



Microsatellites in the Eukaryotic DNA Mismatch Repair Genes as Modulators of Evolutionary Mutation Rate

Dong Kyung Chang, David Metzgar, Christopher Wills, et al.

Genome Res. 2001 11: 1145-1146

Access the most recent version at doi:[10.1101/gr.186301](https://doi.org/10.1101/gr.186301)

References This article cites 9 articles, 5 of which can be accessed free at:
<http://genome.cshlp.org/content/11/7/1145.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

Microsatellites in the Eukaryotic DNA Mismatch Repair Genes as Modulators of Evolutionary Mutation Rate

Dong Kyung Chang,¹ David Metzgar,² Christopher Wills,^{2,3}
and C. Richard Boland^{1,4}

¹Department of Medicine and Cancer Center, ²Department of Biology, and ³Center for Molecular Genetics, University of California San Diego, La Jolla, California 92093-0688, USA

All “minor” components of the human DNA mismatch repair (MMR) system—*MSH3*, *MSH6*, *PMS2*, and the recently discovered *MLH3*—contain mononucleotide microsatellites in their coding sequences. This intriguing finding contrasts with the situation found in the “major” components of the DNA MMR system—*MSH2* and *MLH1*—and, in fact, most human genes. Although eukaryotic genomes are rich in microsatellites, non-triplet microsatellites are rare in coding regions. The recurring presence of exonal mononucleotide repeat sequences within a single family of human genes would therefore be considered exceptional.

A study of 10,000 random primate (primarily human) coding sequences (8.6 million base pairs) showed that the average-sized sequence coding for a human gene is expected to contain approximately 0.1, 0.03, and 0.006 mononucleotide runs of length 7 bp or more, 8 bp or more, and 9 bp or more, respectively (Metzgar et al. 2000). Together, the four minor MMR genes contain three 7-bp runs, four 8-bp runs, and one 9-bp run. By nature of their form and position within the genes, these runs do not appear to be evolutionarily homologous. The probability of finding this many mononucleotide microsatellites by chance in a group of four genes is 3.2×10^{-7} ($0.4^3 \times 0.12^4 \times 0.024$). Controlling for gene length yields a probability of 1.3×10^{-6} . Therefore, the human minor MMR genes contain a significant excess of mononucleotide repeats. The probability would be even lower if the calculation included the *MED1* (*MBD4*) gene, considered by some to be another minor MMR gene, which has an exonal 10-bp polyadenosine sequence.

We surveyed MMR gene sequences in organisms that ranged in complexity from prokaryotes to humans and found that the constellation of microsatellites in the coding re-

gions of the minor MMR genes is a general feature among eukaryotes (Table 1). Although in some species, 7-bp mononucleotide runs are sporadically found in the major MMR genes, the longer runs, which are exponentially more mutable, are exclusively present in the minor MMR genes. Microsatellite sequences are exceptionally vulnerable to spontaneous insertion or deletion mutations, and nontriplet microsatellites, when located in coding sequences, are expected to introduce frameshift loss-of-function mutations at high frequency (Moxon et al. 1994). It is a paradox that the MMR system, which limits mutations in microsatellite sequences, would be particularly vulnerable to mutation by virtue of having microsatellites in its own coding regions. In the minor MMR genes, these otherwise rare and generally negatively selected sequences may have evolved by providing a survival advantage for eukaryotic species. What evolutionary advantage could microsatellites provide in the coding regions of the minor MMR genes? We hypothesize that the exceptional density of microsatellites in the minor MMR genes represents a genetic switch that allows the adaptive mutation rate to be modulated over evolutionary time.

The demonstration of a surprisingly

high frequency of mutator strains among bacterial pathogens appears to illustrate this paradigm, which is an example of the possible advantage of a high mutation rate within a population of organisms subjected to marked changes in environmental conditions (LeClerc et al. 1996). Under adverse experimental conditions, MMR-deficient bacterial mutators outcompete nonmutators and are common in natural and laboratory populations subjected to intensive selection (Rosenberg et al. 1998). Theoretical work also supports the advantage of mutator phenotypes under conditions that require rapid evolution (Sniegowski et al. 1997; Taddei et al. 1997). Does this paradigm apply to eukaryotes? We propose that the existence of microsatellites in MMR genes might be evidence that extends this concept to higher organisms (Fig. 1).

MMR proteins correct replication errors and actively inhibit recombination between diverged sequences (Chen and Jinks-Robertson 1998; Kolodner and Marsischky 1999), thus controlling rates of mutation and evolutionary adaptation. From yeast to humans, the MMR system is composed of homologs of the *MutS* and *MutL* genes of *Escherichia coli*. Either *MSH3* or *MSH6* het-

Table 1. Microsatellites in the Coding Regions of the DNA MMR Genes

	<i>MSH2</i>	<i>MSH3</i>	<i>MSH6</i>	<i>MLH1</i>	<i>PMS2</i> ^a	<i>MLH3</i>
<i>Homo sapiens</i>	A7	A8, A7	C8, A7, T7	none	A8	A9, A8
<i>Mus musculus</i>	none	none ^b	T7	none	none	— ^c
<i>Saccharomyces cerevisiae</i>	none	A10, A7	A8, T7, T7	T7, T7	A8, A7	none
<i>Saccharomyces pombe</i>	A7, A7	A7 ^d	T8, A7, T7	A7, A7	A7	—
<i>Drosophila melanogaster</i>	none	—	—	A7	A7	—
<i>Arabidopsis thaliana</i>	A7	none	none	none	—	—
	<i>MutS</i>			<i>MutL</i>		
<i>Escherichia coli</i>	none			none		

^aIn yeast, *PMS1*.

^bAAGAAAAA in place of human A8 homologous sequence.

^c(—) No available data.

^d*Swi4* gene, which has homology to *MSH3*.

⁴Corresponding author.

E-MAIL crboland@ucsd.edu; FAX (858) 822-0301.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.186301>.

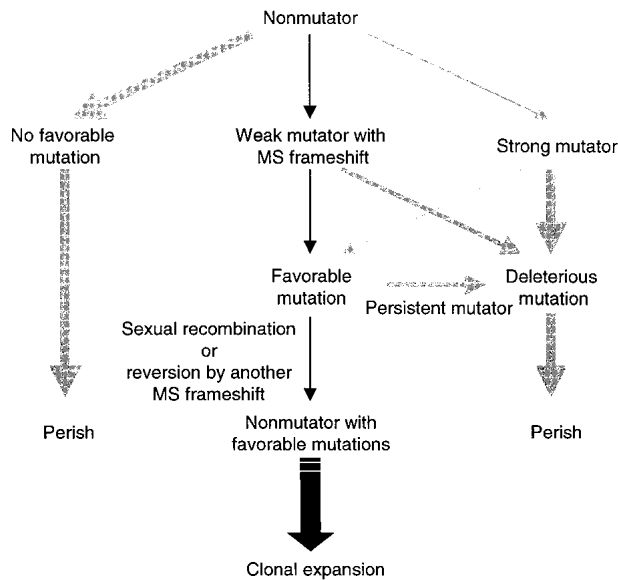


Figure 1 Model of the survival advantage under conditions requiring rapid evolution conferred by the weak mutator phenotype with a microsatellite (MS) sequence in coding regions of the minor DNA MMR genes.

erodimerizes with MSH2, which creates MutS- α or MutS- β complexes, respectively. Similarly, PMS2 (PMS1 in yeast) heterodimerizes with MLH1 to make MutL- α , and MLH3 is also presumed to combine with MLH1. When the function of either MSH2 or MLH1, the major MMR proteins, is impaired, all MMR activity is lost. However, when one of the minor proteins is impaired, MMR function is retained or only partially perturbed because MSH3 and MSH6 have some functional redundancy and can compensate, at least in part, for the loss of the missing proteins. The case may be similar with PMS2 (PMS1 in yeast) and MLH3. As a result, the loss of activity in any one of these minor proteins generates a weaker mutator phenotype than occurs with loss of the major MMR proteins. A high rate of frameshift mutations that inactivate the minor MMR genes would provide a eukaryotic lineage with a subpopulation of individuals that exhibit mildly increased mutation rates. These individuals would be expected to suffer decreased fitness owing to increased deleterious mutations. Nonetheless, under circumstances favoring rapid evolution, a hypermutable subpopulation would also experience an increase in the rate of necessary adaptive mutations. If these microsat-

ellite sequences had appeared in the coding regions of the major DNA MMR genes, a strong mutator phenotype would be likely to occur; this phenotype might be less capable of overcoming the hazardous effects of the accumulation of deleterious mutations.

When strains acquire “favorable” mutations during a transient hypermutable state, they need to revert to a nonmutator phenotype to keep the newly-obtained adaptive characteristics. Among the strains that have acquired favorable mutations, only those that revert to the nonmutator phenotype are likely to survive. Asexual eukaryotes could achieve the nonmutator phenotype through direct reversion of the loss-of-function mutation, because microsatellite mutations can revert at a high rate. In contrast to other types of mutations, such as transitions or transversions, deletion/insertion mutations within repeated sequences have a higher probability of wild-type reversion, because only one additional cycle of DNA polymerase slippage is required to regain the original sequence. In sexually reproducing eukaryotes, recombination would rapidly separate these adaptive mutations from the (otherwise deleterious) mutator phenotypes with which they would initially be associated. Many of these mutator

alleles may in fact be removed by selection, but recombination allows this to happen while simultaneously retaining (or even fixing) beneficial mutations that arose as a result of the mutator alleles’ activity. In the minor DNA mismatch repair genes, the microsatellites may therefore act as an evolutionary switch that modulates the mutation rate under conditions that require rapid evolution.

Is the pool of mutator individuals large enough to be relevant in a putative crisis? In evolutionary terms, even a rare, more fit phenotype might be sufficient if these individuals survive and their progeny increase in number. For example, in humans, there is a non-negligible number of *MSH6* minor MMR gene germ-line mutations in families with a weakly penetrant form of hereditary colon cancer. Kolodner et al. (1999) estimated that germ-line mutations in *MSH6* might be present in 1.5% of all the United States colorectal cancer cases, because familial colon cancer clusters that do not fulfill the Amsterdam criteria for hereditary nonpolyposis colorectal cancer (HNPCC) account for approximately 18% of colorectal cancers, and 8.1% of these probands harbor *MSH6* mutations. Moreover, in 4.5% of the familial non-HNPCC probands, the germ-line mutations have been reported to be located in the exonal polycytosine repeat of the *MSH6* gene (Shin et al. 1999). Perhaps these families represent a case in point of the process we have described.

REFERENCES

- Chen, W. and Jinks-Robertson, S. 1998. *Mol. Cell Biol.* **18**: 6525–6537.
- Kolodner, R.D. and Marsischky, G.T. 1999. *Curr. Opin. Genet. Dev.* **9**: 89–96.
- Kolodner, R.D., Tytell, J.D., Schmeits, J.L., Kane, M.F., Gupta, R.D., Weger, J., Wahlberg, S., Fox, E.A., Peel, D., Ziogas, A., et al. 1999. *Cancer Res.* **59**: 5068–5074.
- LeClerc, J.E., Li, B., Payne, W.L., and Cebula, T.A. 1996. *Science* **274**: 1208–1211.
- Metzgar, D., Bytof, J., and Wills, C. 2000. *Genome Res.* **10**: 72–80.
- Moxon, E.R., Rainey, P.B., Nowak, M.A., and Lenski, R.E. 1994. *Curr. Biol.* **4**: 24–33.
- Rosenberg, S.M., Thulin, C., and Harris, R.S. 1998. *Genetics* **148**: 1559–1566.
- Shin, K.H., Ku, J.L., and Park, J.G. 1999. *J. Human Genet.* **44**: 18–21.
- Sniegowski, P.D., Gerrish, P.J., and Lenski, R.E. 1997. *Nature* **387**: 703–705.
- Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., and Godelle, B. 1997. *Nature* **387**: 700–702.