Available online at www.sciencedirect.com
ScienceDirect
Mutation Research 640 (2008) 89–96
Fundamental and Molecular Mechanisms of Mutagenesis
www.elsevier.com/locate/molmut
Community address: www.elsevier.com/locate/mutres

Sequence-dependent effect of interruptions on microsatellite mutation rate in mismatch repair-deficient human cells

Jayne C. Boyer[a,*], Joshua D. Hawk[a], Lela Stefanovic[a], Rosann A. Farber[a,b,c]

[a] *Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, CB #7525, Chapel Hill, NC 27599, United States*
[b] *Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States*
[c] *Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States*

Received 10 October 2007; received in revised form 21 November 2007; accepted 11 December 2007
Available online 23 December 2007

Abstract

Although microsatellite mutation rates generally increase with increasing length of the repeat tract, interruptions in a microsatellite may stabilize it. We have performed a direct analysis of the effect of microsatellite interruptions on mutation rate and spectrum in cultured mammalian cells. Two mononucleotide sequences (G_{17} and A_{17}) and a dinucleotide [$(CA)_{17}$] were compared with interrupted repeats of the same size and with sequences of 8 repeat units. MMR-deficient (MMR^-) cells were used for these studies to eliminate effects of this repair process. Mutation rates were determined by fluctuation analysis on cells containing a microsatellite sequence at the 5′ end of an antibiotic-resistance gene; the vector carrying this sequence was integrated in the genome of the cells. In general, interrupted sequences had lower mutation rates than perfect ones of the same size, but the magnitude of the difference was dependent upon the sequence of the interrupting base(s). Some interrupted repeats had mutation rates that were lower than those of perfect sequences of the same length but similar to those of half the length. This suggests that interrupting bases effectively divide microsatellites into smaller repeat runs with mutational characteristics different from those of the corresponding full-length microsatellite. We conclude that interruptions decrease microsatellite mutation rate and influence the spectrum of frameshift mutations. The sequence of the interrupting base(s) determines the magnitude of the effect on mutation rate.

Keywords: Mutation rate; Microsatellite instability; Mononucleotide repeats; Dinucleotide repeats; Interruption

1. Introduction

It is estimated that over 50% of the human genome is composed of repetitive DNA sequences [1]. Microsatellites have a repeat-unit size of $n = 1$–6 bp [1]. Microsatellite sequences are abundant in the human genome and occur with a frequency of approximately one per 2 kb of DNA [1]. They consist of about 6–30 repeat units and tend to be highly polymorphic [2], making them useful as genetic markers for linkage mapping of disease genes [3] and in evolutionary [4] and population genetics studies [5]. The phenotype of the familial cancer predisposition syndrome Lynch syndrome, or HNPCC (hereditary non-polyposis colorectal cancer), includes microsatellite instability in the tumors [6]. This syndrome usually results from germline mutations in DNA mismatch repair genes [7]. The microsatellite instability phenotype (MSI) is also observed in a significant fraction of sporadic tumors of the types that are associated with Lynch syndrome [8]. Changes in the size of the microsatellite repeat occur by frameshift mutations that are most likely caused by a DNA polymerase slippage mechanism during DNA replication [9,10]. Most microsatellites are located in non-coding regions of the genome, but some of these repeats are found in coding regions of genes. In MSI tumors, inactivating frameshift mutations have been found in repeats in genes involved in carcinogenesis [11], including *TGFBR2* [12], *BAX* [13], *IGFR2* [14], *MSH3*, and *MSH6* [15].

Changes in microsatellite length can also affect RNA splicing efficiency [16,17] and regulation of gene expression [18,19]. Consequences of MSI in lower organisms include the hypermutability of repetitive sequences (i.e., contingency loci) in pathogenic bacteria and viruses that is thought to aid in their adaptation to the changing environment of their host [20,21]. In

* Corresponding author. Tel.: +1 919 966 2561x292; fax: +1 919 966 3454.
E-mail address: jayne_boyer@med.unc.edu (J.C. Boyer).

0027-5107/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrfmmm.2007.12.005

humans, large expansions of certain microsatellites have been associated with hereditary conditions, such as fragile X syndrome and myotonic dystrophy [22]. The characterization of microsatellite instability and factors that influence it is important in understanding its role in cancer, gene expression, and microbial virulence.

Several characteristics of microsatellites have been shown to influence the extent of their instability. These include repeat-unit length [e.g., A_n vs. $(CA)_n$ vs. $(CAG)_n$] [23], base composition (e.g., A_n vs. G_n) [24–26,57], length of the repeat tract (e.g., A_8 vs. A_{17}) [27–30], and sequence context [25,31–33]. Another factor that can affect microsatellite instability is the degree of "perfection" of the repeat tract. Kunkel [34] first showed that the *in vitro* frameshift mutation frequency at a mononucleotide run was decreased by sequence interruptions. The degree of perfection of microsatellites in the human genome was subsequently shown to correlate with the extent of heterozygosity in the population, suggesting that pure repeats are less stable than interrupted repeats *in vivo* [35]. Repeat stabilization by sequence interruptions has also been observed in trinucleotide-repeat disorders (e.g., fragile X syndrome and spinocerebellar ataxia 1), although the mechanisms involved in these large expansions must be more complex than those associated with the lesser degree of instability associated with most microsatellites [36–38]; similar stabilization of triplet repeats has also been reported in yeast [39–41] and bacteria [42,43]. Mutation rates of dinucleotide repeats are also lowered by interruptions in the DNA of *E. coli* [31] and yeast [44,45]. Length analyses of many microsatellite loci indicate that perfect repeats are significantly more variable than imperfect repeats in humans [46,47], other mammals [48,49] and Drosophila [50,51]. Human meiotic mutation rates are 4 times lower in interrupted tetranucleotide and pentanucleotide microsatellites than in pure repeats [52]. Results of typing of endogenous mutant microsatellite alleles in human tumors [53] and lymphoblastoid cell lines [54] indicate that pure microsatellites are relatively unstable.

Although the above observations show a correlation between the degree of perfection of a microsatellite and its stability, none have measured the mutation rate of a pure microsatellite and its corresponding interrupted sequence. In contrast to these studies, we have directly measured the rates of frameshift mutations in perfect and imperfect microsatellites in cultured human cells. Cells deficient in hMLH1 have been used for these studies in order to focus on the effects of DNA replication errors in the absence of mismatch repair. We have examined the effects of interrupting sequences on the mutation rates and mutational spectra of A_{17}, G_{17} and $(CA)_{17}$-repeats; we also compare these mutation rates with those of pure sequences of half that length [A_8, G_8, and $(CA)_8$], corresponding to one end of the longer interrupted repeats. In general, interrupted sequences had lower mutation rates than perfect ones, but the magnitude of the difference was dependent on the specific interrupting base(s). These data demonstrate the complexities of microsatellite mutagenesis and contribute to an understanding of variation in microsatellite instability throughout the genome.

2. Materials and methods

2.1. Cell culture

H6 is a subclone of the human colorectal cancer cell line HCT116 [55], which lacks mismatch repair activity as the result of the absence of a normal copy of the *hMLH1* gene [56]. It was cultured in DMEM containing 10% iron-supplemented calf serum (Hyclone, Logan, UT). Cells were maintained without antibiotics at 37 °C in 5% atmospheric CO_2 and are free of mycoplasma.

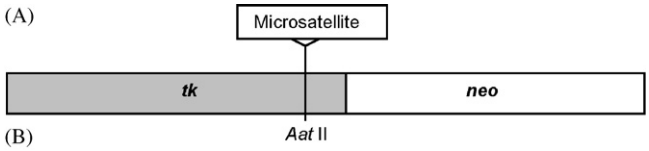
2.2. Plasmid construction

All plasmids except pRTM2 and pJCB1 were made by insertion of annealed oligonucleotides containing repetitive sequences into the parent plasmid, pConπ, which contains a fusion gene consisting of the herpes simplex virus thymidine kinase gene (*tk*) fused at its 3′ end to the 5′ end of a bacterial gene coding for neomycin resistance (*neo*) [57] (Fig. 1A). The oligonucleotides were inserted at an *Aat* II site near the 3′ end of the *tk* gene, such that the *neo* gene is in the −1 reading frame. Revertants with mutations in the repeat sequences that restore the *neo* reading frame are resistant to the neomycin analogue G418. The pJCB1 plasmid [$(CA)_{17}$iT] was constructed by site-directed mutagenesis of the plasmid pRTM2 [$(CA)_{17}$] [58], using the QuikChange™ site-directed mutagenesis kit (Stratagene, La Jolla, CA). Sequences of the microsatellite oligonucleotide inserts are shown in Fig. 1B.

2.3. Transfection and fluctuation analysis

Transfection and fluctuation analysis were performed as previously described [57]. Cells were electroporated with linearized plasmid DNA and transformed clones, which had the plasmid integrated into the genome, were selected with hygromycin B. Independent hygromycin-resistant (hyg^R) clones were isolated from different plates and the cultures were expanded. Most hyg^R clones contained one insert. Clones that have multiple inserts can be identified by the polymerase chain reaction (PCR) product profiles of the DNA from the $G418^R$ clones. Clones with more than one insert exhibit a normal-sized PCR product in addition to the PCR product with the frameshift mutation. The number of inserts can be determined by PhosphorImager analysis of PCR products that are separated on 6% polyacrylamide gels. All of the transformants presented here contained single inserts except where noted.

Ten subcultures were established from each hyg^R clone. Neomycin-resistant microsatellite mutants were selected by plating cells from each subculture in medium containing the neomycin analogue G418. $G418^R$ colonies were isolated



A_8:	GACGTCTCGAGG**(A)$_8$**GATCGAGACGTC
A_{17}iC:	GACGTCTCGAGG**(A)$_8$C(A)$_8$**GATCGAGACGTC
A_{17}iG:	GACGTCTCGAGG**(A)$_8$G(A)$_8$**GATCGAGACGTC
A_{17}:	GACGTCTCGAGG**(A)$_{17}$**GATCGAGACGTC
G_8:	GACGTCTCGA**(G)$_8$**ATCGAGACGTC
G_{17}iT	GACGTCTCGA**(G)$_8$T(G)$_8$**ATCGAGACGTC
G_{17}:	GACGTCTCGA**(G)$_{17}$**ATCGAGACGTC
$(CA)_8$:	GACGTCTCGGAGGA**(CA)$_8$**GATCGACACGTC
$(CA)_{17}$iTA:	GACGTCTCGGAGGA**(CA)$_8$TA(CA)$_8$**GATCGACACGTC
$(CA)_{17}$iTG:	GACGTCTCGGAGGA**(CA)$_8$TG(CA)$_8$**GATCGACACGTC
$(CA)_{17}$:	GACGTCTCGGAGGA**(CA)$_{17}$**GATCGACACGTC

Fig. 1. (A) The target region of the pConπ vector. The *Aat* II site was used to insert oligonucleotides into the vector to construct microsatellite-containing plasmids. (B) Sequences of the oligonucleotide inserts containing the microsatellites. Note that the G_{17} sequence used here is equivalent to that previously named RG_{17} [57]. The original G_{17} sequence [57] and the one used here differ slightly in their flanking sequences and their mutation rates.