Two Modes of Germline Instability at Human Minisatellite MS1 (Locus D1S7): Complex Rearrangements and Paradoxical Hyperdeletion

Ingrid Berg,1 Rita Neumann,2 Håkan Cederberg,1 Ulf Rannug,1 and Alec J. Jeffreys2

1Department of Genetic and Cellular Toxicology, Stockholm University, Stockholm; and 2Department of Genetics, University of Leicester, Leicester, United Kingdom

Minisatellite MS1 (locus D1S7) is one of the most unstable minisatellites identified in humans. It is unusual in having a short repeat unit of 9 bp and in showing somatic instability in colorectal carcinomas, suggesting that mitotic replication or repair errors may contribute to repeat-DNA mutation. We have therefore used single-molecule polymerase chain reaction to characterize mutation events in sperm and somatic DNA. As with other minisatellites, high levels of instability are seen only in the germline and generate two distinct classes of structural change. The first involves large and frequently complex rearrangements that most likely arise by recombinational processes, as is seen at other minisatellites. The second pathway generates primarily, if not exclusively, single-repeat changes restricted to sequence-homogeneous regions of alleles. Their frequency is dependent on the length of uninterrupted repeats, with evidence of a hyperinstability threshold similar in length to that observed at triplet-repeat loci showing expansions driven by dynamic mutation. In contrast to triplet loci, however, the single-repeat changes at MS1 exclusively involve repeat deletion, and can be so frequent—as many as 0.7–1.3 mutation events per sperm cell for the longest homogeneous arrays—that alleles harboring these long arrays must be extremely ephemeral in human populations. The apparently impossible existence of alleles with deletion-prone uninterrupted repeats therefore presents a paradox with no obvious explanation.

Introduction:

Human GC-rich minisatellites are highly variable tandem-repetitive DNA sequences located predominantly in subtelomeric regions of the genome (Jeffreys et al. 1985; Royle et al. 1988). Alleles can be as long as 30 kb, with repeat-unit lengths of 6–100 bp (Jeffreys et al. 1995). New alleles at these highly unstable loci generally arise by germline-specific complex recombination events, including both interallelic gene conversions and intra-allelic rearrangements (Armour et al. 1993; Buard and Vergnaud 1994; Jeffreys et al. 1994; Tamaki et al. 1999), reviewed by Bois and Jeffreys [1999]), with most loci showing a bias toward mutation in the male germline (Henke and Henke 1995; Dubrova et al. 1997). Mutants are also generated via a crossover pathway, but at a lower frequency (Jeffreys et al. 1998b; Buard et al. 2000b). Some loci show polarity, with mutation events targeted to one end of the repeat array (Jeffreys et al. 1994). Analysis of one minisatellite has shown that this polarity correlates with a highly localized meiotic recombination hotspot, adjacent to the repeat DNA, which appears to drive repeat-DNA instability (Jeffreys et al. 1998a). New alleles are also generated in somatic cells but are far less frequent than in the germline and arise by simple intra-allelic duplications and deletions (May et al. 1996; Jeffreys and Neumann 1997; Buard et al. 2000a). Germline mutation frequencies vary not only between minisatellites but also between alleles at a given locus (Monckton et al. 1994; May et al. 1996; Buard et al. 1998; Tamaki et al. 1999), and, in one case, this variation has been linked to a flanking polymorphism that appears to influence hotspot activity (Jeffreys et al. 1998a). The most dramatic variation is observed at CEB1, where sperm mutation rates vary from <0.05% to 25% per allele, the latter being the highest rate yet observed at a minisatellite (Buard et al. 1998).

Minisatellite MS1 (D1S7) is one of the most unstable human minisatellites yet identified and is unusual in having a short repeat unit of 9 bp. It is located in the subtelomeric region of chromosome 1 (1p33-35) and shows alleles ranging from 60 to >1,000 repeats, with an allele-length heterozygosity >99% (Wong et al. 1987; Royle et al. 1988). The germline mutation rate has been estimated in families at 5.2% per gamete (Jeffreys et al. 1988), although this will be an underestimate, given the short repeat unit and long arrays, which make the detection of small mutational changes extremely difficult. In contrast to most minisatellites, mutation occurs with
similar frequencies in both the male and female germline (Jeffreys et al. 1988; Henke and Henke 1995; Dubrova et al. 1997). MS1 has been extensively used in forensic analyses (Wong et al. 1987) and is the only minisatellite known to be unstable in colon cancer cells that show microsatellite instability (Hoff-Olsen et al. 1995; see Mitchell et al. 2002), suggesting that replication or repair errors might contribute to MS1 repeat instability. MS1 instability processes have been studied only in *Saccharomyces cerevisiae*, after integration of the human minisatellite near a recombination hotspot upstream of the yeast *LEU2* locus (Cederberg et al. 1993; Maleki et al. 1997; Berg et al. 2000). In yeast, MS1 shows both mitotic and meiotic instability at high frequencies. This is in contrast to other minisatellites studied in the same system, such as MS32 (Appelgren et al. 1997) and MS205 (He et al. 1999), which mutate at high frequency only in meiosis. In mitosis, MS1 instability is related not only to allele length, with evidence for an instability threshold of 0.75 kb (Maleki et al. 1997), but also to the internal structure of the alleles. Tetrad analysis showed that mutants arising at meiosis are generated by both intra- and interallelic recombination events and that most mutants arise by gene conversions that include the minisatellite plus neighboring DNA (Berg et al. 2000).

The evidence, both from yeast and from tumors, therefore suggests that minisatellite MS1 may be unusual in showing both meiotic and mitotic repeat instability. However, there is no direct information available on MS1 mutation processes in humans. We have therefore used both small-pool (SP) PCR (Jeffreys et al. 1994) and single-molecule (SM) PCR (Yauk et al. 2002) to analyze the incidence of de novo mutant molecules in sperm and somatic DNA. We then used minisatellite variant repeat (MVR) mapping by PCR (Jeffreys et al. 1991; Berg et al. 2000) to characterize the structural basis of mutation within the repeat array. Although this work has shown that MS1 mutation does share some features in common with other GC-rich minisatellites, it has also revealed an entirely novel germline-specific instability process that rapidly eliminates certain classes of alleles, rendering their existence in human populations an enigma.

**Material and Methods**

**Mutant Detection and Purification**

All DNAs were prepared, as described elsewhere, under conditions designed to minimize the risk of DNA contamination (Jeffreys et al. 1994). Genomic DNAs were digested with *MboI* prior to use. DNA concentrations were estimated by use of UV spectrometry, and the number of amplifiable molecules of each MS1 allele was confirmed by Poisson analysis of limiting single-molecule dilutions of digested DNA (Jeffreys et al. 2000). Extreme dilutions of genomic DNA were made in 5 mM Tris-HCl (pH 7.5) and 5 µg/ml carrier herring sperm DNA prior to PCR amplification.

SP-PCR analyses were performed on multiple aliquots of digested genomic DNA, each containing 30–80 amplifiable molecules of each MS1 allele (0.22–0.58 ng DNA per PCR). DNA was amplified by long PCR using *Taq* plus *Pfu* polymerases, as described elsewhere (Jeffreys et al. 1998b), with 0.2 µM primers MS1–420 (5’S-AGGTCTCTAGCATAGCTTTGGGCACG-3’) and MS1+280 (5’S-CATTTCAGATAGGAAAACGTGACGCC-3’). These primers generate an MS1 amplicon containing 410 bp of DNA 5’ to the repeat array plus 310 bp of DNA 3’ to the array. Cycling was for 1 min at 96°C, followed by 25 cycles at 96°C for 20 s, 60°C for 30 s, and 70°C for 2.5 min. PCR products were electrophoresed through a 40-cm 0.8% SeaKem LE (FMC Bioproducts) agarose gel in 44 mM Tris-borate (pH 8.3) and 1 mM EDTA and was then transferred to a nylon membrane (Hybond N, Amersham) and hybridized with a 32P-labeled MS1 probe (Wong et al. 1987). SM-PCR analyses were similarly performed on multiple aliquots of genomic DNA, each containing 4.0 pg DNA (0.55 amplifiable molecule), with DNA amplified for 25 cycles. Positive SM-PCRs were identified by Southern blot hybridization with the MS1 probe; they were then re-amplified for 11–14 cycles, using the nested primers MS1–350 (5’S-CCTTTGGCATTCCATAAAACCGTAC-3’) and MS1–280 (5’S-AAGAAGCATAGTGGCAACCCCATGAGG-3’), and PCR products were electrophoresed side by side on 40-cm gels, as for SP-PCR analyses.

Mutants detected in SP-PCRs were purified away from the progenitor allele by two or three rounds of agarose gel electrophoresis, collection of size fractions around the position of the mutant, limited reamplification of each fraction for 6–15 cycles with primers MS1–420 and MS1+280, and analysis of PCR products by gel electrophoresis and hybridization. Final mutant products were at least 95% pure and were reamplified for 11–24 cycles with the nested primers prior to MVR analysis. Full details of the purification procedure were reported elsewhere (Jeffreys and Neumann 1997).

**MVR-PCR Analysis**

MVR-PCR analysis of progenitor alleles, purified mutants, and PCR products from individual SM-PCR was performed as described elsewhere (Berg et al. 2000), with some modifications. All primers were purchased from Scandinavian Gene Synthesis and were >99% full length. Forward mapping used 1 µM flanking primer 1-HF (5’S-CACGCTATTGGCATTGGATTGATGT-3’), located 263 bp into the 5’ flanking region of MS1, together with
2–5 fg PCR product, 0.05 U/μl Taq polymerase, 10 nM 1-TAG-A primer, 1 nM 1-TAG-B primer, 1 nM 1-TAG-K primer or 10–20 nM 1-TAG-C primer depending on the number of C-type repeats, plus 1 μM TAG (5′-TCATGCCTACCATGTCCGA-3′). The MVR primers 1-TAG-A, -B, -K, and -C are described elsewhere (Berg et al. 2000). DNA was amplified at 96°C for 15 s, 50°C for 35 s, and 70°C for 3.3 min for 4 cycles, and then underwent 13 cycles of 96°C for 15 s, 61°C for 35 s, and 70°C for 3.3 min, followed by a chase at 70°C for 10 min. Reverse mapping was performed as for forward mapping, with the same primer concentrations, using the following primers: 3′ flanking primer 1-HR (5′-AGGACCACCAATCTGGGCTCCCA-3′), located 25 bp downstream of MS1, plus TAG and one of the following A, B, C, or K repeat–specific primers (the synthetic TAG 5′ extension is indicated in lowercase): 1-R-TAG-A (5′-ctagtgcctctagcggctccggaCCCT) [A/G]TCCACCCT[A/G]TCCACCCTA-3′; 1-R-TAG-B (5′-ctagtgcctctagcggctccggaCCCT) [A/G]TCCACCCT[A/G]TCCACCCTCTC-3′; 1-R-TAG-C (5′-ctagtgcctctagcggctccggaCCCT) [A/G]TCCACCCT[A/G]TCCACCCTCTCCA-3′; or 1-R-TAG-K (5′-ctagtgcctctagcggctccggaCCCT) [A/G]TCCACCCT[A/G]TCCACCCTTAA-3′. MVR-PCR products from the four different reactions were analyzed side by side by agarose gel electrophoresis and PCR products from the four different reactions were analyzed side by side by agarose gel electrophoresis and Southern blot hybridization with the 32P-labeled MS1 probe. Forward and reverse MVR codes read from autoradiographs were merged into a complete allele structure. Repeat units that failed to amplify with any of the MVR primers (because of the presence of additional variant[s] that block primer binding) were classified as O-type repeats. MVR maps are oriented as in the article by Gray and Jeffreys (1991).

Results

Detecting MS1 Mutants in Human DNA

Given the small size of the MS1 repeat unit, we focused our attention on short alleles to maximize the resolution of mutants with altered numbers of repeats. We screened a panel of 98 semen donors of northern European origin and identified 11 men who were heterozygous for a large allele plus a short allele of <160 repeats suitable for mutation analysis. Similarly, analysis of 68 Zimbabwean semen donors identified 11 short alleles, with one man heterozygous for two different short alleles.

We used SP-PCR (Jeffreys et al. 1994) to analyze sperm DNA from one of the northern European donors (man 1) for whom blood and buccal DNA was also available. Man 1 was heterozygous for an allele 140 repeats long plus a second allele of ~1,000 repeats, too large for mutation analysis (fig. 1A). Mutants of the short allele involving the gain or loss of at least five repeats (the resolution limit for detecting changes in man 1 by SP-PCR) could be detected at a frequency of 1.5% per sperm cell (54 mutants in 3,500 progenitor molecules screened). Mutational changes of the small allele ranged from a loss of 54 repeats to a gain of 110 repeats. Mutation was biased toward expansion, with 85% of mutants showing gains of repeats. Analysis of blood DNA and buccal DNA (5,100 and 2,900 molecules screened respectively) revealed no detectable mutants, indicating a low frequency of changes in somatic DNA (<0.06% and <0.1% for blood and buccal DNA, respectively; P > .05). Thus, as with other minisatellites, instability at MS1 resulting in large length changes is mainly restricted to the germline.

We used SM-PCR (Yauk et al. 2002) to analyze smaller length changes at MS1. This method uses amplification of minisatellite molecules from extreme dilutions of genomic DNA such that 98% of reactions will contain, at most, only one or two amplifiable molecules. Analysis of 120 molecules from sperm DNA revealed considerable allele length heterogeneity (fig. 1B), with molecules varying in length by as many as four repeats and with some PCRs showing two different length molecules. Although this heterogeneity indicates considerable instability, the small length changes made it impossible to identify progenitor alleles with confidence, preventing the estimation of mutation rate from allele length changes. In contrast, analysis of blood DNA (120 molecules screened; fig. 1B) and buccal DNA (54 molecules analyzed; data not shown) showed no evidence of such small length changes. Thus, as with large mutational changes, these small rearrangements appear to be largely, if not completely, restricted to sperm DNA.

The Structural Basis of Large Mutational Changes in Sperm DNA

We used an MVR-PCR system for analyzing the interspersion patterns of five variant repeat types (A, B, C, K, and O) along MS1 alleles (Berg et al. 2000) to compare the structures of 13 different sperm mutants identified by SP-PCR analysis with the structures of the progenitor alleles (fig. 2). The shorter progenitor allele A of 140 repeats was mapped in its entirety, whereas only the terminal 100 repeats from each end of the larger ~1,000-repeat allele B could be mapped. Comparison of mutant and progenitor alleles showed that most mutants appeared to be derived by intra-allelic rearrangements within the smaller allele. In some cases (e.g., mutant 1; fig. 2), the rearrangement involved a simple and perfect duplication of a block of repeats. In most cases, however, these apparently intra-allelic rearrangements were complex, with multiple imperfect reduplications of blocks of repeats that appeared to be derived solely from the smaller allele (e.g., mutant 4). Four mutants showed possible evidence for interallelic transfer of information, as
Figure 1  Detecting MS1 mutants in genomic DNA. A, Detection of large mutational changes by SP-PCR. Aliquots of blood, buccal, and sperm DNA from the same individual, each containing 50 amplifiable molecules of each allele (140 and ~1,000 repeats), were amplified by PCR, using primers flanking the minisatellite, and the products were analyzed by agarose gel electrophoresis and Southern blot hybridization, with a 32P-labeled MS1 probe. The numbers of repeats gained or lost, relative to the size of the smaller allele, are shown at right. B, Detection of small length changes by SM-PCR. Multiple aliquots of sperm and blood DNA, each containing 0.55 amplifiable molecules of each allele, were amplified and tested for positive reactions, which were then analyzed side by side, as in A. C, SP-PCR analysis of sperm DNA from man 2, who was heterozygous for 104 and 134 repeat MS1 alleles. Each PCR contained 16 amplifiable molecules of each allele.

indicated by rearrangements within the 3′ region of the smaller allele, which consists exclusively of A and C repeats, involving insertion of a block of repeats of unclear origin, including non-A/C repeat types (e.g. mutant 5). The incomplete structure of the larger progenitor allele, however, made it impossible to confirm that these mutation events involved interallelic transfer of repeats.

To investigate large rearrangements in more detail, we analyzed sperm DNA from the Zimbabwean man (man 2) who was heterozygous for two short MS1 alleles, of
104 and 134 repeats. Both alleles in man 2 were MVR mapped in their entirety, revealing very different repeat structures (fig. 3). SP-PCR analysis of 5,000 amplifiable molecules of each progenitor allele yielded 98 sperm mutants with gains or losses of at least four repeats, giving a mean sperm mutation rate of 1.0% per allele, similar to that seen in man 1. MVR analysis of 25 mutants from man 2 showed that 21 appeared to have arisen by sometimes complex intra-allelic rearrangements within one or another allele (fig. 3). The remaining four mutants, however, did show some evidence of interallelic transfer, with short anomalous blocks of repeats within regions
of complex intra-allelic duplication. When the 3’ ends of the two progenitor alleles were aligned, these anomalous blocks matched corresponding blocks in the other allele. Similar, though more definitive, in-register transfers of repeats between end-aligned alleles have been seen during sperm mutation at other minisatellites (Jeffreys et al. 1994).

Intra-allelic rearrangements in both men appeared to
be fairly randomly scattered along the alleles, but putative interallelic transfers were concentrated toward the 3’ end of the repeat array (fig. 4). However, these intra- and interallelic distributions are not significantly different (one-sample runs test and Kolmogorov-Smirnov test, \( P > .05 \)), and the apparent 3’ clustering of interallelic events may be due, in part, to biased ascertainment, particularly in man 1, in whom the restricted repertoire of A and C repeat types near the 3’ end of the allele facilitates the detection of interallelic transfers.

Structural Analysis of Small Mutational Changes in Sperm DNA

The small mutational changes detected in sperm DNA from man 1 (fig. 1B) were further characterized by MVR-PCR analysis of SM-PCR products, after eliminating reactions that contained more than one length class of molecule (fig. 1B). In total, 87 sperm molecules selected at random were analyzed. Comparison of minisatellite structures with the progenitor allele (defined as the allele structure determined from blood DNA) showed that only 22 of the 87 molecules were non-mutant, giving a mutation frequency of 75% for this allele. The 65 mutant molecules carried a total of 81 distinct mutations that altered repeat copy number. All of these mutations were deletions and were highly non-randomly distributed, with 72 located in homogeneous stretches of C-type repeats (fig. 5A) and a further 5 affecting repeat units immediately adjacent to these stretches (results not shown). This clustering in C arrays, which make up only 26% of the allele, is highly significant (\( \chi^2 = 63_{(140)} \), \( P < .001 \)).

The three different C arrays (C₁₉, C₁₁, and C₇) in this allele showed very different levels of instability. The C₁₁ and C₇ arrays each contained five deletion events, which, in all cases, involved a single repeat unit, giving a mutation rate of 6% (5 mutants in 87 molecules) in each of these homogeneous arrays. Instability in the C₁₉ array was far higher, with 62 of the 87 molecules carrying deletions of one to four repeats, giving a mutation rate of 71%. Nine of the molecules showed deletions both in the C₁₉ array and in either the C₁₁ or C₇ array; these double-mutation frequencies are those expected if each class of mutant arises independently in the germline. Similarly, 10 of the 13 large rearrangements characterized in this man (fig. 2) were accompanied by C-array deletions distal to the site of large rearrangement, with 9 of these mapping to the C₁₉ array. The proportion of large mutation molecules carrying C₁₉ deletions (9/13) is not significantly different from the proportion of randomly selected molecules carrying such deletions (62/87) (Fisher’s exact test, two-sided, \( P = 1 \)). This suggests that C-array instability can also occur independently of the larger, more complex rearrangements. It therefore follows that multiple sequential mutations could also occur in the C₁₉ array, perhaps by a process involving

**Figure 4** Distribution of rearrangement breakpoints along alleles. Breakpoints are shown separately for apparently intra-allelic rearrangements (red stars) and for putative interallelic transfers (black stars). The locations of matching regions in the donor allele in interallelic mutants are also indicated (white stars). Locations are defined as the center point of the repeat block involved in the rearrangement. If a mutant showed multiple events, then the central repeat of the entire region involved was taken as the breakpoint location.
the strict loss of a single repeat per mutation event, as is seen in the other C arrays. The observed distribution of length changes in the C_{19} array fits well with this model (fig. 5B) and predicts that an average of 1.3 single-repeat deletion events per sperm cell accumulate in this region. Together with other modes of mutation, the overall level of instability in this allele is extreme, with a combined mutation rate of 80% per sperm cell and perhaps as high as 140%, depending on whether multiple mutations occur within the C_{19} array.

C-array instability in this allele is not atypical. Of the 25 mutants characterized in man 2 (fig. 3), 6 contained single-repeat deletions distal to the site of complex rearrangement. All of these microdeletions were in the 16 mutants derived from allele B, which contains five C arrays ranging from C_{6} to C_{12}; none were in mutants

Figure 5  Hyperdeletion within an MS1 allele. A, Structures of randomly chosen MS1 allele A molecules characterized, by use of SM-PCR, in sperm DNA from man 1. Note that 2 of the 24 nonmutants have lost a repeat unit outside the C arrays (not shown). B, Incidence of deletions in the C_{19} repeat array in this allele. The numbers of single molecules that carry each size of deletion are shown (gray bars), together with the least-squares best-fit Poisson distribution of mutants predicted if each mutation event involves the loss of only a single repeat (blue bars). Under this model, each sperm cell will have accumulated an average of 1.28 mutation events in the C_{19} repeat array.
from allele A, which is devoid of C arrays. All deletions were within the C arrays, with five of the six in the longest C12 array. The deletion rate in C12 is therefore ~31% (5/16), compared with a mean rate of ~1.6% in the other four C arrays. The relationship between C-array length and the frequency of deletion in sperm DNA is shown in figure 6, together with the abundance of C arrays in 13 short MS1 alleles analyzed in human populations.

Discussion

Like other GC-rich minisatelites (Jeffreys et al. 1994; May et al. 1996; Buard et al. 1998, 2000a; Tamaki et al. 1999), MS1 shows a mode of repeat unit turnover that is restricted to the germline and frequently involves complex rearrangements biased toward gains of repeat units and resulting in relatively large changes in allele size. All mutants analyzed have different structures, as predicted for products of meiotic recombination, with no evidence of germinal mosaicism that might signal premeiotic mutation. Most of these large mutations appear to be intra-allelic, involving sometimes complex duplications and deletions of blocks of repeat units. There is, however, evidence in man 2 that some mutation events may involve the transfer of short blocks of repeats between alleles aligned at their 3’ ends, accompanied by complex rearrangements in the recipient allele at the site of insertion. Similar, though more definitive, examples of in-phase interallelic gene conversion have been seen at several other human minisatelites, and, for some but not all loci, represent the major mode of germline mu-

![Figure 6](image)

**Figure 6** Instability of C arrays and their prevalence in MS1 alleles. A, Sperm mutation rates in C arrays deduced from SM-PCR analysis of allele A in man 1, under the assumption that mutation steps always involve the loss of a single C-type repeat. No deletions were seen in the two C2, three C3, and two C4 arrays in this allele, giving a mean mutation rate of <0.005 per array (P > .95). The estimate for C12 arrays was taken from the incidence of deletions in this array in large mutants characterized in man 2 (see fig. 3). B, Abundance of C arrays in MS1 alleles. The distribution was determined from alleles in men 1 and 2 and from 10 additional short alleles (64–155 repeats) from northern European men who were fully typed by MVR-PCR (1,720 repeats typed in total, data not shown). For comparison, the abundance of A-repeat clusters is also shown. The log-linear relationship (gray line) between abundance and short array length (<10 repeats) is the same for C- and A-type repeats.
tation (Jeffreys et al. 1994; Buard et al. 1998). These putative interallelic events at MS1 appear to be targeted toward the 3’ end of the repeat array, reminiscent of a polarized mutation seen at some minisatellites that in one case has been linked to the presence of a flanking recombination hotspot (Jeffreys et al. 1998a). The more frequent intra-allelic rearrangements may also arise by aberrant processing of recombination-initiating events (Biard et al. 1998).

The structure of these putative interallelic events in sperm DNA contrasts with the repeat turnover processes seen when MS1 is integrated near a hotspot for double-strand breaks upstream of LEU2 in S. cerevisiae (Berg et al. 2000). Meiotic instability in yeast frequently results in recombinant arrays consisting of the 5’ end of one allele fused to the 3’ end of the other allele; similar recombinants are also seen at other human minisatellites integrated into yeast at the same position as MS1 (Appelgren et al. 1997; He et al. 1999). However, there are no examples of such recombinants in the 38 sperm mutants characterized in the present study. Similarly, analysis of other human minisatellites has shown that most sperm recombinants involve DNA transfers solely within the repeat array and that relatively few show fused 5’-3’ recombinant structures (Biard and Vergnaud 1994; Jeffreys et al. 1994, 1998a, 1998b; Biard et al. 1998, 2000b; Tamaki et al. 1999). This difference in the meiotic behavior of minisatellites in humans and yeast could reflect either species-specific differences in the way that recombination-initiating events are subsequently processed in tandem repeat DNA or an effect of genome position (e.g., location of the minisatellite with respect to a recombination hotspot).

Unlike all other minisatellites characterized to date, MS1 shows a second novel mode of instability involving repeat unit deletion within sequence-homogeneous regions of alleles. These deletions are targeted to the longest arrays of C-type repeats. The alleles tested for mutation did not contain long homogeneous arrays of other types of repeats—the longest is an A8 array in allele A of man 2—and, although none of the nine mutants characterized from this allele showed mutation within the A8 array, this could still be compatible with a mutation rate as high as 28% (P = .05) (fig. 3). It is therefore unclear whether instability is targeted specifically to C repeats. Indeed, MS1 alleles contain a similar abundance of A and C arrays, declining in frequency with increasing array length (fig. 6B), suggesting that they are subjected to similar population turnover processes and thus that A arrays might also be highly unstable.

These deletion events in C arrays appear to be restricted to the germline and seem to arise independently of large-scale rearrangements. Only 1 of the 13 large rearrangements characterized in man 1 shows a break-
The existence of the MS1 hyperdeletion mechanism creates a paradox concerning the existence of alleles containing long homogeneous arrays. Allele A in man 1 is so unstable that the C_{19} array cannot survive for more than three generations (P < .95) without undergoing deletion, and even a four-repeat deletion (the largest seen in our survey of mutants) would still leave the array prone to further rapid deletion. The length distribution of C arrays in alleles also shows that the C_{19} array is remarkably long (fig. 6B). This leads to the question of how such long arrays can come into existence and persist in human populations (Kovtun and McMurray 2001).

References


Berg I, Cederberg H, Rannug U (2000) Tetrad analysis shows that gene conversion is the major mechanism involved in mutation at the human minisatellite MS1 integrated in Saccharomyces cerevisiae. Genet Res 75:1–12


Acknowledgments

We thank J. Blower, S. B. Kanoyangwa, and volunteers, for providing semen, blood, and buccal samples; S. Mistry, for oligonucleotide synthesis; and our colleagues, for helpful discussions. This work was supported by Swedish Radiation Protection Institute and National Board for Laboratory Animals grants (to U.R.) and by Medical Research Council and Royal Society grants (to A.J.J.).


Maleki S, Cederberg H, Rannug U (1997) Mutations occurring at the human minisatellite MS1 integrated in haploid yeast are similar to MS1 mutations in humans. Mol Gen Genet 254:37–42


