The EMBO Journal vol.13 no.23 pp.5745–5753, 1994

Variable germline and embryonic instability of the human minisatellite MS32 (D1S8) in transgenic mice

Andrew Collick[1], Michael L.Norris[2], Maxine J.Allen, Philippe Bois, Sheila C.Barton[3], M.Azim Surani[3] and Alec J.Jeffreys

Department of Genetics, University of Leicester, Leicester LE1 7RH, [2]Department of Development and Signalling, Babraham Institute, Babraham, Cambridge CB2 4AT and [3]Wellcome CRC Institute of Developmental Biology, Cambridge CB2 1QR, UK

[1]Corresponding author

Communicated by A.J.Jeffreys

Tandem repeat loci such as minisatellites and trinucleotide repeats frequently show instability. We have investigated mutation at human minisatellite MS32 (locus D1S8) transferred to transgenic mice. Three lines of hemizygous transgenic mice were studied. A single-copy line (110D) was seen to be relatively stable, whilst two multicopy lines showed structural instability of the transgene in pedigrees (lines 109 and 110A). For both these lines, mutant structures were detected as a result of mutation events having occurred in the germline or early embryo. Structural changes seen included gain or loss of minisatellite repeat units (110A and 109), alteration of DNA flanking the minisatellite repeat array (109 only) or deletion of the entire transgene (109 only). This work demonstrates that tandem repeat transgenes can show instability and thus provide additional systems for the analysis of repetitive DNA structural change in mice.

Key words: embryo/germline/instability/minisatellite/transgene

Introduction

Minisatellites are tandemly repeated DNA sequences commonly found in many higher eukaryotic species. They are present in ~2000 positions in the mouse and human genome and consist of arrays of 4–90 bp repeat units which can extend for lengths of ≥20 kb (Jeffreys *et al.*, 1985). The most notable feature of minisatellites is their highly unstable nature, caused by the gain or loss of repeat units in the array: germline mutation rates can lie in the range 1–15% per gamete for the most hypervariable loci (Wong *et al.*, 1987; Jeffreys *et al.*, 1988; Vergnaud *et al.*, 1991). This instability creates structural variability which has been successfully used in the areas of forensic casework, kinship determination and linkage analysis.

One outstanding question is the cause of minisatellite variability. It has been argued that minisatellites are recombination hot-spots, as suggested by the similarity of the minisatellite core consensus sequence to *chi*, a well-defined prokaryotic recombination initiator (Jeffreys *et al.*, 1985). Better evidence of a link between minisatellites and recombination is the clustering of minisatellites in the proterminal recombinationally active regions of the human genome (Royle *et al.*, 1988), the presence of minisatellite sequences at or near chiasmata (Chandley and Mitchell, 1988) and directly from the behaviour of minisatellite-containing DNA on reintroduction into mammalian cells (Wahls *et al.*, 1991). Thus, by analysing germline mutation events that change the structure of minisatellite loci it may be possible to explore previously unstudied parts of eukaryotic DNA recombination systems.

The best-studied human minisatellite is MS32 (D1S8). It has a repeat unit sequence of 29 bp (see Figure 1) and a germline mutation rate of ~1% per gamete. The latter has been shown in two ways. The first is pedigree analysis, looking for variations in minisatellite structure detectable by Southern blot analysis after each germline passage (Wong *et al.*, 1987). The second approach is to detect mutant alleles directly in germline (sperm) DNA by single molecule PCR (Jeffreys *et al.*, 1994). The ability to define 'internal structures' of MS32 alleles makes it a powerful experimental system for studying minisatellite biology. Such structures are based upon the natural interspersion of polymorphic repeat unit types throughout the repeat array and their PCR-based detection (minisatellite variant repeat unit PCR or MVR-PCR) (Jeffreys *et al.*, 1991). MVR-PCR analysis has revealed that the changes that occur upon germline mutation are largely restricted to one end of the MS32 array (Jeffreys *et al.*, 1991, 1994). This property of polarity is also seen with two other human minisatellites, MS31 (D7S21) and MS205 (D16S309) (Armour *et al.*, 1993; Neil and Jeffreys, 1993; Jeffreys *et al.*, 1994; see also Buard and Vergnaud, 1994 for an example of more varied mutation events). In addition, MVR-PCR analysis has shown that mutation frequently involves the transfer of repeat units from one allele to another via a complex gene conversion process (Jeffreys *et al.*, 1994). One working model is that minisatellite mutation is driven by *cis*-acting elements which lie adjacent to the repeat array and introduce double-strand breaks into the 5′ end of the repeat unit array (Monckton *et al.*, 1994).

To investigate further minisatellite mutation, and in particular to attempt to define those flanking DNA sequence elements required for instability, we have created transminisatellitic mice transgenic for the human minisatellite MS32. These mice were made by standard microinjection of a PCR product of a human MS32 allele containing 71 repeat units. In addition, a small amount of human flanking DNA (5′ and 3′) was included. If such flanking regions contain all elements required for instability then polar mutation events may be detectable in such transgenic mice. We have studied in detail the structures resulting from eight integration events, including the

isolation of flanking mouse DNA sequences (Allen *et al.*, 1994). Interestingly, five of these transgenes had integrated into mouse gamma-satellite DNA, which is the major centromeric satellite of the mouse (Allen *et al.*, 1994). The breeding of three out of four founder mice gave rise to six separate lines of transgenic mice, two containing multicopy integrations (109 and 110A) and one containing a single-copy integration (110D). The remaining three had rearranged single-copy or partial-copy structures (Allen *et al.*, 1994) and were thought inappropriate for use in this kind of study.

In this report we show that hemizygous multicopy MS32 transgenes can be highly unstable in mouse pedigrees. In addition to high rates of germline mutation (something which was hoped for at the outset, but not certain by any means), both transgenes (109 and 110A) show other, more unusual classes of instability, including early embryonic mutation (109 and 110A), the frequent loss of a single restriction enzyme site outside the tandem repeat array (109) and the common deletion of the entire multicopy transgene (109). In contrast, the single-copy transgene studied (110D) appears to be relatively stable. In summary, our first attempts to model human minisatellite mutation in transgenic mice have led to the creation of novel systems for exploring processes of repetitive DNA instability.

Results

Detection of mutated transminisatellites

Three MS32 transminisatellites were chosen for mutation analysis, the single-copy transgene 110D and the multicopy transgenes 109 and 110A. The structures of these loci are shown in Figure 1. Mice hemizygous for each transminisatellite 109, 110A and 110D were bred and their progeny tested by Southern blot analysis for any changes in the restriction fragment profile indicating transminisatellite mutation events. For mutation analysis the most informative combination of restriction enzymes for each transminisatellite was used. Enzymes *Hin*dIII and *Ssp*I cut the multicopy locus 109 into five DNA fragments, all but one containing a single tandem repeat array corresponding to the progenitor human allele (see Figure 1). *Pst*I cuts locus 110A into five fragment sizes, four of which are single-copy (two flanking fragments, two variant internal fragments) and the fifth being multicopy and derived from the internal copies of the original allele (see Figure 1). For line 110D, the enzyme *Eco*RI releases the single-copy locus on one restriction fragment.

In the pedigrees of both multicopy lines 109 and 110A, mice containing structurally altered transminisatellites were detected (see Table I). Five to ten per cent of transgene-positive progeny showed evidence of mutation. In contrast, the single-copy locus 110D showed no detectable mutation events in any of the 174 transminisatellite-containing progeny studied. Interestingly, mutation at transminisatellite 110A was seen to be clustered in parts of the pedigree. For example, mutation was seen only in the progeny of hemizygous mice mated with wild-type mice and not as a result of crossing hemizygous parents (see Table I).

Germline and embryonic mutation

Mutations were detected as an alteration in size or loss of one or more fragments in the Southern blot profiles of transminisatellite-positive progeny. For example, whilst line 109 mouse 623 (Figure 2A) has a normal Southern blot profile, that of mouse 622 is altered, with the loss of 2.5 and 6.2 kb DNA fragments and a gain of a new fragment of 8.7 kb. No trace of the original fragments remain, consistent with mutation of the transminisatellite having taken place in the parental germline. In contrast, mutants 701 and 707 show a new fragment around 8.7 kb with no loss of any parental fragments. However, the fragments of 2.5 and 6.2 kb are now reduced in intensity, as seen by comparing the profile with the normal pattern shown by mouse 623. These data are consistent with these being similar mutations to that seen in mouse 622 but having taken place after fertilization, generating a mouse mosaic for the normal and mutant transminisatellite types.

Line 109 mouse 625 appears to be similar in that all fragments of the original *Hin*dIII–*Ssp*I profile remain, but that new ones are gained, the most prominent being of 5.8 and not of 8.7 kb (Figure 2A), indicating that this post-zygotic mutation event has altered the structure of the transminisatellite in a different way to that seen in mice 622, 701 and 707. Restriction analysis of DNA from mouse 625 with *Ssp*I alone, which releases the transminisatellite as a single DNA fragment, confirms that this mouse is mosaic for the standard 109 locus 17 kb long plus a mutant form 12.5 kb long (data not shown).

Transminisatellite 110A also displayed evidence for both germline and embryonic mutation events, a complex example of each being shown in Figure 2B. Line 110A mouse 295 contains fragments of the expected sizes at 9.9, 3.9, 2.6, 2.5 and 2.4 kb plus two non-standard fragments of 10.1 and 3.5 kb, suggesting mosaicism. Mouse 837 inherited a mutant version of the 110A transminisatellite which contains the 3.5 kb fragment; the intensity of this fragment increases on transmission, again consistent with mosaicism in the mother. (Note that the 3.9 kb fragment is unchanged in mouse 837, indicating that new embryonic mutant fragments are not necessarily derived from those fragments to which they lie closest in the Southern blot profile of a mosiac parent.) The 10.1 kb mutant fragment in mouse 295 is not transmitted to 837, suggesting that the mother is a triple mosaic, containing the normal allele plus the mutant version transmitted to 837 plus a second mutant form containing the 10.1 kb fragment. In contrast, mouse 836 contains a greatly changed profile. The 5′ flanking 9.9 kb fragment has been reduced in size to 6.7 kb and the two variant internal fragments of 2.4 and 2.6 kb have been deleted. As no trace of the original parental fragments of 9.9, 2.6 and 2.4 kb were detected in mouse 836 and no trace of the 6.7 kb mutant fragment was visible in mouse 295, this mutation was scored as being germline in origin.

Breeding analysis of mice containing post-zygotic mutations

Mice containing a mixture of two transgene types, normal and mutant, make up about half of the total mutations detected at 109 and 110A (see Table I). The presence of high dosages of the newly changed type (see Figure 2) suggests that these mutation events can take place during early embryonic development. This was confirmed by the analysis of progeny from five such mice; germline transmission of both standard and mutant forms was

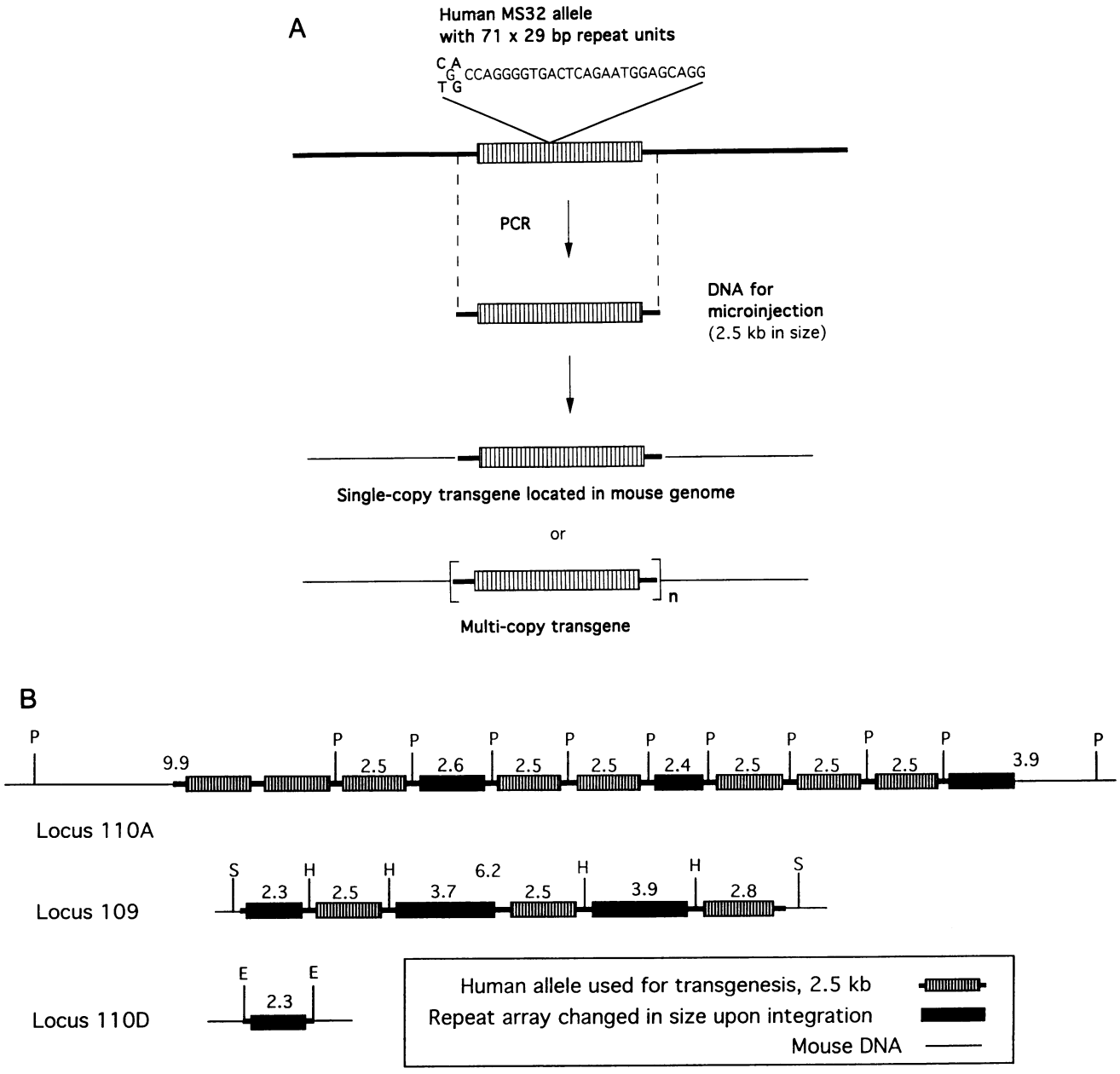


Fig. 1. (**A**) Generation of 'transminisatellitic' mice. A human allele containing an informative mixture of 71×29 bp repeat units was amplified by PCR. Subsequent MVR-PCR allows mutation characterization via the detection of polymorphisms present at the first (C or T) and third (A or G) position of each repeat unit. This amplified allele also contained 212 and 214 bp of 5′ and 3′ human flanking DNA respectively. Thus those regions lying close to the minisatellite which might induce mutation at the minisatellite may be included in this DNA. The amplified allele was used to make transgenic mice (see Allen *et al.*, 1994 for full details). This DNA may integrate to give single-copy or multicopy repeats of the tandemly repeated transgene, as shown. (**B**) Structures of the transgenes studied. The single-copy transminisatellite locus 110D contains a single array of 63×29 bp minisatellite repeat units bordered by human flanking DNA. The multicopy loci 109 and 110A contain six and eleven repeat array units respectively. Those arrays with a size of 2.5 kb were assumed to contain 71 repeat units (shaded grey). Consensus MVR-PCR maps for both transgenes confirm this (data not shown). Those regions changed on integration are shaded black. We were able to show via PCR analysis that non-standard arrays are due to the gain or loss of minisatellite repeat units, most likely occurring at integration (data not shown). Those restriction sites used for mutation analysis are indicated (E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Ssp*I), as are the sizes of minisatellite-containing fragments expected from a non-mutant transgenic mouse. Other sites are present but have been omitted for clarity. The mouse DNA 5′ to loci 109 and 110D is gamma-satellite and that of 110A shows no homology to any sequence present in the EMBL database (Allen *et al.*, 1994).

observed in several cases (Figure 2B and Table II), indicating that these mutations must have occurred early in the embryo, before the separation of germline from soma (days 5–6; reviewed by MacLaren, 1991). Both loci 109 and 110A show early embryonic mutation, despite the fact that the 5′ flanking mouse DNA of 109 is gamma-satellite and that of 110A is apparently single-copy. Thus our limited data suggests that early embryonic mutation may be position-independent for multicopy transminisatellite loci of this type.

Mitotic or meiotic events in the germline?

Data from the analysis of minisatellite MS32 has suggested that in humans the major mutation process is one of gene

Table I. Mutation rates at MS32 transminisatellites estimated by pedigree analysis

Locus	Mating	Positive progeny analysed	Mutants detected	Germline	Embryonic
109	hemi×wt	129	12	6	6
109	hemi×hemi[a]	74	8	4	4
109 total		203	20	10	10
110A	hemi×wt	69	8	3	5
110A	hemi×hemi[a]	96	0	0	0
110A total		165	8	3	5
110D	hemi×wt	94	0	0	0
110D	hemi×hemi[a]	80	0	0	0
110D total		174	0	0	0

Mutations were identified by Southern blot analysis of progeny tail DNA with probe MS32, following restriction analysis with *Hin*dIII plus *Ssp*I (109), *Pst*I (110A) or *Eco*RI (110D). Presumptive germline mutations shared by more than one sibling were presumed to reflect gonadal mosaicism and were scored as a single mutation. This was observed in three pairs of siblings. Hemi, hemizygous; wt, wild-type.
[a]Of hemizygous×hemizygous matings it is expected that 1/3 of transgene-positive progeny mice of line 110A should be homozygous. For line 109 mice, homozygous progeny can be obtained, but at a lower frequency due to reduced parental dosage. For line 110D we have been unable to breed mice homozygous for this locus, thus all 110D progeny should be regarded as hemizygous.

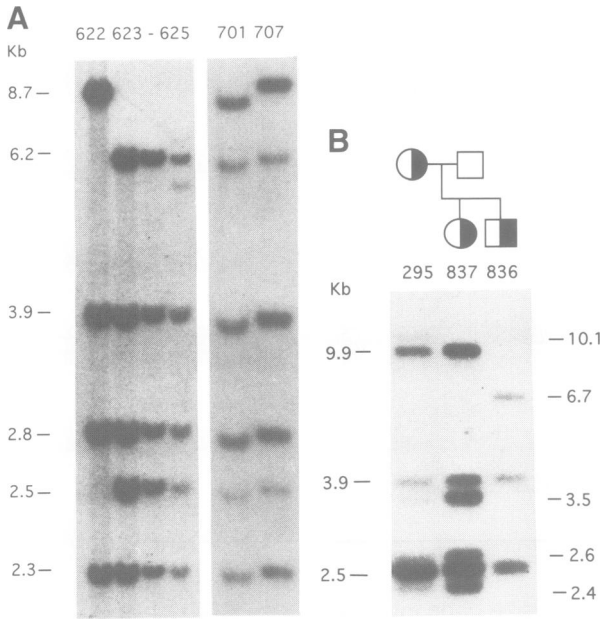


Fig. 2. Examples of germline and somatic mutation detected at multicopy MS32 transminisatellites. (**A**) Analysis of tail DNA samples of line 109 mice showing structural changes. DNA was digested with *Hin*dIII plus *Ssp*I. Mouse 623 contains a normal transgene 109 profile of five fragments of 2.3, 2.5, 2.8, 3.9 and 6.2 kb (see Figure 1) (as does mouse 624 indicated by −). Mice 622, 625, 701 and 707 have changed profiles, the result of transminisatellite mutation (see main text). (**B**) Analysis of tail DNA samples of line 110A mice showing structural changes. Restriction digestion was with *Pst*I. A normal 110A profile is fragments of 2.4, 2.5, 2.6, 3.9 and 9.9 kb (see Figure 1). Mouse 295 contains two non-standard fragments of 3.5 and 10.1 kb. The 3.5 kb fragment was inherited by progeny 837. The increased dosage of the 3.5 kb fragment from mother to daughter, plus breeding data (see Table II), indicate this that this new fragment is the result of embryonic mutation occurring within mouse 295. Mouse 836 inherited a germline mutation from his mother (295) in which the 9.9 kb fragment has been reduced in size to 6.7 kb and other repeat arrays deleted, including the two internal variant repeat arrays of 2.4 and 2.6 kb (see Figure 3).

conversion, acting most likely at the time of meiosis (Jeffreys *et al.*, 1994). We have been able to detect siblings which share similar mutant molecules, the result of mitotic events leading to germline mosiacism (for line 109, mice 1035 and 1036; for line 110A, mice 239, 292 and 307). Restriction fragments specific for each of these events were not detectable in the tail DNA of the parental mice, indicating that mosaicism in the parental germline did not extend to the soma. Such mutation most likely results from a mitotic mutation event. To test this hypothesis we isolated and studied by MVR-PCR the mutant regions of mice 239 and 292. This analysis revealed identical structures in each mouse. In each case seven repeat units had been gained (see Figure 3). This is consistent with both structures being derived from a single mitotic event. Furthermore, this example indicates that mutations which change the length of restriction fragments within a transminisatellite can be the result of the gain or loss of minisatellite repeat units.

Structural analysis of mutated transminisatellites

Knowledge of the structures of loci 109 and 110A has allowed the structural localization of mutation events. This is most clearly seen at locus 109, where the restriction fragments produced by a double digestion with *Ssp*I and *Hin*dIII cuts the locus into five segments, only one of which contains two separate minisatellite repeat arrays (see Figure 1). Thus any change in structure can be mapped to a defined part of the locus. For the larger 110A locus, more complex partial restriction analyses were required to localize mutations (data not shown). Overall we have been able to broadly define the structures of three germline mutants and three early embryonic mutants at locus 110A and 10 germline mutants and six early embryonic mutants at locus 109 (see Figure 3).

At locus 110A it can be seen that all changes appear to involve the gain or loss of tandem repeat units from the transminisatellite, with most of these changes being restricted to a single tandem repeat array (as in mutants 295, 325, p30 and 295; see Figure 3). Larger changes are possible; in mutant 239 two repeat unit arrays were modified (one gaining seven repeats, the other repeat array being deleted) and in mutant 836 six repeat unit arrays were deleted (see Figure 3).

At locus 109 a similar pattern was seen, with some mutation events appearing to be the result of a gain or loss of minisatellite repeat units from a single repeat unit

Table II. Germline transmission frequencies of transminisatellites from mice showing somatic mutational mosaicism

Locus	Mosaic mouse	Mosaicism of tail DNA[a] (%)	Progeny		
			Mutant	Non-mutant	Negative
109	109	20	1	12	15
	38	10	2	7	3
	707	60	2	5	5
	756	30	0	8	16
110A[b]	295(837)	30	4	5	6

[a]Proportion of cells containing the mutant version of the progenitor transminisatellite, estimated from scanning densitometry of the Southern blot profile of each mosaic mouse.
[b]Mouse 295 contains two mutations detectable by Southern blot analysis. Detectable transmission of only one mutation occurred (that found in mouse 837; see Figure 2B).

array and others being larger changes. Mutants 971, 984, 1035 and 1042 appear to be of the apparent gain or loss type with only one repeat unit array being changed (see Figure 3). Mutants 709, 140 and 745 are examples of larger changes with four or two complete minisatellite arrays being deleted from the transminisatellite respectively (Figure 3). In contrast, mutant 1594 has gained an extra repeat unit array (Figure 3).

In addition, other events at locus 109 were seen that did not appear to alter repeat unit content. These changes seem to be restricted to the non-minisatellite human flanking DNA present within the transminisatellite and were detected as the loss of a specific *Hin*dIII site in the second internal junction region (IJ-2) of locus 109. We call such changes 'missing site' events. Such mutations occur at high frequency (6 of 20 mutations detected at 109), both in the germline (four mice) and in early embryo (two mice). We have been unable to determine the structural basis for this kind of event, as the non-minisatellite DNA of this region is somewhat rearranged (data not shown).

Dosage reduction and deletion at locus 109

In the course of breeding large numbers of mice for mutation analysis it was found that transminisatellite 109 showed non-Mendelian inheritance with transmission distortion. For example, only 68 of 205 progeny of hemizygous×non-transgenic parents carried the 109 transminisatellite (33% transmission, a significant deviation from the expected 50%, $P < 0.001$). Individual line 109 animals showed variability in their transmission of the locus, some with rates of around 50%, others with significantly lower rates (see Table III). This effect upon transminisatellite transmission was not seen at the single-copy locus 110D (see Table III).

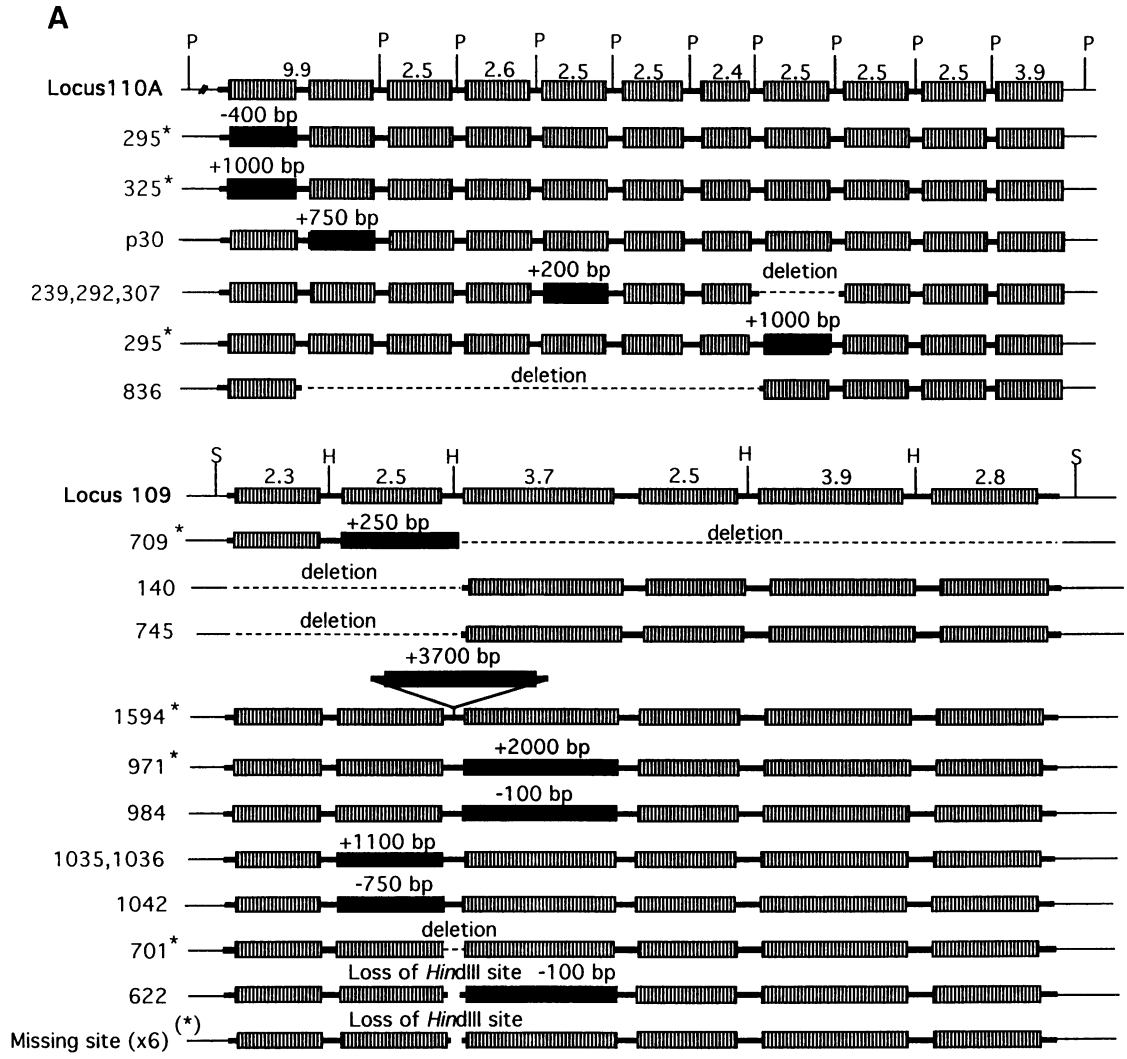
Furthermore, tail DNA analysis revealed low levels of otherwise normal transminisatellite 109 in some mice (see Figure 4). Transminisatellite dosage was quantified by using the mouse LINE-derived 1.3 kb *Eco*RI fragment (Chen and Schildkraut, 1980) as an internal control for the amount of DNA loaded per lane, with levels of signal from both MS32 and *Eco*RI fragment probes being measured via phosphorimaging. Of a randomly selected group of 60 hemizygous line 109 mice examined in this way, 30 showed evidence of reduced levels of transminisatellite (Figure 5). In contrast, 26 line 110A mice similarly analysed did not reveal transminisatellite dosage variation of the extent seen at locus 109 (see Figure 5). Finally, mice with a low dosage of transminisatellite 109 in their tail DNA were also seen to be low transmitters (see Table III). Taken together, these data indicate that significant deletion of transgene 109 can occur during early development.

One further prediction is that each fertilization involving egg or sperm bearing locus 109 will reset the initial dosage levels to 100% of the expected (one copy of the transgene per zygote). Thus low dosage mice should give rise to progeny with normal, intermediate or low levels of the transminisatellite. This is indeed the case, as is shown, for example, by low dosage mouse 144, which is the parent of high dosage mouse 439 (Figure 4).

Discussion

Mutation rate and transminisatellite copy number

Since the human MS32 locus is single-copy, the most appropriate model for human minisatellite mutation in the mouse would be a mutating single-copy transminisatellite. So far we have not seen any mutations in the pedigree of single-copy transgene 110D, and SP-PCR (Jeffreys *et al.*, 1994) analysis of sperm from transminisatellitic mice has shown an ~100-fold reduction compared with the human mutation rate (M.J.Allen, unpublished data). The reasons for this may be lack of appropriate human flanking DNA, the hemizygous rather than heterozygous status of the transminisatellite, possible negative influences from the mouse gamma-satellite DNA 5′ to this locus, the incompatibility of murine and human genetic systems, or genetic background effects of the kind seen possibly to act upon locus 110A. In contrast, the novel multicopy transminisatellites we have created, 110A and 109, show high (hemizygous) mutation rates directly measurable by pedigree analysis, although it remains to be seen how closely the events taking place here resemble those seen at human and mouse single-copy minisatellites. Certainly, the larger changes seen at both 109 and 110A and the smaller missing site mutations observed at 109 appear quite distinct from the normal processes of mutation acting at human MS32 (Jeffreys *et al.*, 1994). Potentially the most similar changes may be those involving fragment length changes, consistent with the gain or loss of small numbers of minisatellite repeat units from one repeat unit array only. It should be possible to define the structural basis of these changes and determine whether such germline mutations involve the complex polar gene conversion events seen at MS32 in man (Jeffreys *et al.*, 1994).



B

Progenitor MVR map (71 repeat units)

aaatataaataaaaaataatataaaatatataaaaaatattaaattaaaaaaaaatttaTTTTATAttttt

Mutant 239

aaatataaataaaaaataatataaaatatataaaaaatattaaattaaaaaaaaatttaTTTTATATTTTATAtttt

Mutant 292

aaatataaataaaaaataatataaaatatataaaaaatattaaattaaaaaaaaatttaTTTTATATTTTATAtttt

Fig. 3. (**A**) Structures of germline and embryonic mutants of multicopy MS32 transminisatellites. The location of altered regions (black boxes) and the approximate extent of each change was established by Southern blot analysis of tail DNA. Asterisks indicate mutations of embryonic origin, all other mutations being germline in origin. 'Missing site' events can occur in either germline or embryo. Mice 1035 and 1036 share the product of a possible germline mitotic event, as do mice 239, 292 and 307. Also, seven missing site mutants were detected, but as two of these were shared by siblings these mutants were assumed to be the product of a single mutation event. Mice 140 and 745 are not siblings and the similar structures they contain have arisen at different places within the line 109 pedigree (data not shown). (**B**) The results of MVR-PCR analysis of the expanded regions of mutants 239 and 292. MVR-PCR was used to type the second polymorphic position of the repeat unit (see Figure 1A). A-type and T-type repeat units differ in that they contain either a G or A at this position respectively (Jeffreys *et al.*, 1991; see Figure 1A). In each case the same seven minisatellite repeat units have been gained, the result of a small internal duplication event towards the 3′ end of the repeat array (highlighted by uppercase lettering). This is consistent with both mutants arising at a common germline mitotic event.

Homozygous and heterozgyous mutation analysis

So far, all of the data presented relate to mutation of hemizygous transminisatellites. We are currently starting the investigation of mutation processes acting in the germline and soma of mice homozygous and heterozygous for transminisatellites 109 and 110A. Interestingly, despite having studied 168 alleles derived from parents homozygous for locus 110A, we have not seen any Southern blot detectable mutation events (A.Collick and P.Bois, unpublished data). Thus it may be that homozygous 110A

Table III. Non-Mendelian segregation of transminisatellite 109 in mice showing low somatic transgene dosage.

Line 109 mouse	Relative tail DNA dose[a]	Progeny		Significance (*P*)
		+	−	
109 (founder)		13	15	
6	0.8–1.0	19	18	
756	0.8–1.0	8	16	
144	0.6–0.8	26	77	<0.001
767	0.4–0.6	4	32	<0.01
625	<0.2	0	10	0.016
Line 110D mice various	1.0	31	29	

[a]The dose of transminisatellite, as estimated from Southern blot hybridization signal in tail DNA. Fully hemizygous mice should have a dose of 1.0 and transmit the locus to progeny with a frequency of 50%.

loci are more stable. Alternatively, this may be further evidence for the non-random mutation of locus 110A within its pedigree. The homozygotes studied were derived from hemizygous×hemizygous crosses which contained no evidence for mutation (see Table I). Thus, perhaps, the transgene has become stabilized with progression through the pedigree. This process is possibly analagous to the irreversible modification of certain other transgenes, where initial 'activity' (in this case gene expression) can become supressed by the transmission of these loci to subsequent generations (Hadchouel *et al.*, 1987; Allen *et al.*, 1990; Engler *et al.*, 1991). A PCR-based analysis of mutation at locus 110A might allow an accurate estimation of the effect of hemizygosity versus homozygosity, irrespective of the influence of other factors, in setting the overall level of mutation rate.

Missing site events and transminisatellite structure

The detection of mutation events which do not appear to change minisatellite repeat unit numbers, but rather the DNA flanking the minisatellite, was unexpected. These events were detected as the simple loss of a *Hin*dIII site from only one internal junction of locus 109 (IJ-2, see Figure 3). The structural basis of these common 'missing site' mutations is not known. Interestingly, mutations which affect repeat unit copy number appear to be clustered around this junction (see Figure 3). One possibility is that a structure formed at transminisatellite integration may promote both kinds of mutation (missing site events and repeat unit changes). This view is supported by the fact that the human minisatellite flanking DNA in the IJ-2 region is structurally abnormal, some sites for PCR primers and restriction enzymes having been deleted (data not shown). Alternative explanations are that transminisatellites such as locus 109 have intrinsically non-random mutation processes, perhaps because of their multicopy nature, their juxtaposition with mouse flanking sequences or their orientation and/or position on the chromosome.

Transgene deletion at locus 109

The high frequency deletion of locus 109 was unexpected. This led to some mice having low dosages of transminisatellite in tail DNA (Figure 4). As such mice also showed reduced transmission from one generation to the next (Table III), we can conclude that deletion can occur in both germline and soma. It is known that locus 109 lies in a region of mouse multi-copy gamma-satellite DNA

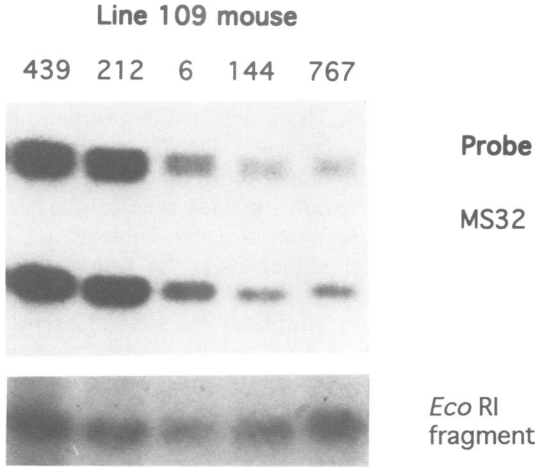


Fig. 4. Dosage variation of transminisatellite 109 in somatic DNA. Tail DNA of five 'hemizygous' mice was digested with *Eco*RI and subjected to Southern blot analysis using MS32 DNA as probe. To measure the amounts of tail DNA present, this filter was then reprobed with the mouse LINE-derived *Eco*RI fragment. Comparison of the signals obtained with each probe indicated that mice 439 and 212 have relatively high amounts of the transgene compared with mice 6, 144 and 767. Transmission of the 109 locus from each of the latter mice has been studied (see Table III).

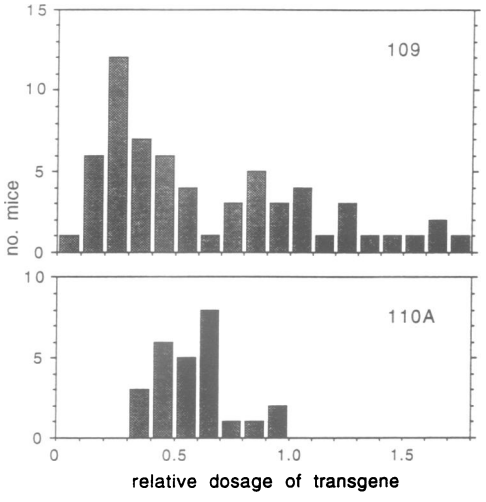


Fig. 5. Dosage variation of multicopy transminisatellites in tail DNA. Sixty mice hemizygous for locus 109 and 26 mice hemizygous for locus 110A were analysed and the relative signal of transminisatellite and *Eco*RI fragment (see Figure 4) compared by phosphorimager analysis. Estimated relative dosage was binned into 0.1 intervals.

(Allen *et al.*, 1994) and this may be responsible for these effects. For example, unequal exchange crossover events between satellite repeat units might lead to deletion of the transminisatellite. Previously, instability has been described for two selectable markers (thymidine kinase and CAD) integrated into centromeric regions of rodent cell lines (Wahl *et al.*, 1984; Butner and Lo, 1986). Thus any foreign sequences integrated into satellite DNA regions may show instability. The non-satellite DNA may actively promote its own deletion by specifically destabilizing the region or may be a passive victim of the normal processes of satellite DNA turnover.

Early embryonic window of instability

Three distinct mutational processes occur during early embryonic development. These are, firstly, the loss of a specific *Hin*dIII site from transminisatellite locus 109, secondly, the apparent gain or loss of minisatellite repeat units from loci 109 and 110A and, thirdly, the complete deletion of locus 109. These observations further expand the list of tandemly repeated DNA sequences which appear to be able to mutate during this 'early window of instability'. Other examples include mouse minisatellites *Ms6-hm* and *Hm-2* (Kelly *et al.*,1989; Gibbs *et al.*, 1993) and the triplet repeat sequences at the human fragile-X and myotonic dystrophy loci (Nelson and Warren, 1993; Reyniers *et al.*, 1993; Worhle *et al.*, 1993; Jansen *et al.*, 1994). Whilst all these loci contain tandem repeats, the response appears to vary with the type of repeat unit present within the array. Minisatellites and transminisatellites do not show the 'run-away' expansion seen at some triplet repeat disease loci. The latter may therefore be a function of small repeat unit size. Another interesting point is why the MS32 locus in man shows very low rates of early embryonic instability (as do six other well-characterized human minisatellite loci; A.J.Jeffreys, unpublished data), only to show higher rates when present in the mouse genome. Perhaps the early embryonic window of the mouse is more potent in inducing tandem repeat instability.

Conclusion

Transgenic mice can play an important role in the analysis of DNA recombination, including analysis of VDJ recombination (Ritchie *et al.*, 1984; Alt *et al.*, 1985; Matsuoka *et al.*, 1991), immunoglobulin heavy chain switching (Gerstein *et al.*, 1991) and germline gene conversion (Ramana Murti *et al.*, 1992). The transminisatellitic mice described here indicate that 'transgenes' containing tandemly repeated sequence arrays can show instability and thus provide additional systems for the study of the complex processes which lead to DNA structural change in higher eukaryotes.

Materials and methods

Breeding of transgenic mice

Transgenic founder mice [background (C57BL/6J.CBA/Ca) F_2] were bred with (C57BL/6J.CBA/Ca) F_1 animals to produce the first generation of transgenic offspring (transgenic F_1). Subsequent generations were produced by crossing these mice and their offspring with C57BL/6J mice (Leicester), although in some cases (C57BL/6J.CBA/Ca) F_1 animals were used (Cambridge). In addition, some F_4 mice were produced by crossing hemizygous F_3 mice. In summary, Southern blot analysis was carried out on a number of mice from each of the F_1, F_2, F_3 and F_4 transgenic generations.

Tail DNA analysis

Tail DNA was prepared as described elsewhere (Allen *et al.*, 1987). Southern blot analysis using MS32 was as described (Jeffreys *et al.*, 1994). MVR-PCR analysis was as described (Jeffreys *et al.*, 1991). The LINE-derived *Eco*RI fragment probe was generated by excision of the appropriate 1.3 kb fragment from an *Eco*RI digest of non-transgenic tail DNA electrophoresed in a low melting point agarose gel (Feinberg and Vogelstein, 1983). The subsequent use of this probe was as for MS32, except that the final stringency of washing after hybridization was to 1× SSC, 0.1% SDS at 65°C. Phosphorimaging analysis was carried out on a Molecular Dynamics PhosphorImager using Image Quant software.

Acknowledgements

We thank Biomedical services (Leicester) for the excellent care given to the mice involved in this study. We also thank Robert Kelly, Wolf Reik, Christine Pourcel and Mark Gibbs for encouragement during this study. The work of A.J.J. was supported in part by an International Research Scholars Award from the Howard Hughes Medical Institute and in part by grants from the Medical Research Council and Royal Society. The work of A.S. was supported by grant number 036481 from the Wellcome Trust.

References

Allen,M.J., Jeffreys,A.J., Surani,M.A., Barton,S., Norris,M.L. and Collick,A. (1994) *Nucleic Acids Res.*, **22**, 2976–2981.
Allen,N.D., Barton,S.C., Surani,M.A.H. and Reik,W. (1987) In Monk,M (ed.), *Mammalian Development: A Practical Approach*. IRL Press, Oxford, pp. 217–233.
Allen,N.D., Norris,M.L. and Surani,M.A. (1990) *Cell*, **61**, 853–861.
Alt,F., Blackwell,T.K. and Yancopoulis,G.D. (1985) *Trends Genet.*, **1**, 213–236.
Armour,J.A.L., Harris,P.C. and Jeffreys,A.J. (1993) *Hum. Mol. Genet.*, **2**, 1137–1145.
Buard,J and Vergnaud,G. (1994) *EMBO J*, **13**, 3203–3210.
Butner,K. and Lo,C.W. (1986) *J. Mol. Biol.*, **187**, 547–556.
Chandley,A.C. and Mitchell,A.R. (1988) *Cytogenet. Cell Genet.*, **48**, 152–155.
Chen,S. and Schildkraut,C. (1980) *Nucleic Acids Res.*, **8**, 4075–4090.
Engler,P., Haasch,D., Pinkert,C.A., Doglio,L., Glymour,M., Brinster,R. and Storb,V. (1991) *Cell*, **65**, 939–947.
Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **137**, 266–267.
Gernstein,R.M., Frankel,W.N., Hsieh,C.L., Durdik,J.M., Rath,S. and Coffin,J.M. (1990) *Cell*, **63**, 537–548.
Gibbs,M., Collick,A., Kelly,R. and Jeffreys,A.J. (1993) *Genomics*, **17**, 121–128.
Hadchouel,M., Farza,H., Simon,D., Ziollias,P. and Pourcel,C. (1987) *Nature*, **329** 454–457.
Jansen,G., Willems,P., Coerwinkel,M., Nillesen,W., Smeets,H., Vits,L., Howeler,C., Brunner,H. and Wieringa,B. (1994) *Am. J. Hum. Genet.*, **54**, 575–585.
Jeffreys,A.J., Wilson,V. and Thein,S.L. (1985) *Nature*, **314**, 67–74.
Jeffreys,A.J., Royle,N.J., Wilson,V. and Wong,Z. (1988) *Nature*, **332**, 278–281.
Jeffreys,A.J., MacLeod,A., Tamaki,K., Neil,D.L. and Monckton,D.G. (1991) *Nature*, **354**, 204–209.
Jeffreys,A.J., Tamaki,K., MacLeod,A., Monckton,D.G., Neil,D.L. and Armour,J.A.L. (1994) *Nature Genet.*, **6**, 136–145.
Kelly,R., Bulfield,G., Collick,A., Gibbs,M. and Jeffreys,A.J. (1989) *Genomics*, **5**, 844–856.
MacLaren,A. (1991) *BioEssays*, **13**, 151–156.
Matsouka,M., Nagawa,F., Okazaki,K., Kingsbury,L., Yoshida,K., Miller,U., Larue,D.T., Winer,J.A. and Sakano,H. (1991) *Science*, **254**, 81–86.
Monckton,D.G., Neumann,R., Guram,T., Fretwell,N., Tamaki,K., MacLeod,A. and Jeffreys,A.J. (1994) *Nature Genet.*, **8**, 162–170.
Neil,D.L. and Jeffreys,A.J. (1993) *Hum. Mol. Genet.*, **2**, 1129–1135.
Nelson,D.L. and Warren,S.T. (1993) *Nature Genet.*, **4**, 107–108.
Ramana Murti,J., Bumbulis,M. and Schimenti,J.C. (1992) *Mol. Cell Biol.*, **12**, 2545–2552.
Reyniers,E. *et al.* (1993) *Nature Genet.*, **4**, 143–146.

Ritchie,K.A., Brinster,R.L. and Storb,U. (1984) *Nature*, **312**, 517–520.
Royle,N.J., Clarkson,R.E., Wong,Z. and Jeffreys,A.J. (1988) *Genomics*, **3**, 352–360.
Vergnaud,G., Mariat,D., Apior,F., Aurias,A., Lathrop,M. and Lauthier,V. (1991) *Genomics*, **11**, 135–144.
Wahl,G.M., Robert de Saint Vincent,B. and De Rose,M.L. (1984) *Nature*, **307**, 516–520.
Wahls,W.P., Wallace,L.J. and Moore,P.D. (1991) *Cell*, **60**, 95–103.
Wong,Z., Wilson,V., Patel,I., Povey,S. and Jeffreys,A.J. (1987) *Ann. Hum. Genet.*, **51**, 269–288.
Wohrle,D., Hennig,I., Vogel,W. and Steinbach,P. (1993) *Nature Genet.*, **4**, 140–142.

Received on July 19, 1994; revised on September 20, 1994