Cell, Vol. 88, 155–158, January 24, 1997, Copyright ©1997 by Cell Press

Mutator Specificity and Disease: Looking over the FENce

Minireview

Thomas A. Kunkel, Michael A. Resnick, and Dmitry A. Gordenin*
Laboratory of Molecular Genetics
National Institute of Environmental Health Sciences
Research Triangle Park, North Carolina 27709

Only a year after the structure of DNA was described, Treffers et al. (1954) discovered the *mutT* E. coli mutator strain, which has a mutation rate uniquely elevated for a specific base substitution. Shortly thereafter, Benzer (1957) described mutational hotspots within a gene, where the probability of detecting a mutation was much greater than average. Four decades later, the study of mutators specific for certain mutations and of mutational hotspots is central to the search for the genes and processes that influence genome stability and the incidence of human diseases. Recent successes in the hunt for disease genes based on mutational specificity include the identification of several cancer susceptibility genes using microsatellite instability as a biomarker (Umar and Kunkel, 1996; Sia et al., 1997) and the identification of several hereditary diseases associated with expanded triplet repeat sequences (Wells, 1996). In this issue of *Cell*, Tishkoff et al. (1997) report that inactivation of the yeast *RAD27* gene, encoding a Flap EndoNuclease (FEN-1), yields a novel mutator phenotype, an elevated rate of duplications of 5 to 108 bp of DNA located between repeated sequences of 3 to 12 bp (Figure 1). This exciting discovery will likely stimulate new investigations into the mutagenic potential of various DNA transactions. It also suggests new approaches for the rapid identification of genome instabilities.

Mutators, Mutational Specificity, and Cancer

The significance of the new study can be appreciated within the context of the extraordinary progress that has been made in the last four years on the relationship between microsatellite instability, postreplication mismatch repair (MMR) and colon cancer. The incidence of cancer in humans is higher than expected, given that several gene mutations are required for tumor development and the spontaneous mutation rate in normal human cells is low. This paradox led to the hypothesis (reviewed in Loeb, 1991) that an early event in tumorogenesis might be a mutation that elevates the mutation rate and therefore the likelihood of changes in multiple genes relevant to tumor development. This hypothesis was supported by the finding that tumor cells from hereditary nonpolyposis colon cancer (HNPCC) patients, as well as some sporadic tumors, have a novel mutator phenotype: genome-wide additions and deletions within simple, tandemly repeated DNA sequences. This microsatellite instability was invaluable for establishing the identity of several cancer susceptibility genes. Most HNPCC families have mutations in one of the three MMR genes, *hMSH2*, *hMLH1*, or *hPMS2*. The products of these genes are required for correcting simple replication errors, including base·base mispairs and small additions/deletions (Figures 2A and 2B). Mutations inactivating *hMSH2*, *hMLH1*, or *hPMS2* lead to microsatellite instability and to greatly elevated rates of base substitution and frameshift mutations. Other known MMR genes include *hMSH3* and *hMSH6* (Figures 2A and 2B). Based on mutational specificity studies in yeast and humans, Msh3 or Msh6 can partner with Msh2 to form heterodimers with partially redundant functions. The Msh2·Msh6 complex is suggested to participate preferentially in repair of single base mispairs (Figure 2A), whereas Msh2·Msh3 preferentially participates in repair of intermediates for small insertions/deletions (Figure 2B).

Progress in understanding the relationship between defective MMR and cancer susceptibility has depended heavily on research in model systems. Studies with E. coli and yeast laid the groundwork for the recent, rapid progress in humans. Each of the five human mismatch repair genes mentioned above is conserved in yeast and mouse, and our current appreciation of their functions is a result of seminal studies in those organisms. Also, studies with model DNA replication reactions in vitro suggest which sequences, and therefore which genes, may be at great risk for inactivation by mutation in MMR-deficient cells. Two of the three processes that contribute to replication fidelity for base addition and deletions—i.e., DNA polymerase selectivity for using correctly aligned template-primers and exonucleolytic proofreading of misalignments—discriminate less efficiently against replication slippage errors in repetitive sequences as the number of repeats increases (Kroutil et al., 1996). This implies that the genes at risk for mutational inactivation in MMR-deficient cells may be those containing long, simple repeats in their coding sequences, especially homonucleotide runs. Mutations in long homonucleotide runs are indeed found in the coding sequences of genes important for the development of colon cancer (Kinzler and Vogelstein, 1996).

HNPCC and sporadic tumors with the microsatellite instability that is typical of a MMR defect account for

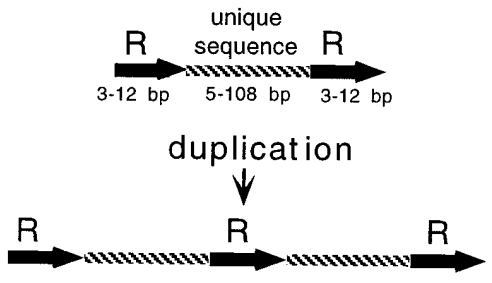

Figure 1. Duplications between Short Repeats Observed in the Yeast *rad27* Mutator

Duplicated are the unique sequence (hatched line) and one of two flanking short direct repeats (R). Such duplications constitute a predominant class of mutations in the *rad27* mutation spectrum (Tishkoff et al., 1997).

*On sabbatical leave from: Department of Genetics, St. Petersburg State University, St. Petersburg, Russia.

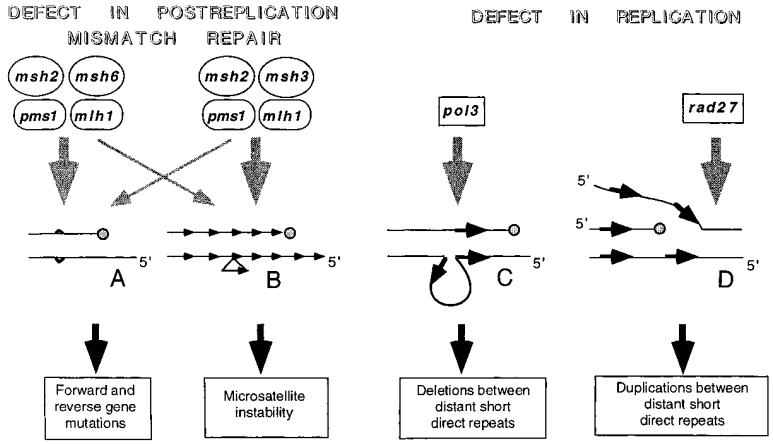

Figure 2. Diagnostic Mutation Events for Yeast and Human Mutators

Shown are proposed mutational intermediates that arise during replication and the gene defects that increase their rates of formation. The indicated genes (yeast designations given) are found in yeast and humans. MMR defects and the mutational consequences have been identified in both organisms, while the other defects and their consequences have only been examined in yeast. (A) Base·base mismatches, leading to substitutions. (B) 1–4 unpaired bases that lead to deletions and additions in microsatellites. The extent of the functional overlap between *MSH3* and *MSH6* in the repair of mispaired and misaligned heteroduplexes is not yet completely understood. (C) Unpaired loops between distant (>30 bp) short (4–9 bp) direct repeats leading to deletions. The origin of deletions in (B) and (C) are proposed to occur via slippage of a 3′ (shaded circle) end of a nascent strand between two small direct repeats (thick arrows) on a template strand. (D) Displaced 5′-flap in the lagging strand occurring in the absence of *RAD27* and leading to duplications (see Tishkoff et al., 1997).

only 15% of colon cancers; the remaining 85% have a high frequency of chromosome rearrangements and loss (Kinzler and Vogelstein, 1996). What processes are responsible for these forms of genome instability? Are there additional mutators that lead to cancer and/or to other diseases? MMR does not repair all types of premutational intermediates generated during replication of normal or damaged DNA. For example, a defect in MMR does not affect the rate of deletions between distant, short-repeat sequences that occur at a 1000-fold increased rate in a yeast *pol3* mutator strain (Tran et al., 1996). This may be due to the inability of the known MMR pathway to repair large unpaired loops generated during replication by DNA polymerase δ (Figure 2C). The study by Tran et al. (1996) also suggested that MMR does not affect the frequency of small deletions in homonucleotide runs if the unpaired intermediate is generated by means other than chromosomal replication, e.g. during synthesis associated with DNA repair. Observations such as these with model systems suggest the existence of additional human mutators distinct from defective MMR genes.

A Yeast Mutator with a Novel Mutational Signature

The search for new mutators in humans may be facilitated by the report of Tishkoff et al. (1997) of a yeast mutator strain with a novel mutagenic specificity. This strain contains a null mutation in the yeast *RAD27* (*RTH1*) gene, a member of a gene family that includes the gene encoding human flap endonuclease, *FEN1*. Flap endonuclease cleaves branched DNA structures, including the 5'-ends of Okazaki fragments, allowing completion of the lagging strand during DNA replication. An earlier study (Johnson et al., 1995) had demonstrated that disruption of *RAD27* yields a strongly elevated rate of two-base additions within a dinucleotide microsatellite. Based on comparing the mutator effect in a *rad27* mutant to that in a *msh2* mutant and in a *rad27-msh2* double mutant, those authors suggested that the *rad27* mutator effect was due to a defect in MSH2-dependent mismatch repair. Tishkoff et al. (1997) likewise examined mutation rates in a *rad27* mutant, but used three different reporter gene systems, a forward mutation system that detects any mutation that inactivates the arginine permease gene (Can^r mutations) and reversion assays that detect mutations that revert a 4 bp insertion in the *LYS2* gene or revert a one A·T basepair insertion in a run of consecutive A·T basepairs in the *HOM3* gene. Like Johnson et al. (1995), they too find that the *rad27* mutant is a strong mutator, comparable to a *msh2* mutator. Surprisingly however, the mutations to Can^r and Lys$^+$ observed in the *rad27* mutant were mostly duplications that range in size from 5 to 108 bp and are flanked by 3 to 12 bp direct repeat sequences. Duplicated nucleotides included both the intervening sequence and one copy of the repeat (Figure 1). In contrast, point mutations predominate in the *msh2* strain and the mutation spectrum in the *msh2-rad27* double mutant (although based on a small sample size) corresponds to the sum of the single-mutant spectra. Collectively, these data suggest that *RAD27* and *MSH2* act in different mutation avoidance pathways. Tishkoff et al. (1997) suggest that aberrant DNA intermediates arising in a *rad27* mutant may be resolved via double-strand break repair. Consistent with this, they observe that mutants defective in both *RAD27* and either *RAD51* or *RAD52*, two genes involved in recombinational repair of double-strand breaks, are inviable. The suggested involvement of double-strand breaks seems reasonable, although alternatives cannot be completely excluded since some phenotypes of *rad52* mutants suggest that *RAD52* may play a role in DNA replication (Tran et al., 1995, and references therein).

The authors propose that the observed duplications are initiated when strand displacement synthesis occurs without the *RAD27*-dependent removal of the resulting flap (Figure 2D). This process for initiating mutations is very different from the mispairs initiated by miscoding or the small misalignments initiated by strand slippage in microsatellites that are normally corrected by MMR. The new results will undoubtedly stimulate additional studies of strand displacement DNA synthesis reactions. Two models are proposed to explain how duplications might arise from the intermediate shown in Figure

2D, long range strand slippage during replication (figure 4, step [i] in Tishkoff et al., 1997) and aberrant annealing of the unexcised flap (figure 4, steps [b]–[h] in Tishkoff et al., 1997). The authors do not favor the former model because some of the observed duplications involve misalignments of a size known (e.g., see Tran et al., 1996) to be corrected by postreplication MMR. However, the data in Tishkoff et al. are too sparse to completely discount the replication slippage model, since only one duplication of less than 10 bp was recovered in the *rad27* mutant and only three duplications of any size were reported from the *rad27/msh2* double mutant.

In addition to its mutator phenotype, the *rad27* mutant also exhibited an increased rate of mitotic crossing over. Double-strand breaks could account for this increased recombination, through interaction of a broken DNA, containing an unexcised flap, with homologous DNA (figure 4, step [e] in Tishkoff et al., 1997) located either in a sister chromatid or elsewhere in the genome. The latter would lead to chromosome rearrangements and to a loss of heterozygosity. Another type of homologous interaction to repair a double-strand break is annealing of overlapping complementary single-stranded DNA tails resulting from an unexcised flap (single-strand annealing, figure 4, step [f] in Tishkoff et al. 1997). This would restore the wild-type sequence. The authors suggest that the duplication mutations are a minor aberrant outcome of single-strand annealing (Figure 4, steps [g] and [h] in Tishkoff et al., 1997). Double-strand break repair could also account for increases in small frameshift mutations caused by the *rad27* mutation, since frameshift mutation rates are very high in the region of gap filling during double-strand break repair (Strathern et al., 1995).

At-Risk Sequences and Diagnostic Instabilities

The increased duplication and recombination rates seen in the *rad27* mutant suggest additional sources of genome instability in humans. Just as defective MMR puts microsatellites at risk of mutation, the yeast *rad27* defect puts nontandem direct repeat sequences at risk. If similar mutators exist in human cells, the number of at-risk sequences is enormous. For example, every 100 bp of random DNA sequence contains on average one pair of 6 base repeats, three pairs of 5 base repeats, and 11 pairs of 4 base repeats (Moore et al., 1984). This gives an estimate of 10 million or more such repeats just in the coding sequences of ~50,000 human genes (Table 1). The target size is further elevated by the fact that the repeats involved in duplication mutagenesis need not be perfect (table 2 in Tishkoff et al., 1997). The authors note that duplications of the kind observed in the mutation spectra of the *rad27* mutator yeast strain have been described in the human *p53* and *APC* genes and as germline mutations in three human diseases (also see Cooper and Krawczak, 1994). The observations by Tishkoff et al. (1997) could be relevant to the large expansions of triplet repeat sequences observed in genes associated with hereditary diseases (Wells, 1996), and also offer an explanation for the origin of the tandem repeats found in many different proteins (Matsushima et al., 1990).

There are two possible sequence arrangements other than microsatellites where changes might be diagnostic of *fen1* (*rad27*) or other mutators distinct from MMR mutators in human tumors (Table 1). One idea would be to use minisatellites, which are not reported to be highly unstable in MMR defective cells. Minisatellites loci, containing 6–100 bp sequences tandemly repeated for hundreds to thousands of base pairs, are common and are used for mapping, DNA typing, and mutation detection by hybridization or PCR (Jeffreys et al., 1995). Based on the observed inability of MMR to prevent *rad27*-dependent duplications between distant short repeats, a high rate of minisatellite changes is predicted in human cells carrying mutators similar to yeast *rad27*. Following the same logic, mutators analogous to yeast *pol3* can reveal increased rates of repeat-associated deletions. In these cases, deletions and duplications might be relatively short (1–5 units). Minisatellite mutations leading to large (in the kilobase range) changes in size could result from recombination rearrangements.

A second category of diagnostic mutations would be duplications and/or deletions between direct repeats separated by random DNA, similar to the mutations actually detected in yeast *rad27* and *pol3* mutators. Although these mutations may occur less frequently than changes in micro- and minisatellites, the target is the entire genome and the mutator effects (in yeast) are very strong, approximately 1000-fold. Furthermore, repeat-associated deletions and duplications create unique sequence junctions that can be specifically identified using short oligonucleotide hybridization probes, perhaps using automated techniques on microchips. One could predict an array of unique sequence junctions, giving highest priority to events expected to be most frequent due to specific DNA structures formed in internal sequences. For example, palindromic sequences between short direct repeats strongly increase the deletion rate in E. coli, as compared with the rate between repeats

Table 1. Examples of At-Risk Sequences for Detecting Human Mutators

At-Risk Sequence		Possible Intermediate		
Type[a]	Occurrence in Human Genome[b]	Loops < 4 nt	Loops > 30 nt	Annealed Short Repeats[c]
Microsatellites (1–4 bp tandem repeats)	≥100,000	+	+	+
Minisatellites (30–100 bp tandem repeats)	≥10,000	−	+	+
Short (4–6 bp) nontandem repeats separated by 30–100 bp	≥10,000,000	−	+	+

[a] Arrangements capable of forming small (<4 nt) and big (>30 nt) loops via replication slippage. These would distinguish between MMR-dependent and MMR-independent mutators (see Figure 2 and text).

[b] For estimated minimal numbers of various microsatellites and minisatellites see Sia et al., 1997, and Jeffreys et al., 1995, and references therein). The number of distant short repeats was estimated according to Moore et al., 1984.

[c] For example, see figure 4, step (g) in Tishkoff et al., 1997.

flanking nonpalindromic DNA. In contrast, intervening palindromic DNA decreases the rate of duplications between direct repeats (Trinh and Sinden, 1993).

FEN-1 Mutations and Disease

Attempts to identify *FEN-1* mutations in human tumors have thus far been unsuccessful (Risinger et al., 1996; Tishkoff et al., 1997). This might reflect a selective disadvantage to a human cell imparted by a mutation in this important replication gene. While yeast *RAD27* (*FEN1*) is not essential for viability, deletion mutants are temperature-sensitive for growth. However, FEN-1 is a multifunctional enzyme that probably participates in several DNA transactions, so mutations that selectively yield a mutator without affecting growth may well exist. It is also logical to search for mutator mutations in other genes that might operate in the same pathway(s) as *FEN1/RAD27*. One obvious candidate is the gene encoding proliferating cell nuclear antigen (PCNA), which physically interacts with and strongly stimulates the endo- and exonucleolytic activities of FEN-1 (Li et al., 1995). No matter what insights are ultimately gained from the article by Tishkoff et al. (1997), it is satisfying to realize that many years after the now classical observations of Treffers and Benzer, the study of mutational specificity still has much to offer in trying to understand the origins of human diseases.

Selected Reading

Benzer, S. (1957). The elementary units of heredity. In The Chemical Basis of Heredity, W.D. McElroy and B. Glass, eds. (Baltimore: The Johns Hopkins Press), pp. 70–93.

Cooper, D.N., and Krawczak, M. (1994). Human Gene Mutation (Oxford, UK: BIOS Scientific).

Jeffreys, A.J., Allen, M.J., Armour, J.A., Collick, A., Dubrova, Y., Fretwell, N., Guram, T., Jobling, M., May, C.A., Neil, D.L., Neumann, R. (1995). Electrophoresis *16*, 1577–1585.

Johnson, R.E., Kovvali, G.K., Prakash, L., and Prakash, S. (1995). Science *269*, 238–240.

Kinzler, K.W., and Vogelstein, B. (1996). Cell *87*, 159–170.

Kroutil, L.C., Register, K., Bebenek, K., and Kunkel, T.A. (1996). Biochemistry *35*, 1046–1053.

Li, X., Li, J., Harrington, J., Lieber, M.R., and Burgers, P.M.J. (1995). J. Biol. Chem. *270*, 22109–22112.

Loeb, L.A. (1991). Cancer. Res. *51*, 3075–3079.

Matsushima, N., Creutz, C.E., and Kretsinger, R.H. (1990). Proteins *7*, 125–155.

Moore, G.P., Moore, A.R., and Grossman, L.I. (1984). J. Theor. Biol. *108*, 111–122.

Risinger, J.I., Umar, A., Boyd, J., Berchuk, A., Kunkel, T.A., and Barett, J.C. (1996). Nature Genet. *14*, 102–105.

Sia, E.A., Jinks-Robertson, S., and Petes, T.D. (1997). Mutation Res. *381*, 61–70.

Strathern, J.N., Shafer, B.K., and McGill, C.B. (1995). Genetics *140*, 965–972.

Tishkoff, D.X., Filosi, N., Gaida, G.M., and Kolodner, R.D. (1997). Cell *88*, this issue.

Tran, T.H., Degtyareva, N.P., Koloteva, N.N., Sugino, A., Masumoto, H., Gordenin, D.A., and Resnick, M.A. (1995). Mol. Cel. Biol. *15*, 5607–5617.

Tran, T.H., Gordenin, D., and Resnick, M. (1996). Genetics *143*, 1579–1587.

Treffers, H.P., Spinelli, C., and Belser, N.O. (1954) Proc. Natl. Acad. Sci. USA *40*, 1064–1071.

Trinh, T.Q., and Sinden, R.R. (1993). Genetics *134*, 409–422.

Umar, A., and Kunkel, T.A. (1996). Eur. J. Biochem. *238*, 297–307.

Wells, R.D. (1996). J. Biol. Chem. *271*, 2875–2878.