J Mol Evol (1985) 22:316–322

Journal of
Molecular Evolution

Evolution of Alu Family Repeats Since the Divergence of Human and Chimpanzee

Ikuhisa Sawada,[1] Cary Willard,[1] Che-Kun James Shen,[2] Barbara Chapman,[3]* Allan C. Wilson,[3] and Carl W. Schmid[1]

[1] Department of Chemistry, University of California, Davis, California 95616, USA
[2] Department of Genetics, University of California, Davis, California 95616, USA
[3] Department of Biochemistry, University of California, Berkeley, California 94720, USA

Summary. The DNA sequences of three members of the Alu family of repeated sequences located 5′ to the chimpanzee $\alpha 2$ gene have been determined. The base sequences of the three corresponding human Alu family repeats have been previously determined, permitting the comparison of identical Alu family members in human and chimpanzee. Here we compare the sequences of seven pairs of chimpanzee and human Alu repeats. In each case, with the exception of minor sequence differences, the identical Alu repeat is located at identical sites in the human and chimpanzee genomes. The Alu repeats diverge at the rate expected for nonselected sequences. Sequence conversion has not replaced any of these 14 Alu family members since the divergence between chimpanzee and human.

Key words: Alu repeats — Primate evolution

Introduction

Members of a family of interspersed repeated DNA sequences are inexact copies of each other and presumably of their ancestral sequence (Britten and Kohne 1968). However, a family of repeated sequences from one species often exhibits specific sequence differences from members of the same family in a divergent species. This curious property of interspersed repeats, that they are heterogeneous sequences within one species but are homogenized for species-specific differences, has long been appreciated (Britten and Kohne 1968; Rice 1972; Deininger and Schmid 1979). The results of more modern cloning and sequencing studies have confirmed and extended these early findings. The two major families of interspersed repeats within primates and rodents, the Alu and L1 families (corresponding to short and long interspersed repeats), have now been studied in detail by sequencing techniques.

There are pronounced structural differences between the human Alu family and the equivalent family of sequences in rodents, often called the B1 family. Most notably, the human Alu sequences are organized as 300-nucleotide (nt)-long dimeric structures, whereas the rodent analogue is typically a 130-nt-long monomer sequence (Schmid and Jelinek 1982). No examples have been found of the human dimer in rodents or of rodent monomers in humans [for a review, see Schmid and Shen (1985)]. The human and rodent Alu families are therefore recognizably different and each has effectively been homogenized with respect to this difference. At least one additional homogenization of the Alu family has occurred since the divergence between humans and prosimian primates. There are two distinct variants of the Alu family in the galago genome, one of which closely resembles the human Alu family (Daniels and Deininger 1983; Daniels et al. 1983). Yet recognizable sequence differences exist between the human and closely related galago Alu families, so that individual Alu repeats isolated from these two organisms are distinctly different.

The L1 family of long interspersed repeats (pre-

Offprint requests to: C.W. Schmid
* *Present address:* Chiron Research Laboratories, Emeryville, California 94608, USA

viously called the Kpn, Bam H1, or MIF family) in rodents and primates illustrates many of these same principles. In exact analogy to the Alu family, the rodent and primate variants of the L1 family are distinct (Singer et al. 1983). Specific sequence differences in the L1 family have also been homogenized in divergent species of mice (Brown and Dover 1981; Martin et al. 1985).

Two models account for the establishment of species-specific differences, often termed concerted evolution, in interspersed repeats (Martin et al. 1985). One involves the conversion of existing sequences into one or more new master copies. In the other, simple turnover of the membership of a family eventually replaces all of its existing members with sequences that were derived from one or more new master copies. Conversion is well documented in, for example, repetitive ribosomal gene families (Arnheim et al. 1980; Dover and Flavell 1984). This conversion is thought to result from homologous recombination. Recent evidence that families of interspersed repeats may be inserted by way of an RNA intermediate (Jagadeeswaran et al. 1981; Van Ardsell et al. 1981) suggests a mechanism for the turnover model. The L1 family has a long open reading frame, so the hypothetical RNA intermediate might be under biological selection (Martin et al. 1984). The Alu family is closely related to functional 7S RNA; thus the postulated RNA precursor to new Alu family repeats also could be selected (Ullu et al. 1982). Biological selection for a master sequence(s) coding for new members of a repeat sequence family could ultimately result in the fixation of the species-specific differences.

We propose to test the conversion model in the Alu family. The conversion of a particular interspersed repeat in one organism could be detected as a change in its sequence as compared with that of the corresponding family member at an identical position in the genome of some divergent species. We employ this approach in a study of human and chimpanzee Alu repeats.

Human Alu family members differ from their consensus sequence by 14% (Deininger et al. 1981; Jelinek and Schmid 1982; Schmid and Shen 1985). Any two Alu family members exhibit independent mutations relative to the consensus sequence, and of the 50 human Alu repeats sequenced to date, each is distinct from all other members of the family. The correction in whole or part of an existing Alu repeat to the sequence of another member of the family should be obvious.

The human alpha-globin gene cluster provides an excellent system with which to identify particular members of the Alu family. At least ten Alu family members are present in the 40-kb region occupied by this gene cluster, including eight that occupy the

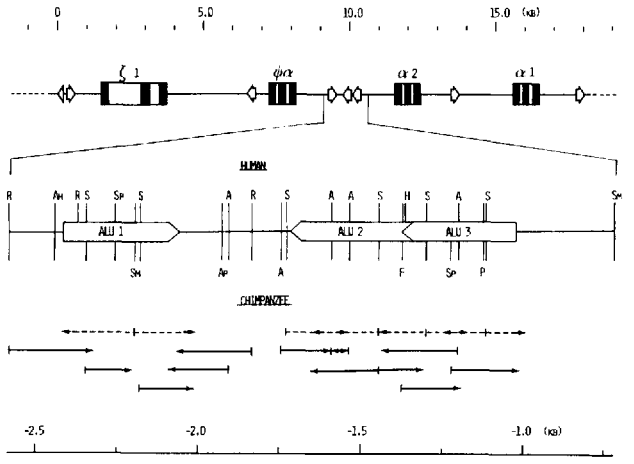

Fig. 1. Alu family members as indicated by open arrows are interspersed throughout the human and chimpanzee alpha globin gene clusters (Lauer et al. 1980; Zimmer et al. 1980; Proudfoot et al. 1982; Hess et al. 1983; Sawada et al. 1983; Willard et al. 1985). The sequences of three Alu family members flanking the chimpanzee alpha 2 gene were determined by m13 subcloning and dideoxy sequencing as indicated by arrows superimposed on the restriction map. The sequences of the three Alu repeats flanking the human $\alpha 2$ gene were determined previously (Hess et al. 1983). For the accuracy required in this study it was necessary to scrutinize possible human–chimpanzee sequence differences by comparing side-by-side human–chimpanzee sequencing gels as indicated by dashed arrows. Additional comparisons of human and chimpanzee Alu sequences were made for those located 5′ to the pseudo alpha gene (Sawada et al. 1983) and for the full-length and partial Alu sequences located 5′ to the $\zeta 1$ gene (Willard et al. 1985). Restriction cleavage sites as follows: A, Alu 1; AH, Aha III, AP, Apa I; F, FnuD II; H, Hpa II; P, Pvu II; R, Rsa I; S, Sau 3A; SM, Sma; SP, Sph I. Restriction site differences are indicated by presence of sites in only the human or chimpanzee map. The numbering for the region sequenced is based on using the transcriptional start site of the $\alpha 2$ gene as position zero

15-kb region depicted in Fig. 1 (Proudfoot et al. 1982; Shen and Maniatis 1982; Hess et al. 1983; Sawada et al. 1983; Willard et al. 1985). Restriction mapping demonstrates that the structure of this gene cluster is very similar in human and chimpanzee (Zimmer et al. 1980). For this reason it is possible to identify the identical Alu repeats in human and chimpanzee for our proposed test of conversion. In this work we report the base sequences of three Alu repeats located 5′ to the chimpanzee $\alpha 2$ gene (Fig. 1). These, the Alu repeat reported in the companion paper (Willard et al. 1985), and previous comparisons of chimpanzee and human Alu family members from the α-like and β-like cluster (Maeda et al. 1983; Sawada et al. 1983) provide a data base of seven pairs of chimpanzee and human Alu repeats.

Materials and Methods

Lauer et al. (1980) described the recombinant DNA containing the human α globin gene cluster. The chimpanzee clone reported

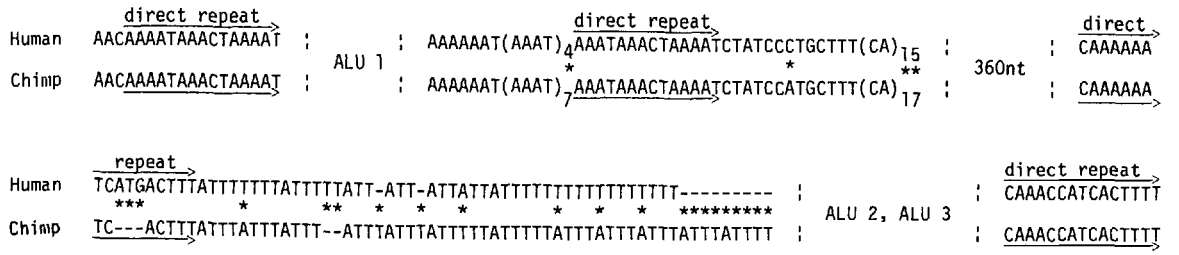

Fig. 2. Comparison of the human and chimpanzee sequences of the regions immediately around the Alu family members located 5′ to the alpha 2 globin gene. Arrows indicate probable direct repeats flanking Alu 1 and the dimeric pair Alu 2–Alu 3. Differences are indicated by asterisks (*), and the lengths of tandem runs in each sequence are indicated by numerical subscripts. The base sequences of Alu 1, 2, and 3 are summarized in Fig. 3

by Sawada et al. (1983) proved to be a mixture of truncated recombinants, several of which extended past the Hind III site of the $\alpha 2$ gene. A full-length sibling clone was reisolated by screening the λ Charon 30 chimpanzee library constructed by J. Slightom, University of Wisconsin. The resulting clone extends through the Hind III site in the $\alpha 2$ gene, as predicted, and has an unrearranged genomic restriction map (Zimmer et al. 1980). DNA sequence determinations were accomplished by the M13 dideoxy method as described in the companion paper (Willard et al. 1985). As indicated in Fig. 1, certain key regions of the human and chimpanzee DNAs were determined using side-by-side sequencing to highlight differences.

Results

The three Alu family repeats located 5′ to the $\alpha 2$ gene are termed $\alpha 2$:Alu 1, 2, and 3, respectively, in the 5′ to 3′ direction. Alu 1 is inverted with respect to Alu 2 and Alu 3, which are fused into a 600-nt dimeric structure (Fig. 1). The base sequence of this region in the chimpanzee was determined by the strategy of Fig. 1. Hess et al. (1983) previously determined the base sequence of this same region in human DNA. This region has been resequenced to achieve the accuracy required in this study (Fig. 1). In particular, the sequence differences reported here were confirmed by running the human and chimpanzee samples side by side. This verification of the human sequence resulted in changes from the published human sequences, mostly at sites of ambiguous readings in the original data.

Although our primary interest is in the evolution of the Alu family members, there are a few noteworthy features in the sequences that immediately flank the Alu family members (Fig. 2). Included in the immediate flanking regions are several runs of simple DNA sequences that have changed in length since the divergence between chimpanzee and human (Fig. 2). $\alpha 2$:Alu 1 is flanked by a tandem array of alternating CA residues. Alternating CA residues have been observed as an evolutionarily conserved family of repeats (Miesfeld et al. 1981). There are two extra CA dinucleotides in the chimpanzee sequence as compared with the human sequence (Fig. 2). As reviewed in the companion paper and as illustrated by the tandem repeat that occupies the second intron of the $\xi 1$ gene, the length of simple sequence runs can undergo abrupt changes in evolution (Jeffreys et al. 1985; Willard et al. 1985). The tandem array of the repeat unit AAAT has increased in length by three units in chimpanzee as compared with human. It is noteworthy that this tandem array of AAATs constitutes the poly(A)-rich tail that is hypothetically associated with the insertion of Alu repeats by way of a cDNA intermediate (Jagadeeswaran et al. 1981; Van Ardsell et al. 1981). Although this hypothesis is quite plausible, the difference in length between the human and chimpanzee A-rich regions suggests two additional considerations. First, the structure of the poly(A) tail is not necessarily identical to that of the original insertion element, but can vary significantly, as observed in other simple sequences (Jeffreys et al. 1985; Willard et al. 1985). Second, the tandem run of AAAT in the present example constitutes part of the short direct repeats that flank the Alu family member (Fig. 2). Since the flanking direct repeat is undoubtedly a duplication of the genomic target site, this implies that at least part of the A-rich tail is derived from preexisting genomic DNA and is not part of the inserted element. These minor details are not explained by current models for the insertion of repeated sequences. Similar considerations arise with respect to the A-rich tail of $\alpha 2$:Alu 2, which, in the orientation of Fig. 2, is a T-rich sequence. The chimpanzee sequence consists of an imperfect homopolymeric run of the element ATTT, whereas the human sequence is more nearly a run of Ts (Fig. 2). The chimpanzee variant includes two additional copies of the repeat unit ATTT. Except for these simple sequence runs, length mutations between chimpanzee and human DNAs are rare (Chang and Slightom 1984; Willard et al. 1985). It is significant that length mutations occur so readily within simple sequence runs; Jeffreys et al. (1985) found that "minisatellite" regions in human DNA are hypervariable.

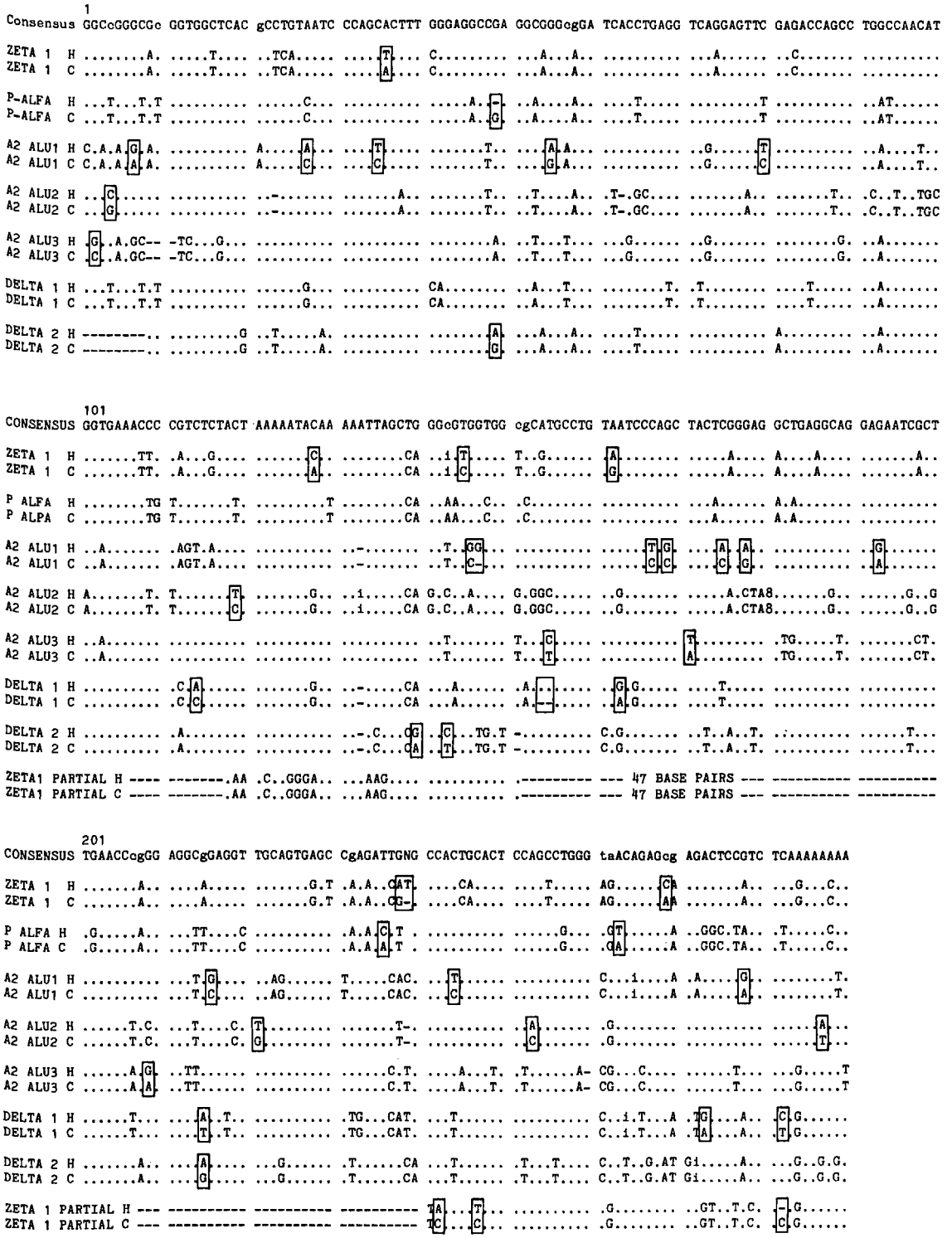

Fig. 3. Sequences of Alu family members in various genes. The consensus sequence is derived from a comparison of human Alu family members (Deininger et al. 1981; Schmid and Shen 1985). Dots indicate agreement between the consensus and a particular Alu repeat. Differences between the consensus and individual Alu repeats are indicated by the appropriate base, deletions by a dash, and insertions by either "i" or a numerical value. In each case a particular human (H) and chimpanzee (C) Alu repeat are compared. The pairs of human and chimpanzee Alu repeats are as follows: ZETA 1, an Alu repeat positioned 5′ to the ζ1 globin gene (Fig. 1) (Willard et al. 1985); P ALFA, an Alu repeat positioned 5′ to the pseudo alpha globin gene (Fig. 1) (Sawada et al. 1983); A2 ALU1, ALU2, and ALU3, the three Alu sequences located 5′ to the alpha 2 gene (Fig. 1) as determined in this work and previously by Hess et al. (1983); DELTA 1 and DELTA 2, Alu repeats positioned near the human and chimpanzee delta globin gene (Maeda et al. 1983); ZETA 1 PARTIAL, an additional region located 5′ to the human and chimpanzee ζ1 genes that has partial homology to an Alu repeat (Fig. 1) (Willard et al. 1985)

Table 1. Divergence of Alu repeats[a]

Alu comparison	Number (N) of differences	Probability of N differences	Probability of N or more differences
ζ1	7	0.14	0.41
ψα	3	0.08	0.95
α2:Alu1	15	0.001	0.025
α2:Alu 2	4	0.13	0.86
α2:Alu 3	4	0.13	0.86
δ:Alu 1	6	0.16	0.58
δ:Alu 2	4	0.13	0.86

[a] The number of differences between the pairs of chimpanzee and human Alu repeats is taken from the sequence comparison of Fig. 3. The probability of observing N differences, P_N, is estimated from the binomial equation assuming an average of 6.14 differences in the 282-nt sequence: $P_N = [282!/(282 - N)!N!]p^N q^{282-N}$, where $p = 6.1/282$ and $q = 1 - p$. The probability of observing at least N differences is taken as 1 minus the sum of the probabilities of observing zero to N − 1 differences

The base sequences of Alu 1, 2, and 3 located 5′ to the chimpanzee α2 gene are compared with the identical Alu repeats located 5′ to the human α2 gene Fig. 3 (see also Hess et al. 1983). Also included in this compilation is an Alu repeat located 5′ to the chimpanzee and human ψα genes and a full-length and a partial Alu located 5′ to both the human and chimpanzee ξ1 genes (Fig. 1) (Sawada et al. 1983; Willard et al. 1985). Maeda et al. (1983) previously compared the base sequences of two pairs of chimpanzee and human Alu repeats located near the δ globin gene (Fig. 3). The sequences of these seven pairs of chimpanzee and human Alu repeats are aligned with respect to a human genomic Alu family consensus sequence (Schmid and Shen 1985). Unlike the consensus sequence derived for renatured DNA samples (Deininger et al. 1981), this new consensus is derived from 25 genomic sequences and provides a more accurate representation of the ends of the Alu repeats, which were "nibbled" by S1 enzyme in the renatured samples. Except for the ends, this new consensus is in excellent agreement with that of Deininger et al. (1981). Individual Alu repeats have an average of about 14% divergence from this consensus (Deininger et al. 1981; Schmid and Shen 1985). The chimpanzee Alu repeats reported in Fig. 3 match the human consensus as accurately as do the corresponding human Alu sequences.

The human and chimpanzee sequences in each pair of Alu repeats share a common set of mutations relative to the consensus sequence (Fig. 3). Of the 50 human Alu family members sequenced to date, no two share a common set of mutations relative to the consensus; each member of the family is distinguishable from the others (Deininger et al. 1981; Schmid and Shen 1985). The shared set of mutations unambiguously identifies each pair of human and chimpanzee Alu's as representing the same member of the family, and each pair is located at the same site in these two species (Fig. 3). Conversion of one member of any pair to a new master sequence following the divergence between human and chimpanzee would have eliminated this pairwise identity. With one exception, shared mutations relative to the consensus are found throughout the length of each pair of Alu repeats (Fig. 3). Partial conversion of one member of a pair to a new master sequence or an adjacent family member would eliminate the identity of a pair within the converted region. A possible example of partial conversion is a 55-nt region (positions 145–200) of the α2:Alu 1 pair (Fig. 3). In this region the chimpanzee and human Alu repeats differ by 7-nt and do not share any mutations relative to the consensus sequence. This Alu may have recombined with an adjacent member of the family in either the human or chimpanzee lineage. Supporting this possibility, Alu family members occupy identical positions 5′ to the duplicate human α2 and α1 globin genes and define an end point in an α globin conversion unit (Hess et al. 1983, 1984). A block of four of the seven differences between the human and chimpanzee α2: Alu 1 sequences (positions 167–178) is present in the Alu flanking the human α1 gene. This conversion could conceivably have resulted from recombination between α2:Alu 1 and the corresponding α1 Alu in chimpanzee. In summary, with the possible exception of a 55-nt region in one pair, conversion has not acted in whole or substantial part on any of these seven full-length Alu repeats since the divergence between human and chimpanzee.

Some mutational change has occurred in each of the pairs of Alu repeats since the divergence between the species (Fig. 3, Table 1). In Table 1, the Alu family member is defined as occupying positions 1–282 and only mutations in this region are scored. The number of mutations in the seven pairs of Alu repeats ranges from 3 to 15 (Fig. 3, Table 1). The average is 6.1 differences, or 2.2% (6.1/282) divergence, for the seven pairs of Alu repeats. Excluding the one pair of Alu repeats with 15 differences, there is an average of 4.7 differences, or 1.7% (4.7/282) divergence. These values agree with the results of a number of studies showing that the average divergence of nonselected human and chimpanzee DNAs is about 1.5–2% (Zimmer et al. 1980; Chang and Slightom 1984; Sibley and Ahlquist 1984; Willard et al. 1985). We conclude that the base sequence of the Alu family is evolving at a rate characteristic of unselected DNA.

Assuming an average of 6.1 differences between

chimpanzee and human Alu's, the distribution of mutations among the seven pairs of Alu repeats is consistent with a random binomial distribution (Table 1). For example, the probability of observing three or fewer differences in the sequences of a pair of Alu repeats is 0.13, whereas the probability of observing 15 or more mutations is 0.025 (Table 1). According to this analysis the pair of Alu's with only three mutations is not especially well conserved compared with a pair with a random distribution of mutations. The observation of a pair, α2:Alu 1, having 15 mutations is also not statistically unusual in a sample of seven pairs. The position of this Alu family member happens to correspond to the end point of a gradient of sequence divergence in the human α globin duplication units (Hess et al. 1984). As discussed above, these differences might result in part from recombinations between α2:Alu 1 and the corresponding Alu located 5′ to the α1 gene.

The partial Alu family member located 5′ to the human and chimpanzee ξ1 genes closely matches the human consensus in a 40-nt region (Fig. 3). The 40-nt homology is flanked by a 47-bp region of nonhomology, but then homology resumes again for an additional 14-bp (Fig. 3). One possibility is that this partial Alu is a vestige of an ancient Alu repeat that is diverging into an unrecognizable sequence. We discount this possibility, because the 40-nt-long partial Alu matches the consensus as well as any full-length member of the family does. However, the sequences adjacent to this 40-nt region are not distantly related to Alu, but rather appear to be an entirely different sequence that is unrelated in origin to the Alu family. This abrupt transition from good consensus homology to totally nonhomologous sequence suggests that the partial Alu is the result of one or more recombination events rather than the product of long-term decomposition. Replacement of Alu family members requires their removal as well as their insertion. This partial Alu and other deleted Alu's (Orkin and Michelson 1980; Jagadeeswaran et al. 1982; Ottolenghi and Giglioni 1982) may be examples of Alu's that are removed from the genome by simple recombination.

Discussion

Both conversion and turnover could homogenize a repetitive sequence family. Their relative frequencies might decide which is the more important process. With the possible exception of a 55-nt region in one Alu family member, α2:Alu 1, conversion has not operated on any of the seven full-length Alu pairs reported here since the divergence between human and chimpanzee. As discussed above, α2:Alu 1 may be exceptional compared with other family members also in that it is an end point of an α globin conversion unit (Hess et al. 1983, 1984). However, for the purpose of discussion we shall assume that this change was a conversion and does reflect the conversion rate of typical Alu repeats. This implies that Alu repeats have exchanged 1.4% (55 nt/14 × 282 nt) of their sequences by conversion since human–chimpanzee divergence. This value, 1.4%, which is probably an upper limit for the true effect of conversion, should be compared with the effect of turnover on the membership of the Alu family.

There are no examples of Alu members that have turned over since human–chimpanzee divergence. However, indirect evidence suggests that the rate of turnover is faster than the rate of conversion. The short direct repeats that flank Alu family members are often inexact copies. Fukumaki et al. (1983) reasoned that mutations accumulating within the short direct repeats are a measure of the age of Alu family members. They estimated that the average divergence between the left and right short direct repeats is 5%. Unfortunately, the definition of the short direct repeats is somewhat subjective. Using a looser definition of the short direct repeats and a larger data base, Schmid and Shen (1985) estimated the average divergence to be as high as 12%. For our present purposes we employ an average of the two values, 8 ± 4%.

As discussed above, nonselected human and chimpanzee DNAs, including Alu family members, differ by 1.5–2% in sequence. Assuming that nonselected sequences diverge at the same rate as the direct repeats flanking Alu, the time since divergence between human and chimpanzee represents a significant fraction (at least 1.5%/8% = 19%) of the average age of Alu repeats in the human and chimpanzee genome. Assuming that at steady state an average age is approximately one-half of the average lifetime, we estimate that ~9% (19%/2) of the Alu family members would have been replaced since the divergence between human and chimpanzee. This value should be compared with the upper limit given above for the conversion of Alu repeats since humans and chimpanzees diverged, 1.4%. We conclude that replacement is more likely to affect the composition of the membership of the Alu family than conversion.

This conclusion assumes that both conversion and replacement are continuing processes and that comparison of their average rates reflects their relative importance. Either might change the composition of the entire family by a "big bang," in which case the present rate comparison would be irrelevant. Conversion might also be very important in determining the sequences of a select group of founder sequences; our finding is that conversion is less important than replacement in determining the

composition of the membership of the family at large.

Assuming that replacement is responsible for homogenizing the Alu family, is the rate of replacement sufficiently fast to account for the homogenization? The ancient divergence between human and galago marks the last recognizable homogenization of Alu repeats within the primary genome (Daniels et al. 1983). Galago and human DNAs differ by about 30% in base sequence (Deininger and Schmid 1979), a value significantly greater than the $8 \pm 4\%$ divergence of the flanking direct repeats. The time since the divergence between human and galago is sufficient for at least one complete turnover of the Alu repeats.

In summary, these observations suggest the simplest possible mechanism for the homogenization of the Alu family: New members are inserted as the products of one or more founder sequences, which may be under biological selection (Jagadeeswaran et al. 1981; Van Ardsell et al. 1981). Members of the family randomly acquire mutations at about the same rate as nonselected DNA sequences do, and after an uneventful existence they recombine out of the genome. At any given time the population of Alu repeats in a given species is in the process of being driven toward a homogenization end point defined by the then-current collection of founder sequences.

Acknowledgment. This investigation was supported by USPHS grants GM 21346 to C.W.S. and AM 29800 to C.-K.J.S. and NSF grant BSR 84-15867 to A.C.W.

References

Arnheim N, Krystal M, Schmickel R, Wilson G, Ryder O, Zimmer E (1980) Proc Natl Acad Sci USA 77:7323–7327

Britten RJ, Kohne DE (1968) Science 161:529–540

Brown SDM, Dover G (1981) J Mol Biol 150:441–446

Chang L-YE, Slightom JL (1984) J Mol Biol 180:767–784

Daniels GR, Deininger PL (1983) Nucleic Acids Res 11:7595–7610

Daniels GR, Fox GM, Lowensteiner DL, Schmid CW, Deininger PL (1983) Nucleic Acids Res 11:7579–7593

Deininger PL, Schmid CW (1979) J Mol Biol 127:437–460

Deininger PL, Jolly DJ, Rubin CM, Friedman T, Schmid CW (1981) J Mol Biol 151:17–33

Dover GA, Flavell RB (1984) Cell 38:622–623

Fukumaki Y, Collins F, Kole R, Stoeckert CJ, Jagadeeswaran P, Duncan CH, Weissman SM (1983) Cold Spring Harbor Symp Quant Biol 47:1079–1081

Hess JF, Fox GM, Schmid CW, Shen C-K J (1983) Proc Natl Acad Sci USA 80:5970–5974

Hess JF, Schmid CW, Shen C-KJ (1984) Science 226:67–69

Jagadeeswaran PMG, Forget BG, Weissman SM (1981) Cell 26:141–142

Jagadeeswaran P, Tuan D, Forget BG, Weissman SM (1982) Nature 296:469–471

Jeffreys AJ, Wilson V, Thein SL (1985) Nature 314:67–72

Jelinek WR, Schmid CW (1982) Annu Rev Biochem 51:813–844

Lauer J, Shen C-KJ, Maniatis T (1980) Cell 20:119–130

Maeda N, Bliska JB, Smithies O (1983) Proc Natl Acad Sci USA 80:5012–5016

Martin SL, Voliva CF, Burton FH, Edgell MH, Hutchison CA III (1984) Proc Natl Acad Sci USA 81:2308–2312

Martin SL, Voliva CF, Hardies SC, Edgell MH, Hutchison CA III (1985) Mol Biol Evol 2:127–146

Miesfeld R, Krystal M, Arnheim N (1981) Nucleic Acids Res 9:5931–5947

Orkin SH, Michelson A (1980) Nature 286:538–540

Ottolenghi S, Giglioni B (1982) Nature 300:770–771

Proudfoot NJ, Gil A, Maniatis T (1982) Cell 31:553–563

Rice NR (1972) In: Smith HH (ed) Evolution of genetic systems. Gordon & Breach, New York, p 44

Sawada I, Beal MP, Shen C-KJ, Chapman B, Wilson AC (1983) Nucleic Acids Res 11:8087–8101

Schmid CW, Jelinek WR (1982) Science 216:1065–1070

Schmid CW, Shen C-KJ (1985) The evolution of interspersed repetitive DNA sequences in mammals and other vertebrates. In: MacIntyre RJ (ed) Molecular evolutionary genetics. Plenum, New York, pp 323–358

Shen C-KJ, Maniatis T (1982) J Mol Appl Gen 1:343–360

Sibley CG, Ahlquist JE (1984) J Mol Evol 20:2–15

Singer MF, Thayer RE, Grimaldi G, Lerman MI, Fanning TG (1983) Nucleic Acids Res 11:5739–5745

Ullu E, Marphy S, Melli M (1982) Cell 29:195–202

Van Ardsell SW, Denison RA, Bernstein LB, Weiner AM (1981) Cell 26:11–17

Willard C, Wong E, Hess JF, Shen C-KJ, Chapman B, Wilson AB, Schmid CW (1985) Comparison of human and chimpanzee ζ1 globin genes. J Mol Evol XX:XXX–XXX

Zimmer EA, Martin SL, Beverley SM, Kan YW, Wilson AC (1980) Proc Natl Acad Sci USA 77:2158–2162

Received July 9, 1985/Revised and accepted September 18, 1985