Volume 17 Number 20 1989 Nucleic Acids Research

# Recombination mediates production of an extrachromosomal circular DNA containing a transposon-like human element, THE-1

Ravi Misra, A.Gregory Matera[1], Carl W.Schmid[1] and Mark G.Rush

Department of Biochemistry, New York University School of Medicine, New York, NY 10016 and [1]Department of Chemistry, University of California, Davis, CA 95616, USA

Received June 2, 1989; Revised and Accepted September 7, 1989

## ABSTRACT

An abundant class of HeLa extrachromosomal circular DNA containing the transposon-like element, THE-1, is shown to arise via site specific recombination. The chromosomal locus from which these circles are derived, however, is single-copy. Northern blot analysis detects homology to two polyadenylated RNAs in HeLa cells. The possible presence of an origin of replication and its role in generating these small polydisperse circles is discussed.

## INTRODUCTION

The 2.3 kilobase (kb) dispersed human repetitive DNA, THE-1, is present in approximately 10,000 copies per haploid genome and has been classified as a proretroviral-like transposable element (1,2,3). THE-1 members (Fig. 1) : a) contain long terminal repeats (LTRs) of about 360 base pairs (bp), each of which begin with TG and end with CA, b) are flanked by 5bp direct repeats, NAPyAC, that presumably represent preferred target site duplications (4), c) are transcribed in both orientations and are homologous to a variety of discrete-length poly $A^+$ RNAs, including one which may represent a full length (1.95kb) transcript (5). THE-1 has inserted into pre-existing non-repetitive transcription units and is capable of providing an LTR specific poly A addition site to at least two of them (5). There exists no recognizable THE-1 homology with retroviral *gag* and *pol* genes. None of the sequenced chromosomal elements contain open reading frames corresponding to identifiable protein products although transcriptional control sites occur at appropriate positions within their LTRs. In addition to the 10,000 copies of 2.3kb THE-1, the haploid human genome also contains about 20,000 solitary LTRs (repeats not associated with the internal THE-1 sequence).

Recently, we identified 1.9kb THE-1 homologous circular DNAs in HeLa cells, the first finding of discrete-length small extrachromosomal DNAs in humans(6). Most eukaryotic cells contain a population of small polydisperse

circular (spc) DNAs ranging in size from a few hundred to many thousands of base pairs (7,8,9). THE-1 specific circles are detected as a sharp band on Southern blots of total HeLa circular DNA and are present, on average, at about 5 copies per cell. Furthermore, *Bam* HI cleavage of this DNA reveals clones that are circular permutations of the linear chromosomal sequence, all of which are related to a single family member (6). Interestingly, the cloned species contain approximately 500bp of sequence that is unrelated to consensus THE-1 elements (see "5R unique", Fig. 1).

To determine the relationship between these extrachromosomal THE-1 circles and the genomic locus or loci from which they were derived, the 5R unique region of HeLa 5R was used as a hybridization probe. Genomic clones from a human placental library and Southern blots of various human DNAs reveal that a single locus is responsible for production of the HeLa 5R circles. Additional experiments detect homology with two poly $A^+$ RNAs in HeLa cells and a comparison of the sequence of this single-copy chromosomal element with its relatively abundant extrachromosomal homologue suggests a recombinational mechanism for these THE-1 related rearrangements.

## MATERIALS AND METHODS

Basic materials and procedures.

All procedures such as restriction endonuclease digestions, $^{32}P$-labeling, Southern blot analyses, plasmid or M13 subcloning, plaque screening and DNA fragment purification were performed using standard methods as previously described (6). High molecular weight DNAs isolated from a CEM T-cell lymphoma cell line, an HTLV-1 containing T-cell line, a human melanoma, and peripheral mononuclear blood cells were generous gifts of Drs. J. Krolweski, R. Della-Favera, D. Zouzias and A. Shih, respectively. A 900bp v-*fos* gene fragment was obtained from Dr. M. Greenberg.

Genomic library screening.

A *Sau* 3A partial digest genomic library of human placental DNA cloned into λ EMBL4 was obtained from Dr. John Hall. Approximately $2.5 \times 10^6$ plaques were screened with a 425bp circular DNA specific *Taq* I/*Sph* I fragment (5R unique, Fig. 1). Thirteen positive clones were identified and isolated following three rounds of screening. After stripping (60% formamide, 0.75M NaCl, 0.075M trisodium citrate, 0.02% polyvinyl-pyrolidone, 0.02% Ficoll at 70° C) the tertiary filters were re-probed with a 1.9kb *Bgl* II fragment from λTHE-1A

(1). As discussed in **RESULTS** , all thirteen clones hybridized to both probes. Isolated plaques were picked and bacteriophage DNA was prepared by standard procedures (10).

Mapping of genomic clone.

Lambda clone 4 was chosen for detailed restriction mapping. Five different subcloned probes (see Fig 5) were used: LTR, a 0.4kb *Hind* III/*Rsa* I fragment from a solitary LTR clone $pBRO_4$ (11); .55 Pvu, a 0.55kb *Pvu* II fragment from THE-1A (1); ND-11, a 0.4kb *Pst* I fragment from THE-1B (1); .51 Hae, a 0.51kb *Hae* III fragment from THE-1A (1); 5R unique, a 0.42kb *Taq* I/*Sph* I fragment from HeLa 5R (Fig. 1 and (6)). In addition to phage and plasmid subclone mapping, fragments were isolated and submapped, where necessary, to position restriction sites.

DNA Sequencing.

Sequencing of single-stranded M13 cloned DNAs was performed by the dideoxy, chain-terminator method using either the Klenow enzyme (12) or a modified T7 DNA polymerase ("Sequenase," USB -- the reagents and protocol of the manufacturer were followed). Termination sequencing was also performed using double-stranded templates such as supercoiled plasmid, total phage DNA and isolated restriction fragments. For linear, double-stranded sequencing, 1 to 3 μg of DNA was boiled for 3 minutes in the presence of an oligonucleotide primer and sequencing buffer in a total volume of 10 μl. The mixture was then chilled quickly to 0° C and used immediately.

Oligonucleotide primers were synthesized by Dr. B. Goldschmidt using an Applied Biosystems 380A DNA sythesizer. Nine different primers ranging in size from 15-18 nucleotides were designed from the sequence of HeLa 5R to ensure extensive overlap. In regions of particular importance, sequencing was performed on both strands. All sequences were obtained from multiple gel readings

Northern Blot Analysis.

Total and poly $A^+$ RNA were prepared by the guanidinium isothiocyanate procedure (10), electrophoresed through 1% agarose, 6.6% formaldehyde, transferred to nitrocellulose and hybridized to $^{32}P$-labeled probes corresponding to THE-1A or the 5R unique region (see above). This fragment is not homologous to consensus THE-1 sequences (6). Prehybridization was performed at 42° C in 0.75M NaCl, 0.075M trisodium citrate, 20mM sodium phosphate, 0.2% (w/v) polyvinylpyrolidone-360, 0.2% Ficoll-400, pH 6.8, 50μg/ml

sonicated salmon sperm DNA for 5hrs and hybridization (16 hrs.) in a similar solution but containing 50% formamide and one tenth the prehybridization concentrations of polyvinylpyrolidone and Ficoll. Filters were washed extensively at 65° C in 0.08M NaCl, 0.008M trisodium citrate, 0.1% sodium dodecyl sulfate.

## RESULTS

Copy number of HeLa 5R unique sequence in human chromosomal DNA.

As shown in Figure 1, an abundant class of THE-1 containing circular DNA (HeLa 5R) is characterized by a 500bp region (5R unique) that is apparently unrelated to the THE-1 family. Hence, a 425bp fragment of this region (see **METHODS**) was used to probe Southern blots and a genomic library of human DNA. The Southern analysis (Fig.2) was performed using seven different human DNAs and demonstrates that the circular DNA specific

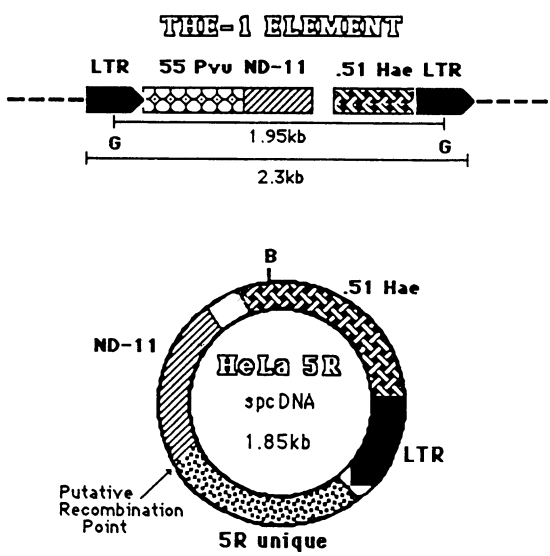



**Figure 1.** Schematic comparison of a consensus chromosomal THE-1 element and an abundant cloned extrachromosomal circular DNA. Top: Linear genomic 2.3kb THE-1 element showing the relative positions of LTRs, flanking chromosomal regions (broken line) and consensus *Bgl* II restriction sites (G). Bottom: HeLa 5R circular 1.85kb extrachromosomal DNA. The single LTR, the non-THE-1 homologous 5R unique 500bp region and the *Bam* HI cloning site (B) are illustrated. The remainder of the internal portion is homologous to about 1.1kb of the element shown above.

fragment has the same genomic structure in each sample and is present at very low (ca. single) copy number. The 2.5kb *Bam* HI band observed here corresponds to the length of the homologous fragment detected in digested genomic λ clones (see Fig. 5). In this experiment, a 900bp v-*fos* gene probe, which shares extensive homology with the single copy human c-*fos* gene, serves as an internal control for copy number (13). Although the autoradiograph was not scanned with a densitometer, it is clear that the c-*fos* hybridizing band is about twice as intense as the circular DNA specific (5R unique) hybridizing band, as would be expected for a single copy chromosomal 5R unique locus.

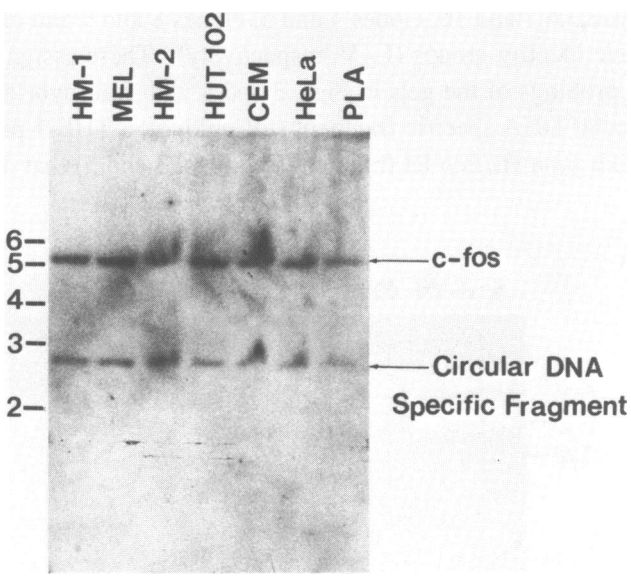
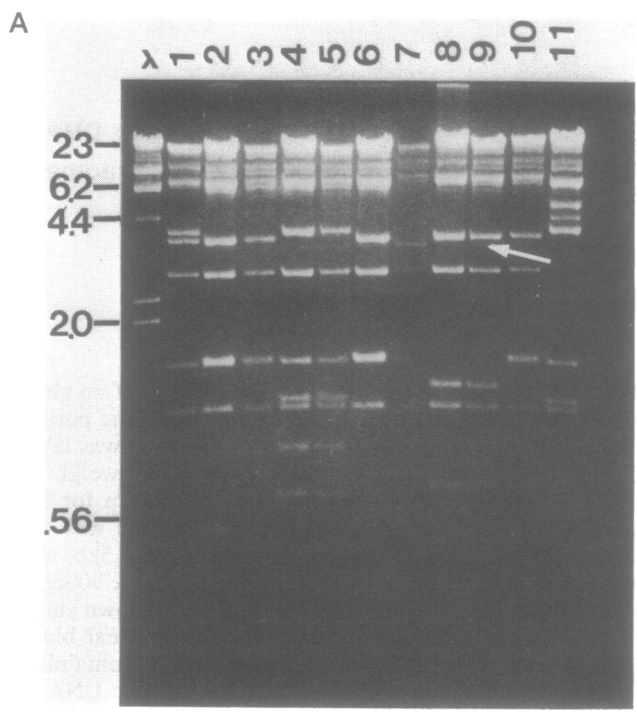



**Figure 2.** Southern blot analysis of the relative copy number of an abundant THE-1 containing circular DNA sequence in various human DNAs. The purified insert (5R unique) from a subcloned region of HeLa 5R (Fig. 1 and text) was labeled and used to probe the indicated DNAs. *Bam* HI digested high molecular weight DNAs (10 μg each) were electrophoresed through 0.8% agarose at 2 volts/cm for 8 hours before transfer to nitrocellulose. The autoradiographic exposure time was 48 hours. The single band detected with the 5R unique probe corresponds to 2.5kb, while the upper band represents an approximately 5kb c-*fos* gene fragment. A 900bp v-*fos* probe was included in the hybridization as an internal control for a known single copy gene. The lanes are: HM-1 and HM-2, human peripheral mononuclear blood cell DNA; MEL, human melanoma DNA; HUT 102, human HTLV-1 containing T-cell line DNA; CEM, T-cell lymphoma line DNA; HeLa, HeLa cell line DNA; PLA, human placental DNA. The sizes of a 1kb ladder (BRL) are indicated on the left.

Genomic clones homologous to HeLa 5R arise from a single locus.

The identification of a probable genomic single-copy sequence within the HeLa 5R extrachromosomal circle allowed for the selection of genomic clones. The 5R unique fragment was used to screen $2.5 \times 10^6$ plaques of a human genomic library and thirteen positive clones were isolated. As anticipated, all thirteen clones also contained THE-1 sequences. A single-copy sequence should be present approximately 20 times in $2.5 \times 10^6$ clones with 20kb inserts. Yet on a random basis, fewer than 1/12 of these clones should contain THE-1 (1,11).

This evident linkage was examined more closely by restriction and Southern analyses of DNAs isolated from eleven of the phage clones (Figs. 3-4). Similar sized restriction fragments are common to all of these DNAs. However, clone 1, clones 2,3,6,7 and 10, clones 4 and 5, clones 8 and 9 and clone 11 appear to represent five identity groups (I - V, respectively). These assignments are confirmed by probings of the gels in Figs. 3 and 4 with four hybridization probes: a circular DNA specific fragment (5R unique), a THE-1 probe, an "adjacent" 3.5kb *Bam* HI/*Eco* RI fragment (see Figs. 3 and 5) that does not probe



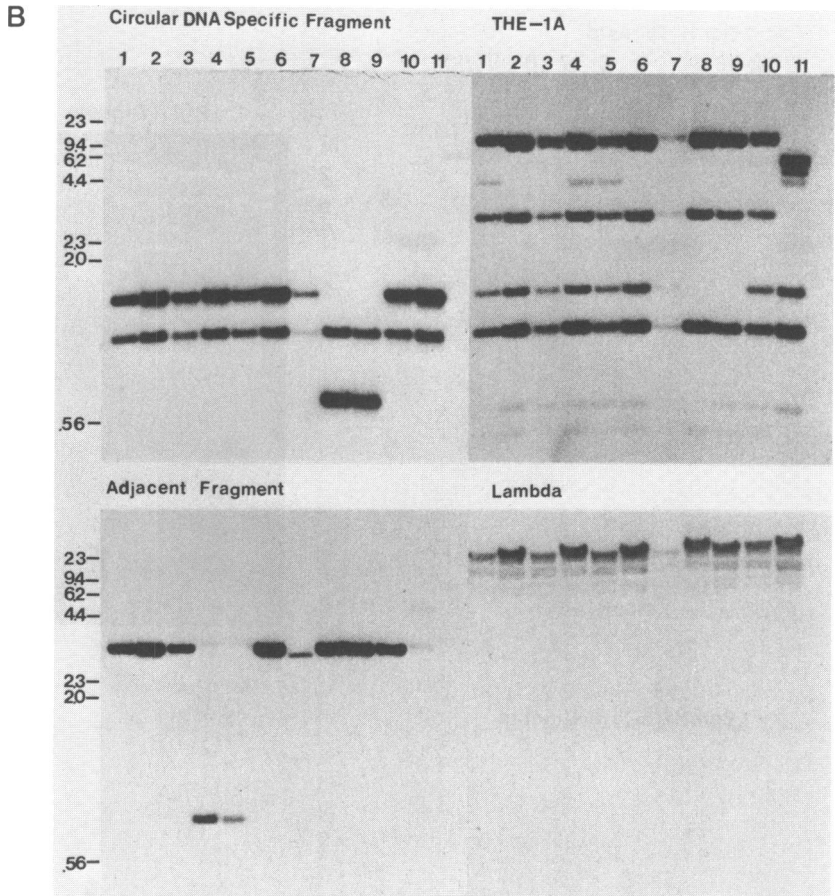
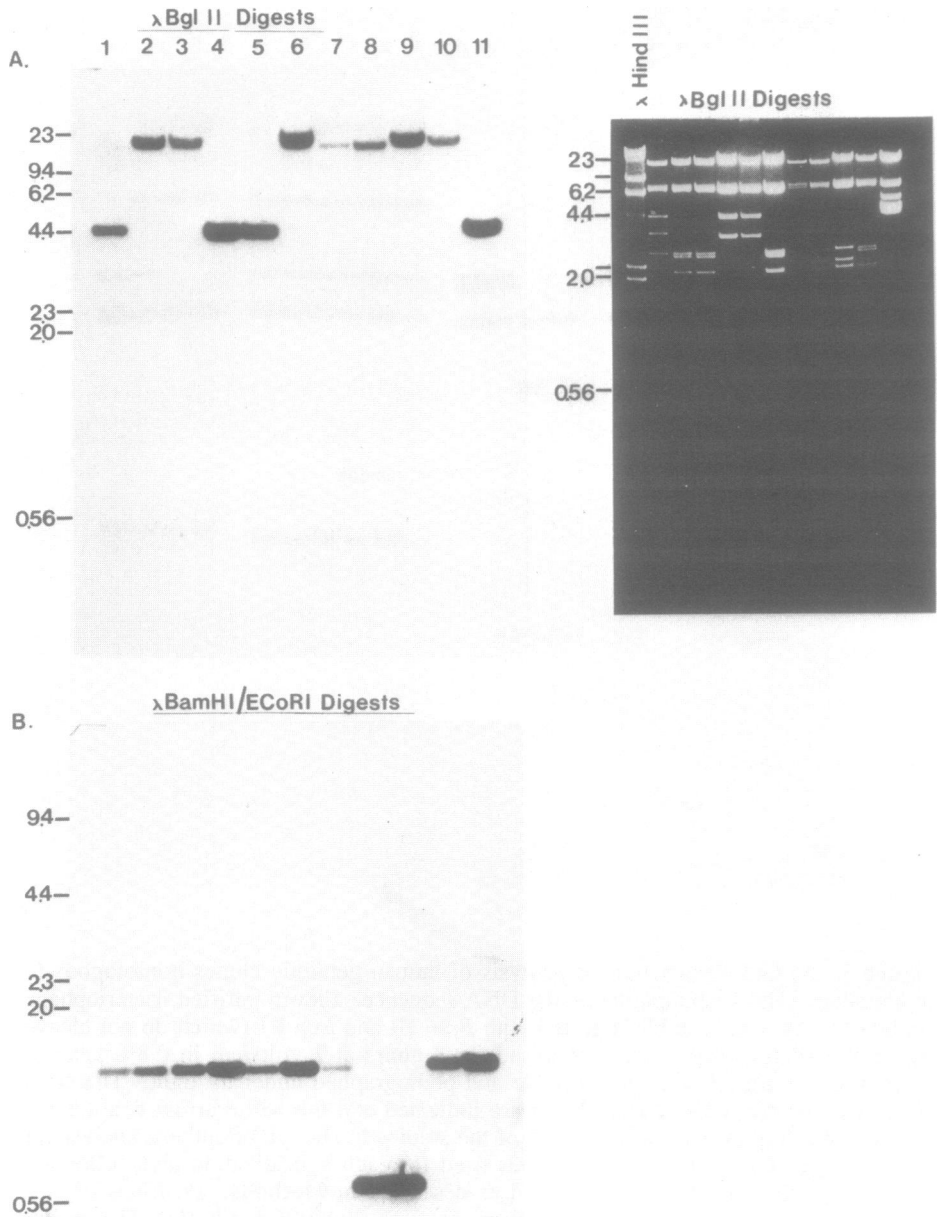



**Figure 3.** A) Gel electrophoresis analysis of human genomic clones homologous to an abundant THE-1 containing circular DNA sequence. Eleven purified bacteriophage lambda DNAs, were double digested with *Bam* HI and *Eco* RI (which do not cleave within the vector arms), electrophoresed for 8 hours at 2 volts/cm in 0.8% agarose containing 0.5 μg/ml ehtidium bromide and photographed under uv light. The sizes of the lambda *Hind* III marker digest are indicated and the white arrow denotes the position of a fragment isolated for one of the studies in (**B**). B) Southern analysis of the gel in part (**A**). The same filter was used for each hybridization and, following autoradiography, signals were removed as described in **Methods.** Positions of size markers are at left and exposure times vary between 20 and 60 minutes. The probes are: Circular DNA Specific Fragment (5R unique), a 425bp fragment from HeLa 5R; THE-1A, a subcloned 1.95kb fragment derived from a genomic THE-1 element; Adjacent Fragment, a 3.5kb band marked with an arrow in part (**A**); Phage lambda, total vector DNA.


λBgl II Digests
A.
1 2 3 4 5 6 7 8 9 10 11
23
9.4
6.2
4.4
2.3
2.0
0.56
λ Hind III
λBgl II Digests
23
6.2
4.4
2.0
0.56
λBamHI/ECoRI Digests
B.
9.4
4.4
2.3
2.0
0.56

within the mapped area, and a 20-mer oligonucleotide (5'AAAAGGGCTGAATCATGCCG3') that lies downstream of recombination point (b) in Fig. 5. To determine the relationship between the circle and its genomic locus, part of lambda clone 4 was mapped and sequenced.

Since the library was constructed from a partial *Sau* 3A digest of chromosomal DNA, a single locus would give rise to many different overlapping clones. These permutations are depicted below the more detailed structure of clone 4 in Fig. 5. Notably, all eleven clones possess an *Eco* RI site within the 5R unique region (Fig. 3), a site which is absent in the HeLa 5R circle. Moreover, the various truncation points are consistent with a single locus. Much of this mapping hinges on the "4kb" doublet in Fig. 3A, clone 1. For example, Identity Group II clones all lack the larger, 3.8kb band of the doublet, as well as the 4.4kb *Bgl* II band (Fig. 4). This places the right endpoint of Group II members within the 0.9kb *Bam* HI/*Bgl* II fragment that is 3' to the "solitary" .51 Hae region (Fig. 5). Additionally, Group III clones lack the shorter 3.5kb band of the aforementioned doublet (the "Adjacent" fragment, Fig.3) placing their left endpoints within this region. Group IV clones lack the 1.5kb *Bam* HI/*Eco* RI band (Fig. 3B). The right endpoint of clones 8 and 9 is thus between the 20-mer and the .51 Hae homolog in Fig. 5. The other truncations all lie outside of the mapped area of clone 4.

Comparison of the HeLa 5R circle and a homologous genomic clone.

A combination of phage, subcloned plasmid and isolated restriction fragment mapping was employed to interpret the structure in Fig. 5. Essentially all of the relevant sequence information is contained within a 4.4kb *Bgl* II fragment. Interestingly, two full-length THE-1 consensus sequences are present in λ clone 4, one of which is immediately adjacent to the circular DNA specific sequence. The other element is situated >10kb upstream of the sequence in Fig. 5 (data not shown). THE-1 members are found in approximately 8% of random

**Figure 4.** Southern analysis of human genomic clones homologous to an abundant THE-1 containing circular DNA sequence using an oligonucleotide probe. Eleven phage clones homologous to HeLa 5R were cleaved with A) *Bgl* II or B) both *Eco* RI and *Bam* HI and electrophoresed as in Fig. 3. In (B) the same filter from Fig. 3 was utilized, after stripping. An end-labeled 20 nucleotide oligomer (see 20-mer in Fig. 5) was used to probe both filters. This oligomer had no homology to either lambda, THE-1A or HeLa 5R, as determined by computer analysis. Positions of size markers are on the left and exposure time was 90 minutes. The panel to the right shows the gel in the *Bgl* II digest under uv light.

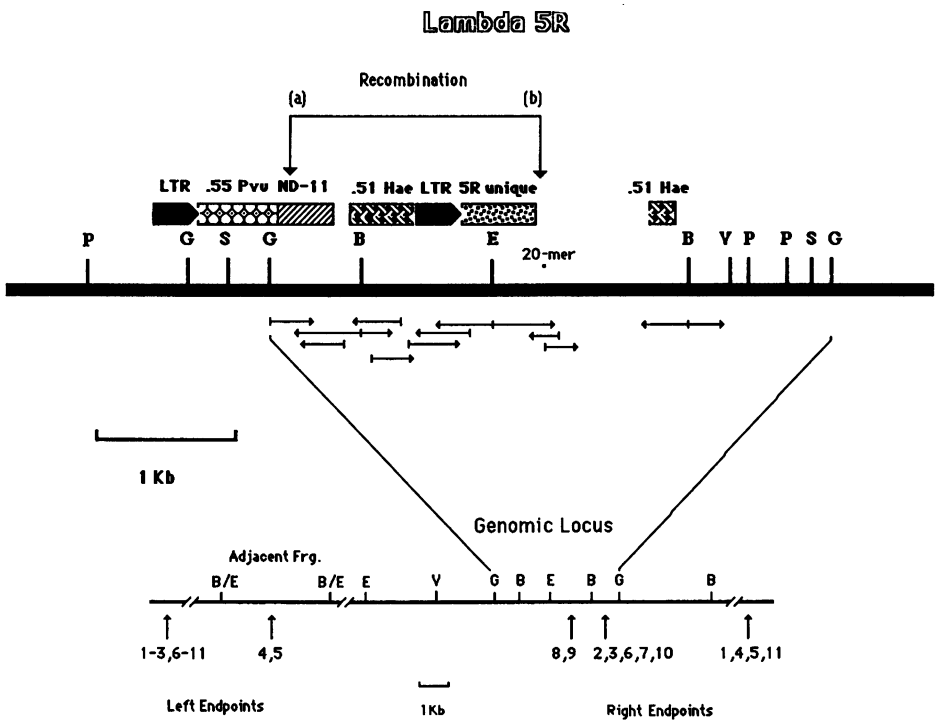



**Figure 5. The Lambda 5R locus. Top: Restriction map of λ clone 4. Boxed regions above the line represent the probes used to deduce the map (see Methods for details). 20-mer marks the position of the oligonucleotide probe used in Fig. 4. Horizontal arrows illustrate the sequencing strategy. Putative recombination sites are indicated at points (a) and (b). Recombination between these positions would generate a circular sequence corresponding to HeLa 5R (Fig. 1). Bottom: A genomic locus schematic of the data in Figs. 3 and 4. Vertical arrows depict the truncations of the various clones, noted below the arrows. The "Adjacent Fragment," used in Fig. 3 is also shown. Restriction sites are: P, *Pst* I; G, *Bgl* II; S, *Sst* I; E, *Eco* RI; B, *Bam* HI; V, *Pvu* II.**

genomic clones; the presence of two THE-1 elements in one clone is not especially noteworthy.

Additionally, the 5R unique region (Fig. 5) is flanked by a region of homology to an internal portion of the THE-1 consensus (.51 Hae). We have not previously identified solitary internal regions (4). Conceivably, this structure is a vestige of a previous THE-1 deletion. The presence of a weakly probing 3.8kb *Bam* HI band in clones 1,4,5 and11 (Fig. 3A, using THE-1A as a probe) raises the possibility of a novel THE-related structure existing downstream of this region.

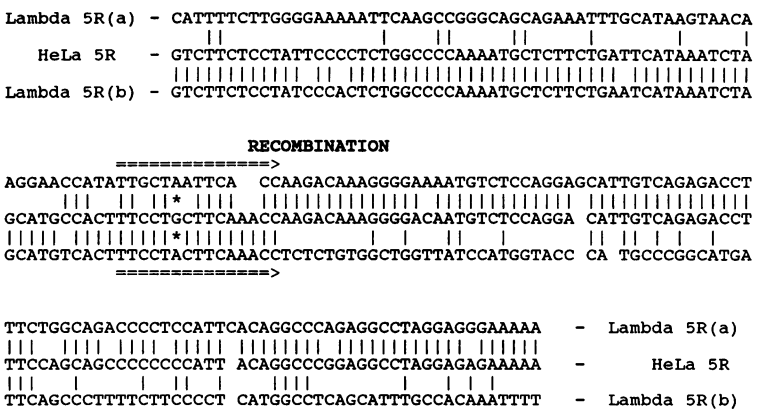



**Figure 6.** Alignment of the regions flanking the putative recombination sites of the chromosomal (points (a) and (b), Fig. 5) and extrachromosomal circular (HeLa 5R, Fig. 1) sequences. Matched nucleotides are bridged by a vertical line. The overlined arrows mark a short direct repeat that occurs precisely at the junction of homology between the circle and the genomic sequences. The * denotes a sequence match within the direct repeat that is not paired with the circle.

However, the obvious possibility that these extrachromosomal circles arise via recombination between the homologs flanking the 5R unique region is disproven by the detailed sequence analysis (see below).

Analysis of 2400bp of λ clone 4 reveals that it contains a homologue of the 1850bp extrachromosomal circular DNA. HeLa 5R can be generated by recombination between sites (a) and (b) in Fig. 5. Approximately 94% sequence similarity is shared between the circle and the chromosomal subclone over this region (data not shown). This relatively large divergence was unexpected, but by confining the comparison to 700 well-resolved nucleotides only 21 veritable differences (3%) were observed. Hence, we attribute half (3%) of the total observed divergence to sequencing artifacts. An alignment of the sequences surrounding the recombination points is compared to the sequence of the HeLa 5R circle in Fig. 6. Note that the regions of homology converge upon an imperfect direct repeat, matching eleven out of thirteen nucleotides plus two unmatched residues (see overlined arrows).

Poly A$^+$transcripts homologous to HeLa 5R.

Transcription of the circular DNA specific fragment (5R unique) was investigated by Northern analysis. HeLa total and poly A$^+$ RNAs were electrophoresed, blotted and hybridized with either a THE-1 probe or the 5R

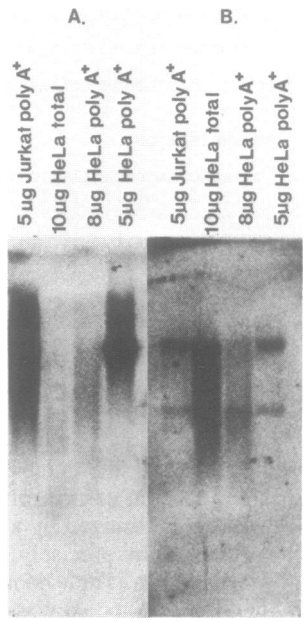



**Figure 7.** Northern analysis of total and polyadenylated RNAs probed with THE-1 and circular DNA specific sequences. The indicated amounts of total or poly A+ RNAs were electrophoresed for13 hours in 1% agarose containing 6.6% formaldehyde at 2.5 volts/cm, transferred to nitrocellulose and hybridized to either A) THE-1A or B) 5R unique probes. The bands in (B) are approximately 2000 and 500nt in length. It should be noted that control experiments detect no homology between the probes. Two different HeLa RNA preparations were analyzed, a rather degraded one in the 8µg lanes and another in the 5µg lanes.

unique fragment. As shown in Fig. 7, both probes detect an approximately 2kb poly A+ species (upper band). The THE-1 homologous material (details of THE-1 transcription are given in (5)) is at least ten fold more abundant than the circular specific RNA. The 5R unique fragment also detects an additional band of about 0.5kb. The precise nature of these transcripts has not yet been determined, but it is clear that HeLa 5R circular DNA and/or its chromosomal homologue are transcriptionally active. Were it not for the data shown in Figs. 5 and 6, the full length transcript could be construed as an ideal intermediate for the generation of HeLa5R by reverse transcription.

## DISCUSSION

The Alu, THE-1 and L-1 families appear to be the most abundant transposable elements in humans (2,3,11,13,14). A variety of characteristics

including the presence of an A-rich 3' tail and variably sized target site duplications have led to the classification of Alu and L-1 sequences as retroposons. Most of the evidence supporting the mobility of these elements is based on structural and evolutionary considerations, but recent data strongly suggest that at least some members of these families are responsible for insertional (transpositional) mutagenesis (3,16-20). The presence of LTRs and uniformly sized short direct repeats has led to the assignment of THE-1 as a proretroviral-like transposable element (1,2).

Our previous studies identified a relatively abundant extrachromosomal circular DNA related to a single member of the THE-1 family (6). Since extrachromosomal species are expected intermediates in retrotransposition, studies were undertaken to understand the relationship between this circular element and its chromosomal and transcriptional homologues. Sequences similar to this abundant spc DNA are shown to be present in two discrete poly $A^+$ RNAs and define an apparently unstable chromosomal locus containing a tightly-linked THE-1 element. Southern blot and limited sequence analysis clearly relates this locus to a homogeneous population of small spc DNAs, the first such region so identified in human DNA.

Recombination produces HeLa 5R.

The sequence of HeLa 5R is a permutation of the genomic one and imperfect short direct repeats are found to span the endpoints of homology. Formation of this spc DNA is proposed to involve a recombinational mechanism occurring within these imperfect repeats. It is not known whether these imperfect "patchy" repeats direct or are related to the proposed recombination event, but their presence is certainly worth noting, and suggestive of a function. Such site specific recombination between direct repeats in the histone gene cluster is known to produce extrachromosomal circular DNAs in *Drosophila* embryos (21). Similarly, a portion of an L-1 transposable element, along with adjacent single-copy sequences, was isolated from HeLa circular DNA (22). The L-1 containing spc DNA is also derived by a site specific recombination between flanking direct repeats.

The HeLa 5R clone described here is different, however, in that it is shown to be an abundant species. A somatic mutation creating an spc DNA should be diluted by cell division unless the circle is capable of autonomous replication and/or is generated by a replicative mechanism.

THE-1 may function as an origin of replication.

The sequence of the HeLa 5R circle, while highly related to the chromosomal locus, is still palpably divergent. It is noteworthy that two of the

twelve replication origin enriched clones isolated by Zannis-Hadjopoulos et al. (23) are related to THE-1 (Zannis-Hadjopoulos, pers. comm.). These clones were isolated by a procedure that purportedly enriches $10^3$- to $10^4$-fold for nascent S phase sequences. Futhermore, one of the two THE-1 related clones is capable of autonomous episomal replication (24). The 6% divergence, mentioned above, between the HeLa 5R circle and the Lambda 5R genomic sequence is too great to be attributed solely to sequencing error. Using only well-resolved sequence and reanalyzing previous data (6), we estimate that nearly 3% of these differences are artifacts. This fact, coupled with the copy number determination of 5 circles per HeLa cell makes for an interesting scenario in which these circles are maintained by the cell and have accumulated point mutations. This intriguing and testable proposition suggests itself for future investigation. On the other hand, it is possible that since the cloned circular DNA was obtained from HeLa cells, and the genomic clones from a placental library, the observed differences could reflect polymorphisms between individuals.

The possibility of a replication origin within HeLa 5R is consistent with the "onion-skin" model proposed for the generation of SV40 extrachromosomal circles (25). It is now documented in several instances, including the present case, that spc DNAs can arise through simple recombination between flanking direct repeats (7,21,22,26,27). Irrespective of the precise mechanism, some reproducible event must be involved in generating this particular THE-1 related circle. This reproducible event may occur at a single genomic locus, or at the extrachromosomal level, the latter implying self-replication. Finally, it should be pointed out that neither of these possibilities is mutually exclusive.

## ACKNOWLEDGEMENTS

We thank Mr. Horace Lozina and Mr. Heinz Annus for expert technical assistance. This work was supported by National Science Foundation Grant PCM8203006(MGR), USPHS NIH grant GM21346(CWS), NIH training grants 5T32-GM07238-08 (RM) and 2T32-GM07377-10 (AGM) , and a University of California Biotechnology grant (AGM).

## REFERENCES

1. Paulson, K.E., Deka, N., Schmid, C.W., Misra, R., Schindler, C.W., Rush, M.G., Kadyk, L., and Leinwand, L. (1985) Nature **316**: 359-361.
2. Weiner, A.M., Deininger, P.L. and Efstradiatis, A. (1986) Ann. Rev. Biochem., **55**: 631-661.
3. Schmid, C.W., Deka, N. and Matera, A. G. in Adolph, K.W. (ed.) **Chromosomes: Eukaryotic, Prokaryotic and Viral**, CRC Press, Boca Raton, in press.

4. Deka, N.,Willard, C.R., Wong, E. and Schmid, C.W. (1988) Nucleic Acids Res. **16**: 1143-1151.
5. Paulson, K.E., Matera, A.G., Deka, N. and Schmid, C.W. (1987) Nucleic Acids Res., **15**: 5199-5215.
6. Misra, R., Shih, A., Rush, M.G. Wong, E. and Schmid, C.W. (1987) J. Mol. Biol. **196**: 233-243.
7. Rush, M.G. and Misra, R. (1985) Plasmid **14**: 177-191.
8. Stanfield, S.W. and Helinski, D.R. (1986) Nucleic Acids Res. **14**: 3527-3538.
9. Fujimoto, S., Tsuda, T., Toda, M. and Yamagishi, H. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**: 2072-2076.
10. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) **Molecular Cloning, a Laboratory Manual**, Cold Spring Harbor Press, Cold Spring Harbor.
11. Sun, L., Paulson, K.E., Schmid, C. W., Kadyk, L., and Leinwand, L. (1984) Nucleic Acids Res., **12**: 2669-2690.
12. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A.**74**: 5463-5467.
13. Curran, T., Peters, G., VanBeveren, C., Teich, N.M. and Verma, I.M. (1982) J. Virol. **44**: 674-682.
14. Rogers, J.H. (1985) Int. Rev. Cytol. **93**: 188-271.
15. Skowronski, J., Fanning, T.G. and Singer, M.F. (1988) Mol. Cell. Biol. **8**: 1385-1397.
16. Willard, C.R., Nguyen, H.T. and Schmid, C.W. (1987) J. Mol. Evol. **26**: 180-186.
17. Economou-Pachnis, A. and Tsichilis, P.N. (1985) Nucleic Acids Res. **13**: 8379-8387.
18. Wiginton, D.A., Kaplan, D.J., States, J.C., Akeson, A.L., Perme, C., Bilyk, I.J., Vaughin, A.J., La Hier, D.L. and Hutton, J.J. (1986) Biochemistry **25**: 8234-8244.
19. Trabuchet, G., Chebloune, Y., Savatier, P., Lachner, J. Faure, C., Verdier, G. and Nigon, V.M. (1987) J. Mol. Evol. **25**: 288-291.
20. Kazazian, H.H., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G. and Antonarakis, S.E. (1988) Nature **332**: 164-166.
21. Pont, G., Degroote, F. and Picard, G. (1988) Nucleic Acids Res. **16**: 8817-8833.
22. Jones, R.S. and Potter, S.S. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 1989-1993.
23. Zannis-Hajopoulos, M., Kaufmann, G., Wang, S-S., Lechner, R.L., Karawya, E., Hesse, J. and Martin, R.G. (1985) Mol. Cell. Biol. **5**: 1621-1629.
24. Frappier, L. and Zannis-Hadjopoulos, M. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**: 6668- 6672.
25. Botchan, M., Topp, W. and Sambrook, J. (1979) Cold Spring Harbor Symp. on Quant. Biol. **XLIII**: 709-719.
26. Krolewski, J.J. and Rush, M.G. (1984) J. Mol. Biol. **174**, 31-40.
27. Fujimoto, S. and Yamagishi, H. (1987) Nature **327**, 242-243.