
Discrete size classes of monkey extrachromosomal circular DNA containing the L1 family of long interspersed nucleotide sequences are produced by a general non-sequence specific mechanism

Christian W.Schindler and Mark G.Rush

Department of Biochemistry, New York University School of Medicine, 550 First Avenue, New York, NY 10016, USA

Received 9 August 1985; Revised 21 October 1985; Accepted 22 October 1985

ABSTRACT

The L1 family of long interspersed nucleotide sequences (LINES) has recently been identified and characterized in the small polydisperse circular DNA (spc-DNA) populations of monkey (1), human (2) and mouse (3) cells. In monkey spc-DNA, the L1 (also known as Kpn I) family is present in discrete size classes (ranging from 300 to 6000 base pairs (bp)) which appear to be generated by non homologous recombination events within chromosomal elements. In this communication it is shown that different regions of the consensus L1 family are present at different frequencies in monkey spc-DNA (as they are in chromosomal DNA), that all regions of the family are present in extrachromosomal DNA, and that each region appears to be represented in an identical discrete spc-DNA size distribution. This size distribution reflects a non-sequence specific mechanism that generates spc-DNA size classes by chromosomal DNA recombination events that are in some way constrained to occur between sites separated by relatively defined lengths.

INTRODUCTION

The L1 family of LINES known as Kpn I in primates and Mif, Bam, or R in rodents is by far the most abundant mammalian long interspersed repeat (for review and references see (4) and (5)). In primates the L1 (Kpn I) family contains about 40,000 members per haploid genome, and is represented not only by full length, 6 kilobase pair (Kb) elements, but also by a greater number of moderately or severely truncated derivatives. Most family members contain polydeoxyriboadenylic acid (polydA) or dA rich regions at their 3' ends, and are bounded in their entirety by relatively short direct repeats (DRs) which have been shown in some cases (and are assumed in others) to represent sites that were duplicated during L1 dispersion. These structural characteristics along with the observation that truncated elements generally lack regions of variable length corresponding to their 5' ends, has resulted in the classification of L1 LINES within the group of transposable elements known as Retroposons or Retrogenes, in analogy with processed pseudogenes and mammalian families of short interspersed nucleotide sequences (SINES) (4-7).

Thus, LINES, SINES, and processed pseudogenes are all assumed to represent at least the end products of transposition events that are mediated by the reverse transcription of an element specific RNA. In the case of the primate Kpn I family, putative RNAs that could serve as transposition intermediates have yet to be indentified, although the presence of low amounts of discretely sized transcripts homologous to Kpn I probes have been reported by some investigators (4,8,9).

The name for the Kpn I repetitive family derives from a characteristic pattern of bands, 1.2, 1.5, 1.8, 2.7, 3.3, and 4.5 Kb, visualized following gel electrophoresis of Kpn I restriction endonuclease digested primate DNAs (10,11). It has been shown that there are four Kpn I sites within a consensus intact (6 Kb) element and that these sites divide the LINE into three fragments with sizes of about 1.8, 1.5, and 1.2 Kb which are ordered, as written, in the 5' to 3' direction (see top of Fig. 1). In addition to the occasional loss of one or more of these four consensus restriction sites and the extensive 5' truncation of others, the Kpn I family also appears to be rather susceptible to sequence rearrangements such as inversions, deletions, and more complex permutations (7,12-14). A compiled sequence of the entire Kpn I repeat has been assembled by M. Singer, J. Skowronski, R. Thayer, and S. Contente and is available from them by request (4).

Recently, sequences homologous to the L1 family were identified in, and cloned from, the spc-DNA populations of monkey (1), human (2) and mouse (3) cells. In the cases of the monkey and human studies it was shown that Kpn I family specific spc-DNAs represented circular permutations of less than full length regions derived from linear chromosomal elements. Moreover analysis of five cloned monkey Kpn I spc-DNAs containing only Kpn I sequences revealed that the recombination events giving rise to them were apparently random and, non homologous, while the analysis of one cloned human Kpn I spc-DNA revealed that the recombination event giving rise to it apparently involved an homologous recombination between two 9 bp repeats (one within the element and the other slightly downstream from its 3' end).

The present interest in spc-DNA containing LINES is due primarily to the fact that extrachromosomal circular and linear DNAs of defined size and structure have been identified for other groups of eucaryotic transposable elements such as retroviral like transposons (15-21), classical transposons (22,23), and SINES (24), and to the fact that at least some of these extrachromosomal species have structures that are consistent with a function as intermediates and/or byproducts of transposition.

In this communication we present a further analysis of Kpn I family containing spc-DNAs isolated from the BSC-1 established cell line of African green monkey kidney cells. It is shown that a) the 3' and 5' ends of the family are present at the same relative frequency, about five to one respectively, in both chromosomal and extrachromosomal DNAs, b) defined regions of the family are not present in unique spc-DNA size distributions (the 5' and 3' ends of the family are apparently present in identical size classes), and c) the regions of the family present in individual cloned circular DNAs reflect the cloning procedure. In short, these Kpn I family containing spc-DNAs appear to be representative of a major class of monkey extrachromosomal DNAs that are produced by a sequence independent mechanism.

MATERIALS AND METHODS

All materials and procedures were as previously described (1). Probes used in the Southern blotting experiments described in RESULTS were prepared from subcloned fragments of the bacteriophage lambda monkey DNA library members described in Fig. 1. (The BSC-1 chromosomal DNA partial Eco RI library was a generous gift from Dr. Maxine Singer (25)). λ Kp-3 contains a 5' truncated Kpn I element while λ Kp-9 appears to contain an intact one. The location of the individual probes was confirmed by numerous hybridization studies against both restriction endonuclease digested chromosomal DNA (data not shown), and against cloned Kpn I element fragments that had been sequenced and compared to a compiled Kpn I sequence (1). Interestingly, sequences present in the 1000 bp fragment derived from the 5' non Kpn I family containing region of λ Kp-9 were found to be located in a similar if not identical position in two other λ Kp (lambda Kpn I element containing) clones (data not shown). (This observation has not been pursued beyond demonstrating that fragment "1000" is dispersed in the BSC-1 genome and that it lacks homology to all regions of the consensus Kpn I LINE, the Alu family of short interspersed nucleotide sequences, and the α -satellite family of tandemly repeated sequences. The significance of its possible association with the Kpn I family requires further investigation.)

RESULTS

Background

Extrachromosomal polydisperse circular DNAs have been isolated and cloned from a wide variety of eucaryotic sources. They all appear to be derived from, or related to, chromosomal DNA and in many cases specific abundant

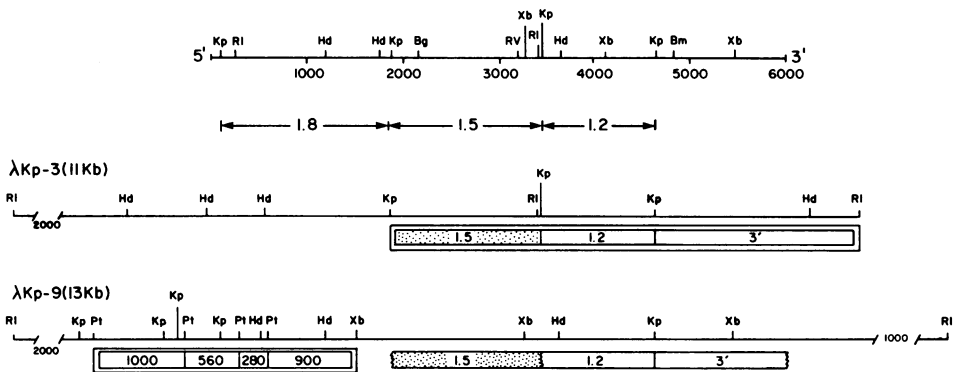


FIG. 1. Restriction endonuclease site maps and homology relationships of two African green monkey bacteriophage lambda library clones containing Kpn I family members. In order to facilitate a comparison between these maps they have been presented in relation to a 6000 bp consensus map of the Kpn I LINE (top), which also indicates the locations of the 1.2, 1.5 and 1.8 Kb, Kpn I canonical fragments. In each map, restriction endonuclease sites are indicated above the horizontal line, while regions of homology of each cloned DNA to seven specific Kpn I family probes are indicated by boxes below this line. Seven fragments, "1.5", "1.2", "3'", "1000", "560", "280", and "900", derived from these clones (shown by double lined boxes) were subcloned, purified, labeled, and used as probes in Southern blot analyses. Boxes with wavy ends indicate that the precise end of a given probe homology cannot be mapped clearly, but that the probe hybridizes to all of the restriction fragments that it overlaps. For example, in clone λKp-9, the "1.5" probe hybridizes in Southern blots to both the Xb-Xb and Xb-Hd fragments with about 80% of the hybridization intensity residing in the Xb-Xb fragment. Boxes with straight ends coinciding with a restriction site indicate that the end of a given probe homology can be mapped by inference, due to the presence of an expected restriction endonuclease site. The numbers 2000 and 1000 in clones λKp-3 and 9 indicate the presence of an additional 2000 or 1000 nucleotides in those regions of the maps. Restriction enzyme sites are indicated by the following letter codes: Bg, Bgl II; Bm, Bam HI; RI, Eco RI; RV, Eco RV; Hd, Hind III; Kp, Kpn I; Xb, Xba I; Pt, Pst I.

members of the population have been shown to be related to repetitive and/or transposable elements (for review and additional references see (1-3), (24), (26-31)). In BSC-1 cells total spc-DNA, a) is present at levels of between 1000-2000 molecules per stationary phase cell, b) ranges in size from a few hundred to thousands of bp, c) can be resolved on polyacrylamide gels into ethidium bromide staining size (abundance) classes of about 0.3, 0.8, 1.2, 1.5 and perhaps 1.8 Kb (29), and d) contains sequences homologous to both unique and repetitive chromosomal DNA (1,24,26).

As previously noted, Kpn I family regions have been identified on

discrete sizes of BSC-1, spc-DNA, predominantly in the abundance classes, and cloned Kpn I containing circular species were found to be circular permutations of the linear chromosomal sequence (1). In addition, sequencing studies of five cloned molecules revealed that apparently non homologous sites within the linear element were joined (recombined) to form the circles, and that no additional (non Kpn I family) nucleotides were introduced during circularization. In contrast to studies of a variety of other repetitive DNA containing spc-DNAs (see DISCUSSION), the properties of the Kpn I, spc-DNAs were not particularly consistent with, or supportive of, their role as intermediates and/or byproducts of either transposition or reproducible chromosomal rearrangements. However, since the size distribution studies were based on Southern blot analyses of total spc-DNA using a probe derived from predominantly the 1.5 Kb Kpn I family fragment, and since the characterized clones were derived from an Eco RI generated spc-DNA library, the possible bias of our previous results has been investigated.

Distribution of Four Regions of the Kpn I Family in Electrophoretically Resolved BSC-1, spc-DNA

Radioactively labeled probes representing four non overlapping regions of the monkey L1 family (see MATERIALS AND METHODS) were hybridized to Southern blotted restriction endonuclease digested and undigested (covalently closed) BSC-1, spc-DNAs. The results, shown in Fig. 2, indicated that each of the four probes was present in similar if not identical patterns in the covalently closed circular DNAs but that 5' sequences were significantly underrepresented. The similarity in the patterns obtained following Eco RI and Bam HI digestions of the spc-DNAs probably reflects the fact that most of them lack the corresponding restriction sites, or contain only one site per molecule. Also shown in Fig. 2 is the result of a Southern blotting experiment in which BSC-1, spc-DNA was probed with BSC-1 α -satellite DNA. (The so-called α -satellite family of primates is found as long arrays of tandemly repeated variants of an approximately 170 bp sequence and constitutes 20% of total African green monkey DNA.) Although the α -satellite probing pattern was known to differ from that of L1 (26) the direct comparison shown here reveals that most of this difference resides in the higher molecular weight range of the spc-DNA population.

It appears that the previously mentioned BSC-1, spc-DNA abundance classes, ranging from 300 to 1800 bp, generally contain rather unselected genomic sequences whose presence as distinct size classes is probably more a reflection of some "general mechanism" of spc-DNA formation than a property

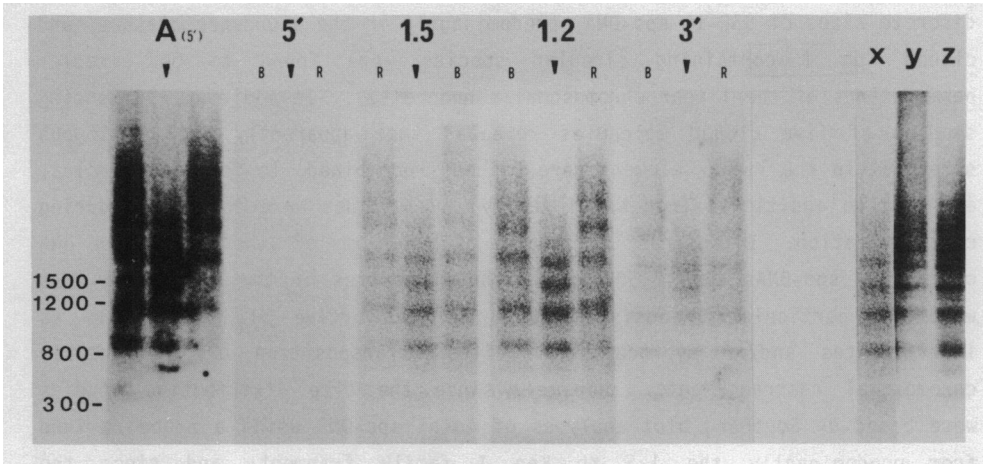


FIG. 2. Comparison of spc-DNA homologies to probes representing different regions of the *Kpn* I family repeat. About 150 ng each of uncut spc-DNA (marked with arrow heads), and *Bam* HI (B) or *Eco* RI (R) restricted spc-DNAs were fractionated on a 1.5% agarose gel, transferred to nitrocellulose, and hybridized with 8.8×10^6 cpm of one of the following *Kpn* I family specific probes: 5' ("900" + "560", 4.0×10^8 cpm/ μ g); "1.5" (3.6×10^8 cpm/ μ g); "1.2" (5.0×10^8 cpm/ μ g); and 3' (4.5×10^8 cpm/ μ g). Autoradiographs were exposed for three days except for panel A which represents a 14 day exposure of the three lanes hybridized to the 5' probe. (On the original autoradiographs the signal intensities of the 3' and 1.2 panels were similar, and that of the 5' panel was faint but clearly visible. The rather low intensities of the 3' and 5' panels shown here are artifacts of the reproduction procedure used to create the composite photograph.) Lanes x, y and z compare the hybridization of 150 ng of uncut spc-DNA with the "1.5" *Kpn* I probe (lane x), or with an α -satellite (monomer) probe (lanes y and z). Lanes y and z contain two different preparations of spc-DNA. The approximate positions of covalently closed circular duplex DNAs of the indicated sizes in bp are shown on the left of the figure.

of the sequences themselves. Interestingly, a very similar abundance class background was also observed when BSC-1, spc-DNA was probed with an *Alu* sequence, although the hybridization pattern in the lower molecular weight range (about 300 bp) was far more intense and diffuse than that observed for either *Kpn* I or α -satellite (30).

Distribution of Two Regions of the *Kpn* I Family in Total BSC-1, spc-DNA

In order to confirm that the 5' end of L1 was underrepresented in both BSC-1 chromosomal (32) and spc-DNAs, a series of comparative dot blot experiments, such as those shown in Fig. 3, were performed. In short, the results of these and other experiments demonstrated that the 3' end of the family was about five fold more frequent in both chromosomal and spc-DNAs than its 5' end. (Interestingly a similar series of dot blot experiments

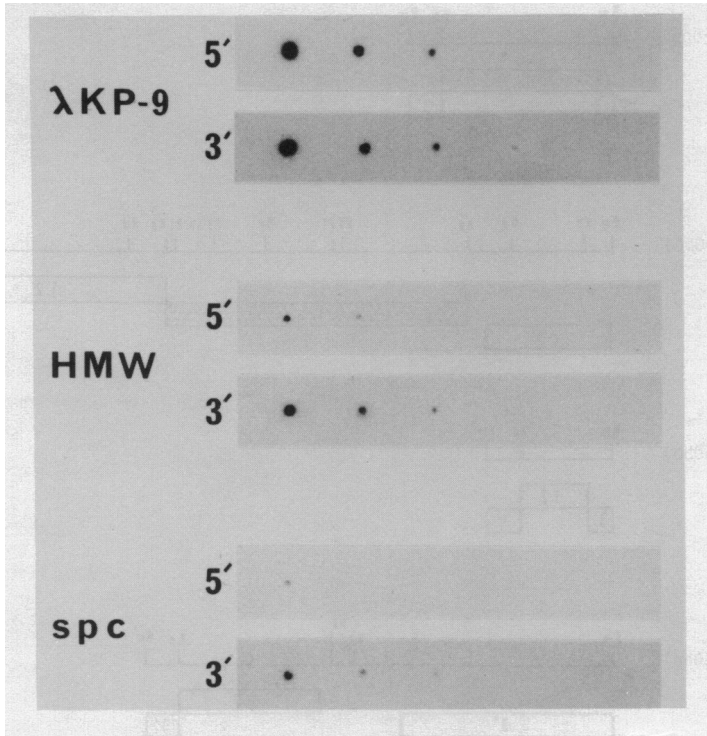


FIG. 3. Comparison of the relative concentrations of the 5' and 3' ends of the *Kpn* I family repeat in BSC-1 high molecular weight (HMW) chromosomal and *spc*-DNAs. λ Kp-9, BSC-1 HMW, and BSC-1, *spc*-DNAs were denatured, serially diluted at ratios of one to four, and spotted in 3 μ l aliquots onto nitrocellulose filters. For λ Kp-9 and BSC-1 HMW DNAs the undiluted dots (at the left) correspond to 4.7 and 6.75 ng of DNA respectively. The precise amount of *spc*-DNA is not known but is between 1 and 3 ng. Each filter was probed with either 3×10^7 cpm of the "3'" (3.8×10^8 cpm/ μ g) or 5' ("900" + "560", 3.8×10^8 cpm/ μ g) probes. Note that the "3'" probe contains about 1.4 Kb of *Kpn* I family sequence homology and the 5' probe about 1.2 Kb.

showed that the 3' end of the family was also about five fold more frequent than its 5' end in BSC-1, cytoplasmic polyA⁺ RNA (R. Misra, unpublished results) a finding similar to that reported for BSC-1, heterogeneous nuclear RNA (4)).

Characterization of Six New *Kpn* I Family *spc*-DNAs

A previous analysis (1) of ten cloned *Kpn* I family circular DNAs isolated from an *Eco* RI generated *spc*-DNA library indicated that all of them were derived primarily from the 1.5 and 1.2 Kb regions of the LINE. However, since the library was screened with 1.5 Kb fragment homologous probes, and since an *Eco* RI site is present near the 1.5-1.2 Kb fragment junction in the

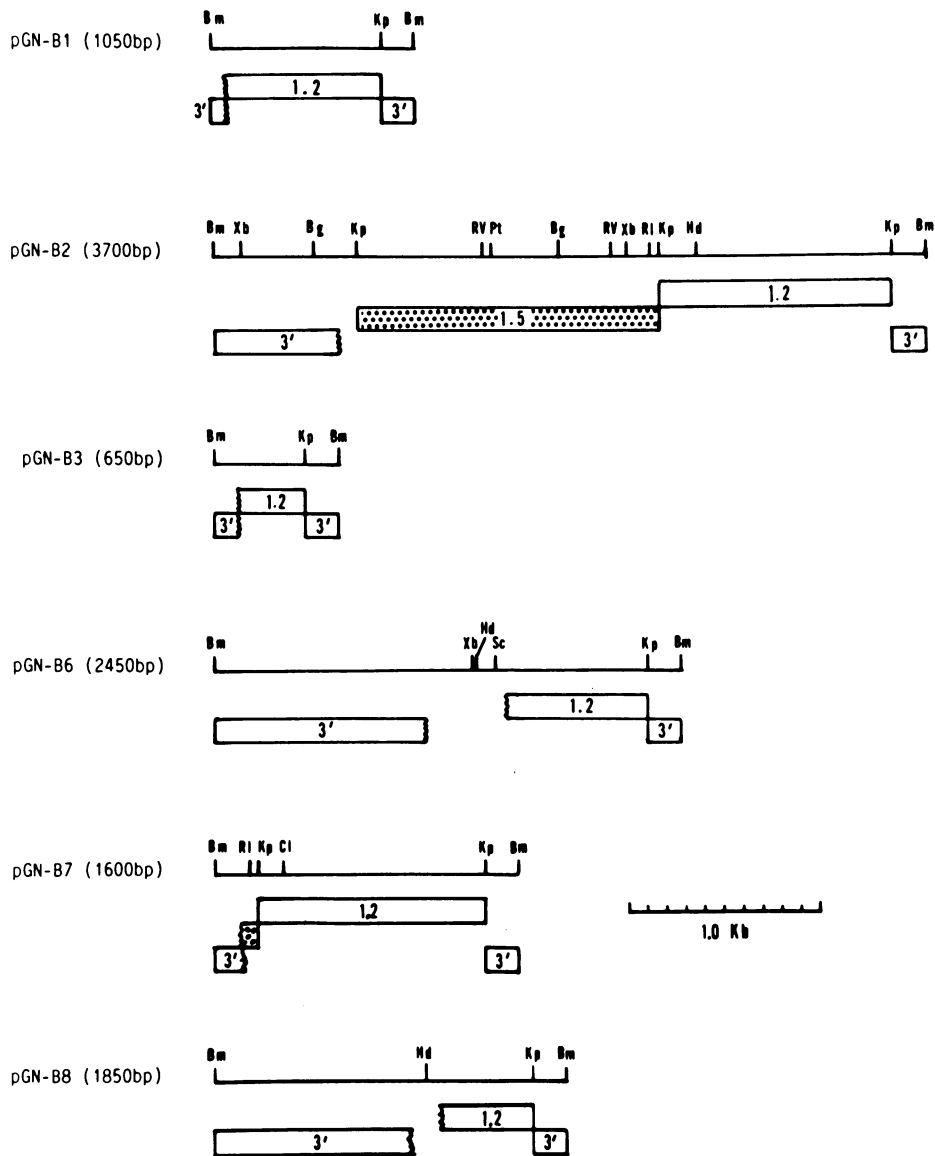


FIG. 4. Restriction maps and homology relationships between six cloned *Kpn* I family-containing spc-DNAs and three defined regions of the *Kpn* I LINE. Blocks of homology within each cloned DNA to three specific *Kpn* I family regions (see Fig. 1 and its legend) are indicated by shaded and clear boxes below the map. These boxed homology regions were determined by Southern blotting, and wavy or straight ends reflect the precision in mapping individual homology blocks (as described in the legend to Fig. 1). Enzyme abbreviations are as for Fig. 1 with the additions of: Sc, *Sac* I; Cl, *Cla* I.

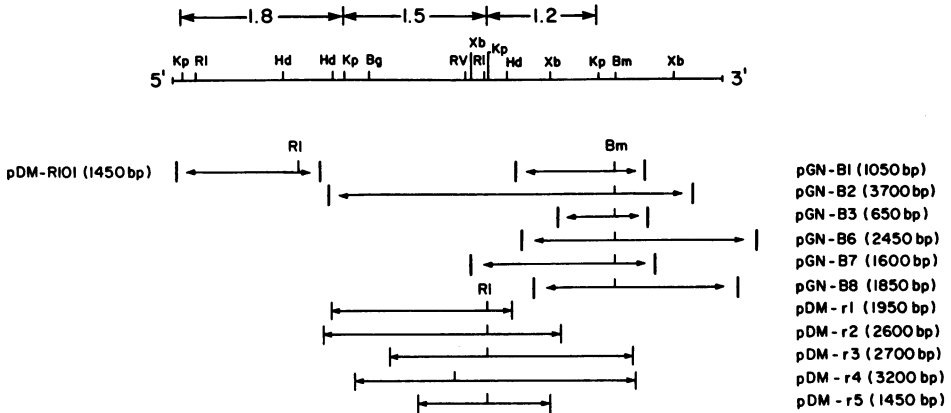


FIG. 5. Orientation of nucleotide sequences in twelve cloned *Kpn* I family spc-DNAs with respect to the consensus chromosomal sequence. Each of these cloned spc-DNAs contains a circularly permuted arrangement of *Kpn* I sequences derived from a defined region of the linear chromosomal element. The relationship of these clones to the *Kpn* I family consensus is indicated by horizontal double headed arrows bounded by vertical lines. The tips of the arrow heads correspond to recombination sites when projected vertically to the *Kpn* I consensus, and the spaces between these tips and their associated vertical lines reflect the imprecision of mapping these sites in the absence of nucleotide sequence data. The restriction site used in cloning each of the circular DNAs is indicated in each clone. The six pGN inserts were isolated from a *Bam* HI library and the six pDM inserts were derived from an *Eco* RI library. Note that the cloning sites for pDM-r4, and pDM-R101 are not in the consensus. The recombination sites of cloned inserts pDM-r1 to r5 were localized by sequencing (1).

consensus L1 (Fig. 1) sequence, it seemed likely that a selected subset of circular species was analyzed. In order to verify this supposition and to determine whether the structures of *Kpn* I containing spc-DNAs derived from other regions of the family were similar to those analyzed previously, a 4,000 member *Bam* HI generated spc-DNA library was screened with a mixture of the seven probes described in Fig. 1, and six positive clones were randomly selected for detailed characterization. The relationships of the cloned inserts to each other, to the *Kpn* I family, and to six *Eco* RI generated spc-DNA library members, were determined by restriction endonuclease digestion and Southern blotting. The results of these experiments are shown in Figs. 4 and 5. (In Fig. 5, the orientations of clones pDM-r1 to r5 are based on published data (1) while the data leading to the orientation of pDM-R101 is unpublished.) The most logical interpretation of the maps presented in Figs. 4 and 5 is that the cloned *Kpn* I containing spc-DNA

inserts were derived from intact circular molecules representing circularly permuted versions of the consensus linear LINE, and that except for clone pGN-B6 all of the nucleotide sequences in the circular species were derived from Kpn I elements. It should also be noted that although localization of the precise ends (recombination sites) of the colinearly oriented spc-DNAs shown in Fig. 5 requires sequence analysis, the boundaries determined by restriction site mapping and Southern blotting are generally quite accurate (1). In summary, the results presented in Figs. 4 and 5 strongly suggest that the specific regions of the L1 family present in cloned BSC-1, spc-DNAs reflect the enzymes used in generating small circular DNA libraries.

DISCUSSION

The experimental results presented in this communication show that a) the 3' and 5' ends of the L1 (Kpn I) family are present at the same relative frequency (5 to 1) in both BSC-1 chromosomal and spc-DNAs, b) different defined regions of the L1 family are present predominantly in the same spc-DNA abundance classes, c) Bam HI linearized L1-spc-DNAs are derived from intact circular molecules that are circularly permuted versions of their linear chromosomal counterparts, and d) all regions of the L1 family can be identified in cloned BSC-1, spc-DNAs. In addition, although the precise ends (recombination sites) of the colinearly arranged Bam generated clones shown in Fig. 5 have not been elucidated, it seems likely that their sequence organization will reflect the same type of non-homologous recombination events that were found for members of the Eco R1 generated clones (1). Taken together, these data strongly suggest that most L1, BSC-1, spc-DNAs are representative of the abundance classes of small circular molecules that can be readily detected following ethidium bromide staining of electrophoretically resolved monkey spc-DNAs (29). The "general mechanism" responsible for the generation of these classes remains to be elucidated, but it will presumably involve sequence independent, predominantly non-homologous chromosomal DNA recombination events that are in some way constrained to occur between sites separated by relatively defined lengths. A relationship of these lengths to the periodic nature of chromatin structure is not obvious from the data, but is at present the most promising potential explanation for their occurrence.

The identification of a relatively large number of chromosomally rearranged primate L1 elements (7,12-14), and the claim that L1 family sequences are enriched in mouse spc-DNA (3), has led to speculations that the

family as a whole is unstable and that extrachromosomal species reflect this instability. It should be noted however, that at present, structural analyses of L1 containing spc-DNAs have not shed light on this speculation, and that quantitation of Kpn I family sequences in spc-DNAs is limited by difficulties in accurate measurements of total spc-DNA concentration. The observation that almost all of the cloned L1 spc-DNAs diagrammed in Fig. 5 contain only L1 family sequences (with the exception of pGN-B6 which may be similar in structure to a cloned human spc-DNA (2)) might be considered supportive of some LINE family directed circularization event, but could also be related to the cloning strategy.

The failure to identify L1 family spc-DNAs of reproducible size and/or structure contrasts markedly with results obtained for some other eucaryotic transposable elements. Specifically, extrachromosomal DNAs of defined size and structure have been characterized for retroviral like transposable elements in Drosophila (16-19, 21), yeast (15), and humans (20), for so called "classical" transposable elements in the round worm C. elegans (22,23), and for the retroposon Alu family of SINES in monkeys (24). Such circular species are assumed to be intermediates and/or byproducts of chromosomal DNA rearrangements, including transpositions, and they are generally part of a larger ill defined population of spc-DNAs. For example, Southern blot analyses of BSC-1, spc-DNA probed with the THE-1 family of retroviral like transposable elements (20) revealed a few intense sharp bands corresponding to circular DNAs with sizes ranging from 3 to 15 Kb, while less intense broad bands corresponding to the BSC-1 abundance classes (0.3 to 1.8 Kb) were also observed (R. Misra and M. Rush, unpublished results). In short, it appears that both general and specific mechanisms are involved in the synthesis of spc-DNAs, and that L1 family containing species are synthesized predominantly by the general pathway.

ACKNOWLEDGEMENTS

We thank Horace Lozina and Heinz Annus for technical assistance, and Dr. Warren Jelinek for many helpful discussions. This work was supported by grant no. PCM-8203006 from the National Science Foundation. C.W.S. is a Medical Scientist trainee supported by NIH training grant 5T32 GM07308 from the National Institute of General Medical Sciences.

REFERENCES

1. Schindler, C.W. and Rush, M.G. (1985) *J. Mol. Biol.* 181, 161-183.
2. Jones, R.S. and Potter, S.S. (1985) *Proc. Natl. Acad. Sci. USA* 82,

- 1989-1993.
3. Fujimoto, S., Tsuda, T., Toda, M. and Yamagishi, H. (1985) *Proc. Natl. Acad. Sci. USA* 82, 2072-2076.
 4. Singer, M.F. and Skowronski, J. (1985) *TIBS*, 10, 119-122.
 5. Rogers, J.H. (1984) *Int. Rev. Cytol.* 93, 231-279.
 6. Sun, L., Paulson, K.E., Schmid, C.W., Kadyk, L. and Leinwand, L. (1984) *Nucl. Acids Res.* 12, 2669-2690.
 7. DiGiovanni, L., Haynes, S.R., Misra, R. and Jelinek, W.R. (1983) *Proc. Natl. Acad. Sci. USA* 80, 6533-6537.
 8. Kole, L.B., Haynes, S.R. and Jelinek, W.R. (1983) *J. Mol. Biol.* 165, 257-286.
 9. Schmeckpeper, B.J., Scott, A.F. and Smith, K.D. (1984) *J. Biol. Chem.* 259, 1218-1225.
 10. Shafit-Zagardo, B., Maio, J.J. and Brown, F.L. (1982) *Nucl. Acids Res.* 10, 3175-3193.
 11. Grimaldi, G. and Singer, M.F. (1983) *Nucl. Acids Res.* 11, 321-338.
 12. Lerman M.I., Thayer, R.E. and Singer, M.F. (1983) *Proc. Natl. Acad. Sci. USA* 80, 3966-3970.
 13. Thayer, R.E. and Singer, M.F. (1983) *Mol. Cell. Biol.* 3, 967-973.
 14. Potter, S.S. (1984) *Proc. Natl. Acad. Sci. USA* 81, 1012-1016.
 15. Ballario, P., Fileticci, P., Junakovic, N. and Pedone, F. (1983) *FEBS Lett.* 155, 225-229.
 16. Flavell, A.J. (1984) *Nature* 310, 514-516.
 17. Flavell, A.J. and Ish-Horowicz, D. (1983) *Cell* 34, 415-419.
 18. Georgiev, G.P. (1984) *Eur. J. Biochem.* 145, 203-220.
 19. Mossie, K.G., Young, M.W. and Varmus, H.F. (1985) *J. Mol. Biol.* 182, 31-43.
 20. 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.
 21. Shepherd, B.M. and Finnegan, D.G. (1984) *J. Mol. Biol.* 180, 21-40.
 22. Rose, A.M. and Snutch, T.P. (1984) *Nature* 311, 485-486.
 23. Ruan, K.S. and Emmons, S.W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 4018-4022.
 24. Krolewski, J.J. and Rush, M.G. (1984) *J. Mol. Biol.* 174, 31-40.
 25. McCutchan, T., Hsu, H., Thayer, R.E. and Singer, M.F. (1982) *J. Mol. Biol.* 157, 195-211.
 26. Bertelsen, A.H., Humayun, M.Z., Karfopoulos, S.G. and Rush, M.G. (1982) *Biochemistry* 21, 2076-2085.
 27. Kunisada, T., Yamagishi, H. and Sekiguchi, T. (1983) *Plasmid* 10, 242-250.
 28. Stanfield, S.W. and Helinski, D.R. (1984) *Mol. Cell. Biol.* 4, 173-180.
 29. DeLap, R.J., Rush, M.G., Zouzias, D. and Khan, S. (1978) *Plasmid* 1, 508-521.
 30. Krolewski, J.J., Bertelsen, A.H., Humayun, M.Z. and Rush, M.G. (1982) *J. Mol. Biol.* 154, 399-415.
 31. Kunisada, T. and Yamagishi, H. (1984) *Gene* 31, 213-223.
 32. Grimaldi, G., Skowronski, J. and Singer, M.F. (1984) *EMBO J.* 3, 1753-1759.