Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 3875–3879, June 1986
Evolution

Insertion and/or deletion of many repeated DNA sequences in human and higher ape evolution

(molecular evolution/higher primates/mobile genetic elements/*Alu* family/*Kpn* I L1 family)

HUEY RU HWU*, JOHN W. ROBERTS*†, ERIC H. DAVIDSON*, AND ROY J. BRITTEN*‡

*Division of Biology, California Institute of Technology, Pasadena, CA 91125; and ‡Carnegie Institution of Washington, Washington, DC 20005

Contributed by Roy J. Britten, January 27, 1986

ABSTRACT **The total numbers of copies of two repeat families, L1 (*Kpn* I) and *Alu*, have been measured in the DNA of four higher apes by an accurate titration method. The number of members of the *Alu* family repeats in the four genomes are as follows: human, 910,000; chimpanzee, 330,000; gorilla, 410,000; orangutan, 580,000. For the *Kpn* I family (3′-ward higher frequency region) the number of copies in these genomes are as follows: human, 107,000; chimpanzee, 51,000; gorilla, 64,000; orangutan, 84,000. Thermal stability measurements show that, although the families of repeats are moderately divergent in sequence, little net sequence change has occurred during the evolution of the higher apes. Most or all of the members of these families of repeats are interspersed throughout the genome. Therefore, a large number of events of insertion and/or deletion of these DNA sequences has occurred during higher primate evolution.**

There are large interspecies differences in the number of copies of repeated DNA sequences in *Drosophila* (1–4) and sea urchins (5–7), and there may be comparable differences among vertebrate genomes (8). Families of repeats are apparently subject to large scale evolutionary changes both in sequence and frequency and have been described as a fluid component in the genome (9). In *Drosophila*, transposable elements are known to be a major source of the evolutionary changes but additional processes may also exist. In general, the relationship between interspersed repeated sequences and transposable genetic elements has not been established. Other types of repeated sequences and other mechanisms of insertion, deletion, or copying may be involved.

The work reported here is part of an investigation of the potential relationship between repeats and transposons and includes an examination of the magnitude of the recent evolutionary changes in two primate families of repeated sequences (L1 or *Kpn* I and *Alu*). These families are familiar and there exist recent reviews (10–14), so we briefly list only their primary characteristics.

L1Hs is a long interspersed repeat. We are using the recently proposed terminology (14), in which L1 stands for the first family of long interspersed repeats and Hs (*Homo sapiens*) is the animal species in which it occurs or from which a probe has been derived. Previous names have included *Kpn* I and Line 1. It was first studied in human DNA (15) and has been extensively studied in primate genomes (16–19); a related family occurs in mouse DNA (20, 21). A region near the 3′ end is homologous to many more repeats than the 5′ end and is included in the probe used in this work. Long terminal repeats are absent (14), a 3′-terminal adenine-rich region is present, and the term retroposon has been proposed for such elements (22–24) on the supposition that an RNA intermediate is involved in the mechanism of transposition.

Alu is a very high frequency short interspersed repeat of human and other primate DNA (ref. 25; reviewed in ref. 11). It is ≈300 nucleotides long and consists of an imperfect duplication of a shorter sequence, has an adenine-rich terminal region as well as a shorter internal adenine-rich sequence, includes a *Pol* III promoter, and also has been termed a retroposon (22–24).

METHODS

Libraries and Hybridization. A human Charon 4a partial digestion library (*Hae* III/*Alu*) was provided by Tom Maniatis and co-workers (26) and a gorilla *Eco*RI partial digestion library prepared by Alan Scott was provided by M. Weiss. Screening was done with nick-translated total human DNA or plasmid pRB1.8 (27) according to Benton and Davis (28) as modified by Anderson *et al.* (29). Hybridization was by standard methods (30), usually at reduced criterion [60°C; 4× SET (0.6 M NaCl/0.12 M Tris·HCl, pH 8.0/0.008 M EDTA)].

DNA Preparations. Human DNA was purchased from Calbiochem. The other primate DNAs were isolated from blood samples (31) obtained from Yerkes Regional Primate Research Center (Atlanta, GA). All of the DNA preparations were melted in an Acta III spectrophotometer, and the DNA concentrations were calculated from the optical density and hyperchromicity. In preparation for the titration measurements, the DNA was sheared by multiple passage through a 26-gauge needle.

Titration Probes. The source of our probe for the L1 family is a recombinant from the human λ library that was shown in earlier work (J.W.R., unpublished observations) to contain a 1.5-kilobase (kb) *Eco*RI/*Hin*dIII L1-related fragment. This fragment hybridizes with the 3′ end of the L1Ca repeat (clone pCa4A10, from the monkey *Cercopithecus aethiops*, a gift from Maxine Singer; see ref. 13). While we have not determined the precise overlap with pCa4A10, our 1.5-kb clone is cut by *Hin*fI into four fragments (0.60, 0.41, 0.26, and 0.18 kb), all of which hybridize to pCa4A10.

The 1.5-kb fragment was recloned into M13 mp10, labeled by primer extension, and isolated by restriction and gel electrophoresis. To remove complementary strands, the fragment preparation was incubated overnight in 0.6 M phosphate buffer at 60°C and passed over hydroxyapatite (0.12 M phosphate buffer; 50°C). About 50% of the label was single stranded and this fraction, designated Lhh, was used as the probe in the titration measurements. It hybridized almost 100% with primate genomic DNA and contained <1% complementary strands.

The probe used for the *Alu* family was derived from a 1.8-kb clone (pRB1.8; ref. 27), which was recloned into

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

Abbreviations: kb, kilobase(s); t_m, melting temperature.

†Present address: California State University at Dominguez Hills, Carson, CA 90747.

vector M13 mp10, extension-labeled, and purified on hydroxyapatite as described above. The labeled single-stranded preparation used as the titration probe is designated A1.8.

The amounts of the probes present in titration measurements were evaluated in two ways: first, from the specific radioactivity of the tracer in the extension labeling and second, by carrying out a titration measurement with linearized replicative form of the M13 clones. The two methods agreed to within 15%.

Titration Measurements. For the titration procedure, a fixed quantity of labeled probe and carefully controlled amounts of genomic DNA were reassociated to completion and assayed on hydroxyapatite to measure the amount of probe present in duplexes. The ratio of DNAs ranged over a factor of 10,000, from large excess of sites to large excess of probe. Incubation was carried out in 0.6 M phosphate buffer at 60°C to at least $C_0t = 0.3$, as calculated for the genomic sites or probe, depending on which was dominant. The sample was diluted to 0.12 M phosphate buffer and loaded on a hydroxyapatite column at 50°C. After thorough washing of the column with 0.12 M phosphate buffer, the temperature was raised to 95°C to elute the duplex fraction.

Least-Squares Analysis of Titration Measurements. The concentrations of the as yet unhybridized components of the reaction are defined as follows: Y, the single complement probe; S, the genomic strand with which it can form duplexes; and X, the other genomic strand. A is defined as the concentration of reassociated genomic duplexes (X:S) and B is defined as that of duplexes of the single complement probe with the appropriate complement of the genomic DNA (Y:S). The rate constants for formation of duplexes apparently differ and f is that for reassociation of the genomic sites, while h is that for hybridization of the probe with genomic sites; their ratio is $g = f/h$. At any time in the reaction, the rate of formation of duplexes is the product of the concentrations of the two single strands and the appropriate rate constant, thus:

$$dA/dt = fSX \text{ and } dB/dT = hSY.$$

Dividing, we obtain

$$dA/dB = gX/Y. \quad [1]$$

If (X_o) and (Y_o) are the initial concentrations, we have $X = X_o - A$, and $Y = Y_o - B$, and thus $dA/(X_o - A) = g \cdot dB/(Y_o - B)$.

Integrating, we obtain

$$\ln(X_o - A) = g \cdot \ln(Y_o - B) + C.$$

Using the initial conditions $A = 0$ and $B = 0$ to evaluate the constant of integration, $C = \ln(X_o) - g \cdot \ln(Y_o)$, and rearranging gives

$$\ln(1 - A/X_o) = g \cdot \ln(1 - B/Y_o). \quad [2]$$

The completion of the hybridization for each ratio of probe to genomic DNA yields complete duplex formation of the genomic DNA or $S = 0$ and $A = X_o - B$; therefore, Eq. 2 becomes $\ln(B/X_o) = g \cdot \ln(1 - B/Y_o)$. If we write the final unpaired fraction of the probe as $(U) = Y/Y_o$ and the initial ratio as $(R) = Y_o/X_o$, then $\ln[R(1 - U)] = g \cdot \ln U$, or

$$R = U^g/(1 - U) = D/W, \quad [3]$$

where D is the ratio of genomic DNA to probe and W is the value of this ratio when the number of sites is equal to the number of probe fragments. This equation reduces to the standard form

$$[(U/(1 + U)] \text{ if } g = 1. \quad [4]$$

Since it is awkward to express $U(R)$ for this transcendental equation, the least-squares analysis was done by calculating deviations in D for each data point by using guessed values of g and W. The best fit was obtained by a successive approximation method, calculating the rms error for all nine combinations of a small displacement d [i.e., g, $g(1 + d)$, $g(1 - d)$, W, $W(1 + d)$, $W(1 - d)$]. W and g were shifted by the combination giving the least error and the process was repeated until no improvement occurred; then d was reduced and the process was repeated until the displacement was negligibly small. The method converged rapidly for reasonable guesses.

RESULTS

Titration Measurements of the L1Hs Family of Repeats. Titration was carried out by incubating known quantities of a single complement probe and denatured genomic DNA to a sufficient C_0t (see *Methods*). The amount of hybridized probe was assayed for a range of DNA ratios from probe excess (where little hybridization is observed) to genomic DNA excess (where essentially 100% hybridization occurs), as shown in Fig. 1 for human and chimpanzee DNA. Since both complements are present in the genomic DNA, the expectation is that ≈50% of the probe will be duplexed when the quantity of complementary genomic DNA "sites" is equal to the quantity of probe. The probe used to assay the L1 (*Kpn* I) family was Lhh (see *Methods*), which represents a high-frequency region near the 3′ end of the L1Hs family of repeats in human DNA. To prepare this probe, a restriction fragment of an L1Hs family member was recloned into M13 mp10, so that a labeled single strand of well-defined length could be prepared by gel purification.

Table 1 lists the results of measurements using the Lhh probe with genomic DNA from four species, and the right-hand column shows the calculated number of copies relative to the number of copies in human DNA. Each of the other apes shows a significantly smaller L1 family repeat frequency than that of human DNA. The quantities of each of the

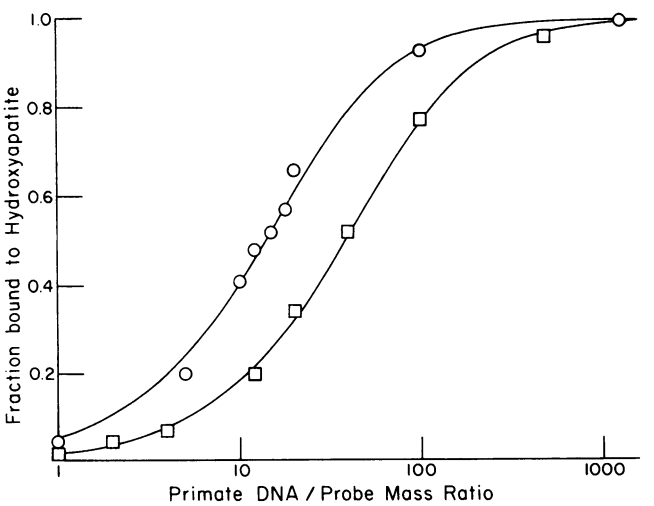

FIG. 1. Titration measurement of the number of L1 repeat family members in human (○) and chimpanzee (□) DNA. Probe Lhh was reassociated with total lightly sheared genomic DNA of the two species to $C_0t > 0.3$ (calculated for genomic site or probe depending on which was in excess). Sample was passed over hydroxyapatite in 0.12 M phosphate buffer at 60°C and the bound fraction was eluted in 0.12 M phosphate buffer at 95°C. The curves are the best fits for Eq. 3 with the rate ratio g set to 0.68. The values of W for these best fits are listed in Table 1.

Table 1. Titration measurements of *Alu* and L1 frequencies

Species	W*	No. of copies relative to human	No. of copies†
	Lhh probe		
Human	21.7	1.0	107,000
Chimpanzee	45.5	0.48	51,000
Gorilla	35.7	0.60	64,000
Orangutan	27.8	0.78	84,000
	A1.8 probe‡		
Human	1.82	1.0	910,000
Chimpanzee	5.00	0.36	330,000
Gorilla	4.00	0.45	410,000
Orangutan	2.86	0.64	580,000

*Least-squares solution for the ratio W of μg of genomic DNA to μg of probe DNA when equal numbers of genomic sites and probe sequences are present. The equations used are given in *Methods*. The ratio g of the rate of reassociation of genomic sites with each other relative to the rate of probe reassociation with genomic DNA was fixed at 0.68, which is the best fit value for six of the eight sets of measurements. When the rate ratio was made a free parameter, the frequency values changed by <10% in six of the eight cases and by <30% in two cases in which there was greater scatter in the data (Lhh on gorilla and A1.8 on orangutan). The best-fit frequencies are quite insensitive to the ratio of the rates g. If g is fixed at 1.0, all of the frequencies are slightly increased, and the maximum change is 15% in the number of copies of either repeat relative to human DNA in any species. It appears that the uncertainty in the determination of the relative number of copies in the different species is ≈15%.

†Assuming a genome size of 3×10^9 nucleotide pairs for all four species. The number of copies is calculated as follows: $N = G/F \times W$; where G is the genome size, F is the probe fragment size, and W is the ratio of the genomic to probe DNA when the number of sites and probe molecules are equal, as determined by least-squares analysis.

‡Only 65% of the A1.8 tracer reassociates with the primate DNAs, while almost 100% reassociates with replicative form DNA containing the A1.8 insert. We assume the low reassociation is due to breakage of probe fragments, since only 300 nucleotides out of 1.8 kb is *Alu* sequence. Another source of low reaction with primate DNA might be vector contamination, and in that case the *Alu* frequencies could be slightly overestimated. Uncertainty in the relative frequencies does not arise from this source.

genomic DNAs used in the measurement are based on optical density at 260 nm and are corrected for the hyperchromicity of each DNA preparation (see *Methods*). We conclude that there are about twice as many sequences that hybridize with the 3′ high-frequency region of the L1 family in the human genome than in the other ape genomes.

Table 1 also presents least-squares estimates of the frequencies (see *Methods*) and the footnote indicates our estimate of the accuracy. A new procedure was developed because systematic deviations were observed from the usual Eq. **4**, which assumes that the rates of reassociation are equal, and our results suggest that duplexes between genomic sequences form at ≈70% of the rate at which probe/site duplexes form. However, the rate of formation only has a minor effect on the results of the titration method and correction for it barely changes the relative L1 frequencies in the DNA of different species.

The genomic DNA reassociation is presumably retarded as a result of interspersed repeats that form networks (8) since the fragments were long enough so that each contained several repetitive sequences. Higher primate DNA is particularly likely to form such structures because of the high-frequency *Alu* repeat (10, 11). If there are several repeats per fragment, branched structures form early in the reassociation and steric hindrance retards the formation of subsequent genomic DNA duplexes. Since the probe is shorter, it is presumably less affected by the steric hindrance and reassociates at a faster rate with available sites, accounting for the fact that the ratio of the rates (g; see *Methods*) is <1.0.

Titration Measurements of the *Alu* Family of Repeats. For the studies of the *Alu* family of repeats, a previously cloned and sequenced member of the *Alu* family (27) was recloned into M13 mp10 and the 1.8-kb single complement, gel-purified after extension labeling, is identified as A1.8. About 300 nucleotides of this probe actually consist of the *Alu* repeat and the remainder is single-copy DNA. A series of titration measurements were made by using probe A1.8, with results similar to those described above for the probe Lhh, except that due to the higher frequency of the *Alu* family less genomic DNA was required. The right hand column of Table 1 shows the number of copies of the *Alu* sequence relative to human DNA for the three other species. The conclusion is that the number of copies of the *Alu* sequence is about 2-fold higher in the human genome than in the genomes of the other higher apes.

Evolutionary Sequence Divergence of the Two Repeat Families. Individual members of a family of repeated sequences could, in principle, change in sequence by base substitution to such an extent as to fail to hybridize with a given probe sequence, and such processes might contribute to the apparent evolutionary frequency change, assayed by hybridization. To assess this possibility, we have used thermal stability as a measure of the average interspecies sequence differences for the L1Hs and *Alu* families.

Probes Lhh and A1.8 were incubated to high C_0t with an excess of genomic DNA of each of the four species and the duplexes bound to hydroxyapatite. The bound fraction was eluted at a series of increasing temperatures, and the midpoint of the resulting melting curves (t_m) is shown in Table 2. The last line of the table is a control showing the t_m of the probe self-reassociated with the replicative form of the clone. The L1Hs family in human DNA shows a 9°C reduction in t_m compared to the control, or ≈9% divergence of the typical family member from the cloned sequence. There is a very slight additional t_m reduction for the L1 families in the other three species, indicating little change in the typical L1 sequence during the evolution of the apes.

Table 2. Evolutionary sequence divergence of repeat families shown by thermal stability of hybrids of probes to genomic DNA

	Family: Probe*		Single copy,§
	L1 (line 1): Lhh†	*Alu*: A1.8‡	Δt_m
Human	78	73.1	0
Chimpanzee	75.7	72.5	1.63
Gorilla	76.5	72.8	2.25
Orangutan	77	73	3.57
Control¶	87	88	

*Single-stranded probes cloned in M13 (labeled by primer extension and isolated by restriction and gel electrophoresis) were hybridized in solution (0.6 M phosphate buffer at 60°C) with an excess of genomic DNA and the thermal stability of duplexes was assayed with hydroxyapatite (see *Methods*). Listed is the temperature (°C) for 50% elution of bound duplexes (t_m).

†Probe Lhh (see *Methods*) is a 1.5-kb region of human DNA inserted into vector M13 mp10, labeled by extension, and reisolated. It contains some or all of the high-frequency 3′ region of one copy of L1Hs.

‡Probe A1.8 (see *Methods*) is derived from pRB1.8 (27) by recloning into M13 mp10.

§Single-copy DNA sequence divergence for comparison, relative to human DNA based on hydroxyapatite thermal stability measurements (Charles Sibley, personal communication). Listed is the median reduction in thermal stability (°C) of interspecies single-copy DNA hybrids.

¶Driver in this case is the complete replicative form of M13 mp10 containing, respectively, the insert Lhh or A1.8, sheared and denatured, including sea urchin DNA as "carrier."

There is a somewhat larger divergence from our probe (≈15%) for the typical *Alu* family member in human DNA and no significant additional interspecies divergence. Thus, sequence divergence contributed little if at all to the frequency change measurements (Table 1) for either repeat family.

For comparison purposes, the right-hand column shows current estimates of the single-copy divergence that has occurred during the evolution of these species (Charles Sibley, personal communication). It appears that the two repeat families showed less sequence change than the typical single-copy DNA sequences. This is not unexpected, since all previous studies indicate slow evolutionary repeated sequence change (5–7, 32–34) compared to single-copy change. There appears to be no correlation of the slight sequence divergence of these repeat families with the single-copy divergence or the probable relationshps of these species.

Interspersion of the Families of Repeats. The interspersion of these families of repeats is important in the interpretation of the large evolutionary frequency changes. We have screened the human library with the Lhh probe and estimate that 120/350 = 34% of recombinant clones contain a sequence homologous to the 3′ high-frequency end of the L1Hs family. This confirms previous estimates (50,000) (35) and gives a slightly higher total frequency (64,000), which is consistent with our estimate from titration (107,000). This comparison indicates that a large fraction of the copies of the L1Hs family are interspersed throughout the genome. There is no published evidence of which we are aware that any clusters or tandem patterns of the L1Hs family occur in human DNA.

Table 3 shows the results of a much more detailed examination that was made for the *Alu* family using the probe pRB1.8. Screening the human library showed that only 6% of recombinants failed to hybridize with this probe. Sixteen examples of the clones that did hybridize to the probe were picked at random for further analysis. DNA from each of the clones was digested with *Eco*RI, the fragments were separated by gel electrophoresis, transferred to nitrocellulose, and hybridized with pRB1.8. As the table shows, in one case five bands were observed to hybridize and in two cases four bands hybridized.

Table 3. Interspersion of the *Alu* family by restriction analysis of λ library inserts

No. of fragments* hybridizing to *Alu* probe	No. of cases	
	Observed†	Best fit‡
0	(1)	0.32
1	4	3.0
2	5	5.9
3	4	4.5
4	2	1.8
5	1	0.5
6	0	0.09

*Sixteen recombinant phage that hybridized to the *Alu* probe (pRB1.8) were digested with *Eco*RI, and the fragments were electrophoresed and transferred to nitrocellulose filters. Number of different fragments per insert that hybridized to pRB1.8 is shown.

†Observed number of inserts with various numbers of hybridizing fragments. The first entry (cases with 0 hybridizing fragments) is estimated from data, which show that 6% of inserts fail to hybridize with pRB1.8.

‡Computer modeled distribution of the number of *Eco*RI fragments per insert that contain *Alu* repeats. The average insert was set at 16.5 kb (matching the average of the 16 inserts) and contained 5.4 fragments. The average number of *Alu* repeats per insert was varied and the best fit was 3.8 for the example shown, implying a spacing of 4.3 kb or 700,000 interspersed copies. There was no statistically significant difference between the observed and best fit distributions, suggesting a random distribution of *Alu* repeats throughout most of the human genome.

The expected distribution for a random arrangement of *Alu* repeats is not Poisson, since individual fragments occur in a large range of sizes. A computer model of the distribution was used that allowed for the variation in fragment size, and the results of the best fit are shown in the right-hand column, suggesting an average spacing of 4.3 kb or 700,000 interspersed copies, consistent with the results of the titration measurements. Therefore, it appears that most copies of the *Alu* sequence are interspersed through the genome and the pattern is consistent with a random distribution. A small fraction of the *Alu* sequences are present in human satellite I (36), apparently occurring once per 2.47-kb repeat, thus interspersed in satellite DNA.

Interspecies Frequency Comparison by λ Library Screening. Human and gorilla total genomic λ libraries were screened with the *Alu* family probe pRB1.8 and with labeled total human DNA fragments. As mentioned above, 6% of the human library recombinants do not hybridize with the *Alu* probe, while 4% fail to hybridize with the total human DNA probe, consistent with the interspersed repeat frequency data above.

The data for the gorilla library are very different. Only 67% of the recombinants hybridize with the *Alu* probe, while 88% hybridize with the total human repeat probe. Taken together, these two observations show that the gorilla genome has a much smaller number of interspersed high-frequency repeats than the human genome. It has to be considered that the gorilla library could contain clones without gorilla DNA. Nevertheless, the 21% (88 minus 67) that hybridize to the total human DNA probe and do not hybridize to the *Alu* probe do contain gorilla DNA but no *Alu* sequence. This result shows that the number of interspersed *Alu* sequences in gorilla DNA is much less than in human DNA and is consistent with the titration data of Table 1. Due to the lower *Alu* and L1 frequencies in gorilla DNA, it is quite possible that 12% of the recombinants contain no high- or middle-frequency repeats and thus all of the clones may include gorilla DNA.

DISCUSSION

The evidence summarized in Table 1 shows that during higher ape evolution there have been large changes of the number of copies of the high-frequency 3′ region of the L1Hs family and of the *Alu* family. Previous evidence, confirmed by observations in this paper, shows that the members of these families of repeats are primarily interspersed throughout the genome. The conclusion is that many interspersed copies have been added or deleted from the higher primate genomes in the last few million years of evolution.

The very existence of the interspersed repeats indicates that there have been many insertion events, but we do not know if they occur regularly or in irregular bursts. In addition, our observations do not distinguish between insertion and deletion processes. For example, a large event of multiplication could have occurred in the genome of an ancestor of the higher apes and there could have been differences in the rates of deletion of repeats among the ape lineages. On the other hand, insertion processes could have been dominant and the interspersed frequency might have increased at a different rate in each of the ape lineages. Most likely, a combination of insertion and deletion events have continued to occur during the evolution of these genomes. When closely spaced copies exist in the genome there may be a high probability of deletion or multiplication by unequal crossover. For unequal crossover to balance a rapid insertion process, natural selection would have to favor deletion over duplication and intervening stretches of DNA would be sacrificed.

There is some evidence regarding insertion and deletion processes for animals other than primates. In the case of sea urchins, the total sets of repeated sequences are very similar in quantity for different species of the genus *Strongylocentrotus* but the genomes of each species are dominated by different high-frequency families of repeats (5–7), implying a balance between deletion and insertion processes. Furthermore, in *Drosophila* the majority of the repeat families in each of the sibling species *D. melanogaster*, *D. simulans*, and *D. erecta* are absent or very nearly absent from each of the other species' genomes (3, 4), suggesting that both insertion and deletion processes are active, although not necessarily balanced.

The frequency differences predict that many repeats will be present in the genome of one ape and absent from the homologous location in the genome of another ape, yet no such local differences between primate genomes have been directly observed. The seven *Alu* family sequences in the globin-gene region that have been compared between chimpanzee and human are in identical locations (37). It may be that there is selection against changes in sequence organization in the globin-gene region.

The minimum number of insertions or deletions since the separation of the lineages leading to chimpanzee and human is ≈56,000 for sequences homologous to the 3′ end of *Kpn* I and 580,000 for the *Alu* family, as shown in Table 1. This is a lower limit, since many insertions could have been compensated by deletions. For comparison, the number of base substitution differences between the genomes of these species is ≈32,000,000 (38). However, each event in the *Alu* family affected ≈300 nucleotides and thus a minimum of 150,000,000 nucleotides was inserted or deleted since these lineages separated. A comparable calculation cannot now be made for the L1Hs family since we do not know the typical length that is inserted or deleted. This comparison with base substitution leaves no doubt that insertion/deletion of repeats is a major source of evolutionary change in these genomes.

The events of insertion/deletion and the frequency change of families of repeats probably occur generally among eukaryotes. For example, rapid evolutionary frequency changes of interspersed repeats occur in echinoderm genomes (5–7) and insect genomes (1–4). It has been argued (8) that such changes occurred during the evolution of many vertebrate genomes. Furthermore, the interspersed repeats may lead to frequent events of unequal crossover. It is an open question as to how much significant variation and evolutionary change in the phenotype are caused by these processes of genomic rearrangement and the classes of change they might cause. If they are not always destructive in their effects and if a minority cause changes in gene regulation, they could be a prominent cause of eukaryotic evolutionary change.

We express appreciation to Maxine Singer for providing clones of L1 repeat and to Alan Scott for the gorilla library. Primate blood samples were supplied by the Yerkes Regional Primate Center, Atlanta, GA. This work was supported by National Institutes of Health Grant GM34031. H.R.H. was supported by Kempner Fellowship from University of Texas Medical Branch at Galveston and Procter & Gamble Postdoctoral Fellowship from Caltech.

1. Potter, S. S., Brorein, W. J., Jr., Dunsmuir, P. & Rubin, G. M. (1979) *Cell* **17**, 415–427.
2. Strobel, E., Dunsmuir, P. & Rubin, G. M. (1979) *Cell* **17**, 429–439.
3. Dowsett, A. P. & Young, M. W. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4570–4574.
4. Dowsett, A. P. (1983) *Chromosoma* **88**, 104–108.
5. Moore, G. P., Scheller, R. H., Davidson, E. H. & Britten, R. J. (1978) *Cell* **15**, 649–660.
6. Moore, G. P., Pearson, W. R., Davidson, E. H. & Britten, R. J. (1981) *Chromosoma* **84**, 19–32.
7. Moore, G. P., Constantini, F. D., Posakony, J. W., Davidson, E. H. & Britten, R. J. (1980) *Science* **208**, 1046–1048.
8. Britten, R. J. (1984) *Carlsberg Res. Commun.* **49**, 169–178.
9. Young, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6274–6278.
10. Schmid, C. W. & Jelinck, W. R. (1982) *Science* **216**, 1065–1070.
11. Schmid, C. W. & Shen, C.-K. J. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 323–358.
12. Singer, M. F. (1982) *Cell* **28**, 433–434.
13. Grimaldi, G., Skowronski, J. & Singer, M. F. (1984) *EMBO J.* **3**, 1753–1759.
14. Singer, M. F. & Skowronski, J. (1985) *Trends Biochem. Sci.* **10**, 119–122.
15. Adams, J. W., Kaufman, R. E., Kretschmer, P. J., Harrison, M. & Nienhuis, A. W. (1980) *Nucleic Acids Res.* **8**, 6113–6128.
16. Thayer, R. E. & Singer, M. F. (1983) *Mol. Cell. Biol.* **3**, 967–973.
17. Grimaldi, G. & Singer, M. F. (1983) *Nucleic Acids Res.* **11**, 321–338.
18. Lerman, M. I., Thayer, R. E. & Singer, M. F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3966–3970.
19. Miyake, T., Migita, K. & Sakaki, Y. (1983) *Nucleic Acids Res.* **11**, 6837–6846.
20. Fanning, T. G. (1983) *Nucleic Acids Res.* **11**, 5073–5091.
21. Singer, M. F., Thayer, R. E., Grimaldi, G., Lerman, M. I. & Fanning, T. G. (1983) *Nucleic Acids Res.* **11**, 5739–5745.
22. Jagadeeswaran, P., Forget, B. G. & Weissman, S. M. (1981) *Cell* **26**, 141–142.
23. Van Arsdell, S. W., Denison, R. A., Bernstein, L. B. & Weiner, A. M. (1981) *Cell* **26**, 11–17.
24. Rogers, J. H. (1985) *Biochim. Biophys. Acta* **824**, 113–120.
25. Houck, C. M., Reinehart, F. P. & Schmid, C. W. (1979) *J. Mol. Biol.* **132**, 289–306.
26. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
27. Baralle, F. E., Schoulders, C. C., Goodbourn, S., Jeffreys, A. & Proudfoot, N. J. (1980) *Nucleic Acids Res.* **8**, 4393–4404.
28. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
29. Anderson, D. M., Scheller, R. H., Posakony, J. W., McAllister, L. B., Trabert, S. G., Beall, C., Britten, R. J. & Davidson, E. H. (1981) *J. Mol. Biol.* **145**, 5–28.
30. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
31. Graham, D. E. (1978) *Anal. Biochem.* **85**, 609–613.
32. Harpold, M. M. & Craig, S. P. (1978) *Differentiation* **10**, 7–11.
33. Hall, T. J., Grula, J. W., Davidson, E. H. & Britten, R. J. (1980) *J. Mol. Evol.* **16**, 95–110.
34. Scheller, R. H., Anderson, D. M., Posakony, J. W., McAllister, L. B., Britten, R. J. & Davidson, E. H. (1981) *J. Mol. Biol.* **149**, 15–39.
35. Schmid, C. W., Fox, G. M., Dowds, B., Lowensteiner, D., Paulson, K. E., Shen, C.-K. J. & Leinwand, L. (1983) in *Perspectives on Genes and the Molecular Biology of Cancer*, eds. Robberson, D. L. & Saunders, G. F. (Raven, New York), pp. 35–41.
36. Frommer, M., Prosser, J. & Vincent, P. C. (1984) *Nucleic Acids Res.* **12**, 2887–2900.
37. Sawada, I., Willard, C., Shen, C.-K. J., Chapman, B., Wilson, A. & Schmid, C. W. (1985) *J. Mol. Evol.* **22**, 316–322.
38. Sibley, C. G. & Ahlquist, J. E. (1984) *J. Mol. Evol.* **20**, 2–15.