 (Glaser *et al.*, 1993), which, as for pJM103, confers Cm^R upon integration into the *B. subtilis* chromosome, giving rise to plasmid pFG2764A (Fig. 1 and Table 2). The chromosome walking technique was used to clone additional fragments further downstream of *cotE*. In a first step, *Sac*I restriction and plasmid rescue carried out on chromosomal DNA containing integrated pFG2764A resulted in plasmid pFG2767. Its derivative pFG2769 was used for an additional walking step, generating plasmid pFG2771.

The nucleotide sequence of the cloned region showed the presence of *cotE* and, downstream of it, two ORFs of 2577

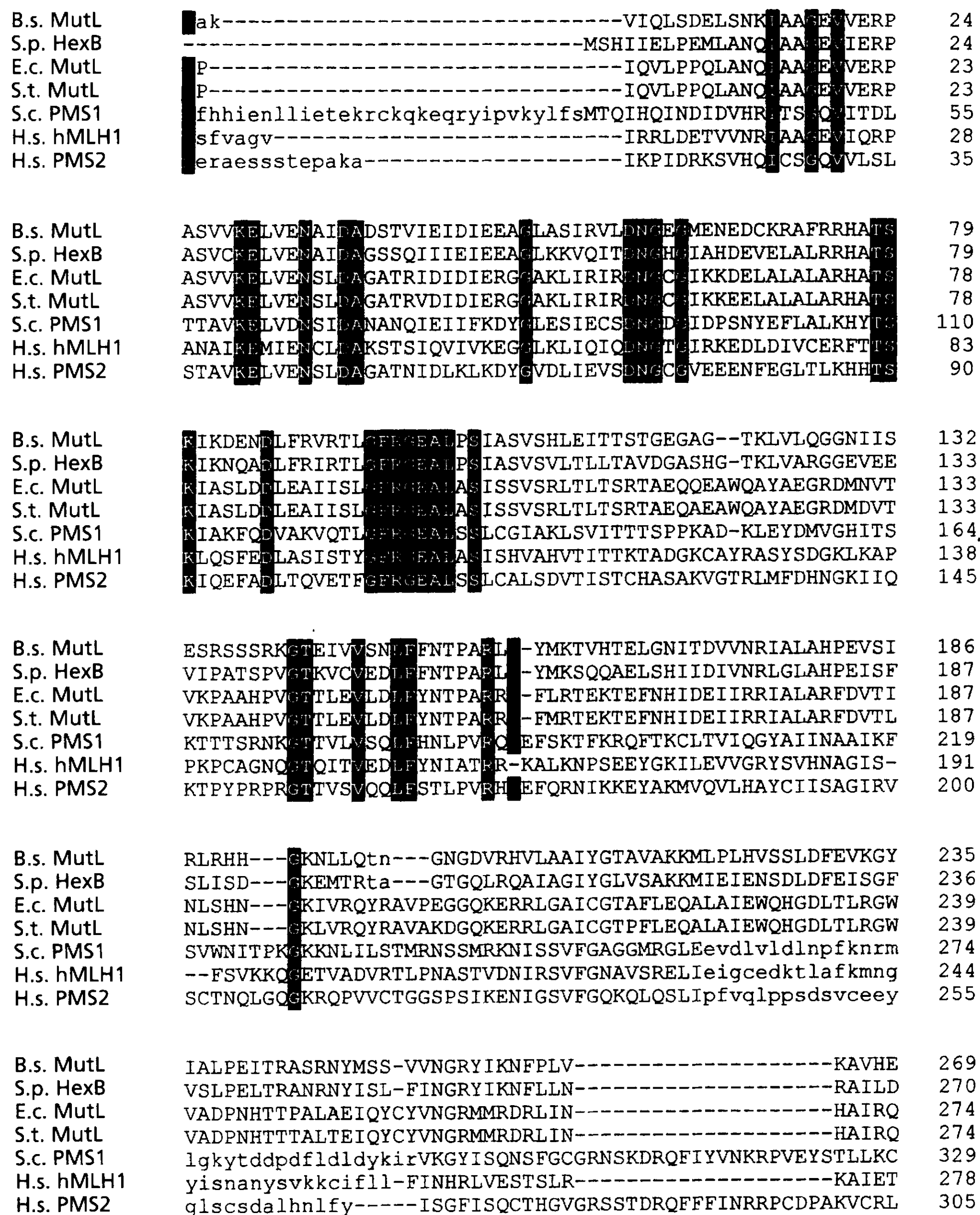

Fig. 3. Multiple alignment of the amino acid sequence of MutL proteins. Shaded boxes indicate universally conserved amino acids. Only the amino terminus of the proteins is shown. Species abbreviations used and the DNA or protein sequence accession numbers are: B.s., *B. subtilis* (U27343); S.p., *Strep. pneumoniae* (P14161); E.c., *E. coli* (P23367); S.t., *Sal. typhimurium* (P14160); S.c., *Sacch. cerevisiae* (P14242); H.s. hMLH1, *H. sapiens* (P23367); H.s. hPMS2, *H. sapiens* (2017356B).

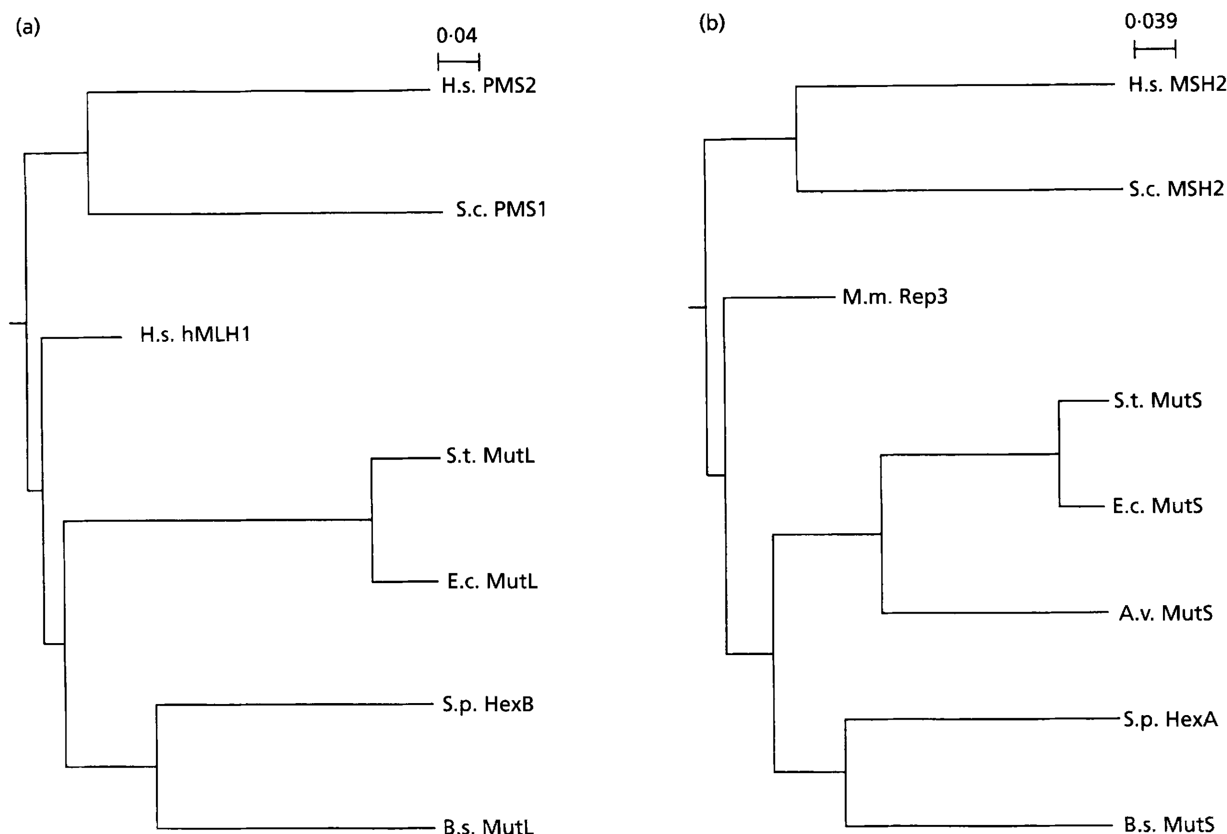

Fig. 4. Phylogenetic analysis of MutS (a) and MutL (b) proteins graphically displayed with the NJPLOT software of M. Gouy (University of Lyon, France). The scale bars indicate the distances as calculated from the multiple alignment (Thompson *et al.*, 1994). Abbreviations are defined in the legends to Figs 2 and 3.

and 1884 bp, respectively. The first ORF starts 151 bp from the end of *cotE*. The ATG start codon is preceded by a sequence that could act as a ribosome binding site. A search for transcription control signal sequences in the intercistronic region showed the presence downstream of a hairpin loop with a ΔG of -110 kJ mol^{-1}, a putative terminator of *cotE* and a potential σ^{A}-dependent -10 sequence, perfectly matching the consensus. The putative -35 region was hardly detectable. The second ORF starts 16 bp after the stop codon of the first one. No terminator site or recognizable promoter was identified between the two ORFs, suggesting that they belong to the same operon. The operon is followed by a strong factor-independent transcription termination signal ($\Delta G - 140{\cdot}6$ kJ mol^{-1}).

With the FASTA (Pearson, 1990) and BLAST programs (Altschul *et al.*, 1990) we compared the two deduced protein sequences to other proteins available in the data banks. For the 2577 bp ORF, significant homology was found to *Strep. pneumoniae* mismatch protein HexA (49·4 % identity score, i.s.), to *A. vinelandii*, *E. coli* and *Sal. typhimurium* mismatch protein MutS (38·4, 40·2 and 39·4 % i.s., respectively), and to yeast MSH2 and human hMSH2 (stretches with 31·9 and 32·4 % i.s., respectively). Multiple alignment revealed a significant degree of similarity over the whole polypeptide sequence with the prokaryotic homologue, particularly with HexA; with the eukaryotic homologue the similarity was particularly striking in the last third of the polypeptides starting from the ATPase consensus site as shown in Fig. 2. The 1884 bp ORF was found to be homologous to *Strep. pneumoniae* HexB (45·6 % i.s.), to *E. coli* and *Sal. typhimurium* MutL (30·7 and 30·1 % i.s., respectively) and to yeast and human DNA mismatch repair proteins PMS1, hMLH1 and PMS2 (stretches with 28·1, 32·7 and 30·1 % i.s., respectively). As for MutS, the *B. subtilis* polypeptide is similar to the prokaryotic counterparts over the whole sequence, whereas the similarity to the eukaryotic homologues is restricted to the amino-terminal half (Fig. 3).

The high level of identity observed for both deduced proteins led us to name the genes *mutS* and *mutL*, respectively. The evolutionary relatedness, as suggested by comparison performed with CLUSTALPPC (Thompson *et al.*, 1994) and NJPLOT for the construction of phylogenetic trees (Fig. 4), suggests a common evolutionary origin for the prokaryotic and eukaryotic mismatch recognition functions of the Mut pathway. As expected, the Gram-positive and Gram-negative bacteria form two different branches of the tree.

As discussed above, the organization of *mutS* and *mutL* in *B. subtilis* strongly suggests that the two genes are part of a single operon. Among the bacterial systems described, this is the first instance of such an organization that could be advantageous to the cell leading to the co-ordinate

expression of two proteins involved in mismatch repair. The hypothesis of common regulation for the *Strep. pneumoniae* mismatch recognition functions HexA and HexB, was first proposed by Prudhomme *et al.* (1989) following the observation of conserved sequences in the −10 promoter and Shine–Dalgarno region of the two genes. As suggested by the authors, the co-ordinate expression of the MutS and MutL functions could be of special importance, not only to avoid the mutator effect of unbalanced production of one of the two elements, but also in view of the constraints imposed by the differentiated stage of competence to transformation. Such a co-ordinate expression seems to be obtained in *Strep. pneumoniae* by the presence of a common putative regulatory sequence upstream of the two genes and in *B. subtilis* by the clustering of *mutS* and *mutL* in a single operon. Nevertheless inspection of the *B. subtilis* sequence upstream of *mutS* showed good homology with the sequences present at the promoters of *hexA* and *hexB* genes (Prudhomme *et al.*, 1989) (data not shown). If the observed sequence is indeed involved in regulation, the presence in the *mutSL* operon of *B. subtilis* may indicate that its expression is co-ordinately regulated with the expression of other, as yet undetected, genes or operons.

Genetic map

The *cotE* gene has been mapped by PBS1-mediated transduction between *pyrD1* and *thyA1*, at about 150° on the genetic map (Zheng *et al.*, 1988). To narrow the location of *cotE*, we performed transductions with PBS1 lysates from PB1841 (*cotE–cat*) used to transduce to ts^+ two temperature-sensitive *polC* and *nrdA* mutants (PB1814 and PB1679, respectively). Two-factor crosses indicated that *cotE* (*cat*) had a linkage of 70% to *polC* and 50% to *nrdA*. To obtain information about the direction of transcription, we did a three-factor cross with *cotE–cat*, *mutL–erm* and *polC*. Strain PB1814 was transduced to ts^+ with a PBS1 lysate from PB5180 (*cotE–cat*, *mutL–erm*). *cat* was co-transduced with *polC* in 56 of 110 transductants; most Cm^R transductants were also Em^R (51 of 56), and only five were Em^S. No Cm^S Em^R transductants were obtained, indicating that the gene order is *polC–cotE–mutL*. Thus the direction of transcription of *cotE* and *mutSL* is toward the terminus of DNA replication.

Construction of a *mutSL* deletion mutant

To prove the effective role of the two genes in the repair process, we constructed a deletion mutant in which both MutS and MutL could not be synthesized. Two fragments, obtained by PCR amplification, were cloned on either side of the Cm^R determinant of the vector pJM105A (Perego, 1993). The first fragment, 687 bp long, contained the *mutS* promoter and the first 572 bp of the gene, whereas the second represented the last 419 bp of the *mutL* gene (Fig. 1). The ensuing plasmid, pFG2791, was linearized with *Kpn*I and used to transform competent cells of strain PB1831, selecting for Cm^R. The chromosomal DNA of the transformants was analysed by *Hin*dIII digestion and Southern hybridization with a suitable probe. One deletion-bearing strain (PB1856) was used to test the role of the two genes in preventing, through mismatch repair, the accumulation of mutations. We compared the spontaneous mutation rate of strain PB1831 to that of PB1856. We found that in the mutant strain the frequency of rifampicin-resistant (Rf^R) mutants was from 40–60 times higher than that observed for the parental strain (Table 3). To further confirm the role of *mutS* and *mutL*, we compared the reversion of phenylalanine auxotrophy between the two strains. The number of Phe^+ revertants in strain PB1856 was 16·5-fold greater than that of the control strain (Table 3). In addition, colonies of strain PB1856 showed an altered morphology on SM. Transparent sectors were clearly visible after 2–3 d incubation, suggesting the formation of sporulation-deficient clones.

Table 3. Spontaneous frequencies of mutation of parental and *mutSL*-deleted strains

Strain		Frequency of Rf^R mutants
PB1831 (parental)	Exp. 1	$0{\cdot}22 \times 10^{-6}$
	Exp. 2	$0{\cdot}08 \times 10^{-6}$
	Exp. 3	$0{\cdot}05 \times 10^{-6}$
PB1856 (Δ*mutSL::cat*)	Exp. 1	$14{\cdot}1 \times 10^{-6}$
	Exp. 2	$3{\cdot}7 \times 10^{-6}$
	Exp. 3	$4{\cdot}6 \times 10^{-6}$
		Frequency of Phe^+ revertants
PB1831 (parental)		$0{\cdot}6 \times 10^{-8}$
PB1856 (Δ*mutSL::cat*)		$9{\cdot}9 \times 10^{-8}$
		Frequency of Ile^+ revertants
PB1868 (parental)		$0{\cdot}46 \times 10^{-6}$
PB1869 (Δ*mutSL::cat*)		$8{\cdot}0 \times 10^{-6}$

An increase in the number of revertants was also observed for the *ilvC1* marker, when present in a *mutSL*-deleted background (Table 3). The *trpC2* and *metB4* markers were not affected in their reversion by the absence of the mismatched repair system (data not shown).

In addition to the mutator phenotype, *Strep. pneumoniae hex* mutants show identical transformation efficiencies for different markers that in the wild-type strain have different frequencies of transformation, due to the different molecular basis of the mutation characterizing the marker and consequently the heteroduplex molecule produced during recombination (Lacks, 1970).

We tested for such a phenotype associated with the *mutSL* deletion in *B. subtilis*. We choose two markers (*ilvC1* and *metB4*) which were previously shown to give different relative frequencies of transformation, probably due to the differences in the efficiency of integration of the donor marker into the chromosome (Barat *et al.*, 1965). The *ilvC1* marker, at a non-saturating DNA concentration, is transformed less efficiently than the *metB4* marker. In *mutSL*-deleted strains, the two markers are transformed with the same efficiency (Table 4), supporting the concept

Table 4. Frequency of transformation of *ilvC1* and *metB4* markers relative to *trpC2*

The pairs of numbers refer to the results of two independent experiments.

Marker	*mutSL*	Δ*mutSL**
ilvC1	0·38, 0·33	1·20, 0·87
metB4	1·00, 1·43	0·90, 0·91

* The ratio was calculated after subtraction of the background due to reversion.

that different efficiencies of integration are caused by the discrimination targeted to the incoming donor strand by the mismatch repair system recognizing certain types, and not others, of heteroduplex.

Expression of the *mutSL* operon

To study the expression of the *mut* operon, we constructed a *mut–lacZ* fusion. A 333 bp fragment containing the *mutS* promoter region was amplified via PCR and cloned upstream of a promoterless *E. coli lacZ* gene in the expression vector pDH32 (Shimotsu & Henner, 1986). The plasmid obtained (pFG2783S) was used after linearization to transform strain PB1831. Stable double cross-over integrants in the *amyE* locus were checked for an inability to degrade starch. β-Galactosidase assays were then performed with the new strain using different media.

In rich media (SM or PY) the *mut* operon is constitutively expressed during vegetative growth. While the cells undergo sporulation, the gene is progressively turned off (data not shown). The levels of expression are in general rather low: the β-galactosidase levels never exceeded 60 Miller units. Growth in minimal medium showed an interesting aspect. A slight increase in expression was constantly observed when the culture entered stationary phase at about t_{-1} (Fig. 5). This may be interpreted as due to the fraction of cells entering the competent state. An active mismatch repair complex may be required during the competent phase to prevent mispairing and it could be responsible for the observed correction of heteroduplexes in the course of transformation (Bresler *et al.*, 1968).

The possible correlation between the increased expression of *mutSL* and competence deserves further investigation, especially in consideration of the reported lack of induction of mismatch repair genes in competent cells of *Strep. pneumoniae* (Humbert *et al.*, 1995).

Expression of the *mut–lacZ* fusion was studied in various mutant backgrounds. It was not affected by deletion of *comK* or *comA–P* which encode regulatory proteins of late competence genes. In a *recA* mutant strain the expression was not affected during exponential growth, whereas the subsequent decrease observed during stationary phase was relieved and the level of β-galactosidase remained constant until 4 h after t_0 (Fig. 5). This observation suggests a possible regulatory role of the RecA protein on the expression of the mismatch repair system. At present we cannot discriminate between a direct or indirect effect of the *recA* mutation.

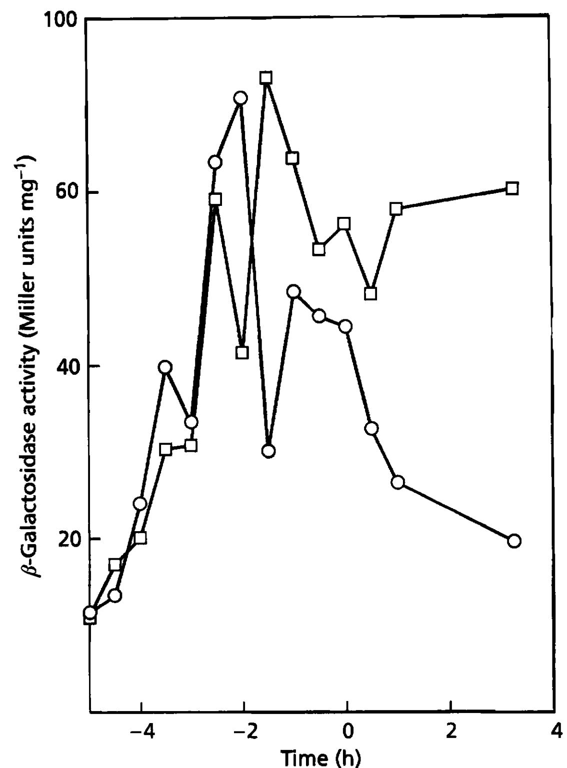

Fig. 5. Expression of *mutSL–lacZ* fusion. The specific activity of β-galactosidase was monitored in minimal medium in a wild-type strain (○) or in a *recA* mutant (□) carrying a *mutSL–lacZ* transcriptional fusion at the *amyE* locus.

The fact that the mismatch repair system is not constitutively expressed but may be under genetic and physiological control, could shed some light on diverse phenomena such as adaptive mutation and spontaneous mutation accumulations in cancers (Longerich *et al.*, 1955).

ACKNOWLEDGEMENTS

We thank Lingbiae Zheng for kindly providing plasmid pLZ100. This work was supported in part by the European Commission under the Biotechnology programme (contract Bio2-CT93–0272) and by Consiglio Nazionale delle Ricerche (Grants 92.02183.CT 174 AND 93.02009.CT14).

REFERENCES

Alani, E., Reenan, R. A. G. & Kolodner, R. (1994). Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. *Genetics* **137**, 19–39.

Albertini, A. M. & Galizzi, A. (1990). The *Bacillus subtilis outB* gene is highly homologous to an *Escherichia coli ntr*-like gene. *J Bacteriol* **172**, 5482–5485.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic Local Alignment Search Tool. *J Mol Biol* **215**, 403–410.

Barat, M., Anagnostopoulos, C. & Schneider, A.-M. (1965). Linkage relationships of genes controlling isoleucine, valine and leucine biosynthesis in *Bacillus subtilis*. *J Bacteriol* **90**, 357–369.

Bresler, S. E., Kreneva, R. A. & Kushev, V. V. (1968). Correction of molecular heterozygotes in the course of transformation. *Mol Gen Genet* **102**, 257–268.

Bronner, C. E. and others (1994). Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary nonpolyposis colon cancer. *Nature* **368**, 258–261.

Claverys, J. P. & Lacks, S. A. (1986). Heteroduplex deoxyribonucleic acid base mismatch repair in bacteria. *Microbiol Rev* **50**, 133–165.

Connoly, D. M. & Winkler, M. E. (1992). Structure of *Escherichia coli* K12 *miaA* and characterization of the mutator phenotype caused by *miaA* insertion mutations. *J Bacteriol* **173**, 1711–1721.

Ferrari, F. A., Nguyen, A., Lang, D. & Hoch, J. A. (1983). Construction and properties of an integrable plasmid for *Bacillus subtilis*. *J Bacteriol* **154**, 1513–1515.

Fishel, R., Lescoe, M. K., Rao, M. R. S., Copeland, N. G., Jenkins, N. A., Gaber, J., Kane, M. & Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**, 1027–1038.

Glaser, P. and others (1993). *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region from 325° to 333°. *Mol Microbiol* **10**, 371–384.

Haber, L. T., Pang, P. P., Sobell, D. I., Mankovich, J. A. & Walker, G. C. (1988). Nucleotide sequence of the *Salmonella typhimurium mutS* gene required for mismatch repair: homology of MutS and HexA of *Streptococcus pneumoniae*. *J Bacteriol* **170**, 197–202.

Hoch, J. A., Barat, M. & Anagnostopoulos, C. (1967). Transformation and transduction in recombination-defective mutants of *Bacillus subtilis*. *J Bacteriol* **93**, 1925–1937.

Humbert, O., Prudhomme, M., Hakenbeck, R., Dowson, C. G. & Claverys, J. P. (1995). Homologous recombination and mismatch repair during transformation in *Streptococcus pneumoniae*: saturation of the Hex mismatch repair system. *Proc Natl Acad Sci USA* **92**, 9052–9056.

Kramer, W., Kramer, B., Williamson, M. S. & Fogel, S. (1989). Cloning and nucleotide sequence of DNA mismatch repair gene PMS-1 from *Saccharomyces cerevisiae*: homology of PMS-1 to procaryotic MutL and HexB. *J Bacteriol* **171**, 5339–5346.

Lacks, S. (1970). Mutants of *Diplococcus pneumoniae* that lack deoxyribonucleases and other activities possibly pertinent to genetic transformation. *J Bacteriol* **101**, 373–383.

Le, O., Shen, B., Iismaa, S. E. & Burgess, B. K. (1993). *Azotobacter vinelandii mutS*: nucleotide sequence and mutant analysis. *J Bacteriol* **175**, 7707–7710.

Leach, F. S. and others (1993). Mutations of a *mutS* homolog in hereditary nonpolyposis colorectal cancer. *Cell* **75**, 1215–1225.

Longerich, S., Galloway, A. M., Harris, R. S., Wong, C. & Rosenberg, S. M. (1995). Adaptive mutation sequences reproduced by mismatch repair deficiency. *Proc Natl Acad Sci USA* **92**, 12017–12020.

Mankovich, J. A., Mcintyre, C. A. & Walker, G. C. (1989). Nucleotide sequence of the *Salmonella typhimurium mutL* gene required for mismatch repair: homology of MutL and HexB of *Streptococcus pneumoniae* and to PMS1 of the yeast *Saccharomyces cerevisiae*. *J Bacteriol* **171**, 5325–5331.

Modrich, P. (1991). Mechanisms and biological effects of mismatch repair. *Annu Rev Genet* **25**, 229–253.

Modrich, P. (1994). Mismatch repair, genetic stability, and cancer. *Science* **266**, 1959–1960.

Nicolaides, N. C. and others (1994). Mutations of two PMS homologues in hereditary nonpolyposis cancer. *Nature* **371**, 75–80.

Papadopoulos, N. and others (1994). Mutation homolog in hereditary colon cancer. *Science* **263**, 1625–1629.

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol* **183**, 63–98.

Perego, M. (1993). Integrational vectors for genetic manipulations in *Bacillus subtilis*. In *Bacillus subtilis and Other Gram-positive Bacteria*, pp. 615–624. Edited by A. L. Sonenshein, J. A. Hoch & R. Losick. Washington, DC: American Society for Microbiology.

Petit, M. A., Dimpfl, J., Radman, M. & Echols, H. (1991). Control of large chromosomal duplications in *Escherichia coli* by the mismatch repair system. *Genetics* **129**, 327–332.

Priebe, S. D., Hadi, S. M., Greenberg, B. & Lacks, S. A. (1988). Nucleotide sequence of the *hexA* gene for mismatch repair in *Streptococcus pneumoniae* and homology of HexA to MutS of *Escherichia coli* and *Salmonella typhimurium*. *J Bacteriol* **170**, 190–196.

Prudhomme, M., Martin, B., Méjean, V. & Claverys, J. P. (1989). Nucleic acid sequence of the *Streptococcus pneumoniae hexB* mismatch repair gene: homology of HexB to MutL of *Salmonella typhimurium* and to PMS1 of *Saccharomyces cerevisiae*. *J Bacteriol* **171**, 5332–5338.

Prudhomme, M., Méjean, V., Martin, B. & Claverys, J. P. (1991). Mismatch repair genes of *Streptococcus pneumoniae*: HexA confers a mutator phenotype in *Escherichia coli* by negative complementation. *J Bacteriol* **173**, 7196–7203.

Rayssiguier, C., Thaler, D. & Radman, M. (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. *Nature* **342**, 396–401.

Reenan, R. A. G. & Kolodner, R. D. (1992). Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. *Genetics* **132**, 963–973.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Schaeffer, P., Millet, J. & Aubert, J. (1965). Catabolite repression of bacterial sporulation. *Proc Natl Acad Sci USA* **54**, 704–711.

Schlensong, V. & Boeck, A. (1991). The *Escherichia coli fdv* gene probably encodes MutS and is located at minute 58.8 adjacent to *hyc–hyp* gene cluster. *J Bacteriol* **173**, 7414–7415.

Shimotsu, H. & Henner, D. J. (1986). Construction of a single-copy integration vector and its use in analysis of regulation of the *trp* operon of *Bacillus subtilis*. *Gene* **43**, 85–94.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

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**, 3238–3241.

Zheng, L., Donovan, W. P., Fitz-James, P. C. & Losick, R. (1988). Gene encoding a morphogenetic protein required in the assembly of the outer coat of the *Bacillus subtilis* endospore. *Genes Dev* **2**, 1047–1054.

Received 3 October 1995; revised 28 February 1996; accepted 11 March 1996.