Volume 14 Number 9 1986 Nucleic Acids Research

Analysis of LINE-1 family sequences on a single monkey chromosome

Theresa N.H.Lee+ and Maxine F.Singer

Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892 and +Division of Metabolism and Endocrine Drug Products, Food and Drug Administration, Rockville, MD 20857, USA

Received 16 January 1986; Accepted 31 March 1986

ABSTRACT

The structure of LINE-1 (L1Ca) family members present on African green monkey chromosome CAE-19 is compared with that of the entire set of L1Ca sequences present in the monkey genome. The analysis involved annealing of cloned subsegments of monkey L1 family members to DNA-blots containing restriction endonuclease digests of either total monkey liver DNA or DNA isolated from a monkey/mouse somatic cell hybrid carrying the single monkey chromosome. In addition, L1Ca segments cloned from hybrid cell DNA were characterized by restriction endonuclease mapping and hybridization. The data indicate that, taken as a whole, the set of L1Ca sequences on CAE-19 tends to differ in characteristic ways from the set present in the total monkey genome.

INTRODUCTION

Members of the LINE-1 or L1 family (previously KpnI family in primates and BamH1 or M1F1 in mice) are dispersed 10^4 or more times in mammalian genomes (1-8 and T.F. Fanning, personal communication). At present, the primate (e.g., L1Hs or L1Ca for Homo sapiens and Cercopithecus aethiops, respectively) and mouse (L1Md, Mus domesticus) sequences are best characterized. Individual family members show substantial intraspecies complexity in both length and sequence organization. While some segments are 6 to 7 kbp long and colinear (see Figure 1 for a typical L1Ca family member), the large majority are truncated and generally lack varying amounts of sequence from the end conventionally designated the 5'-end (see reference 9 for review). Some truncated members contain internal rearrangements of the typical sequence, including deletions and inversions (10,11,12,13). No terminal repeats, either direct or inverted, are found within the segments themselves although some family members are flanked by direct repeats that are target site duplications (4,14,15). Many family members have, close to the 3'-end of one strand, a polyadenylation signal followed by an A-rich stretch. Recent evidence suggests that some L1 units contain long open reading frames

and may encode protein (9,11,16,17,29). Typically, the open reading frames end with a translational stop codon several hundred base pairs from the 3'-end.

Like many other families of interspersed repeated sequences, the L1 elements within a species are more similar to one another than they are to L1 sequences in closely related species or to L1 sequences in different mammalian orders (9,12,8,9,20,21). Such families are said to undergo intra-species concerted evolution, or to be homogenized (21,22). This implies a mechanism whereby a relatively small number of "founder" copies correct mutations that collect in individual family members. Such corrections could occur through deletion of family members and insertion of new copies of the founder sequences or by non-reciprocal homologous recombination such as gene conversion or by a combination of both mechanisms. These possibilities have been extensively discussed (e.g. 20,21,22). One question regarding the mechanism(s) of concerted evolution in mammals is whether it homogenizes the L1 sequences on a single chromosome to a greater extent than it does the entire genomic family. Therefore, we analyzed those family members that reside on the single monkey chromosome, CAE-19 {the homologue of human chromosome 17 (23)}, that is present in a mouse-monkey somatic cell hybrid (24). A preliminary report of these results has been presented (25).

EXPERIMENTAL PROCEDURES

Materials

Cell clone 22, a hybrid line containing a single monkey chromosome, CAE-19, in a mouse background (B82 cells, tk^-) was described previously (24); this chromosome contains the TK gene. The hybrid was stored frozen in liquid N_2 and revived and carried in HAT medium (100μM hypoxanthine, 1μM aminopterin, 16μM thymidine, Sigma) for two weeks before DNA was prepared. African green monkey genomic DNA was isolated from liver. Other materials and the preparation of DNA were as previously described (24).

A series of subcloned probes representing different regions within the 6 kbp long L1Ca unit was used in these studies (Figure 1B). The preparation of most of these has been described (4). The plasmids themselves were used as probes in all but one case. The 2.8 kbp insert in pUC2.8 was removed from the plasmid with endonuclease KpnI and purified by gel electrophoresis prior to labeling; this probe is referred to in the text as pUC2.8. Two plasmids which together encompass an entire 6 kbp L1Ca unit, p4A12 (left end) and p4A10 (right end), were subcloned from the λ-recombinant λCaF_1 (4) into pBR322 by G. Grimaldi.

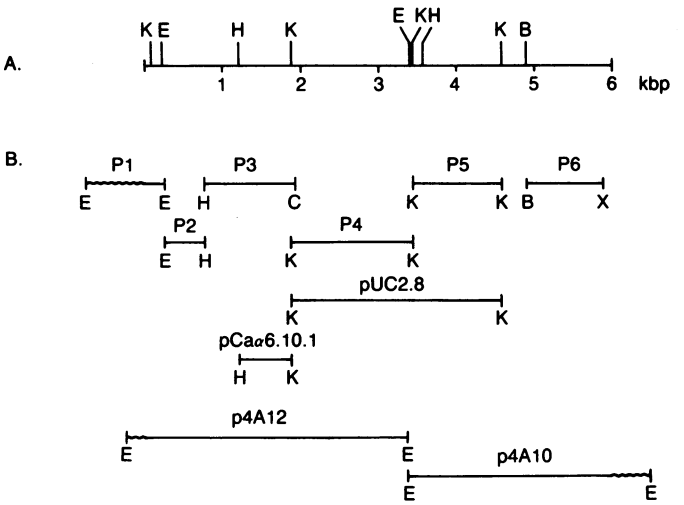

Figure 1. A. Schematic diagram showing a typical L1Ca family member about 6 kbp long. Not all the restriction endonuclease sites shown occur on all family members. The 3'-terminal A-rich stretch found on many family members is at the right.
B. Cloned subregions of the 6 kbp L1Ca unit used as probes in this study. The L1Ca region contained in each plasmid clone is indicated along with the restriction endonuclease sites used for cloning. P1 through P6 were described previously (4) as was pCaα6.10.1 (12). pUC2.8 contains the 2.8 kbp KpnI fragment originally cloned in pCaα6.5 (12) now recloned in pUC8c.2 (kindly supplied by R.E. Thayer). Plasmids p4A10 and p4A12 were subcloned from the genomic segment in λCaF_1 (4) into the EcoR1 site of pBR322. P1, p4A10 and p4A12 each contain several hundred base pairs of low copy number genomic sequences that flanked the L1Ca unit from which they were derived. Restriction endonucleases: K, KpnI; E, EcoR1; H, HindIII; B, BamH1; C, ClaI; X, XhoI.

Methods

All methods were as previously described unless indicated otherwise (4,24). Standard hybridization conditions were 3xSSC (SSC is 0.15 M NaCl, 0.015 M sodium citrate) at 65°. Washing was in 0.1 x SSC at 52°. Restriction endonuclease digests of genomic DNA were monitored for completion by simultaneous digestion of an identical sample to which 0.25µg of λ-phage DNA was added. Complete digestion of the phage DNA was ascertained on ethidium bromide stained gels after electrophoresis.

A genomic library of hybrid cell DNA was prepared in λCharon4A (26). Hybrid cell DNA was partially digested with endonuclease RsaI under conditions designed to give a high yield of approximately 20 kbp fragments. Twenty batches of individually digested DNA were pooled and the fragments

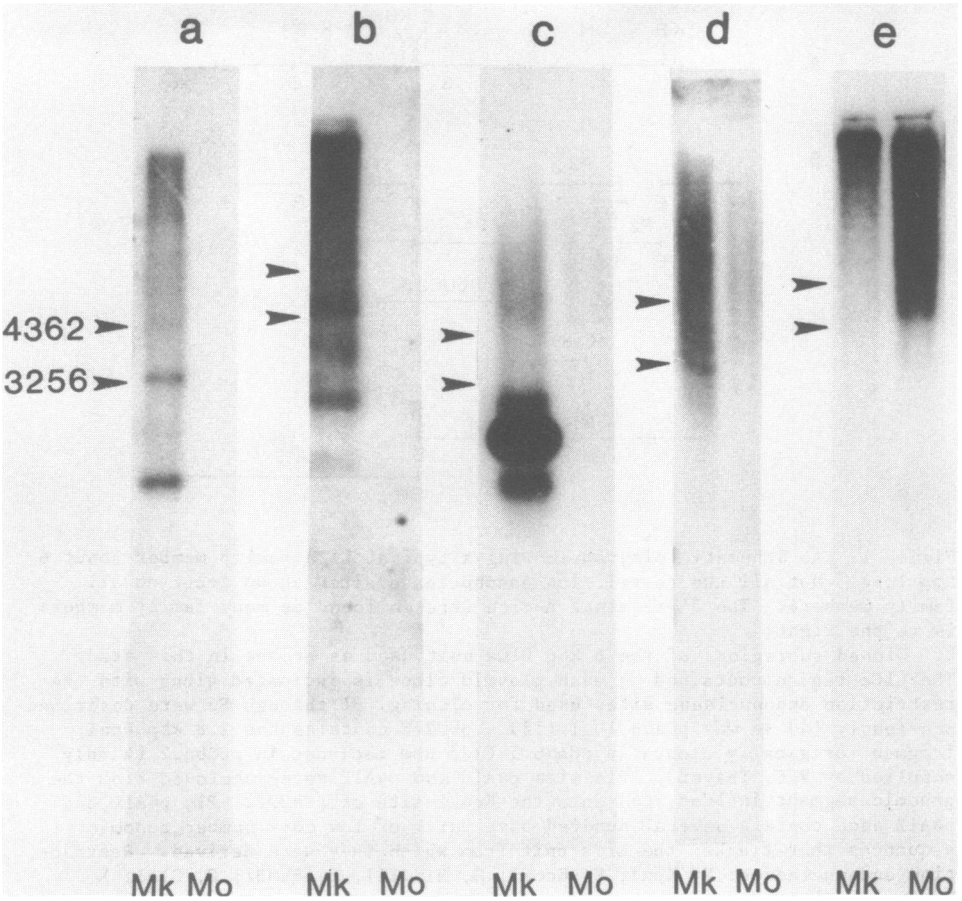

Figure 2. Annealing of subcloned regions of L1Ca to restriction endonuclease digests of monkey (Mk) and mouse (Mo) genomic DNA. The digests were electrophoresed on 1% agarose gels and blotted. Panels: a, b and e, KpnI; panel c, HindIII; panel d, EcoR1. The probes and amounts of DNA used were: panel a, P2, 0.24 µg monkey and 15 µg mouse DNA; panel b, pCaα6.10.1, 0.03 µg monkey and 20 µg mouse DNA; panel c, P4, 0.24 µg monkey and 20 µg mouse DNA; panel d, pUC2.8, 0.1 µg monkey and 20 µg mouse DNA; panel e, P6, 0.03 µg monkey and 20 µg mouse DNA. pBR322 cleaved by restriction endonucleases was used to provide size markers (indicated on the left).

about 20 kbp long were purified by sucrose density gradient fractionation. EcoR1 linkers were added and the fragments were ligated to λCharon4A arms by standard procedures. Approximately 10^6 phage were obtained from the *in vitro* packaging reaction.

RESULTS

Gradient of homology within monkey and mouse L1 units. Before analyzing the L1Ca sequences in the hybrid cell DNA, a series of control experiments was carried out to establish the level of cross-hybridization between various parts of L1Ca and L1Md units. Blots of restriction endonuclease digested genomic DNA from African green monkey liver and B82 mouse cells (the parent of the hybrid cell line) were prepared and annealed to probes representing different regions within the 6 kbp L1Ca (Figure 2). Probe 6 from close to the 3'-end of L1Ca anneals very well to mouse DNA. This result is expected since comparison of the primary nucleotide sequence at the 3'-ends of L1Ca and L1Md (Mus domesticus) indicated 60 to 70 percent homology in part of the region covered by P6 (18). Hybridization of mouse DNA with probes pUC2.8, and P4 gives decreasing amounts of annealing as the probes become more distant from the 3'-end of L1Ca. Finally, no hybridization was detected between pCaα6.10.1 or P2 and mouse DNA. These results were obtained with the standard hybridization conditions. When conditions were made more stringent by raising the washing temperature from 52° to 57°, cross hybridization of pUC2.8, and P4 with mouse DNA diminished relative to that seen with monkey DNA. Under even more stringent hybridization conditions (i.e. 40% formamide, 57°, 4xSSC) annealing between P6 and mouse DNA also disappeared.

These data are consistent with previous reports indicating extensive homology between L1Md and L1Ca units starting about 2 kbp from the left end of long L1Ca units and extending to a point about 220 bp and 750 bp from the 3'-termini of L1Ca and L1Md, respectively (8,9,27,28). Recently acquired primary sequence data confirm this conclusion (4,14,29,34).

Restriction endonuclease mapping of L1 sequences on a single monkey chromosome. Various genomic restriction endonuclease fragments that are produced from the L1Ca units in monkey DNA have been described (4,12). The experiments reported in Figure 3 compare the relative abundance of fragments generated from total monkey DNA (60 chromosomes) with that of fragments generated from the DNA of a single monkey chromosome, CAE-19, contained in the hybrid cells. Overall, the liver DNA and the CAE-19 DNA give similar patterns; however, differences are observed with each restriction enzyme and each probe. For example, the 2.8 kbp KpnI band from liver DNA that anneals well with pUC2.8 is barely visible in the hybrid cell DNA (Figure 3c). When the blot was exposed for longer periods (not shown) so that the 1.2 and 1.5 kbp bands in hybrid cell DNA were at least as intense as they are in the monkey DNA shown in Figure 3c, the relative intensity of the 2.8 kbp band in the hybrid

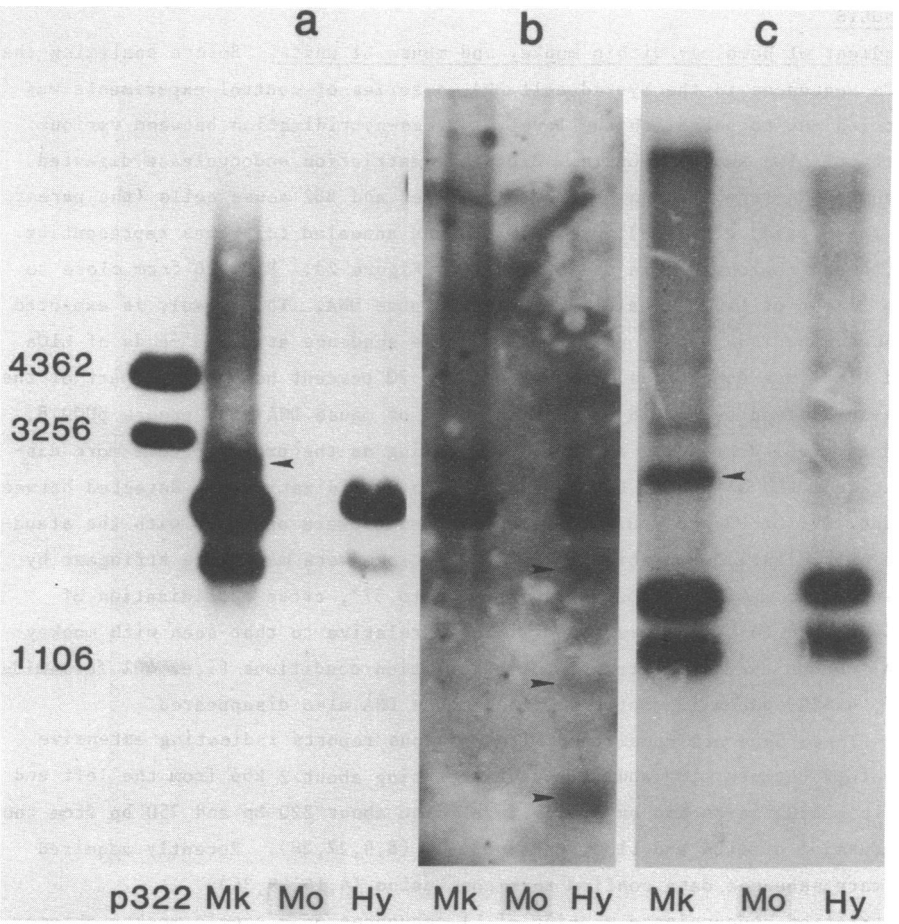

Figure 3. Annealing of subcloned regions of L1Ca to restriction endonuclease digests of monkey (Mk), B82 mouse cell (Mo) and hybrid cell (Hy) DNA. Digests were electrophoresed on 1% agarose gels and blotted. Panels a and b, HindIII; panel c, KpnI. The enzymes, probes and amounts of hybrid cell or B82 (they were the same in each case) DNA used were, respectively: panel a, HindIII, P4, 15 µg; panel b, HindIII, P3, 15 µg; panel c, KpnI, pUC2.8, 20 µg. In each case 0.2 µg of monkey DNA was used. The lane marked p322 contained, as markers, a restriction endonuclease digest of pBR322. The arrow heads indicate bands that are discussed in the text: panel a, 2.9 bkp; panel b, 1.9, 1.0 and 0.45 kbp; panel c, 2.8 kbp.

cell DNA did not increase (unpublished data). Note that the two bands at 3.4 and 4.6 kbp serve as internal controls for blotting efficiencies. In the hybrid cell DNA they are of about equal intensity to the 2.8 kbp band while in monkey DNA they are much weaker than the 2.8 kbp band. The 2.8 kbp KpnI

band arises from family members that lack the KpnI site at about 3.4 kbp on the map but contain the KpnI sites at approximately 1.9 and 4.6 kbp (Figure 1). Such family members appear to be rare in chromosome CAE-19 compared to their abundance in the total genome. Thus, almost all of the CAE-19 family members that contain the KpnI sites at 1.9 and 4.6 kbp also contain the site at 3.4 kbp and yield the 1.5 and 1.2 kbp fragments seen in Figure 3c. Other family members in both monkey DNA and CAE-19 DNA contain only one or none of the 3 sites and yield the minor bands and smear seen above 3 kbp in the digests of both monkey and hybrid cell DNA (Figure 3c).

As predicted from earlier results, a 2.9 kbp HindIII band is seen in monkey DNA (just above the very intense 2.5 kbp band) with probe P4 but no such band is seen in the hybrid cell DNA (Figure 3a). Shorter exposures of the gel shown in Figure 3a indicated that the 2.9 kbp HindIII band was clearly detectable in monkey DNA even when the major 2.5 kbp HindIII band was of an intensity comparable to that of the 2.5 kbp HindIII band in hybrid cell DNA on the figure. Thus, some restriction endonuclease fragments that are readily observed in the total set of genomic L1Ca segments are relatively infrequent on chromosome CAE-19. The reverse situation also occurs. Probe P3 reveals HindIII fragments of 1.9, 1.0 and 0.45 kbp in the hybrid cell DNA but these fragments are at least 5 fold less abundant in monkey DNA (Figure 3b).

These experiments indicate that the abundance of certain restriction endonuclease sites in the subset of L1Ca units on CAE-19 diverges significantly from that in the total L1Ca population. In order to confirm this interpretation, cloned L1Ca units were isolated from a genomic library of hybrid cell DNA and analyzed.

Cloning of genomic segments containing L1Ca units from chromosome CAE-19. Approximately 2×10^5 recombinant phage in a genomic library of hybrid cell DNA were screened with a mixture of p4A10 and p4A12 (see Figure 1B). The positive plaques numbered 134. Of these, 52 were picked, plaque purified and screened with various subcloned probes. Five of the 52 annealed to subcloned probes covering the full 6 kbp unit and were characterized in detail by restriction endonuclease mapping and hybridization. The results are summarized in the maps of the phage inserts in λH1, λH3, λH5, λH6 and λH13 in Figure 4. Each of the phage contains one approximately 6 kbp stretch of DNA in which the subcloned probes anneal in a manner consistent with the basic order shown in Figure 1A. All the L1Ca units are bordered by unrelated genomic DNA on both sides, with the possible exception of λH5 where the

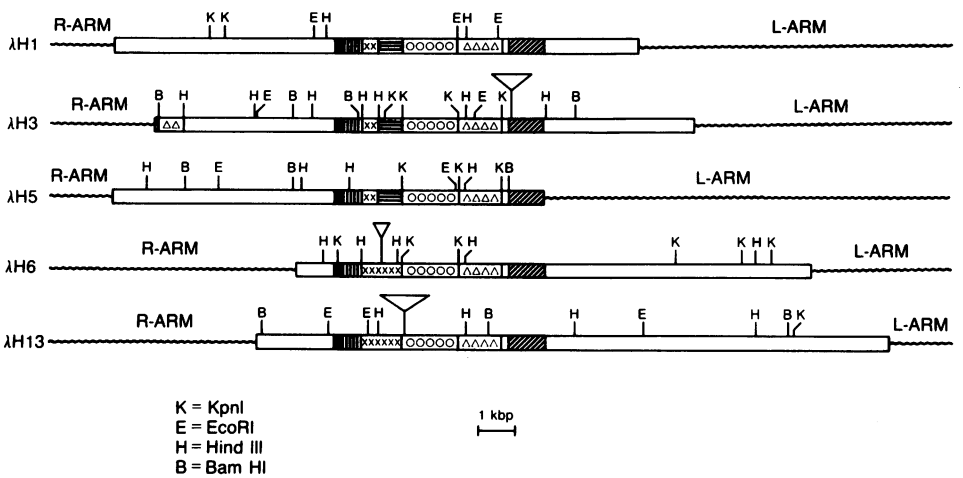

Figure 4. Restriction endonuclease and hybridization maps of L1Ca units in recombinant phage. Phage were isolated from a genomic library constructed from hybrid cell DNA as described in the text. The bars represent the cloned insert and the wavy lines the arms of the λ-vector (L, left and R, right arms). Restriction maps are shown for KpnI (K), EcoRI (E), HindIII (H) and BamHI (B). Each L1 segment is divided by vertical lines into regions that correspond to those contained in the cloned probes. The regions that anneal with the subcloned L1Ca probes shown in Figure 1b are marked as follows: filled bar, P1; vertical lines, P2; x, P3; horizontal lines, pCaα6.10.1; circles, P4; triangles, P5; slanted lines, P6; empty bar, sequences not homologous to L1 probes. Note that a small region within L1 (small open bar between regions covered by P5 and P6) is not contained in any of the cloned probes. In some cases (e.g. the large EcoRI fragment in λH1) a restriction endonuclease fragment annealed with several of the probes but was not conveniently subdivided by other restriction sites; the order and size of the various regions is then arbitrarily shown as corresponding to those of the probes in the canonical sequence (Fig. 1A). No data are available on the annealing of pCaα6.10.1 to λH6 and λH13 DNA although this probe gave positive hybridization to the plaques. The inverted triangles show the length and position of DNA segments that are not consistent with the map of a typical L1Ca segment. The size (kbp) of the total inserts in each phage are: λH1, 14.5; λH3, 16.0; λH5, 12.0; λH6, 14.5; λH13, 18.5.

insert ends very close to the 3'-end of L1Ca. In λH3 an additional L1Ca unit containing about 1 kbp from near the 3'-end of L1Ca appears; it is joined on one side to the right arm of λ and is separated from the 6 kbp unit by about 4 kbp of unrelated genomic DNA sequences.

Each of the five cloned L1Ca units has a distinctive map of KpnI, HindIII, EcoRI and BamHI sites, the restriction enzymes used in the analysis, and they all differ from the typical map deduced from analysis of L1Ca elements in monkey DNA (Figure 1A). For example, neither λH1 nor λH13 has any KpnI site at all.

Although the order of the various L1Ca subsegments in all 5 clones is the same as that in the typical genomic L1Ca units, several of the units from CAE-19 appear to include more than 6 kbp of DNA. In λH3 the segment annealing with P6 is about 1 kbp longer than expected from a typical genomic element. Similarly, λH6 has an extra 0.5 kbp in the region where P3 sequences occur. The region of λH13 that anneals with P3 and P4 contains 1.4 kbp more DNA than expected from the typical map. We do not know if these 'extra' base pairs are internal repetitions of L1Ca sequences or unrelated genomic sequences. A variety of internal L1 rearrangements have been noted in the past and these have no discernable pattern.

DISCUSSION

The L1Ca units on CAE-19 are basically similar in organization and sequence to those in the total genomic family. They anneal with the probes used in our experiments, although the cloned probes were randomly chosen from total genomic DNA. Mapping of the cloned L1Ca units from CAE-19 indicated that the order of subsequences is as it is in most genomic copies. We assume that CAE-19 also contains L1Ca members that are truncated at the 5'-end, since such members are at least 5 times more abundant in the total genome than are 6 kbp units (4). However, the phage clones described here were specifically selected to contain long family members. Rearranged units are also likely to occur on CAE-19 and some of the phage described in Figure 4 may contain internal repetitions.

Although basically similar to the total genomic set of L1Ca members, the set on CAE-19 tends to have distinctive features, as seen on the genomic blots. The maps of the cloned elements are consistent with this conclusion. Certain restriction endonuclease fragments that are readily observed in monkey DNA are either missing or in low abundance in hybrid cell DNA. For example, hybrid cell DNA yields a very small amount of the 2.8 kbp KpnI fragment (represented in probe pUC2.8) (Figure 3c) compared to monkey DNA and none of the λ-phage clones give rise to the 2.8 kbp KpnI fragment (Figure 4). Similarly, the 2.9 kbp HindIII fragment that anneals to P4, P3 and pCaα6.10.1 in digests of total monkey DNA (Figure 3a and b and unpublished data) is barely detectable in hybrid cell DNA nor is it present in any of the cloned phage. On the other hand, other restriction endonuclease fragments that are quite rare in monkey DNA are more frequent in hybrid cell DNA. For example, a 0.45 kbp HindIII fragment that anneals with P3 is more abundant in clone 22 DNA than in total DNA and is present in λH3. This results from cleavage at

HindIII sites at approximately 0.8 and 1.2 kbp from the left end of the L1Ca element. Thus, the frequency of L1Ca family members that contain both these sites appears to be higher in hybrid cell DNA than in the total genome. The site at 1.2 kbp is frequent in L1Ca members, as evident from the very abundant 2.4 kbp HindIII fragment (1.2 kbp to 3.6 kbp) seen in liver and hybrid cell DNA genomic digests (see Figures 1 and 3), although three out of the 5 phage in Figure 4 lack the HindIII site at 1.2 kbp. In addition to the 0.45 kbp fragment, P3 annealed with 1.0 and 1.9 kbp HindIII fragments in hybrid cell DNA that are relatively less abundant in the monkey DNA digests (Figure 3b); phage λH6 contains both these fragments. An additional indication that CAE-19 has a set of L1Ca units that differ significantly from the total genomic family is the lack of any KpnI sites at all in 3 out of 8 cloned segments studied (two of the five shown in Figure 4 and one other not studied further). Altogether, these data indicate that the L1Ca sequences on CAE-19 tend to have characteristic differences from the entire genomic set of L1Ca units.

The relative homogeneity and distinction of the L1Ca sequences on CAE-19 suggests that at least one of the one or more mechanisms responsible for intraspecies homogenization of the family is more effective within CAE-19 than between CAE-19 and other chromosomes. This may or may not be true of the L1Ca elements on other chromosomes or in other species. For example, Schmeckpeper and coworkers have reported that there is no detectable difference between L1Hs family members on the human X-chromosome and those on autosomes (30). These diverse observations indicate that homogenization in mammals is likely to involve multiple mechanisms controlled by a variety of unknown factors. In yeast, where some of the homogenization mechanisms proposed for mammalian genomes can be studied directly, it is already apparent that multiple complex processes and outcomes occur. For example, a yeast Ty retrotransposon recombines more frequently with other Ty elements on the same chromosome than with those on separate chromosomes (31). On the other hand, meiotic gene conversion of a yeast HIS3 gene occurs at about the same rate when the sequences are on homologous or non-homologous chromosomes (32). And the efficiency of interaction between mating type loci varies, depending on which individual chromosomes are involved (33). The situation in mammals is likely to be at least as complex.

ACKNOWLEDGEMENTS

We thank Ronald E. Thayer and Anthony Furano for critical readings of this manuscript and M.H. Edgell and Y. Sakaki and their colleagues for making

data available prior to publication. Carolyn Ray prepared the manuscript with patience and skill.

REFERENCES

1. Adams, J.W., Kaufmann, R.E., Kretschmer, P.J., Harrison, M. and Nienhuis, A.W. (1980) Nucleic Acids Res. 8, 6113-6128.
2. Voliva, C.F., Jahn, C.L., Comer, M.B., Hutchison III, C.A. and Edgell, M.H. (1983) Nucleic Acids Res. 11, 8847-8859.
3. Bennett, K.L. and Hastie, N.D. (1984) EMBO J. 3, 467-472.
4. Grimaldi, G., Skowronski, J. and Singer, M.F. (1984) EMBO J. 3, 1753-1759.
5. Katzir, N., Rechavi, G., Cohen, J.B., Unger, T., Simoni, F., Segal, S. Cohen, D. and Givol, D. (1985) Proc. Natl. Acad. Sci. USA 82, 1054-1058.
6. Witney, F.R. and Furano, A.V. (1984) J. Biol. Chem. 259, 10481-10492.
7. Soares, M.B., Schon, E. and Efstratiadis, A. (1985) J. Mol. Evol., 22, 117-133.
8. Burton, F.H., Loeb, D., Voliva, C.F., Martin, S.L., Edgell, M.H. and Hutchison, III, C.A. (1985) J. Mol. Biol., in press.
9. Singer, M.F. and Skowronski, J. (1985) Trends Biochem. Sci. 10, 119-122.
10. Lerman, M.I., Thayer, R.E. and Singer, M.F. (1983) Proc. Natl. Acad. Sci. 80, 3966-3970.
11. Potter, S.S. (1984) Proc. Natl. Acad. Sci. USA 81, 1012-1016.
12. Grimaldi, G. and Singer, M.F. (1983) Nucleic Acids Res. 11, 321-338.
13. Gebhard, W. and Zachau, H.G. (1983) J. Mol. Biol. 170, 255-270.
14. Miyake, T., Migita, K. and Sakaki, Y. (1983) Nucleic Acids Res. 11, 6837-6846.
15. Thayer, R.E. and Singer, M.F. (1983) Mol. Cell. Biol. 3, 967-973.
16. Manuelidis, L. (1982) Nucleic Acids Res. 10, 3321-3219.
17. Martin, S.L., Voliva, C.F., Burton, F.H., Edgell, M.H. and Hutchison III, C.A. (1984) Proc. Natl. Acad. Sci. 81, 2308-2312.
18. Singer, M.F., Thayer, R.E., Grimaldi, G., Lerman, M.I. and Fanning, T.G. (1983) Nucleic Acids Res. 11, 5739-5745.
19. Manuelidis, L. and Biro, P. (1982) Nucleic Acids Res. 10, 3221-3239.
20. Martin, S.L., Voliva, C.F., Hardies, S., Edgell, M.H. and Hutchison III, C.A. (1985) Mol. Biol. Evol. 2, 127-140.
21. Brown, S.D.M. and Dover, G. (1981) J. Mol. Biol. 150, 441-466.
22. Arnheim, N. (1983) in Evolution of Genes and Proteins, Nei, M. and Koehn, R.K. Eds., pp. 38-61, Sinaver, Sunderland, Massachusetts.
23. Creau-Goldberg, N., Turleau, C., Cochet, C., Huerre, C., Junien, C. and de Grouchy, J. (1984) Human Genetics 68, 333-336.
24. Lee, T.N.H. and Singer, M.F. (1982) J. Mol. Biol. 161, 323-342.
25. Lee, T.N.H. and Singer, M.F. (1985) Federation Proceedings 44, p. 1084.
26. Maniatis, T., Hardison, R.C., Lacy, E., Laver, J., O'Connel, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) Cell 15, 687-701.
27. Fanning, T.G. (1983) Nucleic Acids Res. 11, 5073-5091.
28. Burton, F.H., Voliva, C.F., Edgell, M.H. and Hutchison III, C.A. (1983) DNA 2, 82.
29. Loeb, D.D., Padgett, R.W., Hardies, S.C., Shehee, W.R., Comer, M., Edgell, M.H. and Hutchison, III, C.A. (1986) Mol. Cell. Biol., 6, 168-182.
30. Schmeckpeper, B.J., Willard, H.F. and Smith, K.D. (1981) Nucleic Acids Res. 9, 1853-1872.
31. Roeder, G.S., Smith, M. and Lambie, E.J. (1984) Mol. Cell. Biol. 4, 703-711.

32. Jinks-Robertson, S. and Petes, T.D. (1985) Proc. Natl. Acad. Sci. 82, 3350-3354.
33. Haber, J.E., Rowe, L. and Rogers, D.T. (1981) Mol. Cell. Biol. 1, 1106-1119.
34. Hattori, M., Hidaka, S. and Sakaki, Y. (1985) Nucleic Acids Res. 13, 7813-7827.