Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 3320–3324, May 1987
Genetics

A family of short, interspersed repeats is associated with tandemly repetitive DNA in the human genome

(*Mst* II family/variable tandem repeats/short interspersed repeat sequences/minisatellites/mutation rate)

BRION MERMER, MARK COLB, AND THEODORE G. KRONTIRIS*

Department of Medicine (Hematology/Oncology), Tufts-New England Medical Center, Boston, MA 02111; and Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, MA 02111

Communicated by Victor A. McKusick, January 6, 1987

ABSTRACT **A family of short, interspersed repeats in the human genome, designated the *Mst* II family, is described. The canonical structure of the repeat consists of a 220-base-pair (bp) left arm joined to a 160-bp right arm by a 39-bp junction sequence. The right arm is absent in some isolates. Some homology with the "*O*" and "*THE*" (transposon-like element) families of repeats was observed, suggesting that the *Mst* II elements could be a subgroup of a *SINE* superfamily. The 39-bp junction sequence is tandemly repeated in one of our clones. The association of tandemly repetitive sequences with *Mst* II elements or the putative superfamily is probably nonrandom; a search of DNA sequence data bases revealed that approximately 80 bp of the *Mst* II left arm occurs immediately adjacent to the tandem repeat that comprises the human homologue to the BK virus enhancer. The fortuitous occurrence of a gene duplication event involving an *Mst* II repeat has allowed us to estimate a mutation rate for human DNA.**

The structural basis for genetic variation in many highly polymorphic regions of the human genome is the tandem repetition of 14- to 65-base-pair (bp) nucleotide sequences (1–5). The variable tandem repeat (VTR) approximately 1.5 kilobases (kb) downstream from the cellular homologue of the Harvey rat sarcoma virus oncogene, c-Ha-*ras1* (*HRAS1* in human gene nomenclature) (2), generates at least 25 allelic forms in which a 28-bp consensus sequence is reiterated 30–100 times (6, 7). Many of the rare alleles at this locus are detected only in cancer patients (6, 7). Since such hyperallelism might be exploited in quantitating the degree of genomic instability in human populations, we have sought to identify other regions related to the *HRAS1* VTR.

As a result of this search, we identified a highly polymorphic region distinct from the *HRAS1* VTR (7). Sequence analysis of this region, designated 4.1, revealed the head-to-tail aggregation of a 35-bp monomer (8). When 4.1 was used as a probe for other polymorphic clones, 0.5–1% of the phage in a human DNA library were recognized by *in situ* plaque hybridization. We report here the characterization of the repeat family responsible for this result, including the observation of its intimate association with VTRs.

MATERIALS AND METHODS

Genomic Library, Vectors, and Bacterial Strains. A human genomic library in Charon 4A (9) was kindly provided by Tom Maniatis. It was propagated in *Escherichia coli* LE392. Phage lysates were also prepared from this strain. Plasmid subclones were propagated in HB101 or JM83 with pBR322 or pUC9, respectively.

DNA Transfer Hybridization. DNA was transferred to nitrocellulose filters (Schleicher & Schuell) by methods previously described for phage plaques (10) or agarose gels (11). Low-stringency hybridization was performed at 50°C in 5× SET buffer (1× SET buffer is 0.15 M NaCl/10 mM EDTA/10 mM Tris, pH 7.5) and 0.5% $NaDodSO_4$. Three or four washes were performed at 37°C (unless otherwise stated) in 2× SET buffer/0.5% $NaDodSO_4$ for 30 min each. The radioactive probe was prepared by nick-translation (12).

Nucleotide Sequence Analysis. Maxam–Gilbert DNA sequencing was performed by the modified three-reaction procedure (13).

Computer Analysis. DNA sequence comparisons were performed with an interactive dot-matrix program provided by John Coffin. Sequences in the GenBank and European Molecular Biology Laboratory data bases were analyzed by the GenBank MATCH and FASTN programs.

RESULTS

Identification of the *Mst* II Family. We were interested in determining whether human DNA contained a family of homologous VTR elements related to the VTR located just downstream from the *HRAS1* gene (2). As a probe we used an *Msp* I restriction fragment that was composed primarily of 29 copies of the 28-bp monomer of the *HRAS1* VTR (ref. 2; Fig. 1, probe A). We screened a human genomic library at low stringency and isolated one recombinant phage, 4.1. This phage did contain a VTR with approximately 25 copies of a 35-bp monomer (8). The 4.1 monomer contained a 14- of 15-bp match with one region of the *HRAS1* VTR monomer (see below).

Since our *HRAS1* probe did not identify any other members of a putative VTR family, we wished to determine if the 4.1 VTR would be more useful. Plasmid pBBg3 (Fig. 1), a subclone of 4.1, was isolated after complete digestion of 4.1 with *Bam*HI, partial digestion with *Bgl* II (which cuts many times within the VTR), and ligation of the *Bam*HI/*Bgl* II partial digest products into the *Bam*HI site of pBR322. Probe B was generated by complete digestion of pBBg3 with *Bam*HI and *Bgl* II. The resulting 500-bp fragment included four VTR monomers, since the first three monomers lacked the *Bgl* II recognition site. The remainder of probe B was unrelated to the VTR. When we screened the genomic library with probe B, we found that 0.5–1% of the phage was identified. This frequency suggested the existence of an interspersed, moderately repetitive sequence present at about 5,000-copies per genome. This frequency was much greater than that reported for other VTR families (5).

Abbreviations: VTR, variable tandem repeat; *SINES*, short interspersed repeat sequences (elements); *THE*, transposon-like human element; BKV, BK virus; Hu-BKV, human homologue of BKV enhancer.

*To whom reprint requests should be addressed at: Division of Hematology/Oncology, New England Medical Center Hospitals, 750 Washington Street, Boston, MA 02111.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

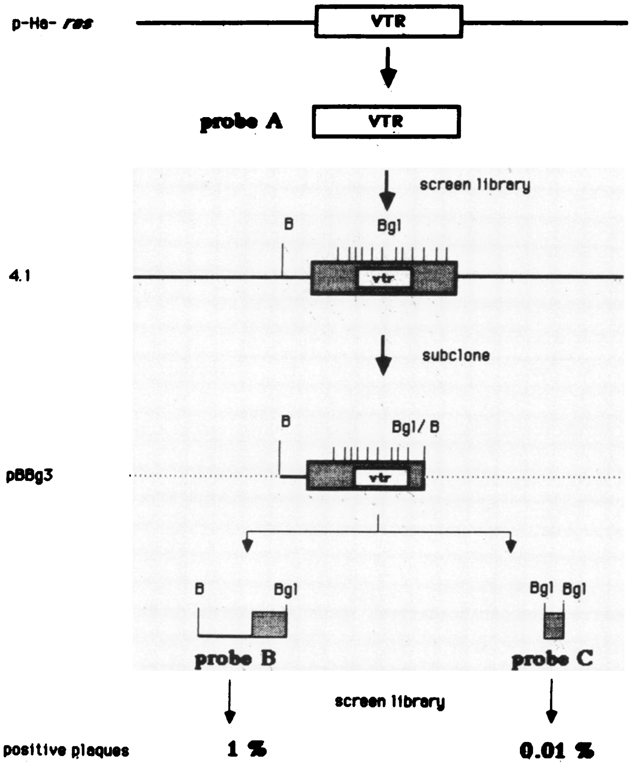

FIG. 1. Derivation of probes. The *Msp* I fragment containing the *HRAS1* VTR was purified from a plasmid containing the *HRAS1* (p-Ha-*ras*) gene (probe A). The recombinant phage, 4.1, was obtained from the library screen under low stringency. The VTR region of 4.1 was subcloned into pBR322 by insertion of a 700-bp *Bam*HI/*Bgl* II partial digest product into the pBR322 *Bam*HI site. This subclone, designated pBBg3, was completely digested with *Bam*HI and *Bgl* II to yield probe B. Limit digestion with *Bgl* II and electrophoresis through polyacrylamide gels resulted in the purification of the 35-bp oligonucleotide, probe C. B, *Bam*HI; Bgl, *Bgl* II.

To determine if the VTR portion of probe B was responsible for the large number of positive plaques, we prepared a VTR-specific probe. *Bgl* II digestion of pBBg3 reduced most of the 4.1 VTR to a 35-bp oligonucleotide. The gel-purified, end-labeled oligonucleotide (probe C) was used to screen the same genomic library. In reconstruction experiments, probe C readily detected phage λ clones bearing either the *HRAS1* or 4.1 VTR in a mixture of library bacteriophage. However, on rescreening, only 0.01% of the library phage plaques were recognized by probe C, indicating that a distinct, interspersed repetitive element had indeed been identified by probe B. From this preliminary result, the interspersed repeat appeared to be homologous to the region between the *Bam*HI site and the VTR in pBBg3 (Fig. 1).

Comparative Sequence Analysis of the *Mst* II Family. Several phage identified by probe B were purified and analyzed by Southern transfer hybridization to permit isolation of fragments containing the 4.1-related sequences. Restriction mapping (not shown) indicated that the human DNA inserts in these phage were different; the repetitive element was located on distinct restriction fragments for each phage examined. A hybridization probe prepared from the 4.1-related fragment of one phage, 15.1, identified the 4.1-related fragments in all other phage, confirming that all clones must have been identified by a common *SINE* sequence in probe B (Fig. 2 *Left*). To our surprise, however, the clones isolated by homology to probe C, the 35-bp oligomer, also reacted with 15.1 (Fig. 2). Probe C was supposed to be a VTR-specific, not a *SINE*-specific probe. The reason for this homology was evident when sequence data became available (see below).

Several isolates detected by probe B—13.3, 15.1, and 37.1—were analyzed further. Restriction fragments homologous to probe B were subcloned, and the nucleotide sequence of the 4.1-related region was determined in each case (Fig. 3). A conserved region of approximately 220 bp was present in each clone and corresponded to the sequence immediately upstream from the 4.1 VTR (nucleotides 0–210). We now refer to this region as the "left arm" of the repeat element (see below). Based upon a conserved *Mst* II site at nucleotide 140, we designated this collection of repetitive elements the "*Mst* II family."

Comparison of the *Mst* II elements in isolates 13.3, 37.1, and 4.1 gave the striking result that the monomeric unit of the 4.1 VTR was, in fact, related to the 39 bp immediately adjacent to the left arm of both 37.1 and 13.3 (Fig. 3). Therefore, these elements appeared to contain only one copy of a sequence that was tandemly repeated in 4.1. (See also below.) Confirmation that the 39-bp unit of the 4.1 VTR was merely an extension of the *Mst* II repeat was obtained from isolate 2.1, which was selected from the human library on the basis of homology with the "VTR-specific" probe C (Fig. 1). The 2.1 clone also contained the *Mst* II left arm (Fig. 3) as well as one copy of the contiguous 39-bp region tandemly repeated in 4.1 (Figs. 3 and 4). Significant differences in the 39-bp junction must exist from element to element in the *Mst* II

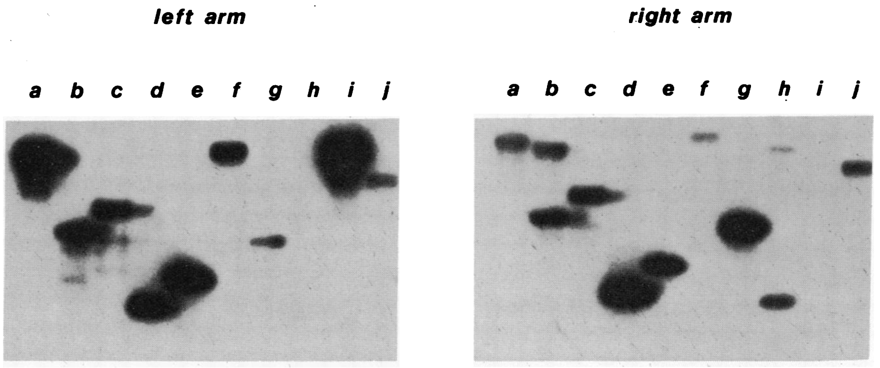

FIG. 2. Representation of the left arm (*Left*) and right arm (*Right*) in *Mst* II element clones. DNAs from recombinant phage containing independent isolates of the *Mst* II repeat (see Table 1) were digested with the indicated restriction enzymes and subjected to Southern blotting. The filter was first hybridized to the left-arm probe, an *Eco*RI fragment of 15.1. After erasure, the filter was rehybridized to the right-arm probe, the *Bgl* II fragment of 2.1 described in the text. Washes were performed at 50°C for the right-arm probe. Lane i may be compared in each panel for the completeness of erasure. Only one *Mst* II element is present in isolates 14.2 (lane b) and 2.3 (lane h); these particular elements contain restriction sites for the enzymes used to digest the phage (restriction maps not shown). Lanes: a, 10.2, *Kpn* I; b, 14.2, *Kpn* I; c, 13.3, *Kpn* I; d, 2.1, *Pst* I; e, 2.5, *Eco*RI/*Hin*dIII; f, 2.5, *Kpn* I; g, 2.1, *Eco*RI; h, 2.3, *Hin*dIII; i, 15.1, *Kpn* I; j, 10.2, *Eco*RI.

```
                    .         .         .               .          .50        .                 .         .         .        .100
CON       TGTATTAGTCTGTTCT CGCA TTGCTATAAAGAA       ATACCTGAGACTGGGTAA TTTATAAAGAAAAGAGGT     TTAATTGGCTCATGGTTCTGCAGGCTGTACA
4.1       ......CA........ .A.. C.........A..       ............CT.... ...T...GA....A....     ...............................
13        ..C.CC..G......C A... ....C........       .....A.........A.G .C.........C......     ............C..................
15        ..C.....G.CA.... T... .......G.....       ...A....TC..T..... ................ AG    .GT.......T......G.......C.....
2.1       GGCCTCCCAAAGTGCTGGGGTTATAGATGTGAGCCACTGCGCCCAG.CA.......... ..................     ......T..CT......TG...A........
37.1      ....C........... T... .............       ......C........... ................C..    .............CA............T...
Cε        ..AG.......A..TGTGTT. C.AGA.A...A..     AA.A.A.ATGA.TCT..... ..................(33bp)...C.A.....CA.........C. A.T.C
Cε2       ...G.....T...    GTT. C.AG .A...A..       ....A.... GCTA.... ..............GT..(136bp)..GG.CGGGGGA.AG.A.A..AGCAG.AG
O         ..............T. .A.. C...,GAT..A..CATAATTC.T..........          .GG.G.A.........     ........ACTTACA..TCCACATGG
Hu-BKV    ......CC.TCA.... ..ATGC......G..A..       .A................A...............T...     .....A....    --ND--

                    .         .      .      .      .150        .         .         .         .               .200      .
CON       GGAAGCATGG TGCTGGCATCTGCTTNGCTTCTGGGGAGGCCTCAGGAAGCTTCCAATCATGGTGGAAGGTGAAGGGGGAGCAGGCATCTCACATGG      CAAGAGAGGAAGCA
4.1       .....    .A GAG...T........    ........................T.....G...........................TGC.........      ...A....... ..
13        .......... ..................TC...........................................CAG....A...TG.....G.........      .C....CA.G....
15        ......G...CAC.G...........CA..........A.CA.T. .........A..G.G..................... ...................      T.....T..G...T
2.1       .........A .................G.............A..........A...A....T..............CA...T............CA.....C..      .C.C..CA.G..G.
37.1      ....A..... .A.............C.N........ .......................A..G..........................CA.........      .C....CA.G..AG
Cε        ....A..C.. CC..............     ....C...........G..A..G......C...A.....C.....................G.........      ...... ...G ..
Cε2       ....A..C.. CC..............     ....T...........G..A..G.........CA.....C.....................G.........      ...... ...G ..
O                                         ............A...     ......A.C.AG.......A..C      A.T...T......TGGTGG......GCCT.TTC

                    .       .        .         .250       .          .         .         .          .300       .       .
CON       AG AAAGAGAGGGGGGAGA CGCCACACACTTTTAAACAACCAGATCTCATGAGGACTCACTGACTGAGATGTAGTCAGTACC AAGGGGATGGTGCTAGGCCATTAATGAGG
4.1       .. ......                - - - - VTR - - - -
13        .. .C..CAG.....AG.TG.CA......................G.........A....T..   ..ATC ..C..        - - - ND - - -
2.1       .. .GT......A...... ...........................G.......G....C...ATC....G.G....... .................AAT..........
37.1      .. C...TCTT.....GA.GA..T.T..T........TG...........AAGTA....GGA.CAA...T. C.C......... ...A.TG...CC.A..C.AT.C.CAAG..
Cε        ..AG......A.T..A     ..T.C..G............TG..C.G...CAG.....A.....GG....AC.C.CTCA.CA.CTG..A.................C......
Cε2       ..AG..A...A.T..C.    ..T.C..G.A..G..TT....GG.C.G..TCTG.....A.....GG....A   C...TC... TG..A......,...........C......
O         GC ..TC.TGA.AAT.GC. G.G                                                                                   ..A..TC

                    .         .           .350        .         .         .         .          .400       .        .
CON       AGTCCACCCCCATGATCCAATCACCTCCCACCCA TAGGCCCCACCTCCAACACTGGGGATTACAATTCGACATGAGATTTGG TGGGGCAACGATCCAAACCATA
2.1       .AG............G....................A........A...A.......................................AC.........G.....
37.1      .TC.T...T..........     ...T.. ....   T.T.................A............. .ACGT........     ....GAC.........A..
Cε        G.....G.,.........T..............T    ATTTT................A..C...TC..A.......A...... A.A.A..CAC........T..T
Cε2       G....TG...........T...............     .... ......G...........          - -  deletion - -
O         ...     T........T....TG.T....CC     .G..T...T..CA.....TGT..AGA..T.....A.GT.........AA.....A CAC.G..........
```

FIG. 3. DNA sequence comparison of independent *Mst* II repeat clones. DNA sequences from isolates 4.1, 13 (=13.3), 15 (=15.1), 2.1, and 37.1 are depicted. The consensus sequence (CON) is derived solely from our sequence data. Sequence data obtained from GenBank for the immunoglobulin-associated *Mst* II repeats in the immunoglobulin ε heavy chain (*IGHE*) constant regions ($C_ε$, and $C_ε2$) and from the European Molecular Biology Laboratory data bank for the *O* family repeat (O) are also provided. The final line Hu-BKV gives the sequence adjacent to the human homologue of the BK virus (BKV) enhancer. The underlined nucleotides (223-261) represent the *Mst* II sequence homologous to the 4.1 VTR monomer. Numbers in parentheses indicate the length of insertions. ND, not determined. Additional 3′ sequence data for 15.1, which were completely unrelated to the junction fragment and right arm of the consensus, are not shown.

family, since probe C recognizes far fewer library phage than probe B.

Comparison of sequences immediately downstream from the 39-bp unit in 37.1 and 2.1 revealed a further extension of the *Mst* II repeat by approximately 160 nucleotides, although the degree of homology was weaker than that observed in the left arm. The presence in other isolates of this portion of the *Mst* II repeat element, which we now designate the "right arm," was confirmed by Southern transfer hybridization (Fig. 2 *Right*). The *Bgl* II fragment of 2.1, which contains the right arm (beginning at nucleotide 250 of Fig. 3) and approximately 250 bp of 3′ flanking DNA, was used to probe other

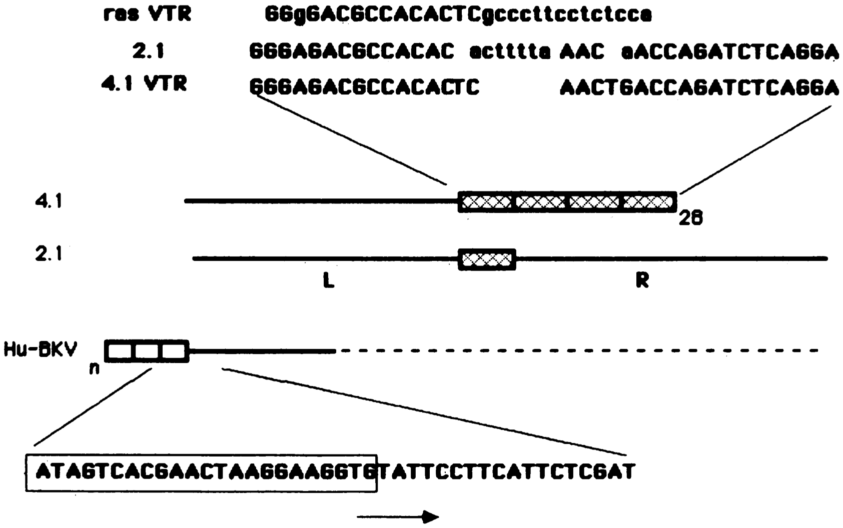
ras VTR
GGgGACGCCACACTCgcccttcctctcca
2.1
GGGAGACGCCACAC actttta AAC aACCAGATCTCAGGA
4.1 VTR
GGGAGACGCCACACTC AACTGACCAGATCTCAGGA
4.1
28
2.1
L
R
Hu-BKV
n
ATAGTCACGAACTAAGGAAGGTGTATTCCTTCATTCTCGAT

FIG. 4. Relationship of VTRs to the *Mst* II consensus element. Schematically depicted are the *Mst* II elements present in 4.1, 2.1, and the human homologue of the BKV enhancer (Hu-BKV). 2.1 is chosen as portraying the canonical form of the repeat, with left (L) and right (R) arms joined by the 39-bp fragment (crosshatched box). The tandem repetition of sequences from the 39-bp box is indicated on the line representing 4.1. Actual sequences from the boxed regions of 2.1 and 4.1 are listed at the top, together with the *HRAS1* VTR monomer consensus for comparison. (Upper-case letters signify sequence identity.) The Hu-BKV monomer (open box) is tandemly repeated at the 5′ end of the left arm. Actual sequences are depicted at the bottom of the figure. 2.1 contains a short deletion at the extreme 5′ end of the left arm. Sequence data beyond the 5′ half of the *Mst* I left arm in Hu-BKV are not available (dashed line).

clones. Only the 15.1 subclone (our left-arm probe) lacked homology to the right arm (Fig. 2, lane i), a result that corresponds to the sequence data showing the deletion of *Mst* II sequences in 15.1 beyond the left arm terminus at nucleotide 210. Although 4.1 was not included in these particular blots, direct sequence analysis of the region 3′ to the 4.1 VTR and other blots similar to Fig. 2 confirmed the absence of a right arm in this isolate (not shown).

In summary, sequence data and Southern blotting have identified a repeat family whose consensus element contains a 220-bp left arm coupled to a 160-bp right arm by a 39-bp junction sequence. In two members of the family that we have analyzed, 15.1 and 4.1, the right arm is absent. The first 40 bp of the left arm are deleted in another clone, 2.1. In the 4.1 isolate, the 39-bp junction is tandemly repeated (Fig. 4). Table 1 compares the properties of clones obtained with probes A, B, and C.

Other *Mst* II-Related Sequences in the Human Genome. A computer search of the GenBank data base revealed two additional members of the *Mst* II repeat family located in the gene for the immunoglobulin ε heavy chain (*IGHE*) constant region and in its closely linked pseudogene (14). These members share extensive homology with the consensus sequence, as shown in Fig. 3. Because one repeat clearly derives from a duplication event involving the other, this particular pair of *Mst* II elements provides interesting material for examining the rate of divergence of two derivative, noncoding DNA sequences (see below). Analysis of the European Molecular Biology Laboratory data base revealed that the *Mst* II family also shared some sequence homology with the *O* (15) and *THE* (16) repeat families. [*THE* repeats are transposon-like structures in which the terminal direct repeat is an *O* element (16)]. The homology was most significant at the ends of the *Mst* II repeat (Fig. 3), although some internal homology, characterized by many insertions and deletions, was also present. This result signified a relatively distant but clear-cut evolutionary relationship. Perhaps *Mst* II repeats represent one subgroup of a *SINE* superfamily that also includes *O* and *THE* elements.

Association of the *Mst* II Family with Tandem Repeats. As noted above, the *Mst* II repeat element in 4.1 was truncated and contained a tandem repeat immediately adjacent to its truncated end. A computer search of the GenBank data base identified another member of this repeat family within Hu-BKV (17). At this site, a tandem repeat occurred at the opposite end of the *Mst* II element compared with the arrangement in 4.1 (Fig. 4 and ref. 17). As indicated in Figs. 3 and 4, the nucleotide sequence published for Hu-BKV did not extend far enough to span the entire *Mst* II repeat element; in fact, the sequence ended at the point at which *Mst* II family members and *O* elements diverge. However, the degree of homology between Hu-BKV and the consensus sequence of the *Mst* II repeat allowed assignment of the former either to the *Mst* II family or to the proposed superfamily (Fig. 3).

The tandem repeat of the human locus homologous to the BKV enhancer demonstrates enhancer activity (17). Although we have not observed enhancer activity for the 4.1 VTR, several motifs within the VTR (8) demonstrate homology with enhancer core sequences. All but two of the VTR monomers contain the sequence GCCACA, which is a five-of-six match with one enhancer consensus. Two monomers also contain the sequence CTTTCCA, which matches exactly the consensus sequence CLLLCCA (L = T or A) (18). Further studies will be required to elucidate the relationship of VTRs and enhancer activity with *Mst* II repeats.

Table 1. Characteristics of DNA clones containing an *Mst* II repeat

			Arms‡		
Isolate	Probe used*	Polymorphism†	Left (L)	Right (R)	Comments§
4.1	A	VTR *Bam*HI	+	−	
10.2	B	ND	+	+	
13.3	B	ND	+	+	
14.2	B	ND	+	+	
15.1	B	*Msp* I	+	−	L arm probe
37.1	B	ND	+	+	
2.1	C	ND	+	+	R arm probe
2.3	C	ND	+	+	
2.5	C	ND	+	+	

*See Fig. 1.

†ND, none detected.

‡+, Present; −, absent.

§Of the nine isolates, only 4.1 demonstrated a VTR.

DISCUSSION

Relationship of Clone 4.1 and Hu-BKV to Other VTR Loci. We have described the novel association of tandemly repetitive DNA with a family of short, interspersed repeats. At two different loci, the tandem repeat occurs at the end of an *Mst* II element. In 4.1, the first VTR subunit corresponds exactly to the position of the 39-bp junction sequence in the *Mst* II consensus (underlined in Fig. 3; see also Fig. 4). In Hu-BKV, the first two bases of the *Mst* II left arm are incorporated into the first two bases of the VTR repeat unit (Fig. 4). We have examined, where available, the 5′ and 3′ flanking sequences of VTRs associated with *HRAS1* and the genes for insulin and ζ-globin, as well as the flanking regions of "minisatellites" whose association with coding sequences is as yet unknown (5). To date, we have documented two occurrences of *Mst* II repeats in 13 published VTR sequences. These results may indicate that two classes of VTRs occur, which are distinguished by the presence or absence of the *Mst* II *SINE* adjacent to the tandem repeat. Each of these classes may have arisen by different means.

The actual mechanism by which tandem repetition occurs is, of course, uncertain. 4.1 was selected on the basis of the homology between its VTR and the *HRAS1* VTR. The homology consists of a 14- of 15-bp core, GGGGACGCCACACTC, common to the two otherwise different VTR sequences (Fig. 4). It is possible that this core sequence accelerates the formation of tandem repetition. This model has been proposed to explain the generation of other families of tandem repeats (5). Perhaps the core sequence serves as a recognition site for a protein, such as a nicking enzyme, which mediates recombinational breakage and rejoining.

However, the association of two structurally dissimilar tandem repeats, 4.1 and Hu-BKV, with the *Mst* II repeat family provides an alternative explanation for the origin of tandem repeats in these isolates. We propose that the interspersed repeat is itself instrumental to the genesis of some VTRs. In this interpretation, the sequence similarities between the 4.1 and *HRAS1* VTRs are probably fortuitous, although selection for function such as enhancement may be superimposed on whatever process generates the repeats.

Frequency of Mutation in Noncoding DNA. The *IGHE* locus is associated with two pseudogenes in human beings, one immediately within the heavy chain constant region gene cluster (pseudogene 1, *IGHEP1*) and one with an unknown, but presumably unlinked, map position (pseudogene 2, *IGHEP2*) (19). Only pseudogene 2 is present in chimpanzees, implying that pseudogene 1 arose by duplication since the time of speciation. The appearance of *Mst* II elements in corresponding locations of the *IGHE* constant region and pseudogene 1 (14) implies that one repeat was derived from the other at the time of this duplication event. We can roughly estimate the time since speciation, 0.5–1 × 10^7 years (20). Therefore, it is possible to calculate the minimum mutation rate for this noncoding DNA, assuming neither selection nor gene conversion has taken place. There are 61 point differences in the 375 base pairs shared by the two *Mst* II repeats within the IGHE constant region, corresponding to a mutation rate of 0.8–1.6 × 10^{-7} per bp per generation. Since the time of duplication relative to the time of speciation is obviously not known, we must emphasize that this is a limit value.

We are grateful to Leslie Groom for technical assistance and to Andrew Dorner and John Coffin for helpful discussions. This work was supported by grants from the National Institutes of Health (CA37866) and Seragen, Inc. M.C. received a stipend from postdoctoral training grant CA09429; T.G.K. was a Hartford Foundation Fellow.

1. Bell, G. I., Selby, M. J. & Rutter, W. J. (1982) *Nature (London)* **295**, 31–35.
2. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goedell, D. V. (1983) *Nature (London)* **302**, 33–37.
3. Goodbourn, S. E. Y., Higgs, D. R., Clegg, J. B. & Weatherall, D. J. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5022–5026.
4. Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985) *Nature (London)* **316**, 76–79.
5. Jeffreys, A. J., Wilson, V. & Thein, S. L. (1985) *Nature (London)* **314**, 67–73.
6. Krontiris, T. G., DiMartino, N. A., Colb, M. & Parkinson, D. R. (1985) *Nature (London)* **313**, 369–374.
7. Krontiris, T. G., DiMartino, N. A., Colb, M., Mitcheson, H. D. & Parkinson, D. R. (1986) *J. Cell Biochem.* **30**, 319–329.
8. Colb, M., Yang-Feng, T., Francke, U., Mermer, B., Parkinson, D. R. & Krontiris, T. G. (1986) *Nucleic Acids Res.* **14**, 7929–7937.
9. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
11. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
12. Maniatis, T., Jeffreys, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1184–1188.
13. Bencini, D. A., O'Donovan, G. A. & Wild, J. R. (1984) *Biotechniques* **2**, 4–5.
14. Hisajima, H., Nishida, Y., Nakai, S., Takahashi, N., Ueda, S. & Honjo, T. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2995–2999.
15. Sun, L., Paulson, K. E., Schmid, C. W., Kadyk, L. & Leinwand, L. (1984) *Nucleic Acids Res.* **12**, 2669–2690.
16. Paulson, K. E., Deka, N., Schmid, C. W., Misra, R., Schindler, C. W., Rush, M. G., Kadyk, L. & Leinwand, L. (1985) *Nature (London)* **316**, 359–361.
17. Rosenthal, N., Kress, M., Gruss, P. & Khoury, G. (1983) *Science* **222**, 749–755.
18. Weiher, H., Konig, M. & Gruss, P. (1983) *Science* **219**, 626–631.
19. Max, E. E., Battey, J., Ney, R., Kirsch, I. R. & Leder, P. (1982) *Cell* **29**, 691–699.
20. McHenry, H. M. (1975) *Science* **190**, 425–431.