Somatic versus Germline Mutation Processes at Minisatellite CEB1 (D2S90) in Humans and Transgenic Mice

Jérôme Buard,*,† Andrew Collick,*‡ Jane Brown,† and Alec J. Jeffreys*

*Department of Genetics and †Division of Biomedical Services, University of Leicester, Leicester LE1 7RH, United Kingdom

Received December 7, 1999; accepted February 15, 2000

The most variable human minisatellites show extreme germline instability dominated by complex intra-allelic rearrangements plus a lower frequency of inter-allelic transfers of repeat units. In contrast, little is known about somatic instability at such loci. We have therefore used single-molecule PCR to analyze mutation at minisatellite CEB1 (D2S90) in human blood DNA. Somatic mutants were rare and involved only relatively simple intra-allelic events, with no bias toward expansions, in sharp contrast to the complex gain-biased rearrangements seen in sperm. Somatic and germline mutation processes were further analyzed in mice transgenic for a cosmid insert containing CEB1. Mutant molecules in transgenic sperm and blood were detected but only at the low frequencies seen in human blood and arose mainly by simple duplications and deletions as seen for somatic mutations in human. These data suggest distinct pathways for germline and somatic CEB1 mutations with germline instability involving recombination-based repair of meiotic double-strand breaks and somatic mutation arising by replication slippage or mitotic recombination. The problem of transferring germline-specific features of minisatellite instability from human to mouse suggests, with other recent observations, that long-range chromatin conformation may be required for the recombination-based mode of germline instability at human minisatellites.

INTRODUCTION

Some human minisatellites show extreme allele length variability arising from spontaneous gains and losses of repeats in the germline (Jeffreys et al., 1988; Vergnaud et al., 1991). New length mutant alleles can be detected by pedigree analysis or by using PCR to screen small pools of gametic (sperm) DNA (Jeffreys et al., 1994). The mutation process underlying this instability can be investigated by using minisatellite variant repeat mapping by PCR (MVR-PCR; Jeffreys et al., 1991) to compare the interspersion pattern of variant repeats along mutant and progenitor tandem repeat arrays. Complex events including transfers of repeats between allelic partners as well as duplications and deletions have been found for all six unstable GC-rich minisatellites analyzed thus far (Buard and Vergnaud, 1994; Jeffreys et al., 1994; Andreassen and Olaisen, 1999; Tamaki et al., 1999). Although the nature and complexity of the rearrangements observed are strikingly similar between the minisatellites, polarized inter-allelic transfers predominate at minisatellites MS32 and MS31A (Jeffreys et al., 1994) while complex intra-allelic duplications represent the majority of expansions at the most highly unstable minisatellites CEB1 and B6.7 (Buard and Vergnaud, 1994; Tamaki et al., 1999). Inter-allelic cross-talk strongly suggests a meiotic, recombination-based component of minisatellite mutation. This homologous recombination pathway is further substantiated by the existence of a hotspot of recombination in the immediate vicinity of the MS32 tandem repeat array, the activity of which is correlated with minisatellite instability (Jeffreys et al., 1998). However, intra-allelic mutations could be explained by a mitotic and replication-based pathway operating in the germline. At minisatellite CEB1, these intra-allelic rearrangements occur at an average frequency of 10% per sperm, account for 75% of all sperm mutants, and increase in frequency with the size of the tandem repeat array (Buard and Vergnaud, 1994; Buard et al., 1998).

Little is known about somatic mutation processes at minisatellites such as CEB1 that mutate in sperm at high rates and predominantly via an intra-allelic pathway, despite their potential use as biological sensors of genotoxic damage in somatic DNA, a role that they fulfill for germline DNA (Dubrova et al., 1996, 1997). Analysis of CEB1 in CEPH individuals did not reveal any mosaicism resulting from mutation events potentially occurring in vivo in the soma or in vitro during propagation of lymphoblastoid cell lines (Vergnaud et al., 1996).
al., 1991), as have been seen for microsatellites (Weber and Wong, 1993). Furthermore, small pool PCR (SP-PCR) failed to reveal any authentic somatic CEB1 mutant in blood DNA (Buard et al., 1998). These data suggest a profound suppression of CEB1 instability in the soma relative to sperm but give no information on the somatic mutation rate and process. We have therefore used more sensitive PCR approaches to investigate somatic instability at CEB1 and to see whether this minisatellite shows somatic mutation processes similar to those of MS32, a minisatellite with a quantitatively and qualitatively different mode of germline mutation and showing a simple mode of low-frequency mutation in blood DNA involving intra-allelic duplications and deletions (J effreys and Neumann, 1997).

Animal models of minisatellite instability would greatly facilitate the analysis of somatic and germline mutation, including the characterization of potential cis- and trans-acting factors that modulate repeat turnover, the developmental timing of instability during spermatogenesis, and analysis of the poorly understood mutation processes in the female germline. Mice transgenic for human minisatellite MS32 plus extensive human DNA including the flanking recombination hot-spot showed somatic (blood) instability very similar in rate and process to that seen in humans, indicating that somatic turnover, for this minisatellite at least, can be successfully modeled in mice. In contrast, no instability of the minisatellite was detected in the sperm of transgenic homozygous mice despite screening several million gametes (Bois et al., 1997). These findings suggest a profound inter-species barrier to the transfer of germline instability at CEB1 and to see whether this minisatellite shows somatic mutation processes similar to those of MS32, a minisatellite with a quantitatively and qualitatively different mode of germline mutation and showing a simple mode of low-frequency mutation in blood DNA involving intra-allelic duplications and deletions (J effreys and Neumann, 1997).

Quantification of the number of progenitor molecules contained in each fraction was estimated by PCR amplification and comparison of Southern blot hybridization band intensities with reference samples containing known amounts of progenitor molecules, as described previously (J effreys and Neumann, 1997). Subsequently, the three adjacent fractions containing virtually all progenitor molecules were pooled, and the total number of amplifiable progenitor molecules recovered was more accurately estimated by Poisson analysis of single molecule amplifications of the diluted pool (J effreys et al., 1994).

For the detection of mutant molecules, 7-μl SP-PCRs were seeded with aliquots from each pool of fractions enriched for mutants, with each reaction containing approximately 100 residual progenitor molecules. These reactions contained PCR buffer, Taq and Pfu DNA polymerases (20:1 ratio, 0.05 U/μl), and carrier herring sperm DNA as described elsewhere (J effreys et al., 1990; J effreys and Neumann, 1997) plus a 0.5 μM concentrations of primers CEB1-72G (5’ ogagccagcttgtgagggg 3’) and CEB1-P14 (5’ gattttcctgtgcctctcct 3’). Cycling conditions in an M R Tetrad thermocycler or a Perkin–Elmer 9600 thermocycler were 96°C for 45 s followed by 24 cycles of 96°C for 30 s, 66°C for 20 s, 70°C for 4 min for the human sample (1.4- and 2.3-kb CEB1 progenitor alleles) or 6 min for the mouse samples (3.7-kb progenitor allele). Half of each SP-PCR product was electrophoresed through a 40-cm-long 0.7% SeaKem HGT (FMC Bioproducts) agarose gel, transferred onto a nylon membrane by Southern blotting, and hybridized with a radioactively labeled CEB1 probe (Feinberg and Vogelstein, 1983).

**MATERIALS AND METHODS**

**Generation of transgenic lines.** The 39-kb human genomic insert of cosmid 53 (Vergnaud et al., 1991), containing an 88-repeat CEB1 allele plus 13.1 kb of 5’ and 22.1 kb of 3’ flanking human DNA previously sequenced (Murray et al., 1999) (GenBank Accession No. AF048727), was separated from the cosmid vector pWE15 by NotI digestion followed by gel electrophoresis, recovered by electroelution, and purified on a NACS column (Wende et al., 1985). Approximately 300 copies of the NotI fragment were micro-injected by standard techniques into the male pronuclei of one-cell embryos of C57BL/6 × CBA/cA F1 genotype. Embryos surviving to the two-cell stage were transferred into recipient female mice. Mice transgenic for CEB1 were identified by Southern blot analysis of tail DNA. Four founder mice were bred with C57BL/6 × CBA/cA F1 mice, and transgene-positive F2 males were used as hemizygous sperm donors after being crossed with hemizygous females to yield potential homozygotes. Homozygotes were identified by comparison of CEB1 Southern blot signal intensities and by test breeding with wildtype mice. Two to 6-month-old homozygotes and hemizygotes were subsequently used for mutation analysis.

**Mutation analysis.** The preparation of human and mouse sperm and blood DNA and all subsequent manipulations were carried out under conditions designed to minimize the risk of contamination (J effreys et al., 1990, 1994). SP-PCR analysis of CEB1 in sperm and blood DNA was performed on MboI-digested DNA as described previously (Buard et al., 1998). For size enrichment SP-PCR (SESP-PCR) analysis, 5- to 10-μg samples of blood DNA were digested to completion with MboI, which cleaves 193 bp upstream and 16 bp downstream of the CEB1 tandem array. After electrophoresis in a 0.7% SeaKem HGT (FMC Bioproducts) agarose gel, size-fractionation and recovery of the DNA were performed as previously described (J effreys and Neumann, 1997).

**FIG. 1.** Detection of CEB1 mutant molecules in human blood DNA by size-fractionation and small pool PCR. Blood DNA from an individual heterozygous for 52- and 33-repeat CEB1 alleles was digested with MboI and electrophoretically separated into size fractions ranging from 0.93 to 3.45 kb (18 to 81 repeats). SP-PCR was performed on pools of fractions (A–I) largely depleted in progenitor molecules. DNA derived from 35,000 diploid genomes was screened, and 14 abnormal length mutant molecules were detected by electrophoresis of SP-PCR products, blotting, and hybridization. The size range of authentic mutants for each pool is indicated on the right as the number of CEB1 repeats. Two presumptive PCR artifacts that fall outside these ranges are shown with arrowheads.
FIG. 2. Structures of CEB1 mutant molecules recovered from blood and sperm DNA of an individual heterozygous for 52- and 33-repeat alleles (alleles U and L, respectively). Sperm and blood mutants were isolated and structurally characterized by MVR-PCR to determine the order of the variant repeats along the tandem array. Each letter represents a variant repeat unit. Repeat unit blocks lost are represented by dashes, and a duplicated copy of a repeat unit block (underlined) is shown in boldface type. Recombination between two repeats can create novel hybrid repeats; these are indicated in blue. Unusual repeats that could have arisen by base substitution at one of the three variant base positions within CEB1 repeats are shown in pink. 

(A) Structure of 25 mutants detected in blood DNA and derived from 140,000 amplifiable progenitor molecules. All events could be interpreted as a simple duplication or deletion, with most being perfect and only a minority showing additional minor rearrangements in the duplicated segment. There were two instances of repeated isolation of the same mutant (BL8 and BL9; BL10, BL11, and BL12). 

(B) Structure of 37 sperm mutants from the same individual. Fourteen expansions out of 33 could not be interpreted as the origin of the supernumerary block of repeats (black). Among the remaining 18 interpretable gains, 5 events are associated with an interallelic transfer (green block within red allele), 11 events showed complex duplications, and 2 involved simple duplications.
Mutant molecules were recovered by two rounds of amplification and gel purification. A PCR with 0.5 μM concentrations of primers CEB1-72G and CEB1-P14 was seeded with 0.5 μl of the SP-PCR containing the mutant molecule and cycled for 5 cycles at 96°C for 30 s, 66°C for 20 s, and 70°C for 4 or 6 min. PCR products were electrophoresed, and mutant molecules (not detectable by ethidium bromide staining) were recovered by excision from the gel followed by electrophoresis. A second PCR with a 0.5 μM concentration of nested primers CEB1-4A (5‘ ggcaggagctctgctgaggtg 3’) and CEB1-P14 was seeded with 1 μl of the eluted product and cycled for 18 to 27 cycles, depending on the mutant molecule size and signal intensity on the initial Southern blot, at 96°C for 30 s, 66°C for 20 s, and 70°C for 4 or 6 min. Mutant PCR products, now detectable by ethidium bromide staining, were electrophoresed in a 40-cm 0.8% SeaKem HGT (FMC BioProducts) agarose gel, excised from the gel, and recovered by electrophoresis. MVR-PCR was performed as described elsewhere (Buard and Vergnaud, 1994) except for the slight following modifications. Primer CEB1-4A was used as flanking primer, and a fourth nucleotide variation between repeats, highly informative in the CEB1 88 allele transferred in mice and also found in a number of other unstable alleles in sperm, was typed by MVR-PCR using MVR-specific primers CEB1-v11A (5‘ tcatgcgctcgctcggaggtgcctgcrgaggtc 3’) and CEB1-v11G (5‘ tcatgcgctcgctcggaggtgcctgcrgaggtc 3’), each at a final concentration of 5 nM.

RESULTS

Detection of CEB1 Mutant Molecules in Human Blood

SP-PCR can analyze only up to 100 progenitor CEB1 molecules per reaction. We therefore used a size-fractionation method on genomic DNA that allows mutant molecules to be enriched prior to SP-PCR and to be validated by size (Jeffreys and Neumann, 1997). Blood DNA from a 45-year-old man heterozygous for 52- and 33-repeat CEB1 alleles, with sperm mutation rates of 7 and 5%, respectively, was digested with Mbol, which cleaves near the repeat array, and fractionated by agarose gel electrophoresis. Twenty-four size fractions extending across both progenitor alleles and including DNA fragments containing 18 to 81 repeats were recovered and assayed by PCR for the presence of progenitor molecules. Each progenitor allele was almost entirely confined to three adjacent fractions (data not shown). The number of amplifiable progenitor molecules recovered was estimated by Poisson analysis of single-molecule amplifications of limiting dilutions of pools of these progenitor-rich fractions (Jeffreys et al., 1994; Buard et al., 1998; Tamaki et al., 1999). Other fractions showed only trace amounts of each allele, indicating at least a 100-fold depletion in progenitor molecule. Pools of these fractions were then surveyed for mutants by SP-PCR. Figure 1 shows the result of a survey of selected DNA fractions equivalent to 70,000 haploid genomes that revealed 14 abnormal length mutant molecules, 12 of which had the correct length for the size range of the fraction in which they were found. The two remaining PCR products were smaller than predicted from the fraction size and were therefore presumably PCR artifacts. Repeating this experiment yielded 13 additional mutants plus one possible PCR artifact. The mean frequency of blood mutants detected per progenitor molecule was therefore 1.8 × 10⁻⁴ (25 mutants out of 140,000 progenitor molecules). This will be an underestimate as gains and losses of fewer than 3 repeats for each allele could not be scored because they were contained in the progenitor-rich fractions. However, such small length changes are unlikely to be common given that none has been detected in 4000 progenitor molecules screened by small-pool PCR (data not shown).

Structural Analysis of Somatic and Germline Rearrangements at CEB1

The structural basis of CEB1 mutation was characterized by mapping the location of minisatellite variant repeats along the 25 blood mutant alleles and comparing their internal structures with those of 37 sperm mutants recovered from the same individual by SP-PCR as previously described (Buard et al., 1998). Alignment of mutant allele structures with the large and small progenitor alleles (U and L, respectively) enabled each mutant to be assigned to one or another progenitor and showed that all blood mutation events involved only intra-allelic duplication or deletion (Fig. 2A). Most events were simple, with 5 of 10 different expansions being perfect duplications, 2 showing a minor alteration of the duplicated copy (one repeat deletion in the duplicated portion; mutants BU10 and BU11), and the remaining 3 events being more complex, but still interpretable, duplications. In contrast, and as previously described for other CEB1 alleles, the sperm mutation process is characterized by extreme complexity (Fig. 2B). Fourteen sperm expansions out of 33 were not interpretable, only 2 simple duplications were seen among 13 clear intra-allelic expansions (mutants SL3 and SL9), and at least one fourth of interpretable expansions (5/18) involved inter-allelic transfer of repeats from allele U into allele L. All sperm mutants were different. In contrast, some blood mutants had identical structures (mutants BL8 and BL9; mutants BL10, BL11, and BL12), indicating low-level mosaicism for at least some of these somatic mutants. Eliminating these repeat isolates gave a frequency of different mutants of 1.8 × 10⁻⁴ and 1.3 × 10⁻⁴ per progenitor molecule for alleles U and L, respectively. These frequencies are not significantly different (P > 0.05). As seen previously, sperm mutation is heavily biased toward gains, with 24 expansions of allele L seen among 26 mutants of size intermediate between the two progenitor alleles (35 to 50 repeat units). In contrast, there were only 10 increases among 22 different blood mutants, and only 6/18 different somatic mutants that fall within the same 35- to 50-repeat window were expansions of allele L. The proportion of gain mutants is significantly higher in sperm than in blood (χ² = 17.28; P ≪ 0.0001), and these data suggest a bias toward deletions for somatic mutation.
CEB1 Mutation Processes in Sperm and Blood of Transgenic Mice

A cosmid insert containing an 88-repeat CEB1 allele plus a 22-kb 5’ flanking human DNA and a 13-kb 3’ human DNA previously isolated (Vergnaud et al., 1991) and sequenced (Murray et al., 1999) was injected into the male pronucleus of fertilized mice eggs. Four transgenic lines (G–J) were produced, and the integrity of the transgene in each line was verified by long-PCR and restriction digestion analyzes (data not shown).

Dosage estimation from Southern blots and single-molecule PCR amplifications of CEB1 from transgenic DNA showed that transgene copy number integrated in each line varied from about 3 for line H to approximately 30 for line I (Table 1). Segregation analysis of CEB1 in pedigrees showed that line G carried an integration on the Y chromosome (data not shown). Among the three lines with autosomal integration, two could be bred to homozygosity (line H and line I). Southern blot analysis did not reveal any new length CEB1 alleles among 40–60 offspring for each of the four lines. SP-PCR screening of more than 1000 sperm from hemizygous and homozygous transgenic males showed similarly no evidence for authentic mutants, suggesting a very low germline mutation rate (0.3% per sperm).

We therefore applied the size-enrichment procedure described above to sperm and blood DNA from transgenic mice. On average, 135,000 amplifiable molecules were screened for expansion and contraction of at least four repeats in sperm and blood from hemizygous and homozygous transgenic males. Mutant molecules were detected in each size range selected, with a similar number of gains and losses (Fig. 3). Mutant frequencies per progenitor molecule (i.e., per amplifiable copy of the transgene) were similar between sperm and blood from the same transgenic line, between lines with high and low transgene copy number, and between homozygotes and hemizygotes (Table 1). These frequencies, close to $2 \times 10^{-4}$, were very similar to that estimated in human blood DNA ($1.8 \times 10^{-4}$, not corrected for mosaicism). This strongly suggests that CEB1 somatic instability has been successfully transferred to the transgenic mice and raises the possibility that the mutant molecules observed in transgenic sperm result from the same somatic mutation process as seen in blood.

These issues were further investigated by analyzing

---

**TABLE 1**

Summary of Sperm and Blood Mutation Frequencies of CEB1 in Humans and Transgenic Mice

<table>
<thead>
<tr>
<th>No. repeats</th>
<th>Copy No.</th>
<th>Frequency of mutants ($\times 10^4$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sperm Hemizygous</td>
</tr>
<tr>
<td>Human</td>
<td>Allele L</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Allele U</td>
<td>52</td>
</tr>
<tr>
<td>Transgenic mice</td>
<td>Line G</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Line H</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Line I</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Line J</td>
<td>88</td>
</tr>
</tbody>
</table>

Note. The frequency of mutants involving the gain or loss of 3–30 repeats, a range fully scorable in all analyses, is given per amplifiable progenitor molecule. For each mutant frequency estimated by size enrichment and small pool PCR, the number of mutants scored and the estimated number of amplifiable molecules screened are given in parentheses. In all cases, the total number of progenitor molecules has been estimated by Poisson analysis of single-molecule amplifications from limiting dilutions of the total DNA (SP-PCR) or of the pools containing progenitor-rich fractions (SESP-PCR). Sperm mutation rates in human and in transgenic line J were estimated by SP-PCR. Mutant frequencies are corrected by elimination of repeat isolates of the same mutant; such mosaicism has been seen only in human blood.

---

**FIG. 3.** Detection of CEB1 mutant molecules in sperm of a line I homozygous transgenic mouse. After MboI digestion, electrophoresis, and fractionation of transgenic sperm DNA, pools of fractions depleted in the 88-repeat progenitor allele (pools A–D) were surveyed by SP-PCR. The size range corresponding to each pool is indicated at the right as the number of CEB1 repeats and includes mutant molecules ranging from 65 to 130 repeats. Screening of DNA fractions equivalent to 200,000 progenitor molecules revealed 38 abnormal-length CEB1 molecules, including 3 molecules falling outside the size range screened and representing presumptive PCR artifacts (arrows).
the structural basis of CEB1 instability in transgenic mice. MVR-PCR was used to characterize mutant allele structures in sperm and blood DNA from a transgenic mouse homozygous for CEB1 (Fig. 4). Most duplication and deletion events were perfect, both in transgenic sperm and in blood. Only two of the six sperm expansions showed evidence of greater complexity; mutant Isp5 showed a deletion of 15 contiguous repeats within the duplicated region, and mutant Isp6 showed evidence of a complex multistep rearrangement involving duplication of a block of repeats near the beginning of the array and insertion of this block near the end of the array. None of the sperm mutants showed evidence of the very complex and uninterpretable rearrangements seen in human sperm (Fig. 2B). Similarly, all 15 deletion events were simple except for one sperm rearrangement that involved the loss of two separate blocks of repeats.

**FIG. 4.** Structure of sperm and blood CEB1 mutant molecules in a line I transgenic mouse. The 88-repeat CEB1 allele transferred into mice contains 14 different variant repeat types distinguishable by MVR-PCR and represented as boxes with different colors and patterns. For each rearrangement, the number of repeats gained or lost is indicated, and the arrows suggest the most plausible interpretation for each duplication event. Sperm (Isp) and blood (Ibl) mutants detected in the same transgenic mouse reveal simple rearrangements in both cell types, with most expansions involving perfect duplications (4/6 in sperm, 5/5 in blood). Deletions are represented by an empty box and are all simple except for one sperm rearrangement (Isp12) that involves the loss of two separate blocks of repeats.
with the complexity of the human germline process, further points toward a baseline mitotic mutation process in the male germline of transgenic mice.

DISCUSSION

A General Process for Somatic Mutation at Human Minisatellites: Rare and Simple

We have characterized the level and structural basis of somatic instability of a human minisatellite that shows very high level of instability in sperm arising mainly by complex intra-allelic rearrangements. We show here, by direct comparison of blood and sperm instability of two CEB1 alleles from the same individual, that the frequency of somatic mutants is at least 200-fold lower than the germline mutation rate. This profound suppression of instability in blood relative to sperm strongly suggests that virtually all CEB1 germline rearrangements, including intra-allelic duplications and deletions as well as inter-allelic transfers, are initiated by events occurring specifically during gametogenesis and most likely at meiosis. In contrast to sperm mutation, somatic instability involves only simple intra-allelic duplications and deletions, with no evidence of interactions between alleles.

The somatic mutation process at minisatellite MS32 also involves rare and simple rearrangements (Jeffreys and Neumann, 1997), suggesting that these features are general characteristics of minisatellite somatic instability. This provides further evidence of two completely distinct mutation mechanisms acting in sperm and in blood, most likely involving different initiating events. Meiosis-specific double-strand breaks (DSBs), possibly formed by staggered nicks within the tandem array (Buard and Vergnaud, 1994), probably initiate germline instability. Single-strand annealing (SSA) between the two 3' protruding single-strand ends, followed by synthesis, destabilization of the heteroduplex, and reiterations of this annealing/synthesis process prior to ligation could produce complex intra-allelic duplications, as suggested by our observations and by recent studies in yeast (Pâques et al., 1998). Strand invasion of this "flapping" 3' end into the allelic partner could similarly initiate inter-allelic rearrangements. In contrast, the relative simplicity of most minisatellite rearrangements in blood, together with the recent observation of increased instability of a minisatellite-like element in Saccharomyces cerevisiae strains defective in correct maturation of Okazaki fragments or in DNA polymerization (Kokoska et al., 1998, 1999), suggests a replication slippage-based mechanism for somatic mutation. Alternatively, mitotic DSBs occurring at low frequency and repaired by SSA and synthesis or by sister-chromatid invasion and synthesis, could also lead to intra-allelic duplications and might account for the occasional complex expansion observed in blood DNA. A number of DNA intermediates predicted above, including meiosis-specific broken molecules, remain to be characterized, and a more tractable experimental model of mammalian minisatellite instability could address this issue.

Transfer of Somatic but Not Germline Instability of CEB1 in Transgenic Mice

We have created transgenic mice containing CEB1, and we have investigated the minisatellite mutation process in the sperm and blood cells of these mice. Somatic mutation rates and the simple nature of rearrangements in blood are similar between transgenic mice and humans. In contrast, the high level of instability and the complexity of rearrangements observed in human sperm are not seen in the germline of transgenic mice. Instead, the low rate and simple process of CEB1 instability in transgenic sperm are virtually identical to the somatic mutation rate and process. This is consistent with mutations arising during the proliferation of mitotically dividing spermatogetic stem cells. Curiously, mice transgenic for human minisatellite MS32 show no trace of germline instability, even at the low rate seen in somatic DNA (Bois et al., 1997). The reasons for this discrepancy between CEB1 and MS32 tandem array instability in the male germline of mice remain obscure.

The frequency of CEB1 blood mutants per progenitor molecule is similar between transgenic lines with high and low transgene copy number. This indicates that the number and arrangement of copies of this transgene, as well as the location of the integrant, do not influence minisatellite instability within each copy and further suggests that potential inter-transgene rearrangements, as seen in multicopy transgene constructs containing human minisatellite MS32 plus limited flanking human DNA (Collick et al., 1994), are unlikely to be a major source of mutation. The fact that germline mutation rates are similar between hemizygous and homozygous mice for the transgene suggests that inter-allelic cross-talk is not a major component of CEB1 instability in sperm of homozygous transgenic mice. Finally, the frequency and simple nature of somatic rearrangements seen in both transgenic and human blood strongly indicate that the somatic mutation process of CEB1 has been fully transferred in the transgenic mice. This transgenic model of CEB1 somatic instability could potentially provide a biological tool for monitoring the genotoxic effects of ionizing radiation and chemical compounds on somatic DNA.

Several explanations, most of them discussed previously in the MS32 transgenic study (Bois et al., 1997), could account for the failure to transfer the germline-specific features of CEB1 instability from human to mouse. Although the mutational competence of the 88-repeat CEB1 allele transferred into mice has not been established in human, it seems very unlikely that this particular allele would not be unstable in human sperm, given that none of the 40 CEB1 alleles larger than 30 repeats previously surveyed shows a sperm
mutation rate below 5% (Buard et al., 1998). It is also possible that homozygosity in these transgenic mice reduces the mutation rate. Given the rarity of men homozygous for identical CEB1 alleles, it is not possible in humans to test whether repeat turnover is driven by allele mismatches in heterozygotes. However, homozygosity at the much more stable insulin minisatellite, which shows a complex germ-line-specific mutation pathway similar to that seen at CEB1, has no significant influence on sperm mutation rate or process (Stead and J effreys, 2000). This suggests that transgene instability is unlikely to be substantially affected by homozygosity for these CEB1 constructs.

Recent findings, in human and in yeast, show that germ-line minisatellite instability is tightly correlated with meiotic recombination activity in the flanking DNA (Appelgren et al., 1997; J effreys et al., 1998). While cis-acting elements are important in modulating minisatellite instability (Monckton et al., 1994; He et al., 1999), other studies indicate that the major factors responsible for this activity do not seem to lie in the primary nucleotide sequence surrounding the tandem array (Murray et al., 1999). This strongly suggests that, as for hot-spots of recombination in yeast, a main factor allowing recombination activity is a long-range chromatin effect (Ohta et al., 1994; Nicolas, 1998), possibly consisting here of an open domain that includes the minisatellite element and the putative cis-acting elements involved in initiating mutation. Interestingly, both CEB1 and MS32 show a very similar ratio of sperm:blood mutation rates (200- to 400-fold), despite an order-of-magnitude difference in germline and somatic instability between these two loci. This raises the possibility that the open chromatin domain is constitutive rather than meiosis-specific, with CEB1 being associated with a more open domain providing greater accessibility both to the meiotic recombinational machinery and to the presumed replication/repair systems involved in somatic instability. It remains to be seen whether somatic instability and germ-line instability share other features in common, for example, similar initiating events such as double-strand breaks that are differentially processed in germ-line and soma.

No true endogenous minisatellites showing high levels of germ-line-specific instability have yet been found in the mouse genome (Bois et al., 1998a) but only highly expanded simple tandem repeat arrays that show not only substantial germ-line instability but also somatic instability, particularly during early development (Kelly et al., 1989; Bois et al., 1998b). This could suggest a profound difference in the way the two species process tandem repeats, possibly reflecting the striking difference between the clustering of human minisatellites in the recombination-proficient subtelomeric regions of chromosomes compared with the more random distribution of minisatellites along mouse chromosomes. It remains to be seen whether true unstable minisatellites exist in the mouse genome in association with regions of high meiotic recombination activity. Alternatively, it may be possible to create mouse models of human germ-line minisatellite instability by targeting human repeat arrays to mouse recombination hot-spots.

ACKNOWLEDGMENTS

We thank Biomedical Services, Leicester, for the excellent care given to the mice involved in this study. We also are grateful to colleagues for helpful discussion and to the sperm and blood donor. This work was supported in part by an International Research Scholars Award to A.J.J. from the Howard Hughes Medical Institute and in part by grants from the Wellcome Trust, Medical Research Council, European Union, and Royal Society. J.B. is a MRC research fellow.

REFERENCES


