Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 1872–1876, March 1993
Evolution

The emergence of new DNA repeats and the divergence of primates

(genetic novelties/pseudogenes/irreversible events/evolutionary rates/phylogenies)

PHILLIP P. MINGHETTI*† AND ACHILLES DUGAICZYK*‡

*Department of Biochemistry, University of California, Riverside, CA 92521; and ‡Department of Physiology–Human Genome Mapping Center, University of California School of Medicine, San Francisco, CA 94143

Communicated by Frank W. Putnam, November 30, 1992 (received for review September 29, 1992)

ABSTRACT **We have identified four genetic novelties that are fixed in specific primate lineages and hence can serve as phylogenetic time markers. One *Alu* DNA repeat is present in the human lineage but is absent from the great apes. Another *Alu* DNA repeat is present in the gorilla lineage but is absent from the human, chimpanzee, and orangutan. A progenitor *Xba*1 element is present in the human, chimpanzee, gorilla, and orangutan, but only in the human lineage did it give rise to a transposed progeny, *Xba*2. The saltatory appearance of *Xba*2 is an example of a one-time event in the evolutionary history of a species. The enolase pseudogene, known to be present as a single copy in the human, was found to be present in four other primates, including the baboon, an Old World monkey. Using the accepted value of 5×10^{-9} nucleotide substitutions per site per year as the evolutionary rate for pseudogenes, we calculated that the enolase pseudogene arose ≈14 million years ago. The calculated age for this pseudogene and its presence in the baboon are incongruent with each other, since Old World monkeys are considered to have diverged from the hominid lineage some 30 million years ago. Thus the rate of evolution in the enolase pseudogene is only about 2.5×10^{-9} substitutions per site per year, or half the rate in other pseudogenes. It is concluded that rates of substitution vary between species, even for similar DNA elements such as pseudogenes. We submit that new DNA repeats arise in the genomes of species in irreversible and punctuated events and hence can be used as molecular time markers to decipher phylogenies.**

If it is true that evolution is punctuated with rapid changes and new species arise in saltatory events rather than by gradual accumulation of point mutations (1), then it should be possible to identify at the molecular level the genetic events that underlie these evolutionary punctuations. For example, genomic rearrangements could change the timing of gene expression and thus profoundly affect development of the host; the new phenotype could sufficiently differ to start a new speciation process. Similarly, DNA rearrangements even in nonfunctional regions could disrupt meiotic chromosome pairing, leading to reproductive isolation of a previously interbreeding population. The specific genetic changes that punctuate the evolutionary process have yet to be identified, but perhaps the closest examples of such saltatory events in our genomes is the spreading of repetitive DNA elements. It therefore seems worthwhile to analyze these events in some detail, because they can provide us with genetic markers timing the evolutionary process.

The most prominent among the various repetitive DNA elements are members of the *Alu* family. They are considered to be pseudogenes that are ancestrally related to 7SL RNA (2, 3) or 4.5S RNA (4), and it appears they arose from an as yet unidentified founder gene. They are specific to primate species, where they are found at an estimated 10^6 copies per genome (5–7). They have been known for some time, but they remain somewhat enigmatic, and their usefulness as evolutionary markers has not been fully appreciated. Although early studies on the α- and β-globin gene regions revealed no differences in the location of 14 specific *Alu* elements between human, chimpanzee, and gorilla (8, 9), later studies did reveal differences in the location of specific DNA repeats between various primate lineages. For instance, in the work by Fitch *et al.* (10), involving sequence determination in the β-globin gene region, a truncated L1 sequence and two *Alu* elements were found to specifically distinguish the spider monkey from the human and the apes. Dugaiczyk and coworkers (11, 12) found an *Xba* and an *Alu* DNA repeat in the human α-fetoprotein (AFP) gene that were absent from the gorilla AFP gene. Additional human-specific *Alu* sequences have also been found, and their species-specific presence or absence was demonstrated by DNA hybridization and PCR (13). In addition, a gorilla-specific *Alu* element in the β,δ-globin region was absent from the human, chimpanzee, and macaque (14).

The process of new DNA repeats arising in the genomes continues today, as has been demonstrated by a contemporary *de novo* insertion of an *Alu* repeat into the neurofibromatosis gene (15) and a *de novo* insertion of an L1 sequence into the factor VIII gene (16) in man. We have extended the work on the human-specific AFP *Alu* repeat to include additional species-specific DNA elements and to screen small populations of higher primates for these DNA sequences. The study was undertaken in an effort to establish a correlation between the emergence of new DNA repeats and the divergence of primate species.

MATERIALS AND METHODS

DNA. Human DNA samples from 16 individuals were provided by Richard Gatti (Department of Pathology, University of California Medical School, Los Angeles). DNA samples from individual primate animals were provided by Oliver Ryder (Research Department, San Diego Zoo) and by Jerry Slightom (Division of Molecular Biology, Upjohn, Kalamazoo, MI).

PCR. Cycling conditions used in PCR experiments are described in figure legends.

Cloning of PCR Products. DNA amplified by PCR was cloned into the *Hph* I site of the cloning vector pCR1000 (Invitrogen) according to the manufacturer's protocol. Recombinant clones were selected on kanamycin/5-bromo-4-chloro-3-indolyl β-D-galactopyranoside plates and grown in liquid culture, and plasmid DNA was isolated.

Sequencing of Recombinant Plasmid DNA. The *Hph* I cloning site in pCR1000 is flanked by *Eco*RI and *Hin*dIII restriction sites. These two sites were therefore used to label the DNA with ^{32}P, and the labeled fragments were sequenced by the chemical degradation method (17).

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: AFP, α-fetoprotein; Myr, million years.

†Present address: Department of Medicine, University of Alabama, Birmingham, AL 35294.

Direct Sequencing of PCR-Amplified DNA. DNA amplified in the PCR was concomitantly labeled with ^{32}P as follows. One of the primers was first end-labeled by using T4 polynucleotide kinase and [γ-^{32}P]ATP prior to addition to the PCR mixture. The end-labeled PCR product was purified from unincorporated primers by polyacrylamide gel electrophoresis, eluted from the gel, and sequenced (17).

RESULTS

Human-Specific *Alu* Repeat in the AFP Gene. A specific *Alu* repeat is present in the human AFP gene but absent from the same (orthologous) position in the gorilla (11, 12). To extend these studies to include other primates, and to see whether the presence of this *Alu* repeat is fixed or polymorphic in the human population, genomic DNA from human, chimpanzee, gorilla, and orangutan was amplified by PCR. The primers (D-534 and D-533; see legend to Fig. 1) for this reaction were from intron 4 of the human AFP gene sequence (11). This set of primers should amplify a 530-bp fragment if the *Alu* sequence is present, or 215 bp in the absence of this *Alu* repeat. Human DNA indeed yielded an amplification product of the expected size. On the other hand, amplified DNA products derived from genomic DNA from the chimpanzee, gorilla, and orangutan were small (215 bp), indicative of the absence of this *Alu* character in these primates (Fig. 1*A*).

To verify that the amplified DNA products originate from orthologous sites in the four species, the DNA bands were excised from the gel and their sequences were determined. The sequences surrounding the *Alu* insertion site are shown in Fig. 1*B*. Within the stretch of 85 nucleotides shown, the sequences are the same, except for one mutation in the orangutan. The 11-bp unoccupied target site is identical in three primates, but in the human this target site was duplicated and gave rise to the terminal repeats, conceivably during the event of the *Alu* transposition. The results indicate that the primers specifically amplify orthologous sites in genomic DNA from the four primates. They also demonstrate that the AFP *Alu* repeat is specific to the human genome.

To establish whether the presence of this *Alu* sequence is fixed or polymorphic in the human population, we performed PCR experiments on genomic DNA from 16 individuals (13 Europeans and 3 Americans—1 Black, 1 white, and 1 Indian), using the same (D-534 and D-533) primers. Only the large (530 bp) fragment was obtained from all 16 samples (data not shown), indicating the presence of the repeat. Thus, including the one individual from which the original AFP gene sequence was determined (11), a total of 17 out of 17 humans contain the AFP *Alu* repeat. We conclude from this population sampling that the AFP *Alu* repeat is an established character in the human lineage, while it is absent from other primates.

Gorilla-Specific *Alu* Repeat in the Globin Gene Region. Trabuchet *et al.* (14) reported a gorilla-specific *Alu* element in the β-globin gene region but also suggested that the presence/absence of this character might be polymorphic in the gorilla lineage. This *Alu* repeat is absent from the human, chimpanzee, and macaque. We decided to include the orangutan to our study and to clarify whether the presence of this repeat is indeed polymorphic in the gorilla. The two primers for our PCR experiments (D-820 and D-821; see legend to Fig. 2) were selected from the reported gorilla sequence (14). The two primers should amplify a 532-bp fragment if the *Alu* sequence is present, or a 215-bp fragment in the absence of this *Alu* sequence. Gorilla DNA gave rise to a PCR product of the expected electrophoretic mobility of 540 bp, whereas the human, chimpanzee, and orangutan DNA gave only the 215-bp fragment (Fig. 2*A*). Thus, among the four primates analyzed, only the gorilla contained the globin *Alu* repeat. The amplified products were further identified by sequence determination, part of which is shown (Fig. 2*B*). In an 85-bp

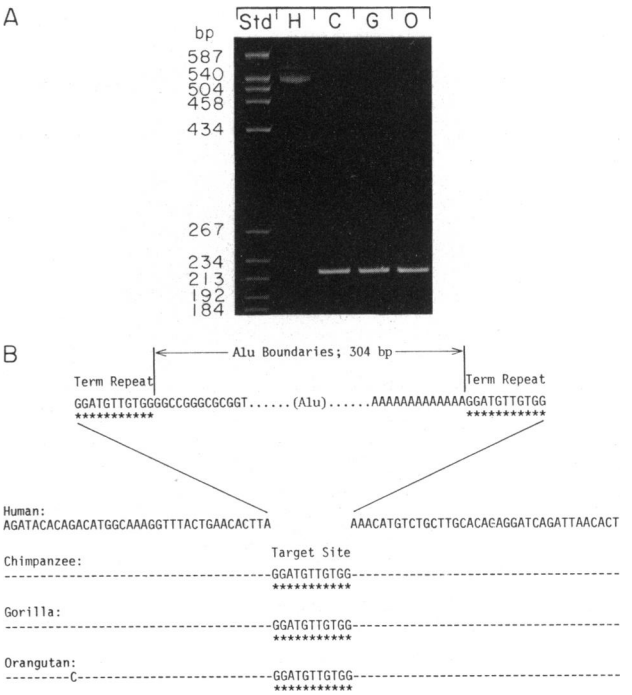

FIG. 1. PCR and sequence analysis of orthologous AFP loci in primates. (*A*) Electrophoretic separation of PCR products. Ingredients for the PCR included (in 100 μl) 20 mM Tris·HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 250 μM each dNTP, 0.2 μM each primer (D-534 and D-533), and 1 μg of genomic DNA. Amplification of DNA was performed in an Ericomp thermal cycler under the following conditions: 98°C for 2 min, pause at 80°C (2.5 units of *Taq* polymerase was added at this point), followed by 30 cycles of denaturation (96°C for 30 sec), annealing (55°C for 30 sec), and elongation (62°C for 30 sec). The last cycle was followed by a final elongation step (65°C for 5 min), and the reaction was stopped by the addition of 10 μl of 0.25 M EDTA at 0°C. Coordinates and sequences (11) of the two primers: D-534 (4965–4986), 5′-GCCTAGAAATAAAAGTTAGAGA-3′; D-533 (5494–5473), 5′-AACCTTCAATTTCCTATCAACA-3′. Samples (10 μl) of the PCR mixtures were electrophoresed in a 5% polyacrylamide gel. Lanes H, C, G, and O represent human, chimpanzee, gorilla, and orangutan PCR products, respectively. A control reaction which contained all ingredients except DNA did not show amplification products (data not shown). Sizes of molecular size standards (lane Std) are indicated in base pairs. (*B*) DNA sequence of regions surrounding the unoccupied and occupied target sites of the *Alu* element within the AFP gene. Dashes represent nucleotides identical to the human sequence; the single difference is indicated (i.e., C in orangutan). The target-site sequence is underscored with stars; duplication of this site generates the terminal (Term) repeats that flank the 5′- and 3′-boundaries of the *Alu* element in the human sequence. The complete sequence of the human *Alu* element was reported previously as part of the AFP gene structure (11).

stretch surrounding the *Alu* repeat, the four sequences are identical except for one mutation in the chimpanzee and one in the orangutan. A 16-bp unoccupied target site is present in human, chimpanzee, and orangutan, and this target sequence is converted to the terminal repeats flanking the gorilla *Alu* element.

Identical PCR experiments on genomic DNA from seven individual gorillas (three of them were related, four were caught in the wild) produced only the large (540-bp) fragment on gel electrophoresis (data not shown), indicating that all seven contained this globin *Alu* repeat. If we include the original animal where the first *Alu* sequence was identified (14), this makes a total of eight out of eight individuals (possibly nine; see *Discussion*) possessing this character. Although the population analyzed is rather small, it appears that the globin *Alu* element is indeed gorilla-specific and fixed in the gorilla lineage.

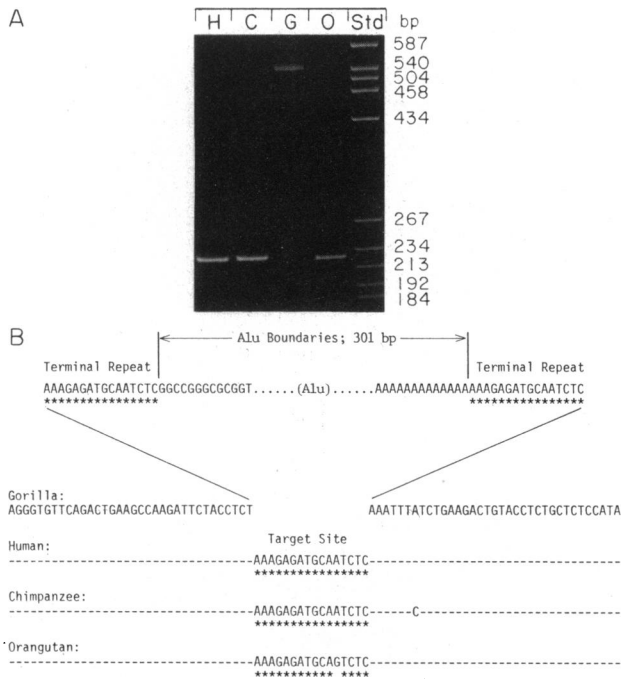

FIG. 2. PCR and sequence analysis of orthologous globin loci from primates. (*A*) Electrophoretic separation of PCR products; conditions were the same as described in Fig. 1, except that the primers differed. Coordinates and sequences (14) of the globin primers: D-820 (544–567), 5′-AGTCTCCTAAATCATCCATCTCTT-3′; D-821 (1075–1051), 5′-CTCAGATTTTCATTCAGAATTAGCT-3′. Lanes H, C, G, and O represent human, chimpanzee, gorilla, and orangutan PCR products, respectively. Sizes of molecular standards are indicated in base pairs. (*B*). DNA sequence of the region surrounding the target site and the *Alu* insertion. The target site is underscored with stars, as are the terminal repeats flanking the gorilla *Alu* element. Dashes represent nucleotides identical to the gorilla sequence; differences are indicated. In the orangutan, one mutation is observed within the unoccupied target site; another mutation is found in the chimpanzee within the 85 nucleotides shown. The complete sequence of the gorilla *Alu* element has been reported (14).

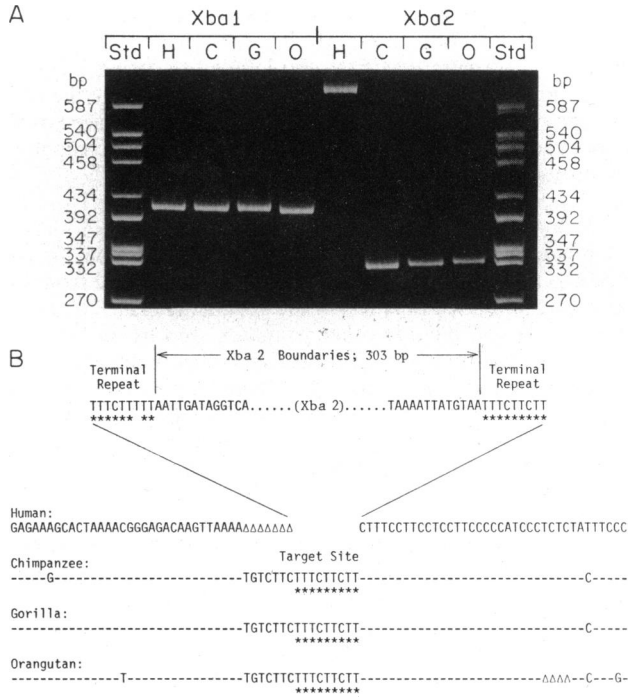

FIG. 3. PCR and sequence analysis of orthologous *Xba*1 and *Xba*2 loci in primates. (*A*) Electrophoretic separation of PCR products; conditions were the same as described in Fig. 1, except that the primers differed. Coordinates and sequences (11) of the primers surrounding *Xba*1: D-860 (11,916–11,939), 5′-TAGAGACATTTG-CAGTACAGTAGT-3′; D-859 (12,298–12,276), 5′-TCAGTTGT-CACTATGTGTCTCCT-3′. Coordinates and sequences (11) of the primers surrounding *Xba*2: D-811 (10,592–10,615), 5′-GAACAAA-CAGGTAAGTGGAAAGGA-3′; D-802 (11,211–11,188), 5′-GTCT-TGTTGAGAACATATGTAGGA-3′. Lanes H, C, G, and O represent human, chimpanzee, gorilla, and orangutan PCR products, respectively. Molecular size standards are indicated in base pairs. (*B*) DNA sequence of regions surrounding the unoccupied and occupied target sites of the *Xba*2 repeats. Target sites and terminal repeats are underscored by stars. Boundaries of the *Xba*2 element are indicated by vertical lines. Open triangles indicate deleted nucleotides, while dashes represent nucleotides identical to the human sequence. The complete sequence of human *Xba*2 has been reported (11).

Progenitor *Xba*1 DNA Gives Rise to Progeny *Xba*2 Repeat. There are two *Xba* elements in the human AFP gene, apparently the only two in the human genome (11). *Xba*1 differs from *Xba*2 by only 3 out of 303 nucleotide positions. Since only *Xba*2 is flanked by direct terminal repeats, this suggested a duplication of *Xba*1 and transposition to a target site at a new location. In the present work, we decided to analyze this unique genomic event in the other primates.

The two primers (D-860 and D-859; see legend to Fig. 3) for the amplification of the *Xba*1 region were selected from the known human AFP gene sequence (11). They should amplify a 383-bp DNA fragment containing the *Xba*1 sequence. When the PCR amplification products of genomic DNA from human, chimpanzee, gorilla, and orangutan were analyzed by gel electrophoresis, they all showed one DNA band with a mobility of ≈400 bp (Fig. 3*A*). Although the electrophoretic mobility of the fragments indicated a somewhat larger size than was expected, their determined sequences were 383 bp long, exactly as predicted from the position of the primers. The PCR-amplified *Xba*1 sequence was almost identical in all four primates (Fig. 4). Compared with the human, the chimpanzee differed by 4, the gorilla by 2, and the orangutan by 9 positions within the entire 303-bp *Xba*1 DNA.

The two primers (D-811 and D-802; see legend to Fig. 3) for the amplification of the *Xba*2 region were also selected from the known human AFP gene sequence (11). They should amplify a 620-bp DNA fragment containing *Xba*2, or 308 bp if *Xba*2 is absent. Human DNA gave rise to the expected 620-bp amplification product, while the three other primates produced only the small fragment, indicating the absence of *Xba*2 (Fig. 3*A*). Because of an additional 7 bp in the unoccupied target site, the size of the PCR-amplified DNA in the chimpanzee and gorilla actually amounts to 315 bp (311 bp in the orangutan), rather than the 308 bp predicted from the position of the primers on the human AFP sequence. These differences are due to small deletions in the human and orangutan DNA sequence. In addition to size, the PCR amplification products were further identified by their sequence determination, and part of the sequence surrounding the *Xba*2 target site is shown (Fig. 3*B*). As was the case with the globin *Alu* and AFP *Alu* elements, the genomic sequence surrounding the *Xba*2 target site is almost identical in the four primates. In the chimpanzee, gorilla, and orangutan, a 16-bp unoccupied target site can be identified. In the human, only part (9 bp) of this target sequence gave rise to the terminal repeats flanking the human-specific *Xba*2 element. The first 7 bp of the target sequence were apparently deleted during transposition of the *Xba* DNA.

Using the same two primers (D-811 and D-802), we amplified by PCR the *Xba*2 region of genomic DNA from 17 human individuals, and they all gave DNA fragments of the same size (620 bp; data not shown), as did the human DNA shown in Fig. 3*A*. From this limited population sample, we

```
Human:                                                                                                    100
AACTGATAGGTCACGTTCTCACTCTATTTGCCTTTAAGGGAAGAAAGCAATCAAGTTAATACGTTTTCCTTCATTGTATAGTATGTAACTATGGACACTATTAGAGGAGGGATTTGTGTAGCACTTAGGACATTATACTTGATAATTTCCAA
Chimpanzee:
--T-C----------------------------------------------------------------------------------------------------------------------------C----------------------------
Gorilla:
--T----------------------------------------------------------------------------------------------------------------------------------------------------------
Orangutan:
-CT----------T-----------------------------------------C-----------------------------------------------------------------------C----------------------------

Human:                                                                                                                                                        300
GGGTCTTTCTAGATTTAAAAGTCTGATTCTAACGTAGTAATAAAAATAAAGGCCCAATTTTCTCTTTAATATTGCCTGAAGATATTACTCTATTATTGCATTAAAATTAAACATTCACACATTGTTTGCACTGCTAAATAAAATTATGTAA
Chimpanzee:
-------------------------------------------------------T--------------------------------------------------------------------------------------------------
Gorilla:
-----------------------------------------------------------------------------------------G----------------------------------------------------------------
Orangutan:
-----------------------------------A---------------T-----------------C-----------------------------------------------G----------------------------------
```

FIG. 4. DNA sequence of the *Xba*1 element from four primate species. The sequence was obtained from the PCR amplification products described in Fig. 3. Dashes indicate nucleotides identical to the human sequence; differences are indicated.

conclude that *Xba*2 is a human-specific character that is fixed in the human genome.

The saltatory transposition of *Xba*1 into its new location is the second known example in the human genome where the progenitor and its progeny sequence have been identified. Previously, an L1 founder has been identified that gave rise to its copy inserted into the factor VIII gene, resulting in hemophilia (16). Although several laboratories (18–20) have intensively searched for the *Alu* "founder" gene that gives rise to new *Alu* transposons, the *Alu* founder gene remains elusive.

Evolutionary Rate and the Age of a Pseudogene. DNA sequences are transposed as repeats or pseudogenes in mechanisms that differ from those that govern nucleotide substitutions. The time of appearance of newly repeated DNA sequences and the rate of nucleotide substitutions should therefore independently measure evolutionary time and, we hope, give similar results. We decided to test this expectation on the human α-enolase pseudogene. This pseudogene has been sequenced, and both the active α-enolase gene and its processed pseudogene were found as single copies on human chromosome 1 (21). Using a published (22) rate for nucleotide substitutions in pseudogenes (5×10^{-9} per site per year), Feo *et al.* (21) calculated that the human enolase pseudogene arose from its active progenitor about 14 million years (Myr) ago. If the calculated date is correct, one should find the pseudogene in the chimpanzee, which diverged from the hominid lineage 6–8 Myr ago, and in the gorilla, which diverged from the hominids and chimpanzees 7–10 Myr ago (23–25). The pseudogene might or might not be present in the orangutan, estimated to have diverged about 16 Myr ago (25–27), and certainly not present in Old World monkeys, estimated to have diverged from the hominoid lineage 25–30 Myr ago (26, 27).

Based on the published sequence for the human α-enolase pseudogene (21), we synthesized two primers (D-929 and D-930; see legend to Fig. 5) that flank the pseudogene sequence. They should amplify a 1.8-kb DNA fragment containing the pseudogene, or 92 bp if the pseudogene is absent. All five primates, including the baboon, yielded the large PCR product, indicative of the presence of the enolase pseudogene (Fig. 5). The baboon PCR product was further identified by cloning it into the pCR1000 vector for partial sequence determination. The 638 bp of sequence data (not shown) that were determined unambiguously identify the enolase pseudogene in the baboon. The identification includes one of the terminal repeats that flank the human pseudogene and part of its surrounding sequence. The baboon sequence differs by 47 of 638 positions from that of the human, amounting to 7.36%. With our observed sequence difference of 7.36% between human and baboon, this would mean a pseudogene divergence rate of 2.73×10^{-9} per site per year for a 27-Myr divergence time. Considering that much higher, as well as lower, rates have been reported for other pseudogenes—for example, 12.6×10^{-9} per site per year for an α-globin pseudogene (28), 4.85×10^{-9} per site per year as an average for six pseudogenes from four species, including human (29), 4.3×10^{-9} per site per year for a lactate dehydrogenase pseudogene (30), or 10^{-9} per site per year for β- and η-globin pseudogenes (31, 32)—one must conclude that evolutionary rates differ significantly even if determined for a similar type of genomic sequences such as nonfunctional pseudogenes. This raises the question of whether the molecular clock can be used to infer divergence times.

DISCUSSION

Old and New Species-Specific DNA Repeats. *Alu* repeats, once inserted, remain stable characters in the genome of the host. Several examples of specific *Alu* elements have been identified as ancient insertions in an ancestor of both humans and apes, and after some 40 Myr of evolution they are found to reside at the same chromosomal position in at least four

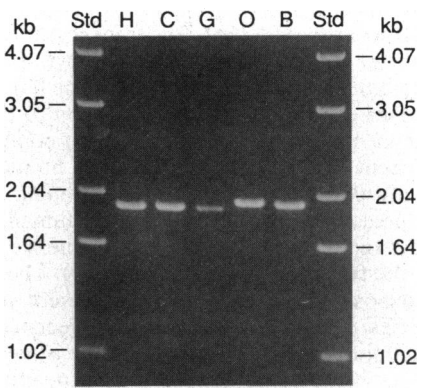

FIG. 5. PCR analysis of orthologous enolase pseudogene loci in primates. Coordinates (taken from ref. 21) and sequences of the PCR primers: D-929 (1–24), 5′-ACCCACAGCATTATATGTAACATC-3′; D-930 (1702–1680), 5′-CCAACAAGACAGGGTCACTTCTC-3′. Fifty nanograms of genomic DNA was mixed with 10 pmol of each primer in a total volume of 10 μl containing 20 mM Tris·HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 250 μM each 4 dNTP, and 0.4 unit of AmpliTaq polymerase. The reaction mixtures were transferred from ice to a preheated 94°C Perkin–Elmer 9600 thermal cycler and left at 94°C for 60 sec, followed by 30 cycles of denaturation (94°C for 15 sec), annealing (62°C for 23 sec), and elongation (72°C for 30 sec). The last cycle was followed by a final elongation step (72°C for 3 min), and the reactions were terminated by ramping the temperature to 4°C. Each reaction mixture was combined with 4 μl of 80% formamide/dye solution and loaded onto a 1.5% agarose gel for electrophoretic separation. Outer lanes contained molecular size standards. Lanes H, C, G, O, and B represent human, chimpanzee, gorilla, orangutan, and baboon PCR products, respectively.

contemporary primates (10). Examples of deletions due to unequal crossover at two *Alu* repeats are known, but there is no evidence for a mechanism of specific removal of single *Alu* elements. Thus, the mechanism of their insertion appears to be an irreversible and ongoing process; the latter is exemplified by contemporary insertions of an *Alu* and an L1 repeat into the human genome (15, 16). However, the most informative repeats for phylogenetic studies are these that are a few million years old and have became established characters in one or a few primate lineages. In the present work, we provide evidence for two such characters (*Alu* and *Xba*) that are human-specific, one *Alu* repeat that is gorilla-specific, and an enolase pseudogene that is present in at least five primate lineages: human, chimpanzee, gorilla, orangutan, and baboon. These results raise the expectation that other DNA repeats will be found that will be specific to small clades of two or three primate species and thus allow their phylogenetic grouping.

Since the gorilla-specific *Alu* element was originally reported to be possibly polymorphic as to its presence in the gorilla lineage (14), it became important to verify this possibility. The reason for the possible polymorphism was based (14) on previous work by Barrie *et al.* (33), who mapped this region and found no difference in the size of restriction fragments between human, gorilla, and baboon. The DNA from the gorilla analyzed by Barrie *et al.* (33) could not be made available to us for sequence determination in the *Alu*-containing region, but we reexamined the published data. In figure 1 of ref. 33, we see in retrospect that the δ (10-kb) DNA fragment from the gorilla shows a slightly larger size than the neighboring human and baboon restriction fragments, indicating the presence of extra DNA. A difference of 0.3 kb (the size of *Alu* DNA) in the resolution of a 10-kb fragment can easily escape detection in a Southern hybridization experiment on total genomic DNA. We think that the globin *Alu* repeat was actually present in the gorilla studied by Barrie *et al.* (33), and thus it was present in nine out of nine animals analyzed. It may be safe to conclude that it is fixed in the gorilla lineage.

Saltatory Appearance of *Xba*2. The transposition of the *Xba* character is one of only two examples known in the human genome where the exact progenitor sequence has been identified; the other example is the identification of the founder for a recent L1 insertion (16). However, in contrast to L1, which is present in ≈100,000 copies in the human genome, there are only two copies of the *Xba* element. *Xba*1, the founder, is present in orthologous loci in human, chimpanzee, gorilla, and orangutan, and its sequence is almost identical in the four primate species (Fig. 4). The target site for the transposition of *Xba*1 is also present in the four primate species: it is a sharp boundary between a purine-rich and a pyrimidine-rich region (Fig. 3). It is an intriguing question as to what triggered a single transposition of *Xba*1 in one primate lineage, while the same *Xba*1 DNA remained inactive for tens of millions of years in other closely related species. Although *Xba*2 arose and became fixed in only the human lineage, the event was probably inconsequential to the host phenotype. Nevertheless, the transposition is an example of a one-time episode in the history of one species, and as such it defies statistical analysis. It seems likely that similar one-time events have occurred in the evolution of genomes, some of them possibly having a direct consequence on the ontogeny and reproduction of the host.

Measuring Evolutionary Time. Mutagenic events are related either to transposition or to point mutations. These two classes of events are basically distinct, obey different molecular rules, and occur at different frequencies. Point mutations have been extensively used to reconstruct phylogenies from the amount of accumulated differences between two species (ref. 28, and references therein). However, there are intrinsic difficulties in the method, and they stem from the fact that the state of a nucleotide is fluctuating due to multiple and reversible mutations occurring at the same site or in two species; the method attempts to reconstruct phylogenies on unstable and shifting grounds of substitutions.

Transpositional events, on the other hand, differ from point mutations in one important aspect: there is no evidence that they revert back to the original state of the character as in the case of point mutations. Although transpositions accumulate at a slower rate and are more difficult to find, the irreversibility of their spreading provides stable evolutionary time markers. Our intention was to identify examples of genetic novelties that can be used to distinguish primate lineages. More data would be needed to decipher the complexities of a phylogeny.

We thank Rita Zielinski for PCR and DNA sequence analysis and Dr. Oliver Ryder and Dr. Jerry Slightom for gorilla DNA samples.

1. Gould, S. J. (1982) in *Evolution Now, A Century After Darwin*, ed. Smith, J. M. (Freeman, San Francisco), pp. 129–145.
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. Schoninger, 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., 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. Fitch, D. H. A., Mainone, C., Slightom, J. L. & Goodman, M. (1988) *Genomics* **3**, 237–255.
11. Gibbs, P. E. M., Zielinski, R., Boyd, C. & Dugaiczyk, A. (1987) *Biochemistry* **26**, 1332–1343.
12. Ryan, S. & Dugaiczyk, A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9360–9364.
13. Batzer, M. A. & Deininger, P. L. (1991) *Genomics* **9**, 481–487.
14. Trabuchet, G., Cheboune, Y., Savatier, P., Lachuer, J., Faure, C., Verdier, G. & Nigon, V. M. (1987) *J. Mol. Evol.* **25**, 288–291.
15. Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, P. E., Glover, T. W. & Collins, F. S. (1991) *Nature (London)* **353**, 864–866.
16. Dombroski, B. A., Mathias, S. L., Nanthakumar, E., Scott, A. F. & Kazazian, H. G. (1991) *Science* **254**, 1805–1808.
17. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
18. Britten, R. J., Stout, D. B. & Davidson, E. H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3718–3722.
19. Matera, G. A., Hellman, U., Hintz, M. F. & Schmid, C. W. (1990) *Nucleic Acids Res.* **18**, 6019–6023.
20. Shen, R. M., Batzer, M. A. & Deininger, P. L. (1991) *J. Mol. Evol.* **33**, 311–320.
21. Feo, S., Oliva, D., Arico, B., Barba, G., Cali, L. & Giallongo, A. (1990) *DNA Sequence* **1**, 79–83.
22. Miyata, T. & Yasunaga, T. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 450–453.
23. Hill, A. & Ward, S. (1988) *Yearb. Phys. Anthropol.* **31**, 49–83.
24. Ruvolo, M., Disotell, T. R., Allard, M. W. & Brown, W. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1570–1574.
25. Tobias, P. V. (1986) *S. Afr. J. Sci.* **82**, 92–94.
26. Sibley, C. G. & Ahlquist, J. E. (1984) *J. Mol. Evol.* **20**, 2–15.
27. Gingerich, P. D. (1984) *Yearb. Phys. Anthropol.* **27**, 57–72.
28. Miyata, T. & Hayashida, H. (1981) *Proc. Natl. Acad. Sci. USA* **85**, 5739–5743.
29. Li, W.-H., Luo, C.-C. & Wu, C.-I. (1985) in *Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 1–94.
30. Fukasawa, K. M., Li, W.-H., Yagi, K., Luo, C.-C. & Li, S. S.-L. (1986) *J. Mol. Evol.* **3**, 330–342.
31. Chang, L.-Y. E. & Slightom, J. L. (1984) *J. Mol. Biol.* **180**, 767–784.
32. Goodman, M., Koop, B. F., Czelusniak, J., Weiss, M. L. & Slightom, J. L. (1984) *J. Mol. Biol.* **180**, 803–823.
33. Barrie, P. A., Jeffreys, A. J. & Scott, A. F. (1981) *J. Mol. Biol.* **149**, 319–336.