Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 9360–9364, December 1989
Evolution

Newly arisen DNA repeats in primate phylogeny

(α-fetoprotein gene/DNA transposition/evolutionary novelties/human–gorilla divergence)

SUSAN C. RYAN AND ACHILLES DUGAICZYK

Department of Biochemistry, University of California, Riverside, CA 92521

Communicated by Frank W. Putnam, September 5, 1989 (received for review July 3, 1989)

ABSTRACT **We discovered the presence of an *Alu* and an *Xba* repetitive DNA element within introns 4 and 7, respectively, of the human α-fetoprotein (AFP) gene; these elements are absent from the same gene in the gorilla. The *Alu* element is flanked by 12-base-pair direct repeats, AGGATGTTGTGG . . . (*Alu*) . . . AGGATGTTGTGG, which presumably arose by way of duplication of the intronic target site AGGATGTTGTGG at the time of the *Alu* insertion. In the gorilla, only a single copy of the unoccupied target site is present, which is identical to the terminal repeat flanking the human *Alu* element. There are two copies of an *Xba* repeat in the human AFP gene, apparently the only two in the genome. *Xba1* and *Xba2*, located within introns 8 and 7, respectively, differ from each other at 3 of 303 positions. *Xba1* is referred to as the old (ancestral) repeat because it lacks direct repeats. The new (derived) *Xba2* is flanked by direct repeats, TTTCTTTTT . . . (*Xba*) . . . TTTCTTCTT, and is thought to have arisen as a result of transposition of *Xba1*. The ancestral *Xba1* and a single copy of the *Xba2* target site are present at orthologous positions in the gorilla, but the new *Xba2* is absent. We conclude that the *Alu* and *Xba* DNA repeats emerged in the human genome at a time postdating the human–gorilla divergence and became established as genetic novelties in the human lineage. We submit that the chronology of divergence of primate lines of evolution can be correlated with the timing of insertion of new DNA repeats into the genomes of those primates.**

We have previously reported that the albumin gene family has been invaded by numerous elements of repetitive DNA (1). These elements have been inserted at random sites within the gene family and, by inference, at different times during their evolutionary history. It would be of interest to find out if a correlation can be established between the emergence of new DNA repeats and the divergence of species, particularly in the time frame of primate evolution.

The most prominent among the various DNA repeats are members of the *Alu* family. These repeats are considered to be pseudogenes that have arisen by retroposition of 7SL RNA (2, 3) or 4.5S RNA (4), and they are presently found at an estimated 500,000 copies per genome in the primates and other vertebrates (5, 6). Although there is some evidence of interspecies differences in the number of *Alu* sequences among closely related primates (7), studies in the α-, and β-globin gene regions revealed no differences in the location of specific *Alu* repeats between humans and two other primates. In the α-globin locus, seven members of the *Alu* family were found in identical positions in human and chimpanzee (8). Furthermore, in the β-globin locus, seven additional *Alu* repeats were also found in identical positions in human and orangutan (9). All of the 14 repeats were thus uninformative with regard to phylogenies of the above primates. However, Trabuchet *et al.* (10) reported the presence of a new member of the *Alu* family in the gorilla β,δ-globin region. This same *Alu* repeat is absent in human, chimpanzee, and macaque, thus suggesting a recent insertion of this element in the gorilla DNA. It is not clear from the report (10) whether this *Alu* repeat is fixed or polymorphic in the gorilla lineage. If more examples are found of species-specific differences in repetitive DNA elements at orthologous sites among higher primates, this would provide the best evidence for the chronology of their insertion and would help in discerning the phylogeny of the species involved.

In determining the complete structure of the human α-fetoprotein (AFP) gene (11), we identified the positions of two *Xba* repetitive elements. We concluded from this work that *Xba1* and *Xba2* were entirely novel repetitive elements. Due to the high degree of sequence similarity between *Xba1* and *Xba2* (99%), we further concluded that their existence in the human AFP gene was of relatively recent evolutionary origin (1, 11). Presently, we have extended our work on the AFP gene to other primates and wish to report a specific difference in an *Xba* and *Alu* repetitive element between the human and gorilla lineages. In addition, we submit that such DNA elements fulfill the criterion for molecular markers of phylogeny.

MATERIALS AND METHODS

Cloning. A gorilla genomic library cloned in a Charon 40 phage vector was graciously provided to us by Jerry Slightom (Division of Molecular Biology, The Upjohn Company, Kalamazoo, MI). Using human AFP gene fragments (11) as probes, gorilla phage clones λGAFP9 and λGAFP22 were isolated and characterized by restriction endonuclease digestion. *Eco*RI fragments of the two phage clones were subcloned into the plasmid vector pBR322. DNA sequence was determined by the method of Maxam and Gilbert (12), frequently strand-separating labeled DNA fragments prior to the chemical degradation.

Southern Hybridization Analysis of Human and Gorilla AFP Genes. Cloned human 9.9-kilobase (kb) and gorilla 9.6-kb *Eco*RI DNA fragments were separated electrophoretically in a 1% agarose gel followed by Southern (13) hybridization to an *Alu* probe, BLUR8 (14). Similarly, 10 μg of human and gorilla genomic DNA was digested with restriction endonucleases, separated electrophoretically in a 1% agarose gel, and hybridized to either a human 6.0-kb *Eco*RI–*Hin*dIII probe (11) or a gorilla 5.4-kb *Hin*dIII–*Eco*RI probe.

RESULTS

BLUR8 Hybridization. Preliminary sequencing results indicated that a gorilla 9.6-kb fragment aligned with the same *Eco*RI sites as the human 9.9-kb fragment, even though it was shorter by 300 base pairs (bp). Both fragments contain a unique *Bam*HI site. *Bam*HI digestion of human 9.9-kb DNA yielded 5.5-kb and 4.4-kb fragments, whereas *Bam*HI digestion of gorilla 9.6-kb DNA yielded 5.5-kb and 4.1-kb frag-

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.

Abbreviation: AFP, α-fetoprotein.

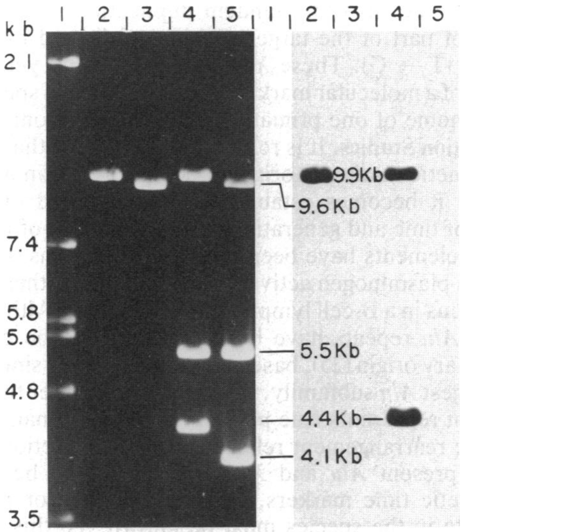

FIG. 1. Hybridization analysis of *Alu* repetitive DNA in the human and gorilla AFP genes. Cloned DNA fragments were separated electrophoretically in a 1% agarose gel followed by Southern (13) hybridization to a BLUR8 (14) probe. (*Left*) Lane 1, DNA size standards; lane 2, human 9.9-kb *Eco*RI DNA fragment; lane 3, gorilla 9.6-kb *Eco*RI DNA fragment; lane 4, partial *Bam*HI digest of human 9.9-kb DNA, giving rise to 5.5-kb and 4.4-kb fragments; lane 5, partial *Bam*HI digest of gorilla 9.6-kb DNA, giving rise to 5.5-kb and 4.1-kb fragments. (*Right*) Autoradiogram of the same gel, showing hybridization to a nick-translated human BLUR8 probe.

ments (Fig. 1). Since the *Alu* repeat within intron 4 of the human AFP gene is located in the 4.4-kb fragment, the 0.3-kb size difference in the small *Bam*HI fragments (4.4 vs. 4.1) between the two species could have been due to an *Alu* repeat. To test this hypothesis, *Bam*HI-digested as well as undigested human 9.9-kb DNA and gorilla 9.6-kb DNA were hybridized to the *Alu* probe BLUR8 (14) (Fig. 1). As expected, the 9.9-kb *Eco*RI fragment and the 4.4-kb *Bam*HI–*Eco*RI fragment from the human AFP gene hybridized to the BLUR8 probe, thus confirming the presence of *Alu* repeats.

The BLUR8 probe did not hybridize to either the 9.6-kb *Eco*RI or the 4.1-kb *Bam*HI–*Eco*RI fragment from the gorilla AFP gene, indicating the absence of *Alu* repeats (Fig. 1).

Sequence Analysis in the *Alu* Region. The sequence of a 5.4-kb *Hin*dIII–*Eco*RI fragment from the gorilla AFP gene was determined, which overlaps the expected site of insertion of the missing *Alu* repeat. The DNA of this region was found to contain a single copy of a 12-bp target site, AGGATGTTGTGG, which is identical to the direct repeats flanking the *Alu* residing within intron 4 of the human AFP gene (Fig. 2). This 12-bp repeat is thought to have arisen during the event of the *Alu* insertion into the human genome and hence the human gene has two copies of this sequence. Full-size *Alu* elements are known to be flanked by direct repeats that are believed to have arisen by means of repair of staggered breaks in genomic DNA at the target site of *Alu* insertion (15, 16). There are also numerous examples of *Alu* elements being deleted from the human genome, but these events produce concomitant deletions of adjacent DNA and are recognized as genetic defects (17–22) when coding DNA is deleted. The *Alu* element in the human AFP gene is flanked by perfect repeats of 12 bp, AGGATGTTGTGG . . . (*Alu*) . . . AGGATGTTGTGG, whereas there is no evidence for DNA deletion in the entire intron 4 in the gorilla AFP gene, as compared to the same intron in the human genome. We submit that this *Alu* element was inserted into the human genome at a time postdating the divergence of the human and gorilla lineages. It is interesting to note that in the neighboring intron 3 of these AFP genes, a 226-bp *Kpn* repeat is found in the orthologous position in the human and the gorilla (Fig. 2). Such an identity suggests that the *Kpn* repeat was inserted into the genome before the divergence of these two species. The gorilla *Kpn* repeat differs in 5/226 nucleotide positions (2.2%) from the human *Kpn* repeat.

Sequence Analysis in the *Xba* Region. We reported earlier (11) that the two *Xba* elements differ in only 3 of 303 positions (1%) and that only *Xba*2 is flanked by direct repeats, TTTCTTTTT . . . (*Xba*2) . . . TTTCTTCTT. This suggests that *Xba*1, located in intron 8, was duplicated and reinserted into intron 7 of the human AFP gene, giving rise to *Xba*2 (Fig. 3). To our knowledge, an ancestral source DNA has not been

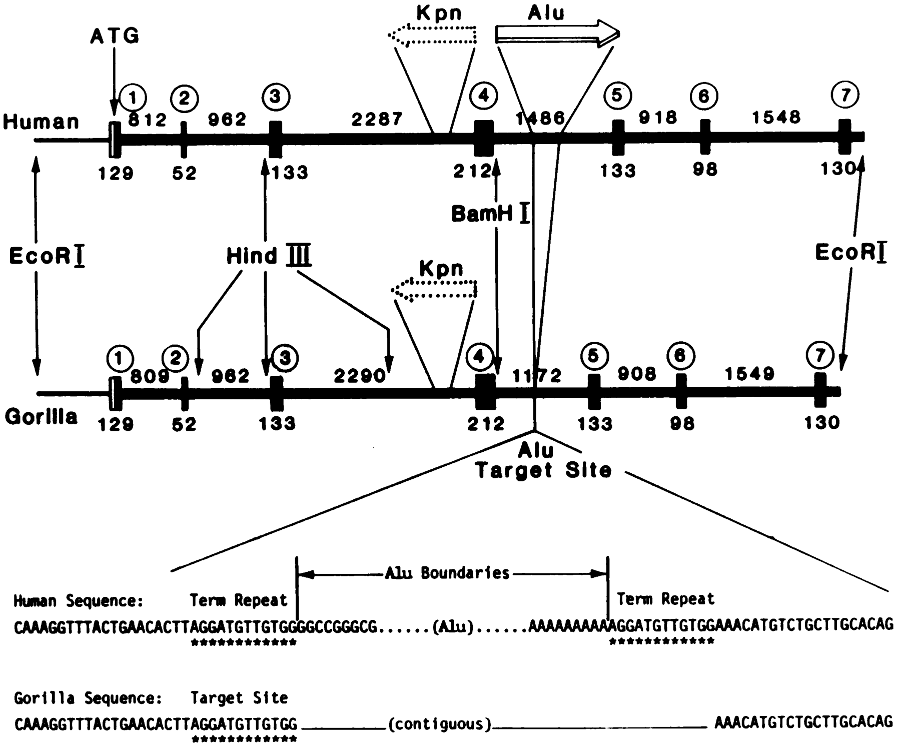

FIG. 2. Comparison of restriction maps for the human and gorilla AFP genes (exons 1–7). The map of the gorilla gene shows the absence of the *Alu* repeat. The human AFP map is taken from Gibbs *et al.* (11). Exons are shown as boxes and introns are shown as heavy lines. The sizes (bp) of exons and introns are also indicated. The gorilla *Kpn* repeat is of identical size and orientation as the human *Kpn* repeat. Also shown is the *Alu* target site in the gorilla gene. In the lower half, the DNA sequence surrounding the *Alu* element is shown for both genes. The terminal repeats (*) flanking the *Alu* element in one species (human) can be recognized as the unoccupied target site (*) in the other species (gorilla); a 12-bp target site in the gorilla is identical to the terminal repeats surrounding the human *Alu* sequence.

identified together with its derived copy previously. Further evidence was provided by a search of GenBank (March 1989) and genomic Southern blot analysis. These results suggest that *Xba*1 and *Xba*2 are the only two members of this family in the human genome. The direct repeats surrounding *Xba*2 are thought to have arisen during the event of *Xba*2 insertion into its new position in the genome. Since then, the 5′ direct repeat has undergone a single base mutation so that presently the two repeats are identical at 8 of 9 positions.

In the human AFP gene, both *Xba* repeats can be isolated on a 5.5-kb *Eco*RI–*Hin*dIII fragment spanning from an *Eco*RI site 5′ of exon 8 to a *Hin*dIII site 3′ of exon 11. The corresponding *Eco*RI–*Hin*dIII fragment in the gorilla AFP gene is only 5.2 kb in size (Fig. 3). This 300-bp size difference could be explained by the absence of a single *Xba* repeat in the gorilla AFP gene. Upon sequencing this 5.2-kb DNA, we discovered that the gorilla AFP gene lacked the *Xba*2 repeat but contained a single copy of the unoccupied target site TTTCTTCTT, which is identical to the 3′ direct repeat flanking *Xba*2 in the human AFP gene (Fig. 3). Sequence analysis has further shown an additional seven nucleotides, TGTCTTC, located 5′ to the unoccupied target site in the gorilla AFP gene that are not present in the human AFP gene. A possible explanation is that the 7-bp sequence was originally present in the human AFP gene but was deleted during the insertion of *Xba*2 into the new site. An alternative is that the TGTCTTC arose by tandem duplication (in the gorilla lineage) of part of the target sequence followed by a point mutation (T → G). These *Xba* repeats provide yet another example of a molecular marker that is present at a specific site in the genome of one primate line but absent from another.

Population Studies. It is reasonable to assume that a newly arisen genetic marker is originally polymorphic in a population, and it becomes established or eliminated only after passage of time and generations. Two examples of polymorphic *Alu* elements have been reported in humans. One is in the tissue plasminogen activator gene (23); the other is in the Mlvi-2 locus in a B-cell lymphoma cell line (24). Although the last two *Alu* repeats have been concluded to be of recent evolutionary origin (25), based on their sequence similarity to the youngest *Alu* subfamily, it is not clear whether the Mlvi-2 *Alu* repeat represents true polymorphism in human DNA or a somatic rearrangement related to tumor induction.

If the present *Alu* and *Xba* repeats are to be used as phylogenetic time markers, their established (or polymorphic) state in the species must be known. To this end, we analyzed restriction digests of genomic DNA from individual humans and gorillas. The presence or absence of repetitive DNA would be reflected in the sizes of such restriction fragments. *Eco*RV digests of genomic DNA from 25 unrelated Caucasians yielded a 3.6-kb fragment containing the *Alu* repeat located within intron 4 of the AFP gene; a 3.3-kb

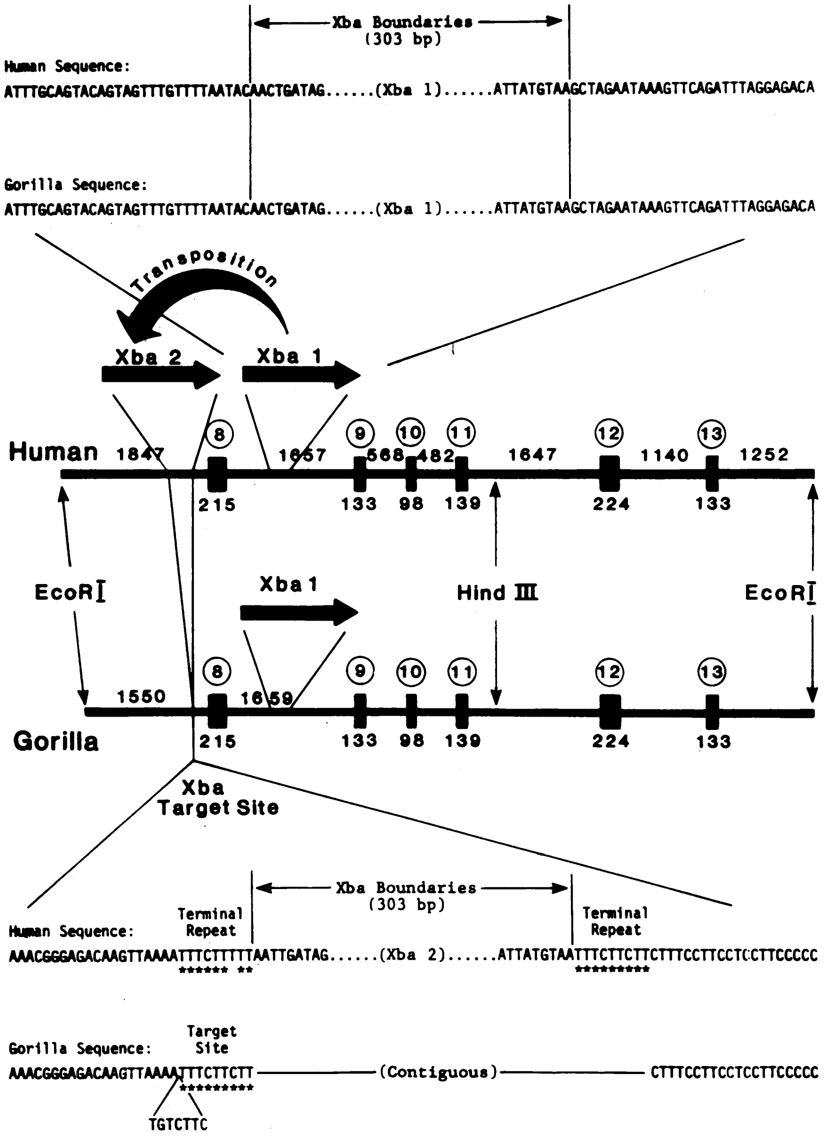
Xba Boundaries
(303 bp)
Human Sequence:
ATTTGCAGTACAGTAGTTTGTTTTAATACAACTGATAG......(Xba 1)......ATTATGTAAGCTAGAATAAAGTTCAGATTTAGGAGACA
Gorilla Sequence:
ATTTGCAGTACAGTAGTTTGTTTTAATACAACTGATAG......(Xba 1)......ATTATGTAAGCTAGAATAAAGTTCAGATTTAGGAGACA
Transposition
Xba 2
Xba 1
Human
1847
8
1657
9
10
11
568
482
1647
12
1140
13
1252
215
133
98
139
224
133
EcoRI
Xba 1
Hind III
EcoRI
1550
8
1659
9
10
11
12
13
Gorilla
215
133
98
139
224
133
Xba
Target Site
Terminal Repeat
Xba Boundaries
(303 bp)
Terminal Repeat
Human Sequence:
AAACGGGAGACAAGTTAAAATTTCTTTTTAATTGATAG......(Xba 2)......ATTATGTAATTTCTTCTTCTTTCCTTCCTCCTTCCCCC
****** **
****
Target Site
Gorilla Sequence:
AAACGGGAGACAAGTTAAAATTTCTTCTT (Contiguous) CTTTCCTTCCTCCTTCCCCC
****
TGTCTTC

FIG. 3. Comparison of restriction maps for the human and gorilla AFP genes (exons 8–13). The map of the gorilla gene shows the absence of *Xba*2. The human AFP map is taken from Gibbs *et al.* (11). Exons are shown as boxes and introns are shown as heavy lines. The sizes (bp) of exons and introns are also indicated. The upper half shows the DNA sequence surrounding *Xba*1 in both genes. The lower half shows the DNA sequence surrounding *Xba*2 in the human lineage, including the flanking terminal repeats (*), which can be recognized as the unoccupied target site in the gorilla lineage. The transposition of *Xba*1 to a new location (*Xba*2) is indicated. The numbering of the human *Xba* repeats is reversed relative to our original publication (11).

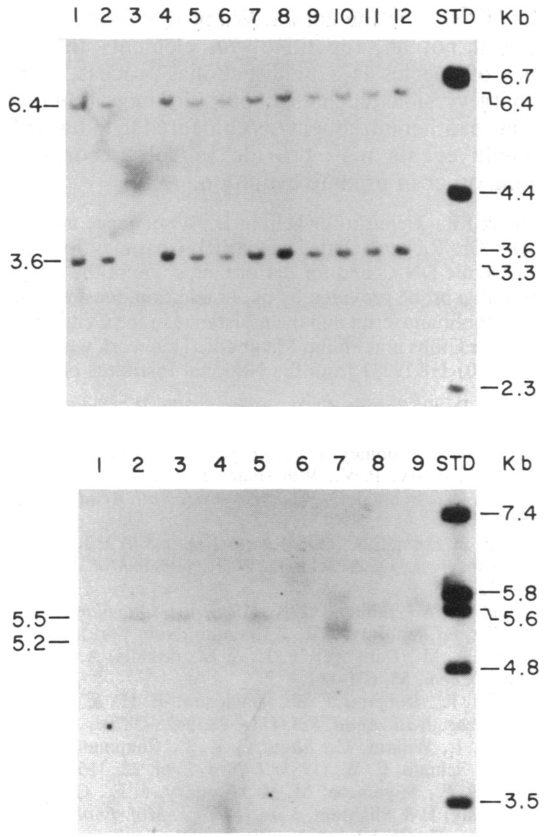

FIG. 4. Southern hybridization of restriction digests of genomic DNA from unrelated individuals. (*Upper*) Autoradiogram of *Eco*RV digests of human genomic DNA hybridized to a nick-translated gorilla (*Hin*dIII–*Eco*RI) 5.4-kb probe. All 12 individuals show hybridization in the 3.6-kb rather than the 3.3-kb region, indicative of the presence of the *Alu* repeat. The probe also recognizes an adjacent region on the AFP map, giving rise to the 6.4-kb band seen in all individuals. (*Lower*) Autoradiogram of *Eco*RI plus *Hin*dIII digests of genomic DNA hybridized to a nick-translated human 6.0-kb probe. Lanes 1–6, 10 μg of digested human genomic DNA; all six individuals show hybridization in the 5.5-kb region, indicative of two *Xba* repeats. Lanes 7–9, 10 μg of digested gorilla genomic DNA; the three individuals show hybridization in the 5.2-kb region, indicative of one *Xba* repeat. The upper bands in lanes 6–9 are due to incomplete *Hin*dIII digests. STD, DNA size standards.

fragment would be expected if the *Alu* repeat was absent. We used the gorilla 5.4-kb *Hin*dIII–*Eco*RI DNA fragment, which encompasses the unoccupied *Alu* target site as the nick-translated probe. The resulting autoradiogram (Fig. 4) shows 12 of the 25 individuals analyzed in this experiment. With regard to the *Xba* repetitive elements, we performed a similar experiment using *Eco*RI plus *Hin*dIII genomic digests of both gorilla and human DNA. This double digest should yield a 5.5-kb fragment if both *Xba*1 and *Xba*2 repeats are present or a 5.2-kb fragment if only the ancestral sequence *Xba*1 is present. A 6.0-kb human fragment encompassing both repeats was used as the nick-translated probe. As can be seen from the resulting autoradiogram (Fig. 4), all 6 Caucasians analyzed possess both *Xba* repeats in their AFP genes. Also, all four gorillas analyzed (including the gorilla from which sequence data were obtained) possess only *Xba*1 in their AFP genes. From these results, it appears that the *Alu* repeat and the two *Xba* repeats are established (fixed) markers in the human lineage and so is *Xba*1 in gorilla. A larger sampling may be required to rigorously prove that they are truly panspecific characters.

DISCUSSION

Time of Insertion of Repeats. Our conclusion about a recent origin of the *Alu* repeat in the human AFP gene draws support from work by Willard *et al.* (26), Britten *et al.* (27), and Jurka and Smith (28). These authors recognized that *Alu* repeats can be subdivided into subfamilies, each being inserted into the host genome at different times during evolution. Based on diagnostic substitutions that are shared among *Alu* subfamilies, Britten *et al.* (27) published a consensus sequence of what they consider the youngest (class IV) subfamily of *Alu* repeats. The *Alu* repeat within the human AFP gene differs from this class IV consensus by only 4 of 283 positions (1.4%), excluding 5 of 283 mutations in noninformative CpG hot spots (29), and it has retained 12 of the 13 mutations considered diagnostic for this youngest *Alu* subfamily (Fig. 5) From this high degree of sequence similarity between *Alu* class IV consensus and the present *Alu*-1, it can be inferred that both belong to the same subfamily, whereas the conclusion about their young age has been reached from two independent considerations.

The *Xba*1 and *Xba*2 repeats in the human AFP gene appear to be the only copies found in the human genome (Fig. 4). Since *Xba*1 lacks direct repeats, it is reasonable to assume that *Xba*1 was duplicated and reinserted into the human genome, resulting in *Xba*2. To our knowledge, both the source (ancestor) DNA and the derived repeat have not been identified previously. The target site for *Xba*2 is at a sharp boundary between a purine- and a pyrimidine-rich region, but we do not know what triggered the movement of the *Xba* element or whether it was an RNA-mediated or a DNA-mediated transposition. The repeats represent a single duplication event of unknown and possibly very novel nature. If it was not for the presence of *Xba*2, *Xba*1 would be difficult to discern from nonmobile parts of chromosomal DNA. Unlike the random retroposition of *Alu* sequences, amplification of the *Xba* element is restricted to the same gene as the parental sequence, although this might be the first event in the rise of a novel repeat element. The human *Xba* elements differ in 3 of 303 positions, whereas human *Xba*1 differs from gorilla *Xba*1 in 2 of 303 positions, indicating a recent separation of the three sequences. The degree of homology (99%) between the two human *Xba* elements suggests that *Xba*1 gave rise to *Xba*2 perhaps less than 6 million years ago, and postdating the time of human–gorilla divergence. It is, *a priori*, possible that *Xba*2 also existed in the gorilla lineage and was subsequently deleted, but the young age of *Xba*2 argues against this possibility.

```
Alu-1             20                  40                  60                  80                 100
GGCCGGGCGCGGTGGCTCACGCTTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACGAGGTCAGGAGATCGAGACCATCCTGGCTAACATGG
:::::::::::::::::::::: :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: ::
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACGAGGTCAGGAGATCGAGACCATCCTGGCTAACACGG
Con-IV                                                                 *   *         *      *    *
                 120                 140                 160                 180                 200
TGAAACCCCGTCTCTACTAAAAATACAAAAAAAATTAGCCGGGCGTGATGGTGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGC
::::::::::::::::::::::::::::::::  ::::::::::::: ::: ::::::::::::::::::::::::::::::::::::::::::::::::
TGAAACCCCGTCTCTACTAAAAATACAAAAAA--TTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGC
                               *                     **        *                                 **
                 220                 240                 260                 280
GTGAACCCTGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCCACTGCACTCCCGCCTGGGCCACAGAGCGAGACTCCGTCTC
:::::::: :::::::::::::::::::::::::::: ::::::::::::::: :::::::: ::::::::::::::::::::
GTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTC
*                  *
```

FIG. 5. Comparison of the *Alu* DNA sequence from intron 4 of the human AFP gene (Alu-1) with the youngest *Alu* subfamily consensus sequence (Con-IV) as described by Britten *et al.* (27). Asterisks (*) show positions of the 13 diagnostic mutations used to distinguish a class IV repeat.

However, it is quite possible that the *Xba* duplication took place prior to the human–chimpanzee separation, rather than being unique to the human lineage.

Reconstructing Phylogenies from Differences in DNA. There is a need for identifying nonrecurrent, irreversible events in the genomes of species that would serve as informative markers of evolution. After years of effort in reconstructing the phylogeny of higher primates from DNA mutations, a number of conflicting phylogenies have been published, leading to an impasse (for review, see refs. 30 and 31). Why can't we decide? is the question of the time (32). The answer appears to be: Because of the mutations' reversibility. Given enough time, consecutive mutations occur at the same site in one species, and parallel mutations occur independently at the same (orthologous) site in two species. This multiplicity of changes makes it difficult to decide whether in a contemporary DNA sequence a given nucleotide represents the ancestral or the derived state of the character. Phylogenetic trees reconstructed from such reversible events are hence statistical trees, and so the uncertainty remains as to whether statistics can reveal the true phylogeny of species.

Mobility of Repetitive DNA Elements. There are transposable elements—e.g., THE-1 (33)—that transfer themselves from one chromosomal site to another. This is accomplished in a controlled mode of movement, in that an enzyme (transposase), encoded by the element, and the element's long (350 bp) terminal repeats participate in the reversible integration/excision mechanism. *Alu* elements belong to a different category. They are pseudogenes that appear to have overrun the genome by means of retroposition from a conserved source gene (27). Over the last 60 million years or so, an impressive number of 0.5 million copies have accumulated in the genome of the human lineage. Multiple examples of deletions due to unequal crossover at two *Alu* repeats have been reported, but we have yet to find a deletion of a single *Alu* element. Although an excision of a circularized *Alu* sequence by means of recombination at its terminal repeats is plausible, it is probably a rare event compared to the vast number of interspersed *Alu* elements in chromosomal DNA. The terminal repeats are short (≈5–15 bp) and diverge in time and apparently are not very efficient recombination sites. The infrequency or lack of single *Alu* deletions can be also inferred from studies in the globin gene locus of numerous primates (34): when a specific *Alu* repeat could be identified as an ancient insertion in an ancestor of apes and humans, it was found to be present at the same chromosomal site in contemporary primates, such as human, chimpanzee, gorilla, and orangutan. Thus, the globin *Alu* element resides at the same position for some 40 million years in at least four lines of evolution. Based on the above arguments, and until evidence to the contrary is found, we submit that inserted *Alu* repeats become rather frozen in the recipient genome. They may be subsequently deleted in unequal crossover events, but it appears they are not excised in a reversible manner.

Another intriguing feature of the *Alu* spread is an apparent lack of control in the process. Viral infections are under the surveillance of the immune system; other transposable elements appear to be under a feedback control that limits their number. The spread of *Alu* repeats (and probably other pseudogenes) seems to be driven by the propensity of DNA to interact with itself. Indications exist (26) that the 0.5 million copies were transposed not at a constant rate in time but rather in sporadic transposition bursts. As we pointed out previously (1), such bursts could have been cataclysmic in that the burden of their mutagenic load could become a way of death in the extinction of species, or perhaps a way of branching off new lines of evolution.

Spread of Repeats as Time Markers in Evolution. In the present work we have demonstrated that two repeated DNA elements are present in the human genome but absent from orthologous sites in the gorilla. We postulate that spreading of most, if not all, repetitive *Alu* elements through the genomes of species is a unidirectional process, stemming from an irreversible mechanism of their integration into new sites. These rather infrequent (as compared to mutations) and irreversible events may provide reliable records of the branching order in primate evolution.

We thank Rita Zielinski for help in DNA sequence analysis. We also thank Dr. Richard Gatti and Patrick Charmley for providing human genomic DNAs and for performing some of the *Alu* hybridizations with a probe provided by us. In addition, we appreciate the critique of this manuscript and the assistance in the GenBank search by Drs. Peter Gibbs and Phillip Minghetti. This work was supported by Grant 1R01 HD19281 from the National Institutes of Health.

1. Ruffner, D. E., Sprung, C. N., Minghetti, P. P., Gibbs, P. E. M. & Dugaiczyk, A. (1987) *Mol. Biol. Evol.* **4,** 1–9.
2. Jelinek, W. R., Toomey, T. P., Leinwand, L., Duncan, C. H., Biro, P. A., Choudary, P. V., Weissman, S. M., Rubin, C. M., Deininger, P. L. & Schmid, C. W. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 1398–1402.
3. Ullu, E. & Tschudi, C. (1984) *Nature (London)* **312,** 171–172.
4. Schoeninger, L. O. & Jelinek, W. R. (1986) *Mol. Cell. Biol.* **6,** 1508–1519.
5. Schmid, C. W. & Shen, C.-K. (1986) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 323–358.
6. Weiner, A. M., Deininger, P. L. & Efstratiadis, A. (1986) *Annu. Rev. Biochem.* **55,** 631–661.
7. Hwu, H. R., Roberts, J. W., Davidson, E. H. & Britten, R. J. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 3875–3879.
8. Sawada, I., Willard, C., Shen, C.-K. J., Chapman, B., Wilson, A. C. & Schmid, C. W. (1985) *J. Mol. Evol.* **22,** 316–322.
9. Koop, B. F., Miyamoto, M. M., Embury, J. E., Goodman, M., Czelusniak, J. & Slightom, J. L. (1986) *J. Mol. Evol.* **24,** 94–102.
10. Trabuchet, G., Chebloune, Y., Savatier, P., Lachuer, J., Faure, C., Verdier, G. & Nigon, V. M. (1987) *J. Mol. Evol.* **25,** 288–291.
11. Gibbs, P. E. M., Zielinski, R., Boyd, C. & Dugaiczyk, A. (1987) *Biochemistry* **26,** 1332–1343.
12. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
13. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
14. Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedmann, T. & Schmid, C. W. (1981) *J. Mol. Biol.* **151,** 17–33.
15. Van Ardsell, S. W., Denison, R. A., Bernstein, L. B. & Weiner, A. M. (1981) *Cell* **26,** 11–17.
16. Jagadeeswaran, P., Forget, B. G. & Weissman, S. M. (1981) *Cell* **26,** 141–142.
17. Jagadeeswaran, P., Tuan, D., Forget, B. G. & Weissman, S. M. (1982) *Nature (London)* **296,** 469–470.
18. Lehrman, M. A., Russel, D. W., Goldstein, J. L. & Brown, M. S. (1987) *J. Biol. Chem.* **262,** 3354–3361.
19. Myerowitz, R. & Hogikayan, N. D. (1987) *J. Biol. Chem.* **262,** 15396–19399.
20. Markert, M. L., Hutton, J. J., Wiginton, D. A., States, J. C. & Kaufman, R. E. (1988) *J. Clin. Invest.* **81,** 1323–1327.
21. Rouyer, F., Simmler, M.-C., Page, D. C. & Weissenbach, J. (1987) *Cell* **51,** 417–425.
22. Huang, L.-S., Ripps, M. E., Korman, S. H., Deckelbaum, R. J. & Breslow, J. L. (1989) *J. Biol. Chem.* **264,** 11394–11400.
23. Friezner Degen, S. J., Rajput, B. & Reich, E. J. (1986) *J. Biol. Chem.* **261,** 6972–6985.
24. Economou-Pachnis, A. & Tsichlis, P. N. (1985) *Nucleic Acids Res.* **13,** 8379–8387.
25. Deininger, P. L. & Slagel, V. K. (1988) *Mol. Cell. Biol.* **8,** 4566–4569.
26. Willard, C., Nguyen, H. T. & Schmid, C. W. (1987) *J. Mol. Evol.* **26,** 180–186.
27. Britten, R. J., Baron, W. F., Stout, D. B. & Davidson, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4770–4774.
28. Jurka, J. & Smith, T. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 475–478.
29. Bains, W. (1986) *J. Mol. Evol.* **23,** 189–199.
30. Lewin, R. (1988) *Science* **241,** 1598–1600.
31. Lewin, R. (1988) *Science* **241,** 1756–1759.
32. Holmquist, R., Miyamoto, M. M. & Goodman, M. (1988) *Mol. Biol. Evol.* **5,** 201–216.
33. Paulson, K. E., Deka, N., Schmid, C. W., Misra, R., Schindler, W., Rush, M. G., Kadyk, L. & Leinwand, L. (1985) *Nature (London)* **316,** 359–361.
34. Fitch, D. H. A., Mainone, C., Slightom, J. L. & Goodman, M. (1988) *Genomics* **3,** 237–255.