Extensive Movement of LINES ONE Sequences in β-Globin Loci of *Mus caroli* and *Mus domesticus*

N. CAROL CASAVANT,[1] STEPHEN C. HARDIES,[2] FRED D. FUNK,[3] MARY B. COMER,[1] MARSHALL H. EDGELL,[1]* AND CLYDE A. HUTCHISON III[1]

Department of Microbiology and Immunology, Program in Molecular Biology and Biotechnology, University of North Carolina, Chapel Hill, North Carolina 27514[1]; Department of Biochemistry, University of Texas Health Science Center, San Antonio, Texas 78284[2]; and Department of Biological Science, Northern Arizona University, Flagstaff, Arizona 86011[3]

LINES ONE (L1) is a family of movable DNA sequences found in mammals. To measure the rate of their movement, we have compared the positions of L1 elements within homologous genetic loci that are separated by known divergence times. Two models that predict different outcomes of this analysis have been proposed for the behavior of L1 sequences. (i) Previous theoretical studies of concerted evolution in L1 have indicated that the majority of the 100,000 extant L1 elements may have inserted as recently as within the last 3 million years. (ii) Gene conversion has been proposed as an alternative to a history of prolific recent insertions. To distinguish between these two models, we cloned and characterized two embryonic β-globin haplotypes from *Mus caroli* and compared them with those of *M. domesticus*. In 9 of 10 instances, we observed an L1 element to be present in one chromosome and absent at the same site in a homologous chromosome. This frequency is quantitatively consistent with the known rate of concerted evolution. Therefore, we conclude that gene conversion is not required for concerted evolution of the L1 family in the mouse. Furthermore, we show that the extensive movement of L1 sequences contributes to restriction fragment length polymorphism. L1 insertions may be the predominant cause of restriction fragment length polymorphisms in closely related haplotypes.

LINES ONE (L1) is a family of long, interspersed, repetitive sequences found in the genomes of mammals (8, 35; for reviews, see references 7 and 32 and C. A. Hutchison III, S. C. Hardies, D. D. Loeb, W. R. Shehee, and M. H. Edgell, *in* D. E. Berg and M. M. Howe, ed., *Mobile DNA*, in press). We will refer to L1 elements from the mouse species *Mus domesticus* and *M. caroli* by the names L1Md and L1Mc, respectively. The L1 elements vary in length, reaching a maximum of about 7 kilobases (8, 35). Full-length elements feature two long open reading frames that have evolved under selective pressure (5, 18, 20). On either end of the segment containing these open reading frames are untranslated regions. Elements that are less than full length are most common and represent random truncation of the 5' end. The copy number increases from 5' to 3', with up to 100,000 copies per haploid genome at the 3' end (10, 35). Most elements are flanked by short direct repeats and have 3' untranslated regions ending in an A-rich sequence. These features have been taken to indicate that L1 is a family of retroposons (36).

The L1 family undergoes concerted evolution as characterized by a greater similarity among the sequences within a given mammalian species than between those in different species (21). Martin et al. (21) characterized the age distribution of L1 sequences in mice and calculated that the majority were less than 3 million years old. Hardies et al. (11) subsequently refined this estimate to 2 million years for truncated L1 elements. Therefore, the concerted evolution of the L1 family is a relatively robust genetic process, requiring a minimum of 50,000 homogenizing events in the last 2 million years.

Two alternative mechanisms have been put forward to explain the rapid concerted evolution of the L1 family, gene

conversion and retroposition (6, 11, 14, 21, 23–25). The gene conversion models generally relegate retroposition to the role of creating the original distribution of the elements and leave their subsequent conce :d evolution to gene conversion. Alternatively, if retropo on occurred often enough, it could be fully responsible for ι.ιe concerted evolution of L1. To achieve the age distribution quoted above, 50,000 copies would need to have dispersed in the last 2 million years (11). It has been observed that the direct repeats surrounding the mouse L1 elements are of a similar average age as the L1 sequences themselves (11). This favors retroposition models over gene conversion.

We report here a direct test di nguishing between gene conversion and retroposition modcls for the concerted evolution of L1 in mice. This was done by comparing the positions of L1 elements within homologous regions of DNA between two species (*M. domesticus* [4, 31, 36] and *M. caroli*) and between different haplotypes. The divergence times represented in these comparisons range from 1 to 6 million years. The retroposition models require that the majority of the L1 elements observed should not be in the same places in the chromosomes that we have compared. In 9 of 10 instances we observed an L1 element to be present in one chromosome and absent at the same site in a homologous chromosome. Therefore, our results support the retroposition models. We conclude that gene conversion is not the primary mechanism of concerted evolution of the L1 family in mice.

L1 insertions cause restriction fragment length polymorphism (RFLP). Owing to the high frequency of L1 retroposition, we anticipated that L1 insertions might be a major contributor of RFLPs in mice. Therefore, we determined the fraction of RFLPs in the β-globin cluster as a result of L1 insertion relative to other causes. We found that under some

---

* Corresponding author.

circumstances, L1 insertions accounted for more than 80% of the observed RFLPs.

MATERIALS AND METHODS

Preparation of an *M. caroli* genomic library. High-molecular-weight DNA was isolated as described by Padgett et al. (26) from a total DNA preparation of an *M. caroli* mouse, kindly provided by V. Chapman. The mouse DNA was cut with *Sau*3A and ligated to Charon 30 (29) cut with *Bam*HI, as described by Padgett et al. (26). To determine the proportion of bacteriophage carrying inserts, we titrated the library on LE392 and Q359. Q359 supports replication only of lambda phage not containing parental Charon 30 stuffer fragments (15). Over 2 million independent recombinant clones were produced, which is equivalent to 10 times a single representation of the haploid mouse genome.

Analyzing the quality of the library. To verify the quality of the *M. caroli* library, we used the technique described by Phillips et al. (28). From the ethidium bromide-stained gel, we observed that there was no striation, which would have been indicative of an underrepresentative library. Southern blots (33) of these gels were probed with a nick-translated *Hbb-y* cDNA (9). The hybridization was carried out at low stringency with 40% formamide as described by Wahl et al. (37). This assay showed that the *M. caroli* embryonic globin genes were present in the library.

Screening the library. We screened approximately 800,000 phage plated at a density of 20,000 per plate on 150-mm plates. Duplicate transfers on nitrocellulose filters (BA 85, 132 mm in diameter; Schleicher & Schuell, Inc.) were made by the method of Benton and Davis (2). These filters were hybridized with the β-globin probes described below. Duplicate positive plaques were purified through three successive platings, with rescreening of each set of plates. Small lysates of the purified plaques and subsequent larger-scale lysates were prepared essentially as described by Maniatis et al. (19).

β-Globin probes. The β-globin cDNA probes used included a 2-kilobase *Hha*I fragment from plasmid pCR1-β$_M$9, which contains a cloned cDNA copy of mouse adult β-globin mRNA (30), and a 1.45-kilobase fragment from an embryonic *Hbb-y* β-globin cDNA plasmid (9). Two genomic globin probes were used: (i) a β*h0*-specific probe, prepared by isolation of a 750-base-pair (bp) *Hin*dIII fragment from pHE100 as described by Brown et al. (3); and (ii) a β*h2* probe, prepared by isolation of a 400-bp *Pst*I-*Bam*HI fragment from pHE112 as described by Brown et al. (3).

Isolating the L1 elements. To determine the location of the L1 elements, we cut the *M. caroli* β-globin clones with *Bam*HI, *Eco*RI, *Bgl*II, and *Hin*dIII. The digested DNA was run on 1% agarose gels and transferred to nitrocellulose filters (33). These filters were hybridized (37) with six separate probes representing the extreme left, right, and middle of L1Md (probes 1, 2, 3, 4, and 5 of Loeb et al. [18]). The sixth probe, a 270-bp *Eco*RI-*Bam*HI fragment, lies between probes 3 and 4 of Loeb et al. (18). The probes were nick translated by the method described by Wahl et al. (37). Southern transfers (33) were performed as described by Jahn et al. (13). The Southern transfer filters were pre-wet, hybridized at low stringency in 40% formamide, and washed as described by Wahl et al. (37). To determine the exact location of the *M. caroli* L1 elements, they were sequenced by using procedures of Bankier and Barrell (1) with modifications as described by Loeb et al. (18). DBSYSTEM programs (34) were used to assemble the data. Sequences

from the *M. domesticus* β-globin cluster were provided by C. Voliva (personal communication).

RESULTS

Cloning the *M. caroli* β-globin gene cluster. The β-globin gene cluster was isolated from *M. caroli* to study the L1 elements present in that region. A total genomic library was made from a single *M. caroli* mouse. The library was screened with a mixture of three *M. domesticus* β-globin probes: an embryonic *Hbb-y* cDNA (9), an adult cDNA pCR1-β$_M$9 (30), and a genomic fragment from β*h0* (3). Restriction maps with *Eco*RI, *Bam*HI, *Bgl*II, and *Hin*dIII were made for 11 clones.

In the process of assembling the clones into a composite map of the locus, it became obvious that two closely related haplotypes were present. We named these 7A and 7B because *Hbb* is on chromosome 7 in *M. domesticus*. The presence of two haplotypes had been anticipated, because the mouse from which the library had been constructed was not inbred for the β-globin locus. The assembly of the 11 clones produced restriction maps for both haplotypes (Fig. 1). However, the assignment of clones λ83-25 and λ83-10 to 7B and the assignment of λ83-5 to 7A is arbitrary, because no one clone connects the single observed polymorphism in these three clones to a more 5' polymorphism.

Identification of the *M. caroli* genes. Hybridization of *M. domesticus* β-globin probes to these clones allowed the *M. caroli* β-globin genes to be localized. A comparison was made of the positions of restriction sites and β-globin genes between *M. caroli* and *M. domesticus* (*Hbb*^{d} from BALB/c). The high congruence of the maps in the area of the nonadult β-globin genes (particularly between the 5'-most gene [y] and β*h1*) suggested that this region represents the embryonic genes of the *M. caroli* β-globin locus including y, β*h0*, β*h1*, and β*h2* (Fig. 1 and 2). The high congruence of the restriction sites radically drops off 5' of y and 3' of β*h1*. The reason for this is discussed in detail below. The identity of the β-globin genes was confirmed by sequencing parts of *M. caroli* y, β*h0*, β*h1*, and β*h2* (data not shown).

Location of L1 repeats in *M. caroli*. A set of hybridization probes derived from *M. domesticus* L1 elements was used to determine the location of L1 elements in *M. caroli*. Four L1 elements were found. Three were located 5' of the β-globin gene cluster, and one was located between β*h1* and β*h2* (Fig. 3 and 4). The L1Mc-3 insertion is the source of the polymorphism among the clones λ83-25, λ83-10, and λ83-5, and, as such, its assignment to chromosome 7B is arbitrary. No other L1 elements were detected by hybridization with the various probes from within L1. An exact location for each L1Mc element was determined by sequencing a fragment containing the repeat (data not shown).

Comparison of L1 positions in different chromosomes. The incidence of sites occupied by an L1 element in one chromosome but not in another is a measure of L1 movement. The number of such nonshared sites is reported in Table 1 for three different pairs of chromosomes. Eight L1 elements from Fig. 3 and 4 are tabulated and discussed below. There is only one example of a shared site (L1Mc-2a and L1Mc-2b) which occurs between the two *M. caroli* chromosomes. The other elements in *M. caroli* (L1Mc-1 and L1Mc-3) occur in only one of the two chromosomes. The unoccupied insertion sites were verified by sequencing (data not shown).

Between *M. caroli* and *M. domesticus*, the site occupied by L1Mc-2a,2b in *M. caroli* is vacant in *M. domesticus*. The unoccupied target site has been sequenced (Voliva, personal
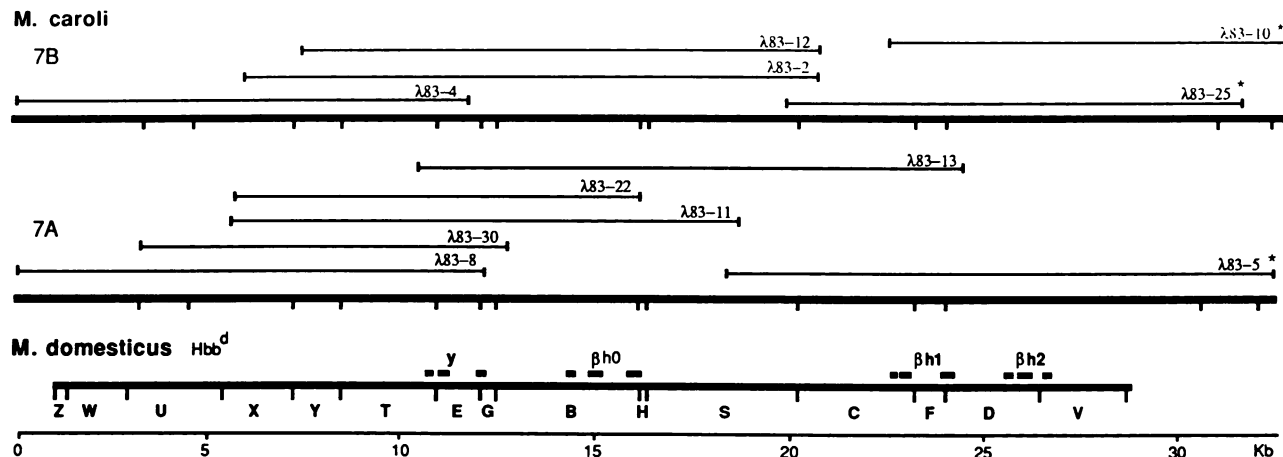
FIG. 1. Alignment of the embryonic β-globin loci of *M. caroli* and *M. domesticus*. 7A and 7B represent two haplotypes of *M. caroli*, and the *M. domesticus* haplotype is *Hbb^d*. EcoRI sites are represented by vertical bars, and the EcoRI fragments are lettered by the system of Jahn et al. (13). The maps were aligned on the basis of similarities of restriction fragments hybridizing to β-globin. Boxes above the *M. domesticus* map indicate β-globin gene exons as determined by DNA sequencing (12, 16, 27). Lambda clones reported in this study are positioned above their respective *M. caroli* haplotypes. The assignment of lambda clones marked with an asterisk to haplotypes 7A or 7B may be reversed. These three clones do not overlap any known polymorphisms in the other clones. Kb, kilobases.

communication). The target sites corresponding to L1Mc-1 and L1Mc-3 are not cloned or have been deleted, respectively, in *M. domesticus*. Therefore, these two repeats are excluded from the comparison between *M. caroli* and *M. domesticus*. Previously characterized repeats, L1Md-1, L1Md-2, and L1Md-3 (36), appear at sites in *M. domesticus* that are vacant in *M. caroli*. The unoccupied target sites corresponding to L1Md-1 and L1Md-2 were sequenced in *M. caroli* (data not shown). L1Md-3 was judged to be missing from *M. caroli* by hybridization and because the corresponding restriction fragment in *M. caroli* is shorter by the length of L1Md-3. L1Md-8 (36) is excluded from the tabulation because it is inserted in another L1 element that is absent in the other chromosomes.

In addition to the two chromosome pairs described above, the *Hbb^d* and *Hbb^s* haplotypes of *M. domesticus* contain three L1Md repeats that are sufficiently characterized for inclusion in this analysis. L1Md-4 (36) and L1Md-6 (4, 31) are present in *Hbb^d* and absent in *Hbb^s*. L1Md-9 is present in *Hbb^s* and absent in *Hbb^d* (4, 31).

Divergence among *M. caroli* 7A, 7B, and *M. domesticus* chromosomes. We estimated the divergence time separating the two *M. caroli* chromosomes and the divergence time

separating the *M. caroli* from the *M. domesticus* chromosomes for use in Tables 1 and 2. This was done by counting differences in DNA sequences determined from the region in the area of L1Mc-2a,2b, and L1Md-1 (data not shown). The two *M. caroli* chromosomes differed at 8 of 865 bp. The compared sequence included L1Mc-2a,2b, which had a similar mismatch frequency to the flanking region. With a neutral mutation rate of 4.6×10^{-9} changes per site-year (17), this corresponds to 1 million years of divergence. *M. domesticus* differed from *M. caroli* at 40 of 704 bp, yielding a divergence time of 6 million years. This value is similar to that estimated by other methods (21).

Percentage of RFLPs due to L1 movement. We wanted to know how the influx of L1s into the globin region was affecting the numbers of RFLPs. To determine this, we counted RFLPs between the two *M. caroli* chromosomes and between chromosome 7A of *M. caroli* and *M. domesticus* (Table 2). To determine the total number of RFLPs between a given pair of chromosomes, we first counted the total number of restriction fragments and then subtracted the number having the same length in both chromosomes. The fusion fragments to the vector were also included. We then examined the maps with the L1 insertions removed (Fig. 3
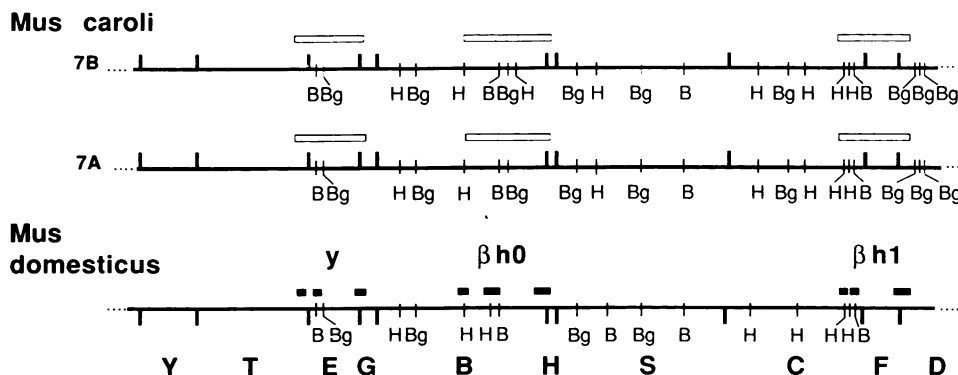

FIG. 2. Detailed restriction map of the region devoid of L1 elements. Symbols: ■■, β-globin exons; ▭, β-globin hybridizing regions. Restriction sites: |, EcoRI; B, BamHI; Bg, BglII; H, HindIII.
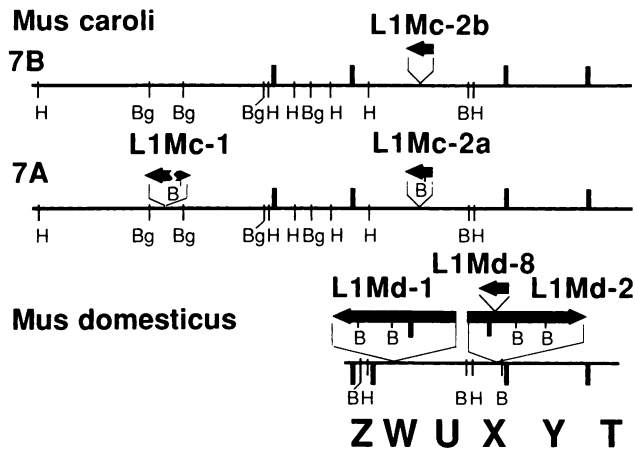
Mus caroli

FIG. 3. Alignment of restriction fragments 5' of the β-globin clusters of *M. caroli* and *M. domesticus* after the excision of the L1s. Alignment of fragments containing L1s was determined by sequencing flanking regions and comparing them with sequence data from the unoccupied chromosomes. The remaining fragments were aligned by length similarities. The L1 elements are indicated as arrows, with the large arrowheads indicating poly(A) regions. L1Mc-1 has an inverted 5' segment. Restriction sites: | , *Eco*RI; B, *Bam*HI; Bg, *Bgl*II; H, *Hin*dIII.

and 4) and repeated the above tabulation. The fraction of total RFLPs that disappeared during this operation is reported as the percentage of RFLPs due to L1 insertion (Table 2). The *M. caroli-M. caroli* comparison basically includes one point mutation and two L1 insertion events. In several cases in the *M. caroli-M. domesticus* comparison, removal of an L1 has no effect, because the fragment is also polymorphic for other reasons (e.g., L1Mc-3). The percentage of RFLPs due to L1s declined from 86% between the two *M. caroli* chromosomes to 22% between the *M. caroli* chromosome 7A and the *M. domesticus* chromosome.

DISCUSSION

Our major conclusion is that concerted evolution of the L1 elements is achieved by their movement. A comparison of
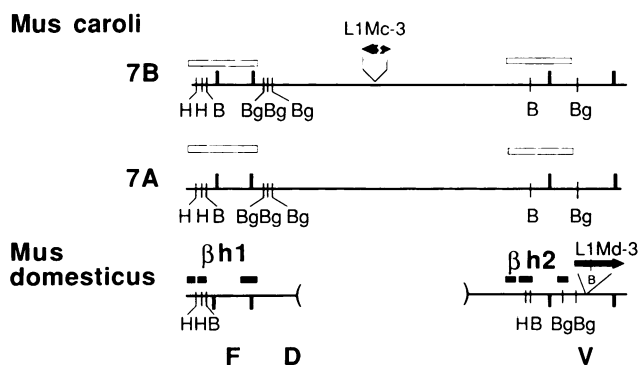

FIG. 4. Alignment of restriction fragments in the 3' end of the embryonic β-globin loci in *M. caroli* and *M. domesticus* after excision of the L1 elements. Alignment between the *Eco*RI restriction sites was done by sequencing around the sites. The parentheses indicate sequence not found in *M. domesticus*. The L1 elements are indicated by arrows. L1Mc-3 has an inverted 5' end. β-Globin hybridizing regions are shown (☐). Restriction sites: | , *Eco*RI; B, *Bam*HI; Bg, *Bgl*II; H, *Hin*dIII.

TABLE 1. Comparison of the theoretical and observed fractions of L1 elements that have moved

Chromosomes compared	Divergence[a] (million yr)	Fraction moved	
		Theoretical[b]	Observed
M. caroli with *M. caroli*[c]	1	0.29	0.67 (2/3)
M. caroli with *M. domesticus*[c]	6	0.87	1.00 (4/4)
M. domesticus with *M. domesticus*[d]	4	0.75	1.00 (3/3)

[a] The divergence time was estimated by counting differences in flanking sequences between the two *M. caroli* chromosomes and between the chromosomes of *M. caroli* and *M. domesticus*. The divergence time between the two chromosomes of *M. domesticus* was previously published (36).

[b] Calculated as $1-[1/2^{(t/2)}]$, where t is the divergence time in millions of years (see text).

[c] Compiled from Fig. 3 and 4.

[d] *Hbb*[d] to *Hbb*[s]. The elements include L1Md-4 (36), and L1Md-6 and L1Md-9 (4, 31).

our observed amount of movement of L1 elements to the amount needed to explain the current rate of concerted evolution in mice is shown in Table 1.

Concerted evolution of mouse L1 was analyzed by Martin et al. (21) and Hardies et al. (11) by comparing sequences of multiple L1 repeats. Concerted evolution was measured as the rate at which newly introduced mutations were spread throughout the L1 family. These observations indicated that L1 elements have an exponential age distribution. Specifically, half of the sequences were introduced within the last 2 million years, and half of the remainder were introduced within the previous 2 million years, etc. The question is whether this was done by gene conversion or by the movement of elements into new sites in the genome. Whereas movement causes repeats to appear at different sites, gene conversion would leave them at the same sites. For the movement model, the amount of movement can be quantitatively predicted from the rate of concerted evolution. For each pair of chromosomes in this study, we made such a theoretical prediction for the number of sites occupied in one chromosome but not the other (Table 1).

Our observed amount of movement (Table 1) is in line with the above prediction. Furthermore, the single incidence of a site occupied in two different chromosomes occurs in the least diverged pair of chromosomes. This is the most likely place for that to happen according to the movement model for concerted evolution. Although the numbers are small, the observations demonstrate a rate high enough to support concerted evolution. Therefore, it is not necessary to invoke gene conversion as a major factor, because the rate of

TABLE 2. Percentage of RFLPs due to L1 movement as a function of divergence time

Chromosomes compared	Divergence (million yr)	Fraction of polymorphic restriction fragments[a]		% RFLPs due to L1
		Including L1s	Excluding L1s[b]	
M. caroli and *M. caroli*	1	0.21 (21/101)	0.03 (3/97)	86
M. caroli and *M. domesticus*[c]	6	0.59 (52/88)	0.46 (36/79)	22

[a] Total number of fragments not of identical length between the two chromosomes divided by the total number of fragments in a combination of *Eco*RI, *Bam*HI, *Bgl*II, and *Hin*dIII single restriction digests.

[b] After adjusting map by excising all L1 repeats.

[c] *M. caroli* 7A to *M. domesticus Hbb*[d].

movement is sufficient to explain the concerted evolution of L1 in mice.

A deletion process for L1 elements has been proposed (11), but it has not yet been experimentally demonstrated. Finding an L1 element present in one chromosome and absent in a homologous chromosome does not directly indicate whether the movement was an insertion or a deletion. If a deletion is responsible, the presence of the element must predate the split between the two chromosomes. However, since the mouse L1 age distribution specifies that most of the L1 sequences are younger than the time separating *M. domesticus* and *M. caroli*, we expect most of the differences between *M. domesticus* and *M. caroli* to be insertions.

The movements discussed above concern the bulk flow of sequence information from the originally mutated element into the genome at large. There are other important roles that gene conversion might be playing. In light of previous observations (11), the movement model should contain only a small number of parental elements in which mutations can fall and then spread into the genome. Even rare gene conversions affecting these parental elements would influence all of the copies subsequently arising. Therefore, even a small amount of gene conversion could have a major impact on concerted evolution. Our observations do not rule out such a role; they just show that gene conversion does not control the rate of spread of new sequence information into the genome.

Contribution of L1 insertion to RFLP. Because of the relatively rapid rate of insertion of L1s, we quantitated their contribution to the total pool of RFLPs. Our results are summarized in Table 2. We found that the percentage of RFLPs due to L1 movement is a function of divergence time. Between the two *M. caroli* chromosomes, representing 1 million years of divergence, the majority (86%) of RFLPs are due to L1 insertions. This declines to 22% between *M. domesticus* and *M. caroli*, representing 6 million years of divergence. This time dependency is understandable in terms of the age distribution of L1 elements. For successively older periods, the contribution of L1 insertions should decline exponentially. In contrast, point mutations accumulate in an essentially linear fashion in this time range. Therefore, L1 insertion initially makes a large contribution to the total RFLP pool. Subsequently, as the genome collects proportionately more point mutations, those mutations become the major contributor of RFLPs.

With respect to the usual use of RFLPs as genetic markers within a species, the comparison of the two *M. caroli* chromosomes is the more applicable example. The knowledge that a high proportion of RFLPs are due to L1 insertion may aid in the search for polymorphic markers. For example, when picking a candidate restriction fragment for detecting RFLPs, one has to avoid including any repetitive DNA. However, by excluding restriction fragments that lie close to L1 elements, one might inadvertently exclude the best candidates for RFLP probes. Consequently, a better strategy might be to pick unique sequences that immediately flank L1 elements.

A number of things other than divergence time influence the fraction of RFLPs due to L1. The fraction in the genome at large may differ from the results in Table 2 because (i) the density of L1s in the β-globin cluster may not be representative, and (ii) the 0.9% sequence divergence between the two *M. caroli* chromosomes would normally be expected to produce more RFLPs due to point mutation than we observed in this instance (22). Also, the number of RFLPs due to point mutations will depend on the choice of restriction

enzyme. On the other hand, each L1 insertion affects all restriction digests. Therefore, screening with a single enzyme would elevate the fraction of RFLPs due to L1 insertion.

LITERATURE CITED

1. **Bankier, A. T., and B. G. Barrell.** 1983. Shotgun DNA sequencing, p. 1–34. *In* R. A. Flavell (ed.), Techniques in nucleic acid biochemistry, vol. B5. Elsevier Scientific, Limerick,, Ireland.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196:**180–182.
3. **Brown, B. A., R. W. Padgett, S. C. Hardies, C. A. Hutchison III, and M. H. Edgell.** 1982. β-Globin transcript found in induced murine erythroleukemia cells is homologous to the βh0 and βh1 genes. Proc. Natl. Acad. Sci. USA **79:**2753–2757.
4. **Burton, F. H., D. D. Loeb, S. F. Chao, C. A. Hutchison III, and M. H. Edgell.** 1985. Transposition of a long member of the L1 major interspersed DNA family into the mouse β-globin gene locus. Nucleic Acids Res. **13:**5071–5084.
5. **Burton, F. H., D. D. Loeb, C. F. Voliva, S. L. Martin, M. H. Edgell, and C. A. Hutchison III.** 1986. Conservation throughout mammalia and extensive protein encoding capacity of highly repeated DNA L1. J. Mol. Biol. **187:**291–304.
6. **Dover, G.** 1982. Molecular drive: a cohesive mode of species evolution. Nature (London) **299:**111–117.
7. **Edgell, M. H., S. C. Hardies, D. D. Loeb, W. R. Shehee, R. W. Padgett, F. H. Burton, M. B. Comer, N. C. Casavant, F. D. Funk, and C. A. Hutchison III.** 1987. The L1 family in mice, p. 107–129. *In* G. Stamatoyannopoulos and A. W. Nienhuis (ed.), Developmental control of globin gene expression. Alan R. Liss, Inc., New York.
8. **Fanning, T. G.** 1983. Size and structure of the highly repetitive *Bam*HI element in mice. Nucleic Acids Res. **11:**5073–5091.
9. **Fantoni, A., I. Bozzoni, E. Ullu, and M. G. Farace.** 1979. Construction of a recombinant bacterial plasmid containing DNA sequences for a mouse embryonic globin chain. Nucleic Acids Res. **6:**3505–3517.
10. **Gebhard, W., T. Meitinger, J. Hochtl, and H. G. Zachau.** 1972. A new family of interspersed repetitive DNA sequences in the mouse genome. J. Mol. Biol. **157:**453–471.
11. **Hardies, S. C., S. L. Martin, C. F. Voliva, C. A. Hutchison III, and M. H. Edgell.** 1986. An analysis of replacement and synonymous changes in the rodent L1 repeat family. Mol. Biol. Evol. **3:**109–125.
12. **Hill, A., S. C. Hardies, S. J. Phillips, M. G. Davis, C. A. Hutchison III, and M. H. Edgell.** 1984. Two mouse early embryonic β-globin gene sequences. J. Biol. Chem. **259:**3739–3747.
13. **Jahn, C. L., C. A. Hutchison III, S. J. Phillips, S. Waver, N. L. Haigwood, C. F. Voliva, and M. H. Edgell.** 1980. DNA sequence organization of the β-globin complex in the BALB/c mouse. Cell **21:**159–168.
14. **Jubier-Maurin, V., P. Wincker, G. Cuny, and G. Roizes.** 1987. The relationships between the 5′ end repeats and the largest members of the L1 interspersed repeated family in the mouse genome. Nucleic Acids Res. **15:**7395–7410.
15. **Karn, J., S. Brenner, L. Barnett, and G. Cesareni.** 1980. Novel bacteriophage λ cloning vector. Proc. Natl. Acad. Sci. USA **77:**5172–5176.
16. **Leder, P., J. N. Hansen, D. Konkel, A. Leder, Y. Nishioka, and C. Talkington.** 1980. Mouse globin system: a functional and

evolutionary analysis. Science 209:1336–1342.

17. Li, W.-H., T. Gojobori, and M. Nei. 1981. Pseudogenes as a paradigm of neutral evolution. Nature (London) 292:237–239.

18. Loeb, D. D., R. W. Padgett, S. C. Hardies, W. R. Shehee, M. B. Comer, M. H. Edgell, and C. A. Hutchison III. 1986. The sequence of a large L1Md elements reveals a tandemly repeated 5' end and several features found in retrotransposons. Mol. Cell. Biol. 6:168–182.

19. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

20. Martin, S. L., C. F. Voliva, F. H. Burton, M. H. Edgell, and C. A. Hutchison III. 1984. A large interspersed repeat found in mouse DNA contains a long open reading frame that evolves as if it encodes a protein. Proc. Natl. Acad. Sci. USA 81:2308–2312.

21. Martin, S. L., C. F. Voliva, S. C. Hardies, M. H. Edgell, and C. A. Hutchison III. 1985. Tempo and mode of concerted evolution in the L1 repeat family of mice. Mol. Biol. Evol. 2:127–140.

22. Nei, M., and W.-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76:5269–5273.

23. Ohta, T. 1985. A model of duplicative transposition and gene conversion for repetitive DNA families. Genetics 110:513–524.

24. Ohta, T. 1986. Population genetics of an expanding family of mobile genetic elements. Genetics 113:145–159.

25. Ohta, T., and G. A. Dover. 1984. The cohesive population genetics of molecular drive. Genetics 108:501–521.

26. Padgett, R. W., D. D. Loeb, L. R. G. Snyder, M. H. Edgell, and C. A. Hutchison III. 1987. The molecular organization of the beta-globin complex of the deer mouse, Peromyscus maniculatus. Mol. Biol. Evol. 4:30–45.

27. Phillips, S. J., S. C. Hardies, C. L. Jahn, M. H. Edgell, and C. A. Hutchison III. 1984. The complete nucleotide sequence of the β-globin like structure, βh2, from the [Hbb]d mouse Balb/c. J.

Biol. Chem. 259:7947–7954.

28. Phillips, S. L., N. C. Casavant, C. A. Hutchison III, and M. H. Edgell. 1985. An empirical method for the evaluation of the quality of genomic DNA libraries. Nucleic Acids Res. 13:2699–2708.

29. Rimm, D. L., D. Horness, J. Kucera, and F. R. Blattner. 1980. Construction of coliphage λ Charon vectors with BamHI cloning sites. Gene 12:301–309.

30. Rochaix, J.-D., F. Rougeon, and B. Mach. 1978. Electron microscope analysis of mouse and rabbit globin and immunoglobulin gene sequences. Gene 3:9–16.

31. Shyman, S., and S. Weaver. 1985. Chromosomal rearrangements associated with LINE elements in mouse genome. Nucleic Acids Res. 13:5085–5093.

32. Skowronski, J., and M. F. Singer. 1986. The abundant Line-1 family of repeated DNA sequences in mammals: genes and pseudogenes. Cold Spring Harbor Symp. Quant. Biol. 51:457–464.

33. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.

34. Staden, R. 1982. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res. 10:4731–4751.

35. Voliva, C. F., C. L. Jahn, M. B. Comer, C. A. Hutchison III, and M. H. Edgell. 1983. The L1Md long interspersed repeat family in the mouse: almost all examples are truncated at one end. Nucleic Acids Res. 11:8847–8859.

36. Voliva, C. F., S. L. Martin, C. A. Hutchison III, and M. H. Edgell. 1984. The dispersal process associated with the L1 family of interspersed repetitive DNA sequences. J. Mol. Biol. 178:795–813.

37. Wahl, G. M., M. Stern, and G. R. Stark. 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. USA 76:3683–3687.