Cell, Vol. 103, 711–721, November 22, 2000, Copyright ©2000 by Cell Press

Evolutionary Implications of the Frequent Horizontal Transfer of Mismatch Repair Genes

Erick Denamur,[1] Guillaume Lecointre,[2] Pierre Darlu,[4] Olivier Tenaillon,[3] Cécile Acquaviva,[1] Chalom Sayada,[5] Ivana Sunjevaric,[6] Rodney Rothstein,[3,6] Jacques Elion,[1] François Taddei,[3] Miroslav Radman,[3,7] and Ivan Matic[3]
[1] INSERM U 458
Hôpital Robert Debré
48 boulevard Sérurier
75935 Paris Cedex 19
[2] Service de Systématique Moléculaire
(GDR CNRS 1005)
Muséum National d'Histoire Naturelle
43 rue Cuvier
75231 Paris Cedex 05
[3] INSERM E9916
Faculté de Médecine Necker-Enfants Malades
Université Paris V
156 rue Vaugirard
75730 Paris Cedex 15
[4] INSERM U 155
80 rue du Général Leclerc
94276 Kremlin-Bicêtre
[5] ACTGene Recherche et Développements
Allée Christophe Colomb
Génopole
91035 Evry Cedex
France
[6] Department of Genetics and Development
Columbia University
College of Physicians and Surgeons
New York, New York 10032

Summary

Mutation and subsequent recombination events create genetic diversity, which is subjected to natural selection. Bacterial mismatch repair (MMR) deficient mutants, exhibiting high mutation and homeologous recombination rates, are frequently found in natural populations. Therefore, we have explored the possibility that MMR deficiency emerging in nature has left some "imprint" in the sequence of bacterial genomes. Comparative molecular phylogeny of MMR genes from natural *Escherichia coli* isolates shows that, compared to housekeeping genes, individual functional MMR genes exhibit high sequence mosaicism derived from diverse phylogenetic lineages. This apparent horizontal gene transfer correlates with hyperrecombination phenotype of MMR-deficient mutators. The sequence mosaicism of MMR genes may be a hallmark of a mechanism of adaptive evolution that involves modulation of mutation and recombination rates by recurrent losses and reacquisitions of MMR gene functions.

[7] To whom correspondence should be addressed (e-mail: radman@necker.fr).

Introduction

Darwinian evolution is the selection of the fittest variants produced by mutation and/or recombination. The contribution of each mechanism to the generation of genetic diversity varies between species. It has been estimated for *Escherichia coli* that, in spite of the clonal structure of its natural populations (Desjardins et al., 1995), any single nucleotide change is about 50 times more likely to have occurred by recombination (with a partner carrying such nucleotide change) than by a de novo mutation (Guttman and Dykhuizen, 1994). The rate of successful gene transfer between bacterial populations in nature depends on numerous factors such as ecological isolation, host range of genetic exchange vectors, activity of the recipient cell nucleases, DNA sequence divergence between recombining molecules, type and efficiency of the recipient cell recombination machinery and, finally, the hybrid fitness (Matic et al., 1996). Some of those factors have been studied in the laboratory, but their impact on gene flow in nature is generally unknown.

Molecular mechanisms involved in the control of homologous recombination in bacteria have been extensively studied. For example, it was shown that the frequency of integration of donor DNA in the chromosome of recipient cells decreases exponentially with increasing sequence divergence between the two recombining DNA molecules (Zawadzki et al., 1995; Vulic et al., 1997; Majewski et al., 2000). The degree and pattern of DNA divergence determine actions of the enzymes that control the outcome of recombination (Radman and Wagner, 1993). The initiation of recombination in bacteria depends on the activity of the RecA protein, which is selective for sequence identity only at the initial stage of the strand exchange process (Shen and Huang, 1986). Once initiated, RecA-mediated strand exchange can occur in spite of large numbers of mismatches and even large heterologies (Bianchi and Radding, 1983; Lichten and Fox, 1984). During the strand exchange process, fidelity of recombination is controlled by the methyl-directed mismatch repair (MMR) system. MMR proteins recognize mispaired and unpaired bases in the joint heteroduplex regions and block RecA-catalyzed strand transfer (Worth et al., 1994). Thus, MMR is a potent inhibitor of recombination between nonidentical DNA sequences (Rayssiguier et al., 1989; Shen and Huang, 1989; Matic et al., 1995; Vulic et al., 1997). Even low divergence is sufficient to impede recombination. For example, transductional recombination between two serovars (Typhimurium and Typhi) of *Salmonella enterica*, whose genomes differ only 1%–2% at DNA sequence level, is increased 10^2- to 10^3-fold in MMR-deficient genetic backgrounds (Zahrt and Maloy, 1997).

MMR system is also involved in the control of replication fidelity. Inactivation of MMR genes increases mutation rates 10^2- to 10^3-fold (Glickman and Radman, 1980; Schaaper and Dunn, 1987). Studies of mutation rates in natural populations of *E. coli*, *S. enterica*, and *Pseudomonas aeruginosa* show that MMR-deficient strains often exceed a frequency of 1% (LeClerc et al., 1996; Matic

Table 1. Characteristics of the Phylogenetic Data from Mismatch-Repair, *mutT*, and *recD* Genes

Gene	Mutator Potency[a]	Nucleotides Total	Nucleotides Variable	Nucleotides Informative for Parsimony[b]	Amino Acids Total	Amino Acids Variable	Non-synonymous Base Substitutions[c] (%)	ILD Test[d] (P value)	Incongruent Taxa[e]	Length of Recombined Sequence[f] (No. of Nucleotides) Range	Length of Recombined Sequence[f] (No. of Nucleotides) Median Value
mutS	10^2–10^3	390	75	55	127	2	1.3	0.001/0.250	8	6–159	64
mutL	10^2–10^3	337	41	33	112	5	8	0.001/0.390	3	12–268	91.5
mutH	10^2–10^3	300	33	27	100	4	12	0.056/0.600	1	108–240	154.5
mutU	10^2–10^3	371	58	53	123	1	1	0.001/0.630	3	72–360	117
mutT	10^2–10^3	333	31	18	111	17	53	0.984/—	—	197	197
recD	0	630	45	34	210	12	25.4	0.160/—	—	15–363	266

[a] This is an approximate average increase in mutation frequency, on different chromosomal targets, over the wild type (for reviews, see Modrich, 1991; Fowler and Schaaper, 1997 and references therein).
[b] For *mutS*, *mutL*, *mutH*, and *mutU* genes, nucleotides informative for the parsimony are given among ingroup members, whereas they are given for all the strains for *mutT* and *recD* genes.
[c] Non-synonymous base substitutions were inferred onto the most parsimonious tree.
[d] The given P values correspond to the P calculated before and after the removal of incongruent taxa, respectively (significance threshold $P = 0.05$).
[e] Corresponds to the number of strains that have to be removed to clearly raise the P-value above the significance threshold.
[f] The length of recombined sequences might extend outside of the sequenced DNA for 2 *mutL*, 3 *recD*, 3 *mutH*, 3 *mutS*, and 6 *mutU* genes.

et al., 1997; Oliver et al., 2000). Because inactivation of MMR greatly increases recombination rates between related strains and species, MMR-deficient strains might have contributed considerably to recombination that has given rise to the observed genomic sequence mosaicism of *E. coli* natural isolates (for reviews see Milkman and McKane, 1995; Guttman, 1997). For most of the recombined genes, it is not possible to know whether a recombination event happened in an MMR-proficient or -deficient background. However, like any other gene, the MMR genes themselves can be subject to horizontal transfer. Therefore, the inactivation of a given MMR gene should increase the probability of reacquisition of its functional allele by horizontal transfer from a related but diverged bacterium. If this were the case, one would expect a high level of recombination of MMR genes reflecting recurrent losses and reacquisitions of MMR functions. This process would result in successions of high and low rates of mutation and recombination, giving rise to genetic diversification.

Horizontal transfers of chromosomal genes between strains can be revealed by incongruence (i.e., disagreement) between individual gene and strain phylogenies (Dykhuizen and Green, 1991; Bull et al., 1993). We have looked for such incongruence by comparing the phylogenies of MMR genes with the strain phylogeny in a representative set of strains from the ECOR collection of *E. coli* natural isolates (Ochman and Selander, 1984) to test the hypothesis that in the course of their evolution, bacteria have been alternating between MMR-proficient and -deficient states. As a control, two more genes have been analyzed: *mutT*, whose inactivation increases spontaneous mutagenesis (just like inactivation of MMR genes) but does not stimulate recombination (Fowler and Schaaper, 1997; this work), and *recD*, whose inactivation increases intra- and interspecies recombination but not mutagenesis (Lovett et al., 1988; Zahrt and Maloy, 1997). We found that MMR genes, in particular *mutS* and *mutL*, exhibit high recombination frequencies among *E. coli* natural isolates, suggesting that the MMR functions have frequently been lost and reacquired in the course of *E. coli* evolution.

Results

Phylogenetic Analysis of Mismatch Repair Genes

The incongruence between gene and strain phylogenies provides evidence for horizontal transfer that has occurred in the relatively recent evolutionary past (Dykhuizen and Green, 1991; Bull et al., 1993). To test for horizontal gene transfer of *E. coli* MMR genes, *mutS*, *mutL*, *mutH*, and *mutU*(*uvrD)* and two control genes (*mutT* and *recD*) were partially sequenced from 30 natural isolates (Table 1). Phylogenies were obtained by the maximum likelihood method (Figure 1) and compared with the whole genome reference tree (Lecointre et al., 1998). The comparison clearly identified incongruences, which range from single (*mutT*) to multiple (*mutS*) (Figure 1). To test whether such taxonomic incongruence is the consequence of recombination or of tree reconstruction artifacts (Maddison, 1997), the incongruence length difference (ILD) test of Farris et al. (1995) was used (see Experimental Procedures).

The ILD test showed significant incongruence of *mutS*, *mutL*, and *mutU* genes with the whole genome reference data set; it was nearly significant for *mutH* gene, but it was not significant for *mutT* and *recD* genes (Table 1). To measure the extent of incongruence due to recombination, single strains or combinations of strains were progressively removed and the analysis repeated until the incongruence was no longer significant. For *mutS, mutL*, and *mutU*, at least eight, three, and three strains, respectively, have to be removed to raise the P-value above the significance threshold (Table 1). From the analysis of the trees and the ILD test, it can be concluded that the MMR gene trees, when compared with the whole genome tree, exhibit significant incongruence most likely as a result of horizontal gene transfer.

Quantification of Horizontal Transfer of Mismatch Repair Genes

It has been reported previously that some genes, such as plasmid loci or the *gnd* gene, show a high level of incongruence (Lecointre et al., 1998). The incongruence

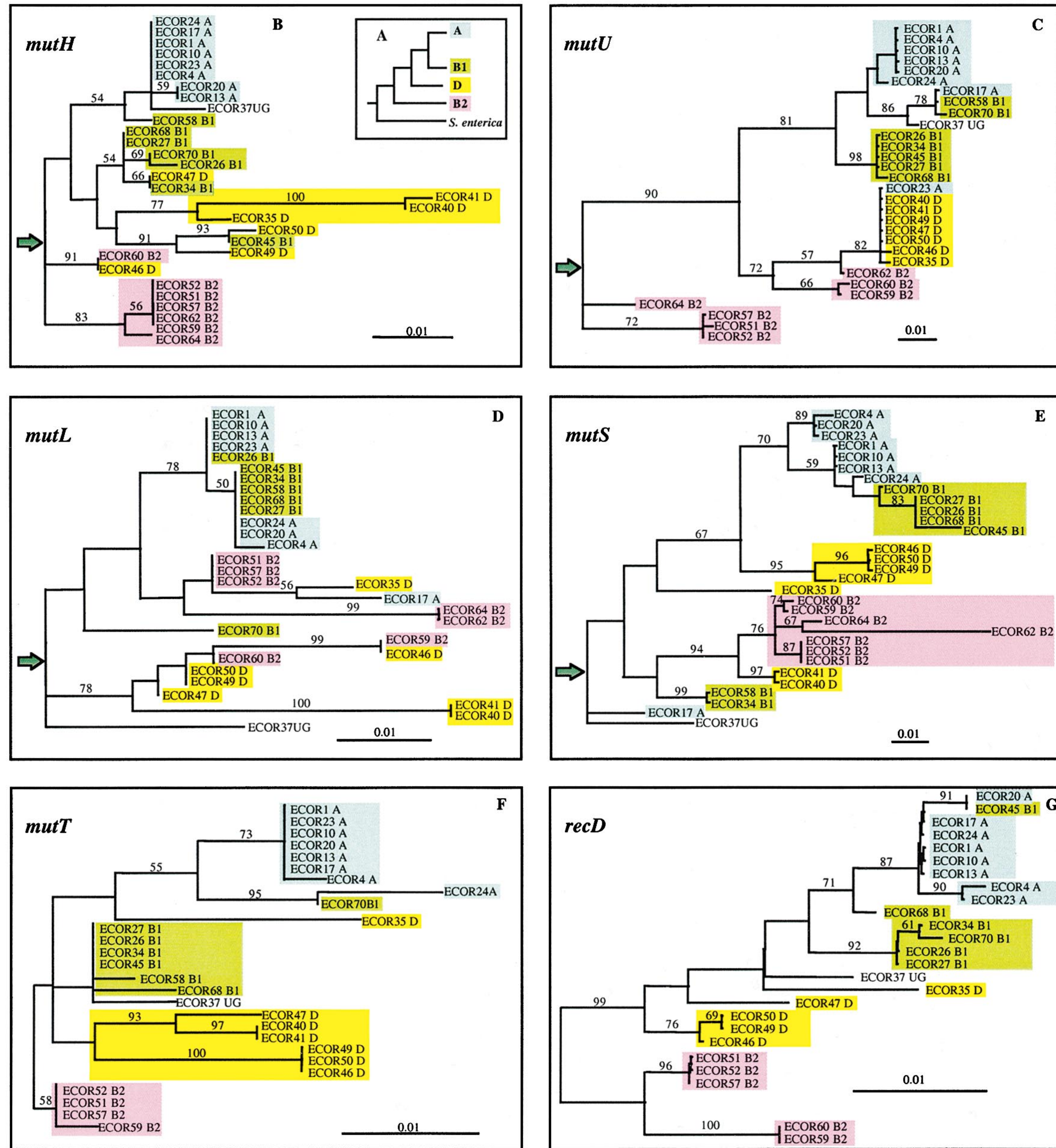

Figure 1. Molecular Phylogenetic Gene Trees of 30 *E. coli* Strains Belonging to Groups A, B1, B2, D, and UG (Ungrouped)

(A) Interrelationships of the main strain groups as inferred in Lecointre et al. (1998) and considered as the reference strain tree. Trees inferred from (B) *mutH,* (C) *mutU,* (D) *mutL*, (E) *mutS,* (F) *mutT*, and (G) *recD* were obtained by the maximum likelihood method. Similar phylogenies were retrieved when the data were analyzed by the parsimony method, providing high consistency and retention indices (data not shown). The arrow shows the branching point of the outgroup (*H. influenzae* for *mutU*, *S. enterica* for *mutH*, *S*, and *L*). *mutT* and *recD* trees were rooted on the B2 group strains. Numbers above the branches correspond to bootstrap values $\geq$ 50 (100 replications). Scale bar indicates the expected number of nucleotide substitutions per site.

of the *gnd* locus reflects a strong selective pressure exerted by the host immune system on the closely linked O antigen complex: 28 out of 34 strains must be removed to relieve *gnd* gene incongruence (data not shown). Clearly, MMR recombinant genes are selected for less often than extremely diversified genes like those coding for surface antigens.

Because in some cases small stretches of transferred DNA cannot be detected by the ILD test (Lecointre et al., 1998), we have developed a complementary approach using computer-assisted analysis of aligned DNA sequences based on previously established phylogeny to determine whether recombination at the MMR loci was significantly higher than that observed for housekeeping genes (see supplemental Appendix at www.cell.com/cgi/content/full/103/5/711/DC1). All the incongruences detected by the ILD test in four MMR, as well as in *mutT* and *recD* genes, were identified as recombinational events over the total or partial length of the studied sequences (Table 1). Additional recombination events

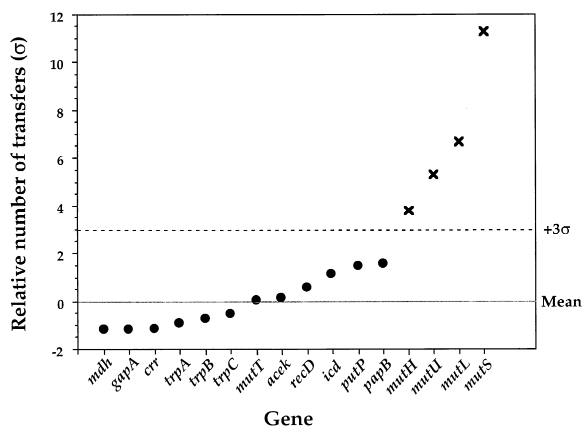

Figure 2. Relative Number of Transfer Events Occurring along the Evolutionary Trees of Different *E. coli* Genes

On this figure, relative number of transfers (t') is expressed in terms of σ, the standard deviation from the mean calculated from the ten housekeeping genes, *mutT* and *recD* (circles) ($m = 2.35 \times 10^{-8}$; $\sigma = 2.15 \times 10^{-8}$). MMR genes *mutH, mutU*, *mutL*, and *mutS* (crosses) are more than three standard deviations from the mean.

were identified that were not detected by the ILD test due to insufficient lengths of transferred stretches of DNA. For all recombination events detected in MMR genes, the phylogenetic group of the donor strain was clearly identified. The majority of strains have detectable traces of recombination of at least one of these antimutator genes.

Ungrouped (UG) strains of the ECOR collection do not belong to any defined phylogenetic group as their phylogenetic position within the *E. coli* species is variable according to the analyzed gene (Lecointre et al., 1998). For this reason, we did not take into account horizontal transfer within UG strains (see supplemental Appendix at the *Cell* website). However, we clearly identified by eye the presence of stretches of DNA sequences originating from various *E. coli* phylogenetic groups, defined by at least three bases, contiguous or not, within MMR genes from ECOR 37 UG strain (*mutS*: four transfers, *mutL*: two transfers), indicating a high level of mosaicism of MMR genes for this UG strain.

The values of observed number of transfers (t) for studied genes are as follows: (1) for control group: *crr*, 0; *mdh*, 0; *gapA*, 0; *mutT*, 1; *trpA*, 1; *pabB*, 3; *putP*, 3; *aceK*, 4; *recD*, 4; *trpB*, 4; *icd*, 5; *trpC*, 8, and (2) for MMR genes: *mutH*, 4; *mutU*, 8; *mutL*, 8; and *mutS*, 17. The observed number of transfers (t) is significantly higher in the group of MMR genes than in the control group (Mann-Withney test, $p = 0.03$). Because t can depend on the number of strains used to build the tree and upon the length of the compared sequences, the relative number of transfers (t'), which corrects for these effects, was calculated. t' is also significantly higher for the MMR genes than for *mutT, recD*, and the housekeeping genes (Mann-Withney test, $p = 0.0055$; Figure 2). When more stringent criteria (four and five mutations; see supplemental Appendix available on the *Cell* website) for detection of recombined fragments were used, t' was still significantly higher for MMR genes than for control group (data not shown). Interestingly, different MMR genes show different numbers of horizontal transfers: *mutS* and *mutL* exhibit higher levels of sequence mosaicism than *mutU* and *mutH* (Figures 2 and 3).

Effect of Mismatch Repair Genes on Horizontal Transfer of Chromosomal Genes

The elevated level of inferred recombination events within MMR alleles does not prove that horizontal transfers of MMR sequences have occurred under MMR-deficient conditions. However, the phenotype of MMR mutants offers a unique opportunity to test such a hypothesis. Whereas MMR mutants show a practically identical mutator phenotype (Glickman and Radman, 1980; Schaaper and Dunn, 1987), their effect on interspecies horizontal gene exchange (generator of incongruence) is distinct and characteristic (Rayssiguier et al., 1989). Therefore, we searched for eventual correlation between the hyperrecombination phenotype of different MMR$^-$ alleles measured in the laboratory and the frequency of inferred (historical) intragenic recombination events within the corresponding MMR genes.

The effect of different MMR genes on homeologous recombination was tested in crosses between an Hfr *S. typhimurium* (official designation, *S. enterica* serovar Typhimurium) donor strain and *E. coli* wild-type, *mutH*$^-$, *mutL*$^-$, *mutS*$^-$, and *mutU*$^-$ recipient (F$^-$) strains (Figure 4). In addition, two more genes, *mutT* and *trpB*, were included as controls: *mutT* because its inactivation increases mutation rate at least as much as that of MMR genes (for review see Fowler and Schaaper, 1997) and because *mutT* mutators have been found among natural isolates (Gross and Siegel, 1981). However, at the onset, it was not known whether *mutT* affects recombination. *trpB* was chosen as a representative housekeeping gene in which sequence mosaicism has been originally reported (Milkman and McKane, 1995; Lecointre et al., 1998). The results show a striking correlation between the frequency of interspecies recombination measured in different mutants (spanning over four orders of magnitude; Figure 4), and the frequency of inferred horizontal transfers within the corresponding MMR genes ($p = 0.0055$, $R^2 = 0.881$; Figure 5). MMR genes show high levels of mosaicism and have hyperrecombination phenotype. In contrast, inactivation of *trpB* and *mutT* genes, which have low level of mosaicism, had no significant effect on the frequency of intra- or interspecies recombination. These results support the hypothesis that the inferred recombination events accounting for the mosaic structure of the MMR genes occurred under MMR-deficient conditions.

Discussion

The diversity of chromosomal DNA sequences present in natural populations results from the activity of molecular mechanisms that control the rate of accumulation of genetic variability as well as from natural selection that acts on this variability. Hence, the frequency of detectable horizontal gene transfers in natural populations is dependent on the rate of transfer from donor to recipient cell, recombination rate, and the fitness of the recombinants. Although it is likely that all chromosomal regions have a similar probability of being exchanged, the selection acting on a given recombinant gene depends on its function and the interactions of its product

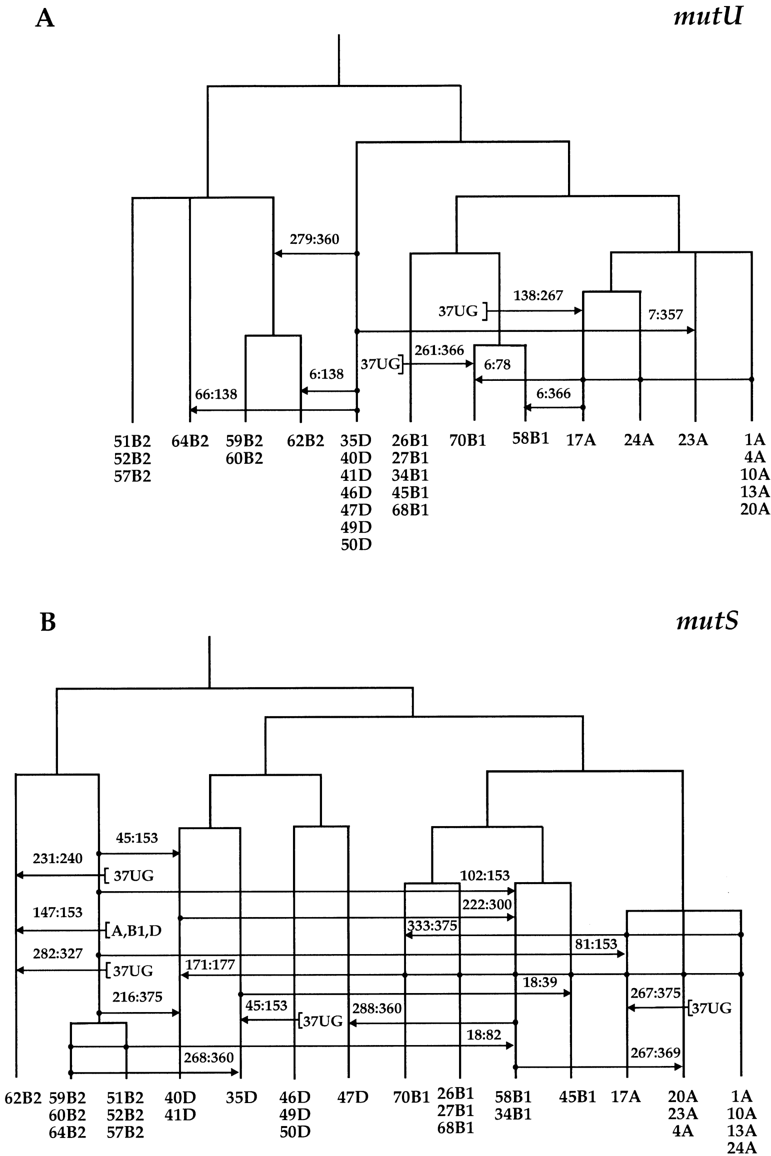

Figure 3. Hypothetical Horizontal Transfers of *mutU* and *mutS* Genes within the *E. coli* Evolutionary Tree

Transfers are schematized by the arrows, the head of the arrow indicating the direction of the transfer. The minimal length of the transferred segment is indicated above the arrow, the two numbers separated by a colon representing the nucleotide numbering of the amplified *mut* gene 5′ to 3′. (A) *mutU*, (B) *mutS*.

with other cellular components. For example, it has been observed that genes encoding proteins participating in transcription and translation are less likely to be successfully transferred than other housekeeping genes (Jain et al., 1999). Housekeeping genes are exchanged at lower rates than genes encoding surface antigens and antibiotic resistance targets whose variability provides an immediate selective advantage to bacteria exposed to the immune system and antibiotics (Maynard-Smith et al., 1991; Marklund et al., 1992; Li et al., 1994). It has been reported previously that some antimutator genes like *mutD (dnaQ)*, *mutH*, and *mutT* may have been acquired by *E. coli* K12 from an organism with a different codon preference (Médigue et al., 1991). However, the rate of horizontal transfer of MMR genes was previously unknown.

The Rate of Horizontal Transfer of MMR Genes

Phylogenetic analysis shows that the *E. coli* MMR gene trees, compared with the whole genome tree, exhibit significant incongruence as a result of horizontal gene transfers (Figures 1–3). In all cases, it was possible to identify unambiguously the phylogenetic origin of acquired sequences. DNA sequence analysis indicates that different MMR genes show different levels of intragenic recombination, but all of them clearly have a higher level of mosaicism than the housekeeping genes (Figures 2 and 3). However, their mosaicism is not as high as that observed for genes whose products are submitted to strong selective pressure for sequence variation, e.g., those encoding antigenic determinants or antibiotic resistance (Maynard-Smith et al., 1991; Marklund et al., 1992; Li et al., 1994). The reason for this difference might be that, in the case of MMR genes, the selective pressure is on restoration of MMR functions rather than on diversification of MMR genes and protein sequences. Indeed, most of the sequence polymorphism found in this phylogenetic analysis of MMR genes is neutral (Table 1). Only 1%–12% of base substitutions in MMR genes result in an amino acid change, which is similar to the average of 11.5% for housekeeping genes analyzed in this study

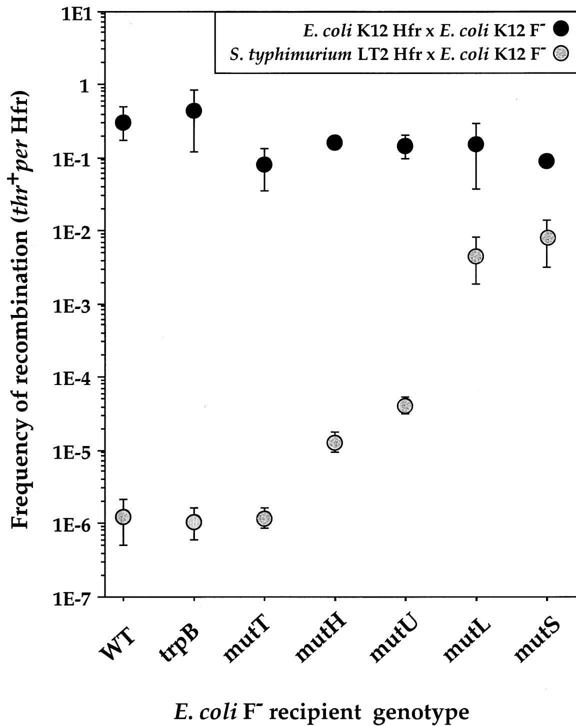

Figure 4. Influence of Mismatch–Repair Deficient Alleles on *E. coli* Hfr × *E. coli* F^- and *S. typhimurium* Hfr × *E. coli* F^- Conjugational Recombination

TrpB⁻ and *mutT*⁻ alleles have been chosen as control markers. The recombination frequencies for the selected marker (thr^+) are expressed per Hfr donor after subtracting unmated revertants. Each number represents the mean of at least three independent experiments.

(data not shown). However, this is in contrast with ~80% amino acid changing mutations observed for dominant surface antigens, such as the major outer membrane protein in *Chlamydia trachomatis*, which is subjected to selective pressure by the immune system (Kaltenboeck et al., 1993). The observation that the natural polymorphism of *mutT* gene sequences changes amino acid in 52% of the cases (Table 1) may indicate that MutT is a rapidly evolving enzyme.

Additional confirmation that MMR genes are more recombined than the housekeeping genes derives from analysis of the size of the recombined segments. It has been reported that the patch size in housekeeping genes is on the order of several kilobases (Milkman and McKane Bridges, 1993). Since the size of DNA recombined in the course of *E. coli* conjugation (up to one megabase, or more) or phage-mediated transduction (up to one hundred kilobases; Miller, 1991), under laboratory conditions, is several orders of magnitude larger, it was not clear how such mosaic patterns consisting of small DNA segments are generated. Milkman and McKane made a proposition, supported by in silico simulation and experiments (Milkman and McKane, 1995), that the cutting and degradation of donor DNA by the recipient cell's nucleases, as well as successive overlapping replacements, might be responsible for this phenomenon. Hence, more frequent recombination events are expected to result in a higher incidence of sequence mosaicism consisting of smaller patch sizes. Indeed, the *mutS* gene has the shortest length of recombined segments (Table 1; for an equivalent length of sequenced DNA, 82%, 75%, 25%, and 25% of detected transfers were shorter than the length of sequenced DNA for *mutS*, *mutL*, *mutU*, and *mutH* genes, respectively), the highest mosaicism, and the strongest hyperrecombination phenotype among MMR genes. Therefore, *mutS* genes are interstrain mosaics of fragments smaller than those found in housekeeping genes, probably as a consequence of higher frequencies of successful intragenic recombination.

The observed correlation between the hyperrecombination phenotype of different MMR⁻ alleles and the frequency of inferred intragenic recombination events within the corresponding MMR genes strongly supports

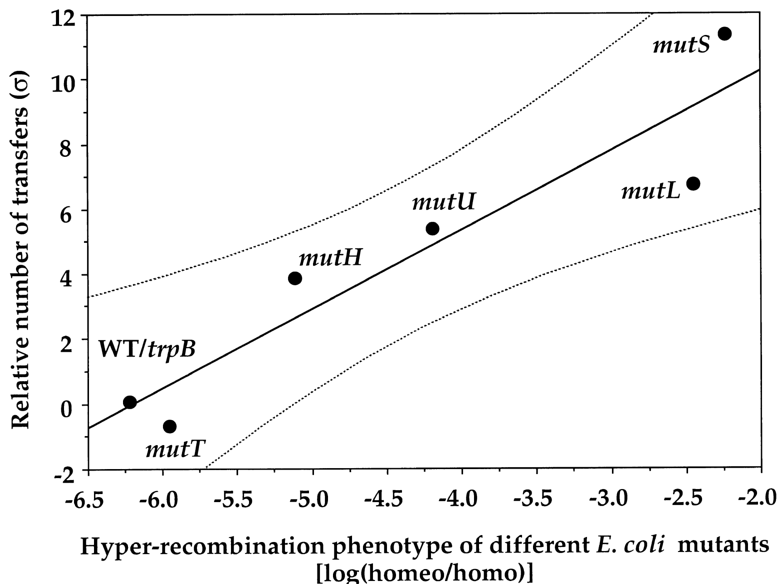

Figure 5. Correlation between the Hyperrecombination Phenotype of Different Alleles Measured In Vitro and the Number of Inferred Transfers of the Corresponding Genes in Natural Isolates

The effect on recombination between divergent DNA sequences is expressed as the logarithm of the ratio between the median frequencies of homeologous (homeo) and homologous (homo) conjugational recombination (see Figure 4). The relative number of transfer events (t') occurring along the evolutionary trees of *mutS*, *mutL*, *mutH*, *mutU*, *mutT*, and *trpB* genes is expressed as the standard deviation (σ) from the mean calculated from 10 housekeeping, *mutT*, and *recD* genes (see Figure 2). The dotted curves show the 95% confidence interval around the regression.

the hypothesis that those recombination events have occurred under MMR-deficient conditions (Figure 5). Mosaic MMR alleles have been observed in strains belonging to all phylogenetic groups. No obvious correlation was established between the degree of MMR gene mosaicism and the presence of pathogenic determinants (Bingen et al., 1998; Boyd and Hartl, 1998), geography, or the bacterial host (Ochman and Selander, 1984). 18 out of 30 analyzed strains showed detectable traces of recombination within MMR genes. Moreover, seven, five, and five strains showed recombination in at least one, two, and three different MMR genes, respectively (data not shown). These observations suggest that the loss and reacquisition of MMR genes may be a general characteristic of the species, and not of a particular group of strains. However, ECOR 37 strain, with the highest level of mosaicism within MMR genes, seems to be a special case. This strain belongs to UG strains that do not have a clear phylogenetic position within the *E. coli* species probably because they have highly mosaic genomes (Figure 1; Lecointre et al., 1998). The highest level of recombination among MMR genes was observed within the *mutS*, which in some strains showed mosaicism originating from at least three different phylogenetic groups (Figure 3B). Clearly, the restoration of MMR functions by recombination with an identical gene (which is expected to be the most efficient recombination) leaves no detectable trace at the DNA sequence level, and may also often erase a mosaic gene by replacing it. Thus, the rates of loss and reacquisition of MMR functions inferred from this study must be greatly underestimated.

The Fate of MMR-Deficient Mutators in the Laboratory and in Nature

The high frequency of detectable recombination events within MMR genes was probably facilitated by the high frequency of MMR-deficient *E. coli* strains in nature (LeClerc et al., 1996; Matic et al., 1997). Such abundance of MMR^- strains can be explained either by: (1) the high frequency of generation of mutator alleles, (2) the fact that mutator alleles are close to neutral and therefore enriched in some bacterial populations through genetic drift, or (3) positive selection for high mutation rates. The third hypothesis seems most plausible since the rate of mutations that inactivate MMR genes is not higher than that inactivating housekeeping genes such as *hisD* and *lacI* (measured in the laboratory) (Drake 1991 and references therein; Boe et al., 2000). In addition, it has been demonstrated in the laboratory that the fitness of mutators decreased much faster than that of wild-type cells (Boe et al., 2000, Funchain et al., 2000), indicating that mutator alleles cannot be considered as neutral.

A possible explanation for the abundance of mutator strains in natural populations has been provided by several experimental and theoretical studies showing that mutator strains could be transiently favored when adaptation is limited by the supply of mutations (for review see Metzgar and Wills, 2000). Computer simulations suggest that, despite the cost caused by the load of deleterious mutations, mutators rise to a high frequency in bacterial populations through their association with the favorable mutations they generate (Taddei et al., 1997; Tenaillon et al., 1999). It has been experimentally demonstrated that the fraction of *E. coli* mutator cells can increase in bacterial populations under very strong or prolonged selection (antibiotic treatments [Mao et al., 1997] and adaptation to a new environment [Sniegowski, et al., 1997]). The study of Oliver et al. (2000) shows that selective conditions predicted by the above-mentioned studies could indeed be responsible for the selection of mutator strains in nature. They reported very high frequency of strong mutator *P. aeruginosa* strains isolated from lungs of cystic fibrosis (CF) patients. Those *P. aeruginosa* populations have been exposed to the challenges of the host immune defenses and antibiotic therapies for a long time. They found no mutators among *P. aeruginosa* strains isolated from acutely infected non-CF patients. This observation confirms the prediction that in the absence of a strong selective pressure, MMR mutator cells should not accumulate. Indeed, in short-term cultures of *E. coli* and *S. enterica* laboratory strains grown without particular selective pressure, the frequency of mutator cells is only 10^{-5} and 10^{-6}, respectively (Mao et al., 1997; LeClerc et al., 1998). Therefore, our data suggest that MMR-deficient alleles are not neutral and that, due to the genetic diversity they generate, they are subject to positive natural selection during adaptation and negative selection when adaptation is achieved.

Peculiarities of MMR Mutator Strains

Inactivation of over 20 different *E. coli* genes can confer mutator phenotypes of different strengths (for reviews see Miller, 1996; Horst et al., 1999), but among strong mutators, only MMR-deficient alleles appear to be present at high frequency in nature. One explanation for this phenomenon is that inactivation of the other genes involved in important aspects of DNA metabolism (e.g., replication) may have too high a cost to be compensated by advantageous mutations. Even different MMR-defective alleles are not found with the same frequency. An example is provided by $mutL^-$ alleles which are less abundant in nature than $mutS^-$ (LeClerc et al., 1996; Matic et al., 1997). Such an underrepresentation also appears to be reflected in the lower degree of *mutL* gene mosaicism, compared to that of the *mutS* gene (Figure 2). Yet, inactivation of these genes has indistinguishable effect on spontaneous mutagenesis and homeologous recombination (Glickman and Radman, 1980; Schaaper and Dunn, 1987; Rayssiguier et al., 1989; Figure 4). In contrast to the underrepresentation of $mutL^-$ strains in nature, $mutL^-$ alleles can be isolated in the laboratory as readily as $mutS^-$ in the absence of natural selection and competition (Brégeon et al., 1999 and references therein). The genomic environment of the *mutL* gene may provide a clue for this disparity. The *mutL* gene is part of a superoperon with a complicated arrangement of seven genes that mediate several important cellular processes, e.g., DNA repair, tRNA modification, proteolysis, and pleiotropic regulation (Tsui et al., 1994). Mutations within the *mutL* gene may be counterselected in nature due to interference with the transcription of other genes of this complex operon.

The specific advantage of MMR-deficient strains over other mutators (for example $mutT^-$), which might also

explain their abundance in nature, is their hyperrecombination phenotype. This phenotype has been neglected in theoretical and experimental studies of the role of mutators in bacterial evolution. Recombination can also increase adaptability by increasing genetic variability (Crow and Kimura, 1965). Consequently, it might be expected that genotypes with increased recombination rates might be selected for by virtue of the favorable genotypes they generate by associating beneficial mutations that have appeared in different individuals (Otto and Michalakis, 1998). To test whether a hyperrecombination phenotype, in the absence of high mutation rates, can be under positive selection, we chose to examine the *recD* gene, whose inactivation increases intra- and/or interspecies recombination efficiency (Lovett et al., 1988; Zahrt and Maloy, 1997). We found that *recD* genes from the ECOR collection show a low level of sequence mosaicism (Figures 1 and 2; Table 1). Although these results suggest that hyperrecombination alone does not account for the mosaicism seen in MMR genes, there are other potential explanations for why *recD* genes show low mosaicism. For example: (1) the hyperrecombination phenotype of *recD*$^-$ allele may not be under positive selective pressure or (2) *recD* defective alleles might be counterselected due to pleiotropic defects, such as loss of protection against phages (Myers and Stahl, 1994), and consequently not abundant enough to influence significantly recombinant frequencies at the population level.

However, irrespective of whether selection of MMR mutators is due to their hyperrecombination or hypermutagenesis phenotype, the cost of high mutation rates remains the same. When selective pressure for increased genetic variability is not present any more, the hyperrecombination phenotype of MMR-deficient strains may facilitate the reacquisition of the antimutator functions. This can "save" the adapted mutator genome from being overburdened with deleterious mutations and allow survival of the recombined MMR gene.

Conclusions

Our study suggests that MMR functions have been repeatedly lost and reacquired during the evolutionary history of *E. coli*, and that the restoration of MMR antimutator functions often occurred via genetic exchange, which occasionally gave rise to mosaic alleles. Since the majority of natural isolates have mosaic MMR genes, our data suggest that they recurrently pass through periods of high mutation and recombination rates. The combination of hypermutation and hyperrecombination phenotypes allows rapid generation of genetic diversity upon which selection acts during adaptation and, at the same time, provides a means for the restoration of lower mutation rates following adaptation. Therefore, the changing proportion of MMR mutators within bacterial populations may provide a sort of genetic "gear-shifting" that can modulate the speed of bacterial evolution. Given the potent genome-wide hyperrecombination effect of MMR deficiency, it may be that these alternations have significantly contributed to the regulation of the rate of horizontal exchange of all chromosomal genes in natural *E. coli* populations. It may be that a significant fraction of genomic diversity of *E. coli* natural isolates (Milkman and McKane, 1995; Guttman, 1997) was generated during the MMR-deficient intermezzos of their evolutionary past.

Experimental Procedures

PCR Amplification and Sequencing

ECOR1, 4, 10, 13, 17, 20, 23, 24, 26, 27, 34, 35, 37, 40, 41, 45, 46, 47, 49, 50, 51, 52, 57, 58, 59, 60, 62, 64, 68, and 70 were studied with the exception of ECOR60, 62, and 64 for *mutT*, and of ECOR40, 41, 58, 62, and 64 for *recD*. The studied strains are representative of the genetic diversity of the ECOR collection of natural isolates (Ochman and Selander, 1984). Partial *mutS* (390 bp), *mutH* (300 bp), *mutU* (371 bp), *mutL* (337 bp), *mutT* (333 bp), and *recD* (630 bp) genes were amplified by the polymerase chain reaction (PCR) from *E. coli* DNA as in Bingen et al. (1998) and sequenced without interim cloning using a 373 sequencer model (Applied Biosystems, Foster City, CA). Positions of sequenced regions and sequences of the primers used are available on the *Cell* website.

Phylogenetic Analysis

For each data set, sequences were aligned by eye using the ED program of the MUST package (Philippe, 1993). The introduction of gaps was not necessary. Mutational saturation was not detected in any of these data sets. Phylogenetic analyses were done using maximum likelihood (DNAmL in PHYLIP [Phylogeny Inference Package], Version 3.5c, distributed by J. Felsenstein, Department of Genetics, University of Washington, Seattle and FASTDNAmL [Olsen et al., 1994]) and parsimony (PAUP 3.1.1. [Swofford, 1993]) methods. The trees were rooted on another bacterial species, either *S. enterica* (*mutS*, *mutL*, and *mutH*) or *Haemophilus influenzae* (*mutU*), or on the *E. coli* B2 group strains (*mutT* and *recD*), which have been shown to be the most basal within the *E. coli* species (Lecointre et al., 1998). Character incongruence was measured between each antimutator gene data set and a large data set representative of the ECOR strain phylogeny (Lecointre et al., 1998). This whole-genome data set includes multilocus enzyme electrophoresis (Herzer et al., 1990), random amplified polymorphic DNA, and *rrn* restriction fragment length polymorphism (Desjardins et al., 1995) data. We applied the ILD test (Mickevich and Farris, 1981; Farris et al., 1995), which tests the null hypothesis of congruence between two or more data sets. This test specifically measures the contradictory phylogenetic signal contained in two data sets (character incongruence). Tree reconstruction artifacts due to homoplasy (character states shared by two or several taxa that occurred independently by convergence, parallelism, or reversion) are unlikely to yield statistically significant character incongruence, whereas recombination will generate such character incongruence. In this test, the null hypothesis of congruence is rejected when the P-value obtained with the test version 4.0.0d64 of PAUP*, written by David L. Swofford, is below a particular threshold, in this case 0.05. This indicates that there is more incongruence between the data sets than would be expected from chance alone. In our analysis, we generated 1000 random partitions for each test and the P value considered correspond to X/1000 (Lecointre et al., 1998).

Quantification of the Transferred Segments

Transferred segments among sequenced genes for a given strain were identified by a computer algorithm based on aligned DNA sequences and the previously established phylogeny (see Appendix available on the *Cell* website). They were defined by at least three bases, contiguous or not, delineating a stretch of DNA with a sequence that is different from those of the strains belonging to the same phylogenetic group, but that is identical to those of another phylogenetic group (choice of this criterion is discussed in more detail in the Appendix available on the *Cell* website). Only the minimal length of the transferred segment was estimated, delineated by a base at each extremity (5′ and 3′) that is different from that of the sequences of the other strains belonging to the same phylogenetic group. The data sets analyzed here exhibit particular properties allowing to infer horizontal transfers. Indeed, some positions along

the DNA sequences show nucleotide substitutions, which do not fit correctly the grouping inferred from the phylogeny. They are rare and, when observed, usually not randomly distributed along the sequences, but clustered together in a stretch favoring another contradictory grouping. This distinctive pattern leads to retaining the hypothesis of a single gene transfer rather than the hypotheses of multiple correlated substitutions. Moreover, since the sequences differ more between main strain groups A, B1, D, B2 than within each group, the donor group can often be unambiguously identified.

Both mutation and recombination events were simultaneously minimized considering that: (1) one transfer event is more parsimonious than at least three independent mutational events occurring within the corresponding DNA stretch and (2) a single ancestral transfer is more likely than multiple individual transfers in each descendant. Ancestral transfers are inferred onto the reference tree (Figure 1A) whose topology was partially relaxed within each of the main groups (A, B1, B2, D). The reasons for this are 2-fold. First, the sequence data are often too variably structured within each strain group to identify stretches of transferred DNA and second, it would be nonparsimonious to infer transfers onto a reference topology that is poorly supported by the reference data.

The relative number of transfers (t') was calculated as the number of inferred transfers (t) divided by the maximum number of possible transfers $(1/2\,(l - 2)\,(l - 1))\,(n - 1)^2$, where the first term gives the total number of possible transfers of DNA stretches longer or equal to three nucleotides within a sequence of l nucleotides, and where the second gives the number of transfers which can occur between two contemporaneous branches of the tree built with n strains. Demonstration of the formula is available on the *Cell* website.

Conjugation Experiments

Hfr strains used in this study were: *S. typhimurium* LT2 SA965 (Sanderson et al., 1972) and *E. coli* K12 P4X *metB*$^+$ (Kahn, 1968). All recipient strains (see Figure 4) were *E. coli* K12 AB1157 NalR (Bachmann, 1972) derivatives: either previously constructed as *mutS215*::Tn*10* and *mutL218*::Tn*10* (Pang et al., 1985) or constructed for this study by introducing *mutH471*::Tn*5* (Pang et al., 1985), *uvrD260*::Tn*5* (Pang et al., 1985), *mutT*::Km^r (B.A. Bridges collection, University of Sussex, Brighton, UK), or *trpB114*::Tn*10* (*E. coli* Genetic Stock Center, Yale University, New Haven, CT) alleles using P1 mediated transductions (Miller, 1972). Because the level of spontaneous reversion of the AB1157 *thr1* allele was very high in *mut*$^-$ genetic backgrounds, the *thr43*::Tn*10* allele (*E. coli* Genetic Stock Center, Yale University, New Haven, CT), with very low reversion rate, has been transduced into *uvrD260*::Tn*5*, *mutH471::Tn5*, and *mutT*::Km^r strains (as well as into AB1157, as a control *mut*$^+$ strain) in order to allow for the measurement of small increases in the frequency of interspecies recombination. Conjugation experiments were performed as previously described (Rayssiguier et al., 1989). The exconjugants were plated on M63 medium supplemented with arginine, histidine, leucine, proline, and tryptophane (100 μg/ml each), thiamine (30 μg/ml), glucose (0.4%), and nalidixic acid ([40 μg/ml] to counterselect the Hfr donor cells), lacking threonine to select for *thr*$^+$ recombinants. Recombinants were scored after 48 hr. Results presented are the median values of three independent experiments.

Supplemental Data

Additional information, including an Appendix by Olivier Tenaillon, regarding the experimental and phylogenetic statistical methods is available online (http://www.cell.com/cgi/content/full/103/5/711/DC1).

Acknowledgments

We would like to thank L. Rachdi for technical assistance and S. Tillier for constant encouragement. We thank H. Le Nagard for his advice concerning computer algorithms. We thank V. Colot, J. Courcelle, E. Friedberg, P. Francino, E. Stewart, M. Vulic, and R. Wagner for critical reading of the manuscript. We also thank the anonymous referees for their helpful criticism. This work was supported by grants from the "Programme Hospitalier de Recherche Clinique—MES", the "Programme de Recherche Fondamentale en Microbiologie et Maladies Infectieuses et Parasitaires—MENRT" (E. D., F. T., G. L., I. M, M. R., O. T., P. D.), the Ligue contre le Cancer (F. T., I. M., M. R.), "Programme Environnement et Santé - MATE" (F. T., I. M., M. R., O. T.), the Association pour la Recherche sur le Cancer (F. T., I. M., M. R., O. T.), and the NIH (R. R.).

Received June 3, 1999; revised October 6, 2000.

References

Bachmann, B.J. (1972). Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. *36*, 525–557.

Bianchi, M.E., and Radding, C.M. (1983). Insertions, deletions, and mismatches in heteroduplex DNA made by RecA protein. Cell *35*, 511–520.

Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J., and Denamur, E. (1998). Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. J. Infect. Dis. *177*, 642–650.

Boe, L., Danielsen, M., Knudsen, S., Petersen, J.B., Maymann, J., and Jensen, P.R. (2000). The frequency of mutators in populations of *Escherichia coli*. Mutat. Res. *448*, 47–55.

Boyd, E.F., and Hartl, D.L. (1998). Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. J. Bacteriol. *180*, 1159–1165.

Brégeon, D., Matic, I., Radman, M., and Taddei, F. (1999). Inefficient mismatch repair: genetic defects and down regulation. J. Genet. *78*, 21–28.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., and Waddel, P.J. (1993). Partitioning and combining data in phylogenetic analysis. Syst. Biol. *42*, 384–397.

Crow, J.F., and Kimura, M. (1965). Evolution in sexual and asexual populations. Am. Nat. *99*, 439–450.

Desjardins, P., Picard, B., Kaltenböck, B., Elion, J., and Denamur, E. (1995). Sex in *Escherichia coli* does not disrupt the clonal structure of the population: evidence from random amplified polymorphic DNA and restriction-fragment length polymorphism. J. Mol. Evol. *41*, 440–448.

Drake, J.W. (1991). A constant rate of spontaneous mutation in DNA-based microbes. Proc. Natl. Acad. Sci. USA *88*, 7160–7164.

Dykhuizen, D.E., and Green, L. (1991). Recombination in *Escherichia coli* and the definition of biological species. J. Bacteriol. *173*, 7257–7268.

Farris, J.S., Källersjö, M., Kluge, A.G., and Bult, C. (1995). Testing significance of incongruence. Cladistics *10*, 315–319.

Fowler, R.G., and Schaaper, R.M. (1997). The role of *mutT* gene of *E. coli* in maintaining replication fidelity. FEMS Microbiol. Rev. *21*, 43–54.

Funchain, P., Yeung, A., Stewart, J.L., Lin, R., Slupska, M.M., and Miller, J.H. (2000). The consequences of growth of a mutator strain of *Escherichia coli* as measured by loss of function among multiple gene targets and loss of fitness. Genetics, in press.

Glickman, B.W., and Radman, M. (1980). *Escherichia coli* mutator mutants deficient in methylation-instructed DNA mismatch correction. Proc. Natl. Acad. Sci. USA *77*, 1063–1067.

Gross, M.D., and Siegel, E.C. (1981). Incidence of mutator strains in *Escherichia coli* and coliforms in nature. Mutat. Res. *91*, 107–110.

Guttman, D.S. (1997). Recombination and clonality in natural populations of *Escherichia coli.* Trends Ecol. Evol. *12*, 16–22.

Guttman, D.S., and Dykhuizen, D.E. (1994). Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. Science *266*, 1380–1383.

Herzer, P.J., Inouye, S., Inouye, M., and Whittman, T.S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-

stranded DNA among natural isolates of *Escherichia coli.* J. Bacteriol. *172*, 6175–6181.

Horst, J.P., Wu, T.H., and Marinus, M.G. (1999). *Escherichia coli* mutator genes. Trends Microbiol. *7*, 29–36.

Jain, R., Rivera, M.C., and Lake, J.A. (1999). Horizontal gene transfer among genomes: The complexity hypothesis. Proc. Natl. Acad. Sci. USA *96*, 3801–3806.

Kahn, P.L. (1968). Isolation of high frequency recombining strains from *Escherichia coli* containing the V colicinogenic factor. J. Bacteriol. *96*, 205–214.

Kaltenboeck, B., Kousoulas, K.G., and Storz, J. (1993). Structures of and allelic diversity and relationships among the major outer membrane protein (*ompA*) genes of the four Chlamydial species. J. Bacteriol. *175*, 487–502.

LeClerc, J.E., Baouguang, L., Payne, W.L., and Cebula, T.A. (1996). High mutation frequencies among *Escherichia coli* and *Salmonella* pathogens. Science *274*, 1208–1211.

LeClerc, J.E., Payne, W.L., Kupchella, E., and Cebula, T.A. (1998). Detection of mutator subpopulations in *Salmonella typhimurium* LT2 by reversion of his alleles. Mutat. Res. *400*, 89–97.

Lecointre, G., Rachdi, L., Darlu, P., and Denamur, E. (1998). *Escherichia coli* molecular phylogeny using the incongruence length difference test. Mol. Biol. Evol. *15*, 1685–1695.

Li, J., Nelson, K., McWhorter, A.C., Whittam, T.S., and Selander, R.K. (1994). Recombinational basis of serovar diverity in *Salmonella enterica*. Proc. Natl. Acad. Sci. USA *91*, 2552–2556.

Lichten, M., and Fox, M.S. (1984). Evidence for inclusion of regions of nonhomology in heteroduplex products of bacteriophage l recombination. Proc. Natl. Acad. Sci. USA *81*, 7180–7184.

Lovett, S.T., Luisi-DeLuca, C., and Kolodner, R.D. (1988). The genetic dependence of recombination in *recD* mutants of *Escherichia coli*. Genetics *120*, 37–45.

Maddison, W.P. (1997). Gene trees in species trees. Syst. Biol. *46*, 523–536.

Majewski, J., Zawadzki, P., Pickerill, P., Cohan, F.M., and Dowson, C.G. (2000). Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. J. Bacteriol. *182*, 1016–1023.

Mao, E.F., Lane, L., Lee, J., and Miller, J.H. (1997). Proliferation of mutators in a cell population. J. Bacteriol. *179*, 417–422.

Marklund, B., Tennent, J.M., Garcia, E., Hamers, A., Baga, M., Lindberg, F., Gaastra, W., and Normak, S. (1992). Horizontal gene transfer of the *Escherichia coli pap* and *prs* pili operons as a mechanism for the development of tissue-specific adhesive properties. Mol. Microbiol. *6*, 2225–2242.

Matic, I., Rayssiguier, C., and Radman, M. (1995). Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. Cell *80*, 507–515.

Matic, I., Taddei, F., and Radman, M. (1996). Genetic barriers among bacteria. Trends Microbiol. *4*, 69–73.

Matic, I., Radman, M., Taddei, F., Picard, B., Doit, C., Bingen, E., Denamur, E., and Elion, J. (1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. Science *277*, 1833–1834.

Maynard-Smith, J., Dowson, C.G., and Spratt, B.G. (1991). Localized sex in bacteria. Nature *349*, 29–31.

Médigue, C., Rouxel, T., Vigier, P., Hénaut, A., and Danchin, A. (1991). Evidence for horizontal gene transfer in *Escherichia coli* speciation. J. Mol. Biol. *222*, 851–856.

Metzgar, D., and Wills, C. (2000). Evidence for the adaptive evolution of mutation rates. Cell *101*, 581–584.

Mickevich, M.F., and Farris, J.S. (1981). The implications of congruence in Menidia. Syst. Zool. *30*, 351–370.

Milkman, R., and McKane Bridges, M. (1993). Molecular evolution of the *Escherichia coli* chromosome. IV. Sequence comparisons. Genetics *133*, 455–468.

Milkman, R., and McKane, M. (1995). DNA sequence variation and recombination in *E. coli*. In Population Genetics of Bacteria, S. Baumberg, J.P.W. Young, E.M.H. Wellington, and J.R. Saunders, eds. (Cambridge, UK: Cambridge University Press), pp. 127–142.

Miller, J.H. (1972). Experiments in Molecular Genetics (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Miller, J.H. (1991). Bacterial genetic systems. In Methods in Enzymology, Vol. 204 (San Diego, CA: Academic press, Inc.).

Miller, J.H. (1996). Spontaneous mutators in bacteria: insight into pathways of mutagenesis and repair. Annu. Rev. Microbiol. *50*, 625–643.

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

Myers, R.S., and Stahl, F.W. (1994). Chi and the RecBCD enzyme of *Escherichia coli*. Annu. Rev. Genet. *28*, 49–70.

Ochman, H., and Selander, R.K. (1984). Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. *157*, 690–693.

Oliver, A., Canton, R., Campo, P., Baquero, F., and Blazquez, J. (2000). High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. Science *288*, 1251–1254.

Olsen, G.J., Matsuda, H., Hagstrom, R., and Overbeek, R. (1994). FastDNAmL: A tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. *10*, 41–48.

Otto, S.P., and Michalakis, Y. (1998). The evolution of recombination in changing environments. Trends Ecol. Evol. *13*, 145–151.

Pang, P.P., Lundberg, A.S., and Walker, G.C. (1985). Identification and characterization of the *mutL* and *mutS* gene products of *Salmonella typhimurium* LT2. J. Bacteriol. *163*, 1007–1015.

Philippe, H. (1993). MUST, a computer package of Management Utilities for Sequences and Trees. Nucleic Acids Res. *21*, 5264–5272.

Radman, M., and Wagner, R. (1993). Mismatch recognition in chromosomal interactions and speciation. Chromosoma *102*, 369–373.

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

Sanderson, K.E., Ross, H., Zeiger, L., and Mäkelä, P.H. (1972). F +, Hfr, and F′ strains of *Salmonella typhimurium* and *Salmonella abony*. Bacteriol. Rev. *36*, 608–637.

Schaaper, R.M., and Dunn, R.L. (1987). Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: the nature of *in vivo* DNA replication errors. Proc. Natl. Acad. Sci. USA *84*, 6220–6224.

Shen, P., and Huang, H.V. (1986). Homologous recombination in *Escherichia coli*: Dependence on substrate length and homology. Genetics *112*, 441–457.

Shen, P., and Huang, H. (1989). Effect of base pair mismatches on recombination via the RecBCD pathway. Mol. Gen. Genet. *218*, 358–360.

Swofford, D.L. (1993). Phylogenetic analysis using parsimony (PAUP). Version 3.1.1. Illinois Natural History Survey, Champaign, IL.

Sniegowski, P.D., Gerrish, P.J., and Lenski, R.E. (1997). Evolution of high mutation rates in experimental populations of *E. coli*. Nature *387*, 703–705.

Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., and Godelle, B. (1997). Role of mutator alleles in adaptive evolution. Nature *387*, 700–702.

Tenaillon, O., Toupance, B., Le Nagard, H., Taddei, F., and Godelle, B. (1999). Mutators, population size, adaptive landscape and the adaptation of asexual population of bacteria. Genetics *152*, 485–493.

Tsui, H.C., Leung, H.C., and Winkler, M.E. (1994). Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12. Mol. Microbiol. *13*, 35–49.

Vulic, M., Dionisio, F., Taddei, F., and Radman, M. (1997). Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. Proc. Natl. Acad. Sci. USA *94*, 9763–9767.

Worth, L., Jr., Clark, S., Radman, M., and 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.

Zahrt, T.C., and Maloy, S. (1997). Barriers to recombination between closely related bacteria: MutS and RecBCD inhibit recombination between *Salmonella typhimurium* and *Salmonella typhi*. Proc. Natl. Acad. Sci. USA *94*, 9786–9791.

Zawadzki, P., Roberts, M.S., and Cohan, F.M. (1995). The log-linear relationship between sexual isolation and sequence divergence in Bacillus transformation is robust. Genetics *140*, 917–932.

GenBank Accession Number

All gene sequences from this work have been deposited in the GenBank database. Accession numbers can be obtained upon request from the authors.