Molecular Evolution of the ψη-Globin Gene Locus: Gibbon Phylogeny and the Hominoid Slowdown[1]

Wendy J. Bailey, David H. A. Fitch,†,[2] Danilo A. Tagle,* John Czelusniak,† Jerry L. Slightom,†,‡ and Morris Goodman†*

*Department of Molecular Biology and Genetics and †Department of Anatomy and Cell Biology, Wayne State University School of Medicine; and ‡Division of Molecular Biology, Upjohn Company

An 8.4-kb genomic region spanning both the ψη-globin gene locus and flanking DNA was sequenced from the common gibbon (*Hylobates lar*). In addition, sequencing of the entire orthologous region from galago (*Galago crassicaudatus*) was completed. The gibbon and galago sequences, along with published orthologous sequences from 10 other species, were aligned. These noncoding nucleotide sequences represented four human alleles, four apes (chimpanzee, gorilla, orangutan, and gibbon), an Old World monkey (rhesus monkey), two New World monkeys (spider and owl monkeys), tarsier, two strepsirhines (galago and lemur), and goat. Divergence and maximum parsimony analyses of the ψη genomic region first groups humans and chimpanzees and then, at progressively more ancient branch points, successively joins gorillas, orangutans, gibbons, Old World monkeys, New World monkeys, tarsiers, and strepsirhines (the lemuriform-lorisiform branch of primates). This cladistic pattern supports the taxonomic grouping of all extant hominoids into family Hominidae, the division of Hominidae into subfamilies Hylobatinae (gibbons) and Homininae, the division of Homininae into tribes Pongini (orangutans) and Hominini, and the division of Hominini into subtribes Gorillina (gorillas) and Hominina (chimpanzees and humans). The additional gibbon and galago sequence data provide further support for the occurrence of a graded evolutionary-rate slowdown in the descent of simian primates, with the slowing rate being more pronounced in the great-ape and human lineages than in the gibbon or monkey lineages. A comparison of global versus local molecular clocks reveals that local clock predictions, when focused on a specific number of species within a narrow time frame, provide a more accurate estimate of divergence dates than do those of global clocks.

Introduction

In the taxonomic classifications traditionally used for primates, the lesser apes or gibbons are separated from humans always at the family level and are separated from the great apes (orangutans, chimpanzees, and gorillas) either at the family level or at the subfamily level. In the former case the gibbons are placed in the family Hylobatidae (Fiedler 1956; Fleagle 1988), and in the later case they are placed in the

1. Key words: ψη-globin gene, noncoding nucleotide sequences, DNA hybridization, primate phylogeny, maximum parsimony, cladistic classification.

2. Present address: Department of Molecular Genetics, Albert Einstein College of Medicine, Bronx, New York 10461.

Address for correspondence and reprints: Morris Goodman, Department of Anatomy and Cell Biology, Wayne State University School of Medicine, Detroit, Michigan 48201.

Mol. Biol. Evol. 8(2):155–184. 1991.
© 1991 by The University of Chicago. All rights reserved.
0737-4030/91/0802-0001$02.00

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

subfamily Hylobatinae, which is then grouped with Ponginae (the subfamily for great apes) into family Pongidae (Simpson 1945). Although the primate classifications in present-day use invariably place gibbons, along with great apes and humans, in the superfamily Hominoidea, some students of primate phylogeny have viewed the gibbons as a transitional group between Old World monkeys and great apes. On the basis of brain convolutions and several aspects of dentition, von Koenigswald (1968) suggested that gibbons might be closer to Old World monkeys than to great apes and humans. Karyotypic evidence also separates gibbons from other hominoids. *Hylobates* has 44 chromosomes (Chu and Bender 1962), while each of the great ape genera [*Pongo* (orangutan), *Gorilla,* and *Pan* (chimpanzee)] has 48, and humans have 46 (Hamerton 1963). The gibbon and Old World monkey karyotypes share a secondary constriction in a pair of autosomal chromosomes, a characteristic absent in great apes and humans (Hamerton et al. 1963; Chiarelli 1966). In addition, both common and siamang gibbons lack a group of satellite-bearing acrocentric chromosomes, distinctive of other apes and humans. The phylogenetic evidence provided by karyotypes is comparable to that provided by morphology and is not always reliable. In contrast to karyotypic evidence (Hamerton 1963; Hamerton et al. 1963; Klinger et al. 1963; Klinger 1963; Chiarelli 1966), the evidence from proteins and DNA has consistently supported a monophyletic grouping of gibbons, great apes, and humans. This macromolecular evidence includes that from protein immunology (Goodman 1963; Hafleigh and Williams 1966; Sarich and Wilson 1967; Goodman and Moore 1971; Dene et al. 1976; Sarich and Cronin 1976), from DNA-DNA hybridization (Hoyer et al. 1972; Kohne 1975; Benveniste and Todaro 1976; Sibley and Ahlquist 1984, 1987; Caccone and Powell 1989), from sequencing mtDNA (Brown et al. 1982; Hayasaka et al., 1988), and from sequencing nuclear DNA, such as that of the involucrin gene (Djian and Green 1990) and portions of the genomic domain containing β-type globin genes (Fitch et al. 1990; Goodman et al. 1990; present study).

We report here, for the common gibbon (*H. lar*), the DNA sequence of an 8.4-kb region encompassing the ψη gene, a nonfunctional member of the β-like globin gene family (Fritsch et al. 1980; Jagadeeswaran et al. 1982; Goodman et al. 1984; Harris et al. 1984). We also present, for galago (*Galago crassicaudatus*), a strepsirhine primate, the complete DNA sequence over the 5′ flanking region of the ψη locus, as well as the previously reported sequence over the exons and introns of the ψη locus (Koop et al. 1989*b*). The ψη-globin gene occurs in a genomic domain called the "β-globin gene cluster." This domain in mammalian genomes arose from a series of tandem gene duplications that began 150–200 Mya and that, by the time of the early eutherian mammals (80–100 Mya), had resulted in five paralogously related genes linked in the order 5′-ε-γ-η-δ-β-3′ (Goodman et al. 1984; Harris et al. 1984). Divergence and parsimony analyses indicate that the η-globin locus originated from an embryonically expressed proto-epsilon gene (Goodman et al. 1984, 1987; Koop and Goodman 1988). It has since maintained embryonic expression in the lineage to artiodactyls, has been lost entirely in the descent to rodents and lagomorphs, and has become a pseudogene in the stem primates (Goodman et al. 1984). Subsequent to η becoming a pseudogene in primates, a paralogous crossover occurred between η and δ in the lemur, resulting in a hybrid ψη/δ locus (Harris et al. 1984; Jeffreys et al. 1982). In galago the 3′ end of the η gene and most of the η-δ intergenic region have been deleted (Tagle 1990).

We have previously established a data set of aligned orthologous noncoding sequences representing >11,000 nucleotide positions of the ψη-globin locus and its

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

flanking DNA (Goodman et al. 1989; Koop et al. 1989*b*). The addition of the gibbon sequence to this data set completes a comparative series of the major hominoid lineages. With this updated data set, which also includes Old World monkey, New World monkey, tarsier, galago, lemur, and goat orthologues, we have reexamined questions concerning the phylogenetic relationships of gibbons and other primates. Furthermore, we have investigated questions concerning whether mutations accumulate at uniform rates or nonuniform rates in the noncoding sequences of different primate lineages and whether divergence dates among these lineages are best estimated by the model of a "global" molecular clock or by a clock model that is only applied to localized regions of the phylogenetic tree.

Material and Methods

Material

Restriction endonucleases were purchased from either Bethesda Research Laboratories or New England Biolabs and were used as specified by the vendor. Radioactive nucleotides [α-^{32}P]dATP (400 Ci/mmol) and [γ-^{32}P]ATP (2,000–3,000 Ci/mmol) were obtained from either DuPont—New England Nuclear or ICN Pharmaceuticals. T4 DNA ligase was purchased from Collaborative. Sequenase sequencing kits were obtained from United States Biochemicals; X-ray XAR-351 roll film was from Kodak, and DNA sequencing gel apparatuses were from Fotodyne, Inc.

Cloning Strategy for Gibbon ψη

Lambda clone Hla Ch40-12.5 was isolated from a recombinant bacteriophage Charon 40 (Dunn and Blattner 1987) library, was constructed using *Sau*3AI (Slightom and Drong 1988), and was propagated in *Escherichia coli* host ED8767 (Murray et al. 1977). The library was screened with a 245-bp *Ava*II-*Eco*RI fragment from human γ gene cDNA clone pJW151 (Wilson et al. 1978). Recombinant clone Hla Ch40-12.5 contains the entire ψη gene region including the 3′ end of the γ^2 gene and 4.8 kb of the intergenic ψη-δ sequence (see fig. 1). Hla Ch40-12.5 was digested with *Kpn*I; this generated two restriction fragments, one of 6.5 kb and one of 6.0 kb, which were subcloned into pUC19. Transformations were performed using competent *E. coli* K12 strain DH5α cells purchased from Bethesda Research Laboratories. Subclone nomenclature designates the species name [Hla (*Hylobates lar*)], the parent clone (Ch40-12.5), the restriction endonuclease used for cloning (*Kpn*I), and the fragment size (6.5 or 6.0) (see fig. 1).

DNA Sequencing

Chemical and enzymatic sequencing techniques described by Maxam and Gilbert (1980) and Sanger et al. (1977), respectively, were both employed. Because of the close similarity between gibbon and human sequences, synthetic 20–25-base oligonucleotides based on the human sequence were synthesized using an Applied Biosystems, Inc., synthesizer. These oligomers were then used as primers for directed dideoxynucleotide sequencing of double-stranded plasmid DNA (Zagurski et al. 1985). Primers were positioned at 500-bp intervals, alternating between each strand to ensure overlap of sequencing reads for both DNA strands. Initial sequence for the galago was obtained with commercially available universal forward and reverse primers designed to prime within the polylinker sites of M13 and pUC cloning vectors. Internal galago-specific primers were then synthesized on the basis of sequences obtained in the first set of reactions. Double-stranded sequencing involved DNA synthesis with T7 DNA

Downloaded from https://academic.oup.com/mbe/article/9/2/155/1134108 by guest on 24 August 2022

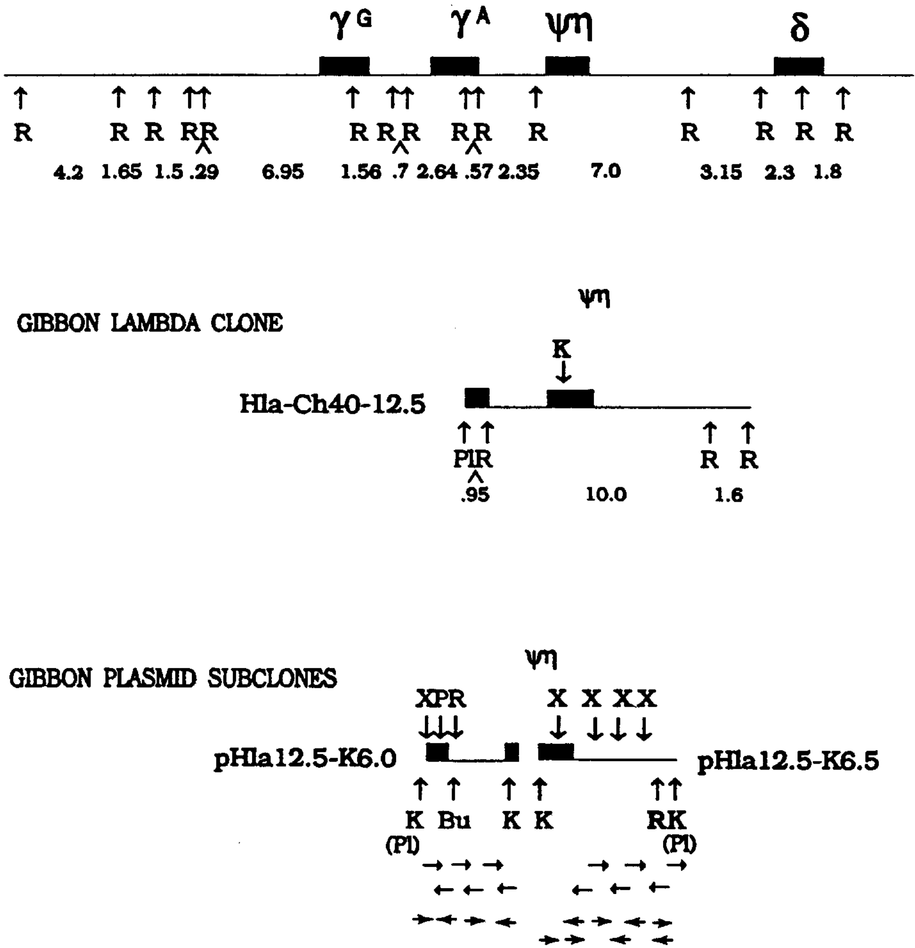


FIG. 1.—Gibbon ψη clones and sequencing strategies. The linear map at the top of the figure shows the organization and some restriction sites present in the human β-globin cluster encompassing the ψη gene. Black boxes denote genic regions for $^{G}\gamma$, $^{A}\gamma$, ψη, and δ. Bacteriophage lambda clone Hla-Ch40-12.5 containing the insert for the gibbon ψη gene (Fitch et al. 1990) is pictured directly below the human map. Digestion of recombinant clone Hla-Ch40-12.5 with restriction enzyme *Kpn*I generated two fragments, one of 6.5 kb and one of 6.0 kb, which were subsequently subcloned into pUC19. Plasmid subclones are shown below their parent clone. Sequencing was done by either 5′ end-labeled restriction fragments (script horizontal arrows) with ^{32}P for chemical sequencing, or synthetic oligonucleotides (horizontal arrows) were used as primers spaced alternately 500 bp apart for dideoxy chain-termination sequencing. Horizontal arrows illustrate the direction and location of individual sequencing runs. R = *Eco*RI; K = *Kpn*I; X = *Xba*I; P = *Pvu*II; Bu = *Bsu*36I; (P1) = restriction-endonuclease recognition sites which within Charon 40 polylinker.

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

polymerase (United States Biochemicals), from a primed site, followed by chain termination by dideoxynucleotide incorporation. More than 97% of the gibbon and galago ψη sequence was confirmed by sequencing both strands; the remaining 2%–3% was sequenced multiple times on one strand.

Sequence Alignment

The pairwise alignment algorithm of Smith and Waterman (1981), as modified by Goodman et al. (1984), was employed to generate initial alignments for the galago

and human ψη sequences. Because of the great similarity among all simian sequences, the gibbon sequence was added directly by hand to the full data set. A parsimony criterion was followed in completing the alignments by hand: gaps were inserted only where they increased sequence similarities that could be attributed to common ancestry while reducing convergencies. Altogether, the full data set contained 13 aligned orthologous sequences. (A copy of the alignment will be provided on request to the corresponding author.) The entire 10.8-kb human sequence (HsaA) and the second human sequence (HsaB) are each composites of two separate homologous regions represented by two alleles each. In HsaA the sequence over positions 1–9032 is from the data of Collins and Weissman (1984) as corrected by Miyamoto et al. (1987), and the sequence from positions 9033–12741 is the allele designated HumanT in figure 2 of Maeda et al. (1988). In HsaB the upstream sequence (positions 2060–5335) is from Chang and Slightom (1984), and the downstream sequence (positions 9033–12741) is the allele designated HumanR in figure 2 of Maeda et al. (1988). Chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), and orangutan (*Pongo pygmaeus*), sequences are from Miyamoto et al. (1987) and Maeda et al. (1988); that of common gibbon (*Hylobates lar*) is from the present paper; that of rhesus monkey (*Macaca mulatta*) is from Koop et al. (1986), Miyamoto et al. (1988) and Maeda et al. (1988); that of spider monkey (*Ateles geoffroyi*) is from Fitch et al. (1988) and Maeda et al. (1988); those of owl monkey (*Aotus trivirigatus*) and brown lemur [*Lemur macaco* (*fulvus*) *mayattensis*] are from Harris et al. (1984); that of tarsier (*Tarsius syrichta*) is from Koop et al. (1989*a*); that of galago (*Galago crassicaudatus*) is from Koop et al. (1989*b*), Tagle (1990), and the present paper; and that of goat (*Capra hircus*) is from Shapiro et al. (1983).

Evolutionary Reconstruction

Estimates of sequence divergence were calculated by using pairwise comparisons of the 13 sequences. These divergence values were then used to construct bifurcating trees both by the UPGMA method (Sneath and Sokal 1973) and by the neighbor-joining (NJ) method (Saitou and Nei 1987). The resulting topologies were used as initial "trees" to generate and test a number of alternative branch arrangements by using two branch-swapping parsimony programs, MPAFE and MPAL8, that swap branches according to procedures described by Czelusniak et al. (1990) for unrooted trees or networks. MPAFE examines all nearest-neighbor single-step changes (NNSSCs) in the network topology; that is, it examines for each two adjacent interior nodes the tree swaps for the four branches originating from the two nodes. For the N exterior nodes of the network, there are $2(N-3)$ alternative trees in each round of swaps. A tree having the lowest nucleotide-substitution (NS) score initiates the next round of swaps. This iterative lowering of NS score stops when a consecutive round of NNSSCs fails to lower the NS count. An option that may be employed in program MPAFE (and that is also in program MPAL8) is to designate subtrees within which no branch swaps are permitted; the maximum-parsimony (MP) solution for each designated subtree's ancestral sequence then serves as a terminal taxon. Program MPAL8 examines all possible trees for eight terminal taxa. A terminal taxon can either be an exterior node (an extant sequence) or a designated subtree. MPAL8 computes the NS score of each of the 10,395 possible trees that eight terminal taxa can form and then lists the trees in order of their distribution, from lowest to highest NS score. In addition to the four letters used in the sequence file for the four nucleotides, a fifth letter was used for gaps. As this fifth letter was employed only once per gap, irrespective of the

Downloaded from https://academic.oup.com/mbe/article/9/2/155/1134408 by guest on 21 August 2022

length of the gap, each indel (insertion or deletion) was scored as a single NS (i.e., for this study we treat indels as if they were NSs).

Branch or link lengths were determined by two different computational methods. The first method employed the NJ algorithm (Saitou and Nei 1987) and used divergence values corrected for superimposed base changes (Jukes and Cantor 1969). The second method is an MP method (Moore et al. 1973; Goodman et al. 1979, 1984; Goodman 1981) using a computer program called TPA (Czelusniak et al. 1990). For the given rooted tree and set of contemporary sequences at the exterior nodes, TPA chooses that set of parsimony ancestral sequences at the interior nodes, with link lengths (numbers of substitutions between adjacent ancestral and descendant sequences) that minimize the sum of the distances on the tree for every pair of exterior nodes. Of alternative parsimony solutions, the one chosen by TPA is the "A" solution of Goodman et al. (1974). Because extensive portions of the contemporary sequences were missing from the full alignment, the link lengths were normalized to represent numbers of sequence changes (NSs + indels) per every 100 sequence positions. The calculation of the normalized link length values shown in figure 2 is described in the legend to that figure.

Results

Molecular Features of the ψη Gene Sequence

The gibbon sequencing strategy is presented in figure 1, and the galago sequencing strategy is presented in figure 37 of Tagle (1990). Sequences for 8.4 kb of the gibbon and for 3.3 kb of the galago ψη-globin gene and flanking regions were aligned against published orthologous sequences from brown lemur, tarsier, owl monkey, spider monkey, rhesus monkey, orangutan, gorilla, chimpanzee, human and goat (for references, see Methods subsection above). The gibbon sequence contains all previously described mutations (Koop et al. 1986; Fitch et al. 1988) that may have contributed to the inactivation of the η-globin gene in an ancestor of extant primates, including two transitions that change the initiation codon from ATG to GTA (position 2445–2447 in the data set of the 13 aligned sequences). Located 980 bp upstream from the initiation codon (positions 1469–1844) in the gibbon sequence is an Alu-element insertion which is also present in the other catarrhines but which is absent in the spider monkey and galago orthologues. Therefore, the date of the Alu insertion can be placed sometime after the New World monkeys diverged from catarrhines but before the Old World monkeys diverged from hominoids, i.e., in the 25–35-Mya range (Gingerich 1984; Pilbeam 1984; Rosenberger 1984). In addition, the gibbon sequence has a unique L1 element (long interspersed nuclear element) insertion beginning at position 10158 and extending 413 bp, to the end of the cloned fragment (position 10571).

The galago ψη sequence has several unique structural features and sequence anomalies. These include two deletion mutations—a loss of a dinucleotide at positions 3260–3261 and a nucleotide at position 3348 that generated, by frameshift, six in-frame termination codons in the downstream region. The GT dinucleotide splice junction at the beginning of intervening sequence 1 is replaced by GC (positions 2537–2538), and the intron 2 splice junction is replaced by TT (positions 3445–3446). The 3′ splice site AG found at the end of intron 1 is replaced by CA (positions 3212–3213). Promoter elements were also altered or lost. A 9-bp and a 39-bp deletion abolished the AATAAA (positions 2359–2365) and proximal CCAAT (positions 2305–2309) elements, respectively. Mutations have also altered the distal CCAAT (positions 2266–2270) and the CACCC (positions 2237–2241) elements, which changed to

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1034008 by guest on 21 August 2022

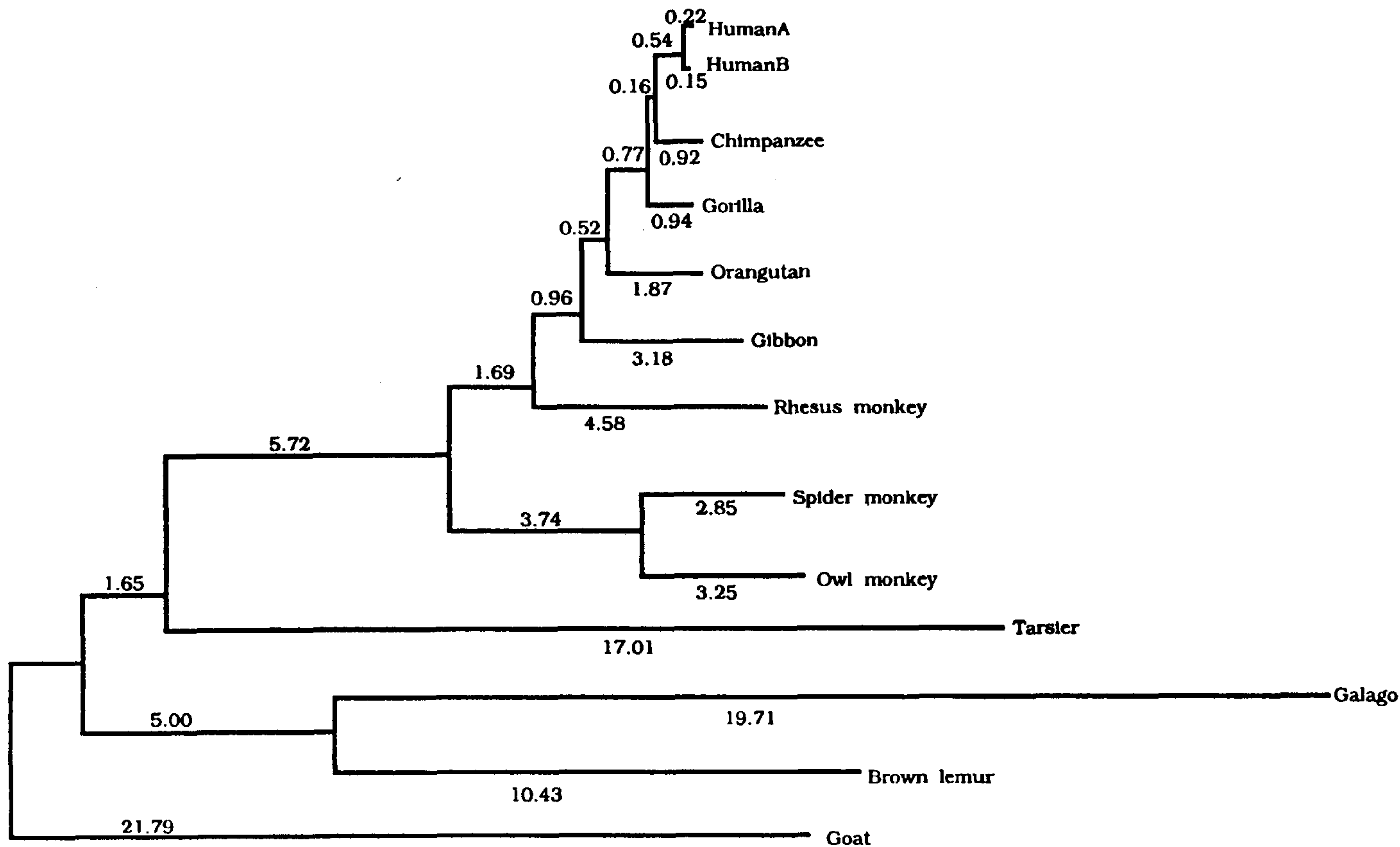
AAAAT and CATTC, respectively. A truncated L1 LINE element insertion is present beginning at position 2612 within intron 1 and extends 3′ for 540 bp. In addition, the galago ψη gene is truncated in its 3′ end by a deletion which begins ~560 bp downstream of exon 2 and encompasses almost the entire ψη-δ intergenic region. Beginning within intron 2 at position 4006, in the galago, there are several truncated L1 elements inserted tandemly which total ~11 kb in length and extend 3′ to just upstream of δ (Tagle 1990; M. A. Stanhope, P. A. Tagle, J. L. Slightom, and M. Goodman, unpublished data). None of these structural defects are shared with other primates, indicating that they occurred in the galago lineage after the initial inactivation event of the η gene in a common primate ancestor.

Phylogenetic Reconstruction

Pairwise comparisons of the ψη sequences over shared nucleotide positions are shown in table 1. Pairwise divergence values for this noncoding genomic region are consistent with those that Bonner et al. (1980), Sibley and Ahlquist (1987), and Caccone and Powell (1989) found for total single-copy genomic DNAs by using DNA-DNA cross-hybridization. After UPGMA and NJ trees were generated from these pairwise divergence values (table 1), the same lowest-length (LL) tree was found by program MPAFE, regardless of whether the branch swaps started from the UPGMA or the NJ tree. Furthermore, exhaustive branch swaps by program MPAL8 did not reveal any trees of equal or lower score than that found by MPAFE. This most parsimonious tree is shown in figure 2 and has a total score of 3,838 (3,433 NSs + 405 indels). The number of substitutions in the UPGMA and NJ trees (3,868 and 3,859, respectively) were, respectively, 30 and 21 greater than those of the LL tree. In the UPGMA tree, lemur and Anthropoidea were closest in common ancestry and were increasingly more remote in common ancestry first from tarsier, next from galago, and last from goat. In the NJ tree, galago grouped with tarsier, and this galago-tarsier branch was next to Anthropoidea while lemur and goat were farthest from Anthropoidea. The MP or LL tree (fig. 2) joins galago to lemur, yielding the stepsirhine branch, while tarsier stays with Anthropoidea, yielding the haplorhine branch. In all three trees (UPGMA, NJ, and MP or LL), Anthropoidea divides into a platyrrhine or ceboid (New World monkey) branch and a catarrhine branch, and the latter divides into a cercopithecoid (Old World monkey) and hominoid branch. These three trees also support cladistically the taxonomic grouping of all extant hominoids into family Hominidae, the division of Hominidae into subfamilies Hylobatinae (gibbons) and Homininae, the division of Homininae into tribes Pongini (orangutans) and Hominini, and the division of Hominini into subtribes Gorillina (gorillas) and Hominina (chimpanzees and humans). In the Discussion section (see below) we justify the use of this cladistic classification of extant hominoids as an alternative to the traditional gradistic classification which places the great apes in subfamily Ponginae of family Pongidae and which places humans alone among extant hominoids in family Hominidae.

Program MPAL8 was used to assess how strongly these ψη sequences support the monophyly both of apes and human and of each of the hominoid groups depicted in figure 2. To arrange the lineages shown in figure 2 into eight terminal branches, the two human sequences were fixed as a single subtree, the two ceboid or New World monkey sequences as another single subtree, and all the nonsimians (tarsier, galago, lemur, and goat) as yet another single subtree which represented the outgroup of the seven simian terminal branches (human, chimpanzee, gorilla, orangutan, gibbon, rhesus, and ceboid). Using the MP method to test each of the 10,395 trees formed by

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1124108 by guest on 24 August 2022



Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134

these eight branches, the MPAL8 program confirmed that the tree shown in figure 2 is indeed the LL tree. In addition, MPAL8 identified the lowest scoring tree or trees that broke up each clade of two or more hominoids. These results are summarized in figure 3 for the human-chimpanzee, human–African ape, human–great ape, and human-ape clades. The numbers on the stems of these hominoid clades (fig. 3) are the number of additional NSs required by the trees that break up these particular clades. Thus the human-chimpanzee clade is first broken by the tree with a chimpanzee-gorilla clade which required an additional eight NSs (fig. 3). The human–African ape clade is broken when orangutan rather than gorilla joins the human-chimpanzee stem, adding an extra 65 NSs. The tree breaking up the human–great ape clade joined gibbon and orangutan branches, adding 26 NSs to the most parsimonious solution. The tree breaking the human-ape clade joined rhesus and gibbon branches, adding 55 extra NSs. The tree which grouped all the great apes into a monophyletic clade (Kluge 1983) scored at 3,918, or 80 more NSs. Grouping human with orangutan and chimpanzee with gorilla, the arrangement favored on morphological grounds by Schwartz (1984), also scored at 3,918. Thus these two contending hypotheses that oppose a human–African ape clade each require adding 80 extra NSs to the most parsimonious tree.

The Phylogenetically Informative Positions

Over the 12,741 nucleotide positions of the alignment for the ψη sequences there are 12 different single or contiguous positions where the sequence changes uniquely group *Homo* and *Pan* together, whereas there are only four positions where *Pan* and *Gorilla* could be so grouped, and just three positions for *Homo* and *Gorilla* (table 2). The 12 sequence characters that may be viewed as shared derived characters or synapomorphies supporting a *Homo-Pan* clade consist of the following: three transitions, at positions 6062 (T→C), 10947 (A→G), and 12110 (A→G); three transversions, at positions 583 (A→C), 7075 (T→G), and 8910 (T→G); one insertion at position 4821; and five deletions, at positions 1700, 4576–4579, 4999, 5945–5946, and 9159–9164 (table 2; position 4999 is shown, underlined and in boldface, within positions 4978–5001). In our previous analysis of phylogenetically informative positions (Goodman et al. 1989), the deletions at positions 1700 and 4999 of the present align-

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

FIG. 2.—MP tree constructed from the data set of the 13 aligned sequences. Link lengths represent the most parsimonious number of changes or NSs, defined to also count each gap as 1 NS/100 nucleotide positions during the evolution of the ψη-sequences from hypothetical ancestral sequences. To obtain the data needed to calculate the normalized link lengths, our TPA program for constructing MP ancestral sequences was run not only on the full alignment of 12,741 positions but also on subsets of the full alignment. Each subset spanned all sequenced positions for a particular species or group of species, and another subset spanned the remaining positions; for example, the gibbon sequence spans positions 1–10571, which, except for its unique insertions, it shares with most other simian sequences. The TPA results on this subset of the alignment placed 254 NSs on the terminal link to gibbon, 80 NSs on the link to the HCGOGb ancestral node, and 43 NSs on the link to the HCGO ancestral node. In turn, the TPA results on the remaining subset of the alignment (positions 10572–12741) placed 0, 0, and 25 NSs on the terminal Gb, ancestral HCGOGb, and ancestral HCGO links, respectively. Our first task was to distribute the 25 NSs that the TPA program arbitrarily placed on the ancestral HCGO link between this link and the ancestral HCGOGb link. From the TPA results on the subset represented by gibbon, we used the ratio of 80:43 to distribute 16 of the 25 NSs to the ancestral HCGOGb link (this gave it a total of 96 NSs) and nine NSs to the ancestral HCGO link (this gave it a total of 52 NSs). Since all catarrhine sequences except gibbon were represented by ~10,000 shared orthologous positions, the normalized link length values for ancestral HCGOGb and ancestral HCGO links became 0.96 and 0.52, respectively. In turn, since the gibbon sequence was represented by ~8,000 shared positions, the normalized link length value for the 254 NSs on the terminal Gb link became 3.18.

Table 1
Pairwise Divergence Values among ψη Sequences of 13 Species

Species Compared[a]	No. of Shared Positions	No. of Substitutions				Divergence[d] (%)		DNA-DNA Hybridization	
		Transversions	Transitions	Transitions/Transversions[b]	Gaps[c]	Uncorrected	Corrected[e]	$dT_{50}H$[f]	dT_m[g]
HsaA/HsaB . . .	5,290	7	9	1.29	4	0.38	0.38	. . .	. . .
HsaA/Ptr	10,159	44	100	2.27	17	1.58	1.60	1.63	1.59
HsaA/Ggo	10,176	35	114	3.26	22	1.68	1.70	2.27	2.50
HsaA/Ppy	10,098	90	201	2.23	45	3.31	3.38	3.60	3.49
HsaA/Hla	8,002	108	247	2.29	46	4.98	5.15	4.76	5.04
HsaA/Mmu . . .	9,962	206	457	2.22	67	7.28	7.66	7.34[h]	6.79[h]
HsaA/Age	9,396	300	601	2.00	103	10.57	11.39	11.23[i]	. . .
HsaA/Atr	2,098	66	155	2.35	17	11.26	12.20	. . .	. . .
HsaA/Tsy	1,890	151	293	1.94	47	25.35	30.93	. . .	. . .
HsaA/Gcr	2,512	288	392	1.36	74	29.16	36.92	. . .	. . .
HsaA/Lfu	770	55	111	2.02	15	23.06	27.55	. . .	. . .
HsaA/Chi	1,822	213	283	1.33	71	29.95	38.23	. . .	. . .
HsaB/Ptr	5,286	23	50	2.17	9	1.55	1.57	. . .	. . .
HsaB/Ggo	5,288	19	59	3.11	12	1.70	1.72	. . .	. . .
HsaB/Ppy	5,272	47	106	2.26	21	3.29	3.36	. . .	. . .
HsaB/Hla	3,154	42	109	2.60	13	5.18	5.37	. . .	. . .
HsaB/Mmu . . .	5,233	119	246	2.07	34	7.56	7.97	. . .	. . .
HsaB/Age	5,172	170	352	2.07	58	11.19	12.12	. . .	. . .
HsaB/Atr	2,075	65	154	2.37	19	11.37	12.33	. . .	. . .
HsaB/Tsy	1,868	149	290	1.95	49	25.46	31.10	. . .	. . .
HsaB/Gcr	1,264	130	220	1.69	40	29.91	38.16	. . .	. . .
HsaB/Lfu	770	55	111	2.02	15	23.06	27.55	. . .	. . .
HsaB/Chi	1,800	211	280	1.33	71	30.04	38.38	. . .	. . .
Ptr/Ggo	10,194	43	121	2.81	22	1.82	1.84	2.21	2.55
Ptr/Ppy	10,116	99	205	2.07	43	3.42	3.50	3.58	3.52
Ptr/Hla	8,019	127	255	2.01	47	5.32	5.52	4.85	4.66
Ptr/Mmu	9,980	216	466	2.16	65	7.44	7.83	7.29[h]	7.01[h]

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest

Ptr/Age	9,418	311	610	1.96	104	10.76	11.61	...	...
Ptr/Atr	2,097	69	161	2.33	20	11.81	12.85	...	...
Ptr/Tsy	1,889	152	298	1.96	50	25.79	31.60	...	...
Ptr/Gcr	2,513	289	392	1.36	76	29.24	37.06	...	...
Ptr/Lfu	770	58	109	1.88	15	23.18	27.72	...	...
Ptr/Chi	1,817	213	280	1.31	71	29.87	38.09	...	...
Ggo/Ppy	10,136	85	221	2.60	39	3.39	3.47	3.55	3.57
Ggo/Hla	8,039	113	256	2.23	46	5.13	5.31	4.69	5.15
Ggo/Mmu	10,003	202	465	2.30	65	7.27	7.65	7.18[h]	7.12[h]
Ggo/Age	9,437	297	613	2.06	97	10.56	11.38	...	...
Ggo/Atr	2,099	68	160	2.35	14	11.45	12.42	...	...
Ggo/Tsy	1,888	149	295	1.98	46	25.33	30.91	...	...
Ggo/Gcr	2,512	288	401	1.39	75	29.53	37.53	...	...
Ggo/Lfu	770	55	114	2.07	15	23.43	28.09	...	...
Ggo/Chi	1,821	211	286	1.36	70	29.98	38.28	...	...
Ppy/Hla	7,984	120	252	2.10	43	5.17	5.36	4.83	3.83
Ppy/Mmu	9,990	210	460	2.19	61	7.27	7.65	7.43[h]	7.33[h]
Ppy/Age	9,399	306	608	1.99	96	10.64	11.47	...	...
Ppy/Atr	2,071	71	152	2.14	16	11.45	12.42	...	...
Ppy/Tsy	1,859	149	293	1.97	48	25.69	31.45	...	...
Ppy/Gcr	2,511	292	401	1.37	78	29.78	37.95	...	...
Ppy/Lfu	767	64	112	1.75	14	24.33	29.41	...	...
Ppy/Chi	1,789	212	280	1.32	70	30.26	38.75	...	...
Hla/Mmu	7,868	183	404	2.21	55	8.10	8.57	7.05[h]	6.98[h]
Hla/Age	7,366	264	489	1.85	84	11.23	12.17	...	...
Hla/Atr	2,076	75	159	2.12	16	11.95	13.02	...	...
Hla/Tsy	1,864	152	299	1.97	48	26.10	32.08	...	...
Hla/Gcr	2,509	301	413	1.37	81	30.69	39.47	...	...
Hla/Lfu	769	58	114	1.97	15	24.16	29.16	...	...
Hla/Chi	1,790	211	298	1.41	69	31.09	40.15	...	...
Mmu/Age	9,453	336	704	2.10	90	11.84	12.89	...	...
Mmu/Atr	2,103	83	175	2.11	14	12.85	14.09	...	...
Mmu/Tsy	1,892	164	306	1.87	44	26.54	32.76	...	...
Mmu/Gcr	2,517	295	418	1.42	77	30.45	39.07	...	...
Mmu/Lfu	770	64	107	1.67	16	23.79	28.61	...	...
Mmu/Chi	1,818	231	296	1.28	69	31.58	40.99	...	...

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by

Table 1 (*Continued*)

Species Compared[a]	No. of Shared Positions	No. of Substitutions: Transversions	Transitions	Transitions/ Transversions[b]	Gaps[c]	Divergence[d] (%): Uncorrected	Corrected[e]	DNA-DNA Hybridization: $dT_{50}H$[f]	dT_m[g]
Age/Atr	2,118	37	72	1.95	9	5.55	5.77	...	...
Age/Tsy	1,902	156	307	1.95	49	26.24	32.29	...	...
Age/Gcr	2,490	304	402	1.97	77	30.50	39.15	...	...
Age/Lfu	765	66	109	1.65	18	24.65	29.89	...	...
Age/Chi	1,825	218	298	1.37	74	31.07	40.12	...	...
Atr/Tsy	1,905	154	312	2.03	45	26.21	32.25	...	...
Atr/Gcr	1,262	148	221	1.49	38	31.31	40.53	...	...
Atr/Lfu	766	66	109	1.65	19	24.71	29.98	...	...
Atr/Chi	1,832	223	304	1.36	73	31.50	40.85	...	...
Tsy/Gcr	1,191	139	192	1.38	38	30.02	38.34	...	...
Tsy/Lfu	685	62	114	1.84	22	28.01	35.07	...	...
Tsy/Chi	1,804	227	330	1.45	83	33.91	45.13	...	...
Gcr/Lfu	746	70	113	1.61	20	26.50	32.69	...	...
Gcr/Chi	1,176	189	222	1.17	52	37.70	52.39	...	...
Lfu/Chi	676	70	102	1.46	15	27.06	33.56	...	...

[a] Sequence comparisons were made of the entire data set of aligned sequences. HsaA = entire 10.8-kb human sequence; HsaB = second human sequence; Ptr = *Pan troglodytes;* Ggo = *Gorilla gorilla;* Ppy = *Pongo pygmaeus;* Hla = *Hylobates lar;* Mmu = *Macaca mulatta;* Age = *Ateles geoffroyi;* Atr = *Aotus trivirigatus;* Tsy = *Tarsius syrichta;* Gcr = *Galago crassicaudatus;* Lfu = *Lemur macaco* (*fulvus*) *mayattensis;* and Chi = *Capra hircus.*

[b] In all cases, transitions outnumbered transversions.

[c] Counted as one substitution, regardless of length.

[d] Calculated as the total number of differences, including non–common gaps, divided by the sum of the number of positions which have aligned nucleotides in both species being compared, and of the number of introduced gaps, each gap being counted as one position.

[e] To include superimposed mutations (Jukes and Cantor 1969).

[f] Divergence values from DNA hybridization data by using the HAP method (see Sibley and Ahlquist 1987, table 5).

[g] Divergence values from DNA hybridization data by using the TEACL method (see Caccone and Powell 1989, table 6).

[h] Combined values from several cercopithecine species, including *M. mulatta.*

[i] Source: Bonner et al. (1980).

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 b

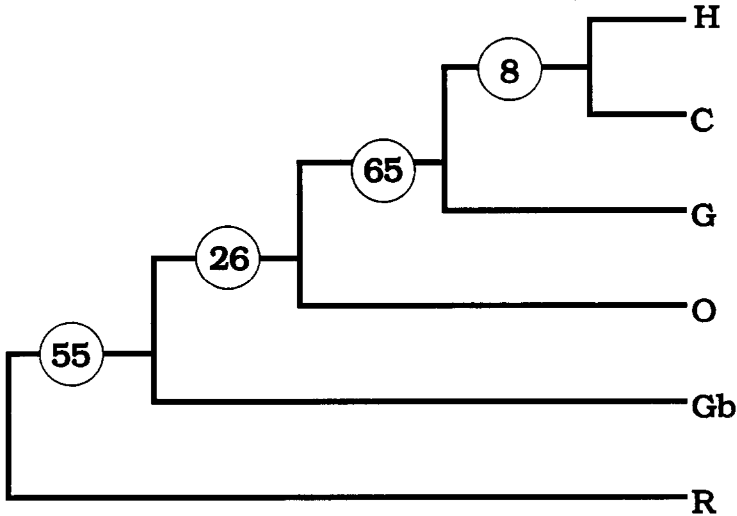


FIG. 3.—Strength of grouping results for hominoid lineages. Numbers within the circles on each stem to a hominoid clade represent the minimum number of NSs that must be added to the MP score to break up that clade. H = human; C = chimpanzee; G = gorilla; O = orangutan; Gb = gibbon; and R = rhesus.

ment were not included as supporting the *Homo-Pan* clade. However, inclusion of the gibbon sequence in our present analysis clearly shows T to be the ancestral hominoid nucleotide at position 1700 and C to be such at position 4999; thus the deletions of T and C are derived states shared by *Homo* and *Pan*. On the other hand, position 1751 from the present data set of 13 aligned sequences that previously supported a *Homo-Pan* clade (formerly position 1693 from fig. 3 of Goodman et al. 1989) no longer supports this clade; the G found in *Homo* and *Pan* is also found in *Hylobates* and thus could be the ancestral hominoid nucleotide, whereas the A found in both *Gorilla* and *Pongo* could have resulted from independent G→A substitutions. As in our previous study (Goodman et al. 1989), only three putative synapomorphies support a *Homo-Gorilla* clade: two transitions, one at position 35 (G→A) and one at position 7986 (T→C), and one transversion at position 11,816 (C→G) (table 2). Similarly, only four putative synapomorphies support a *Pan-Gorilla* clade; three were previously identified (Goodman et al. 1989) and consist of two transitions, one at position 6751 (A→G) and one at position 8746 (C→T), and one transversion at site 6748 (A→C). In addition, inclusion of the gibbon sequence in the present study implicates positions 1777–1782 as supporting a *Gorilla-Pan* clade, in that *Gorilla* and *Pan* share the same deletion (table 2). Since only three putative synapomorphies support a *Homo-Gorilla* clade and since only four support a *Gorilla-Pan* clade, compared with 12 that support the *Homo-Pan* clade, the former may be viewed as false synapomorphies (i.e., homoplasies) resulting from parallel mutations, whereas the latter (those supporting the *Homo-Pan* clade) are more likely to represent true synapomorphies.

No sequence character supports a great-ape clade; that is, no sequence character is uniquely shared by orangutan, chimpanzee, and gorilla. Nor are there sequence characters that would separate a human-orangutan clade from an African-ape clade. However, the human-chimpanzee-gorilla clade is supported by 74 synapomorphic characters (appendix A). The orangutan-human-chimpanzee-gorilla clade that excludes gibbon is supported by 41 synapomorphic characters over that large portion of the

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1031010 by guest on 21 August 2022

Table 2
Sequences Favoring Alternative Groupings

A. Sequence Characters Favoring Hsa-Ptr Grouping

	Position(s)										
Species	583	1700	4576–4579	4821	4978–5001[a]	5942–5951	6062	7075	8910	9159–9164	10947	12110
HsaA	C	*	****	T	************************	ATA**CACAC	C	G	G	******	G	G
HsaB			****	T	GTCC*ACTATGTTTGTACCTA*TG					******	G	G
Ptr ...	C	*	****	T	GTCC*ACTATGTTTGTACCTA*TG	ATA**TACAC	C	G	G	******	G	G
Ggo ..	A	T	TAAT	*	GTCC*ACTATGTTTGTACCTACTG	ATATATACAC	T	T	T	AATATA	A	A
Ppy ..	A	T	TAAT	*	************************	ATA*******	T	T	T	AATATA	A	A
Hla ..	A	T	TAAT	*	GTCC*ACTATGTTTGTACCTACTG	ATA******C	T	T	T	GATGTA		
Mmu	A	*	TAAT	*	************************	********AC	T	T	T	AATATA	A	A
Age ..	A	–	TAAT	*	GTCC*ACTGTGTTCGTACCTGCTG	CTGTAT****	T	T		ACTATA	A	A
Atr ...			TAGT	*	GTCC*ACTGTGTGTGTACCTACTG							
Tsy ..			****	*	GTCT*ATTATGTTGGTGCCTCCTG							
Gcr ..	*	–										
Chi ..			CCAT	*	GGCCTACCATGCTGGTGCCTA***							

B. Sequence Characters Favoring Hsa-Ggo Grouping

	Position		
Species	35	7986	11816
HsaA	A	C	G
HsaB			G
Ptr ...	G	T	C
Ggo ..	A	C	G
Ppy ..	G	T	C
Hla ..	G	T	
Mmu	G	T	C
Age ..	G	T	C
Gcr ..	G		

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest

C. Sequence Characters Favoring Ggo-Ptr Grouping

SPECIES	POSITIONS			
	1777–1795	6748	6751	8746
HsaA	AAAAAAAAAAGAAAGAAAG	A	A	C
HsaB				
Ptr . . .	******AAAAGAAAGAAAG	C	G	T
Ggo . .	******AAAAGAAAGAAAG	C	G	T
Ppy . .	*******************	A	A	C
Hla . .	AAAAAAAAAAAAAA***AG	A	A	C
Mmu	-------------------	G	A	C
Age . .	-------------------	A	A	C
Gcr . .	-------------------			

NOTE.—Species abbreviations correspond to those in table 1. Numbers of nucleotide positions correspond to those used for the full alignment of the 13 sequences. Gaps are coded with an asterisk (*); sections within more extensive gaps are coded with a dash (–); and positions not represented are left blank.

[a] A single-basepair deletion at position 4999 (shown underlined and in boldface) within position 4978–5001 is the shared character state supporting a human-chimpanzee grouping. The full region spanning positions 4978–5001 is shown because the 22- or 23-bp sequence that it encompasses has been independently deleted in rhesus monkey, orangutan, and human allele A. The inclusion of the gibbon sequence in the present study establishes from the parsimony analysis that presence of the 23-bp sequence is the likely ancestral hominoid condition, as well as the likely ancestral gorilla-chimpanzee-human state. Similarly, the presence of the 22-bp sequence is the likely ancestral chimpanzee-human state (i.e., on the stem to the human-chimpanzee ancestor, a single-basepair deletion at position 4999 changed the 23-bp sequence to a 22-bp one). This 23-bp sequence region spanning positions 4978–5001 is tandemly duplicated and is preceded 5′ by a polypyrimidine tract. As such tracts are thought to initiate genetic variation through slippage during replication (Tautz et al. 1986), the occurrence of several independent deletions of the same sequence region need not be viewed as highly improbable.

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 b

alignment (positions 1–10571) where the gibbon is represented (appendix B). In addition, over the 3′ region of the alignment, where orthologous gibbon sequence is missing, another 25 synapomorphic characters support grouping orangutan with the human-chimpanzee-gorilla clade. Over the ~8,000 orthologous nucleotide positions that the gibbon sequence shares with other simian sequences, there are 78 synapomorphic characters which specifically group gibbon with the great ape–human clade (appendix B) and thus support the monophyly of Hominoidea.

Divergence Dates Based on Clock Calculations

Different rates of evolution are evident in the phylogenetic tree shown in figure 2, where branch lengths represent the number of mutations fixed during the decent of the lineages. Because rates of fixation of NS vary considerably from one lineage to the next, a global molecular clock model can estimate inaccurate divergence dates. This is apparent either on using the percentage divergence values from pairwise comparisons of the ψη sequences (table 1) or on using the branch lengths from the MP tree (fig. 2). For an initial set of global clock dates, we assume that all mammalian lineages fix mutations in their "neutral" noncoding DNA at the rate of 5×10^{-9} substitutions/site/year, the rate proposed by Kimura (1983) and others (Hayashida and Miyata 1983; Li et al. Wu 1985) who calibrated their global clock by taking 80 Mya as the reference date for the divergence of rodents and primates. At the rate of 5×10^{-9}, it turns out that the percentages of sequence divergence in the pairwise comparisons are the same as the divergence times in Mya units. Thus, from the estimates in table 1 of pairwise divergence values corrected for superimposed substitutions, the dates for the *Pan-Homo,* Hominina-Gorillina, Pongini-Hominini, Hylobatinae-Homininae, Cercopithecoidea-Hominoidea, and Platyrrhini-Catarrhini divergences are 1.6, 1.8, 3.4, 5.3, 7.9, and 12.3 Mya, respectively (each date is obtained by equally weighting the pairwise divergence values representing the two taxa). These global clock dates are much more recent than the divergence times inferred from paleontological evidence (Gingerich 1984; Andrews 1985, 1986; Fleagle 1988, pp. 257–413); the latter, for example, places the Pongini-Hominini, cercopithecoid-hominoid, and platyrrhine-catarrhine divergence dates at about 15, 25, and 35 Mya, respectively. If we calibrate the global clock with a date acceptable to experts on the primate fossil record—e.g., a date such as a cercopithecoid-hominoid divergence time of 25 Mya—it is not surprising that, for branch points of lineages within Anthropoidea, we obtain dates that are not too far off from those inferred from fossil evidence—but then the clock dates for older branch points are far too ancient; for instance, the dates for the Tarsiiformes-Anthropoidea and Lorisiformes-Lemuriformes divergences are then 101 and 104 Mya, respectively, about twice as ancient as inferred from the fossil record (Gingerich 1984; Beard et al. 1988; Fleagle 1988; Martin 1988, 1990).

On using the branch lengths from the MP tree (fig. 2) to calculate divergence dates by the model of a global clock, we obtain results (table 3) similar to those obtained from the pairwise comparisons in table 1. In these global clock calculations that use branch lengths from figure 2, we equally weight the branches (as evident from values listed for the pairs of taxa in table 3) rather than the individual pairwise comparisons. When the reference date used to calibrate the clock is taken as 25 Mya for the cercopithecoid-hominoid branchpoint, the dates for branch points of lineages within Anthropoidea are not too far from dates inferred from fossil evidence (e.g., ~11 Mya for the Pongini-Hominini branch point, as compared with 15 Mya from fossil evidence), but again the dates for older branch points (such as the 88 Mya for Tarsiiformes-

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

Table 3
Global versus Local Clocks in Deriving Lineage Divergence Dates

Phylogenetic Branch Points	Branch Lengths[a]	Molecular Clock Dates (Mya)	
		Global[b]	Local[c]
Pan-Homo	0.92, 0.72	4.9, 2.5	7.2, 5.7
Hominina-Gorillina	0.98, 0.94	5.8, 2.9	8.4, 6.7
Pongini-Hominini	1.87, 1.73	10.9, 5.5	15.0, 12.0
Hylobatinae-Homininae	3.18, 2.32	16.6, 8.3	19.2, 15.3
Cercopithecoidea-Hominoidea	4.58, 3.71	25.0[d], 12.6	25.0[d], 19.9
Aotus-Ateles	3.25, 2.85	18.4, 9.3	15.9, 12.6
Platyrrhini-Catarrhini	6.79, 5.83	38.1, 19.2	34.2, 27.[illegible]
Tarsiiformes-Anthropoidea	17.01, 12.03	87.6, 44.1	63.0, 50.[illegible]
Lorisiformes-Lemuriformes	19.71, 10.43	90.9, 45.7	55.1, 43.[illegible]
Strepsirhini-Haplorhini	20.07, 16.17	109.3, 55.0[d]	69.1, 55.0[d]

[a] Derived from link lengths in fig. 2. The two lengths for each pair of taxa (the descendant lineages) from a phylogenetic branch point are listed in the same order as the two taxa. In deriving the length of the next higher (i.e., more inclusive) taxon, the lengths of the two member taxa of this more inclusive taxon are first averaged, and then this averaged length is added to the length of the link representing the stem of the two taxa; for example, to obtain the branch length of Hominina (the subtribe for *Pan* and *Homo*), the *Pan* and *Homo* lengths of 0.92 and 0.72 are averaged, and this averaged length (0.82) is added to 0.16 (the link length of the stem of *Pan* and *Homo*) to yield the Hominina branch length of 0.98.

[b] The distance between the two taxa is the sum of the two branch lengths, and this distance is treated as being directly proportional to time.

[c] Each date is derived from the length of the stem of the two descendant branches and the length of the longest of the two descendant branches (the stem length, the longest descendant branch length, and then the sum are all treated as being directly proportional to time; see text).

[d] Reference data based on fossil record.

Anthropoidea and the 91 Mya for Lorisiformes-Lemuriformes) are far too ancient. In turn, since fossils representing true primates first appear in the fossil record during Eocene times (36–55 Mya), the strepsirhine-haplorhine divergence node may be placed at the Paleocene/Eocene boundary, i.e., at ~55 Mya, a date which agrees with that used by Fleagle (1988, pp. 259, 284, and 319). Using this reference date to calibrate the global clock, we obtain relatively reasonable dates for the older branch points but dates for branch points within Anthropoidea that are too recent as judged by fossil evidence.

The global clock dates are listed in table 3, where they are compared with dates obtained, for the same branch points, by local clock calculations. Local molecular clocks focus on localized regions of the phylogenetic tree and are calibrated using one or more well-established paleontological time points within each of these regions (Goodman 1986; Koop et al. 1986; Li and Tanimura 1987). In the application of this local molecular clock to the Anthropoidea region of figure 2, the pattern of variation in branch lengths influenced how we calculated divergence dates. We observed that, from the Anthropoidea ancestral node to the present, branch lengths were longest in the lineages to New World monkeys (the platyrrhines or ceboids) and to the Old World (rhesus) monkey (a cercopithecoid), next longest in the lineage to gibbon (the hominoid lesser ape), and shortest in the lineages to the remaining hominoids. Thus we assumed that the rate at which mutations accumulated in the early catarrhines was closer to the rhesus monkey rate than to any hominoid rate, rather than assuming that the rate has recently picked up in the New World monkeys and Old World monkeys. Similarly, the early-hominoid mutation rate was closer to the gibbon rate

Downloaded from https://academic.oup.com/mbe/article/10/1/... by guest on 1 August 2022

than to the rates for the great apes and human lineages. Using this local molecular clock model and a reference date of 25 Mya for the cercopithecoid-hominoid divergence node (Gingerich 1984; Fleagle 1988, pp. 259 and 363–413), we dated the platyrrhine-catarrhine divergence node by the following calculations: The early-catarrhine link length of 1.69 was added to the rhesus link length of 4.58 (see fig. 2) to give a total branch length of 6.27. Then, with length 4.58 being equated to 25 Mya, length 6.27 equated to 34.2 Mya, the local molecular clock date for the platyrrhine-catarrhine divergence node. Similarly, for calculating the date of the divergence node at which gibbon separated from other hominoids, we added the early-hominoid link length of 0.96 to the gibbon link length of 3.18 to give a total branch length of 4.14; then, with length 4.14 being equated to 25 Mya, length 3.18 equated to 19.2 Mya, the date for the divergence of gibbon from other hominoids. Proceeding by this local clock model, we used 19.2 Mya as the reference time for calculating the date of the next divergence node (that of orangutan from the African ape–human branch) and added 0.52 (the length of the great ape–human stem) to 1.87 (the orangutan link length) to give a total length of 2.39; then with length 2.39 being equated to 19.2 Mya, length 1.87 equated to 15.0 Mya, the date for the divergence of orangutan from the African ape–human branch. Continuing these local molecular clock calculations by focusing at each step only on the next stem and its longest descendant lineage, we dated first the divergence of gorilla from the human-chimpanzee branch and then the divergence of human and chimpanzee branches as being 8.4 and 7.2 Mya, respectively. This same procedure was applied to date the remaining primate branch points. Thus, to date the *Aotus-Ateles* branch point, the platyrrhine-catarrhine divergence time of 34.2 Mya was taken as the reference date and the rate of the *Aotus-Ateles* stem was determined by the longest platyrrhine branch; that is, this stem link length was added to the *Aotus* link length to represent the 34.2 Mya and to yield by proportionality a date of 15.9 Mya for the *Aotus-Ateles* branch point. Similarly, the platyrrhine-catarrhine divergence date of 34.2 Mya, the Anthropoidea stem, and the platyrrhine branch (the longest Anthropoidea branch) dictated that the time of the tarsier-Anthropoidea divergence was 63.0 Mya. In turn, this date, the haplorhine stem, and the tarsier branch dictated that the time of the strepsirhine-haplorhine divergence was 69.1 Mya. Finally, this latter date, the strepsirhine stem, and the galago branch dictated that the time for the lorisiform-lemuriform divergence was 55.1 Mya. These local clock dates for the older primate branch points, especially the lorisiform-lemuriform date, are more ancient than the divergence dates suggested by the available fossil record on these taxa but are not as grossly overestimated as they are in the global clock calculations using a reference date of 25 Mya for the cercopithecoid-hominoid divergence node (table 3).

We carried out a second set of local clock calculations using a reference date of 55 Mya for the strepsirhine-haplorhine divergence node. Again, the model was followed that the longest descending branch of the two branches from a branch point dictated the date for this branch point. The local calculations then placed the lorisiform-lemuriform divergence node at 43.9 Mya (table 3), which is near the range of dates (30–40 Mya) suggested by recent assessments of the fossil evidence on the strepsirhines (Gingerich 1984; Beard et al. 1988; Martin 1988, 1990). Similarly, the local clock date for the tarsier-simian divergence node was then 50.1 Mya (table 3), which is consistent with prevailing views on the haplorhine fossil record (Fleagle 1988, p. 319; Martin 1990). The local clock dates for branch points within Anthropoidea tend to be slightly more recent than those suggested by the simian fossil record but are not as

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1034108 by guest on 21 August 2022

grossly underestimated as are the global clock calculations using as a reference date 55 Mya for the strepsirhine-haplorhine divergence node (table 3).

In summary, with the cercopithecoid-hominoid divergence node at 25 Mya used as the reference for dating branch points in the Anthropoidea region of the MP tree and then with the strepsirhine-haplorhine divergence node at 55 Mya used as the reference for dating the two remaining primate branch points (lorisiform-lemuriform and tarsier-simian), we obtained local clock dates that are in close agreement with paleontological dates. In contrast, global clock calculations either grossly underestimated simian divergence times or grossly overestimated the remaining primate divergence times, as compared with fossil evidence (Gingerich 1984; Andrews 1985, 1986; Fleagle 1988).

Rates of Molecular Change

It is apparent from the lengths of the branches of the MP tree that different primate lineages accumulated mutations at markedly differing rates (fig. 2). The most striking difference in rates is between galago and lemur. From the strepsirhine ancestral node to the present, the lineage to galago fixed mutations two times faster than did the lineage to lemur. Marked differences in rates also occur among haplorhine lineages. For example, from the haplorhine ancestral node to the present, the lineage to tarsier accumulated mutations at a rate that, on average, was one and a half times faster than the lineage which passed through simian, catarrhine, and hominoid ancestral nodes in descent to humans.

To convert the branch length data of figure 2 into rates at which mutations were fixed in primate lineages during different periods of descent (table 4), we used the reference and derived dates from the local clock calculations (table 3) that agree with paleontological assessments of the primate fossil record. In turning to these paleontological assessments, we have assumed (after the model of punctuated equilibrium)

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

Table 4
Rates of Noncoding DNA Evolution in Region of Primate Genomes Spanning ψη-Globin Locus

Primate Lineage To	Age (Mya)	Rate (substitutions × 10^{-9}/site/year)
Strepsirhine ancestor	55–44	4.5
Galago	44–0	4.5
Lemur	44–0	2.4
Tarsier	55–0	3.4
Anthropoidea ancestor	55–34	3.5
Owl monkey	34–0	2.1
Spider monkey	34–0	1.9
Catarrhini ancestor	34–25	1.9
Rhesus monkey	25–0	1.8
Hominidae ancestor	25–19	1.7
Gibbon	19–0	1.7
Orangutan	19–0	1.2
Gorilla	19–0	1.2
Chimpanzee	19–0	1.2
Human	19–0	1.1

NOTE.—Calculations are based on branch length data from fig. 2 (see text).

that the first fossils indicative of an adaptive radiation serve to designate when the last common ancestor of the diverging lineages existed. Thus we placed the strepsirhine-haplorhine divergence node just before the time that tarsier-like and lemur-like fossils appear in the fossil record. If we had used the model of phyletic gradualism, which accounts for the absence of transitional forms in the fossil record by the gross incompleteness of this record rather than by rapid emergence of new differentiated forms, we could have followed Martin (1990) rather than Fleagle (1988, pp. 259, 284, and 319) and would have chosen a date of 90–100 Mya for the strepsirhine-haplorhine divergence node, but then our molecular clock calculations would have produced, for the tarsier-simian and lemuriform-lorisiform divergence nodes, dates much more ancient than those advocated by Fleagle (1988), Martin (1988, 1990), and other students of the primate fossil record (Gingerich 1984; Beard et al. 1988).

Having assigned dates to five divergence nodes spanning the time from the earliest true primates (55 Mya) to the last common ancestor of living hominoids (19 Mya), we proceeded to calculate from the branch length data in figure 2 the rates at which noncoding DNA evolved in the 11 primate lineages during earlier and later periods of descent. The highest rates are in lineages descending from the primate ancestor to galago (4.5×10^{-9} substitutions/site/year), to tarsier (3.4×10^{-9}), and to the Anthropoidea ancestor (3.5×10^{-9}). In contrast, the great apes have accumulated changes at the much lower rate of 1.2×10^{-9}, almost one fourth that of galago. Intermediate rates—2.4×10^{-9} to 1.7×10^{-9} substitutions/site/year—are seen in lemur, spider monkey, rhesus monkey, and gibbon. Thus, adding the gibbon sequence to the ψη data set has corroborated and extended previous observations on a graded slowdown of rates in hominoid phylogeny (Koop et al. 1986; Miyamoto et al. 1987; Fitch et al. 1988; Goodman et al. 1989). This slowdown is also evident for other noncoding sequences from the primate β-globin gene cluster (Koop et al. 1989*b*). It has also been shown to occur in other genes (Goodman 1985; Wu and Li 1985; Britten 1986; Li and Tanimura 1987; Li et al. 1987) and throughout whole genomes, as judged by DNA-DNA hybridization data (Britten 1986).

Discussion

Pseudogenes, since they no longer contribute to the phenotype, are released from purifying or stabilizing selection which eliminates variation and from positive directional selection which is reflected by adaptive variation. Therefore, pseudogenes and other noncoding DNA lacking regulatory or structural functions should accumulate mutations at rates approximating the rates of occurrence of spontaneous mutations (Kimura 1983). In resolving cladistic relationships over short phylogenetic distances, such noncoding DNA offers advantages over coding DNA. The noncoding sequences, compared with the coding sequences, accumulate nucleotide changes at a more rapid rate (Gojobori et al. 1982; Li et al. 1984). Also, they accumulate changes independently of natural selection; therefore, shared changes more frequently reflect changes shared through history, rather than changes through parallel function. In the present study, noncoding DNA that spans the ψη-globin locus was chosen to answer phylogenetic questions regarding placement of the common gibbon among the simian primates, as well as other evolutionary questions regarding molecular clocks and rates of DNA evolution.

A variety of morphological characters place Hylobatinae, the subfamily of gibbons, within the superfamily Hominoidea but separate the lineage to gibbons from the ancestral lineage to the common ancestor of great apes and humans (Schwartz 1986;

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1131108 by guest on 21 August 2022

also see references therein). The molecular studies reviewed in the Introduction (see above) support this sister-group relationship of gibbons to great apes and humans. In addition, the results of our present study on the ψη-globin genomic region provide very strong evidence on the sister-group relationships not only of gibbon but also of orangutans, gorillas, chimpanzees, and humans. Our parsimony results showed that the lowest-scoring tree without a monophyletic Hominoidea joined gibbon to rhesus monkey rather than to other hominoids, but this added 55 extra fixed mutations to the most parsimonious score. Thus, clearly, the gibbon shares more genetic ancestry with the hominoids than with Old World monkeys. These results also showed gibbons to be the sister group of great apes and humans and, in turn, showed orangutan to be the sister group of gorillas, chimpanzees, and humans. This last derived clade, that of African great apes and humans, was the strongest supported clade, in that to break it up required 65 mutations more than were in the parsimony score (fig. 3).

A traditional taxonomic scheme (Simpson 1963), based on a combination of grades and clades, grouped the living African great apes (common chimpanzee, pygmy chimpanzee, and gorilla) into the same genus *Pan,* grouped *Pan* with *Pongo* (the Asian great ape orangutan) into subfamily Ponginae, and grouped Ponginae with Hylobatinae (the subfamily of Asian lesser apes or gibbons) into the ape family Pongidae while placing humans into the separate family Hominidae. The grouping of all the living apes into one family no doubt reflected Simpson's (1945) judgment that, from the perspective of all of mammalian systematics, the degrees of morphological divergence among the species, genera, and subfamilies within his family Pongidae were no greater than those seen within any average mammalian family. In contrast, Simpson (1961, 1963) subsequently argued that we humans have diverged radically in morphology and behavior from apes and thus, on a grade basis, should be placed in our own separate family. At the genetic level of sequenced proteins and sequenced DNA, however, the human lineage appears to be just as conservative or even more conservative than each of the other hominoid lineages. For example, with regard to the ψη DNA sequences, the percentage divergence of human from gibbon (5.3%) is no more than that of each great ape from gibbon (5.3%–5.6%) (table 1). A strictly cladistic classification that reflected recency of common ancestry and the actual degrees of genetic relationship would group not only all apes together but also all apes and humans together in one family. Thus, we have proposed (Goodman 1989; Goodman et al. 1990) the following cladistic classification of extant hominoid genera:

Superfamily Hominoidea
- Family Hominidae
  - Subfamily Hylobatinae
    - *Hylobates:* common and siamang gibbons
  - Subfamily Homininae
    - Tribe Pongini
      - *Pongo:* orangutans
    - Tribe Hominini
      - Subtribe Gorillina
        - *Gorilla:* gorillas
      - Subtribe Hominina
        - *Pan:* common and pygmy chimpanzees
        - *Homo:* human

By adhering to a strictly cladistic classification, this use of taxonomic nomenclature

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

(i.e., subfamily, tribe, and subtribe) most accurately describes the phylogenetic branching arrangements shown in figures 2 and 3. The strength of the grouping results given in figure 3, however, show that the weakest supported part of this hominoid classification is the subdivision of tribe Hominini into subtribes Gorillina for gorillas and Hominina for chimpanzees and humans.

The human-chimpanzee clade depicted by the noncoding sequences spanning the $\psi\eta$-globin locus is corroborated in several other recent DNA sequence studies, studies involving an immunoglobulin pseudogene (Ueda et al. 1989), 28S ribosomal genes (Gonzalez et al. 1990), and the cytochrome oxidase II locus of the mitochondrial genome (Ruvolo et al., accepted). The human-chimpanzee clade is also corroborated by DNA-DNA hybridization studies (Sibley and Ahlquist 1984, 1987; Caccone and Powell 1989; Sibley et al. 1990). However, a recent DNA sequence study of the involucrin locus in hominoids (Djian and Green 1990) does not corroborate the sister grouping of humans and chimpanzees. With orangutan as the closest and gibbon as a more distant outgroup of the gorilla-chimpanzee-human clade, they concluded, on the basis of shared repeats and "marker nucleotides," found within a repetitive coding segment of the involucrin gene, that gorilla and chimpanzee are more closely related than either is to human. Each of these DNA sequence studies could have found a correct allele phylogeny, but, clearly, at least one or the other of these sequence studies has not revealed the correct species phylogeny. Lack of congruence between the branching pattern of the alleles and the branching pattern of the species could result if polymorphic alleles had existed in the stem-species that was the last common ancestor of the descendant species-lineages. Suppose, for example, that allele A was fixed in species-lineages 1 and 2 while allele B was fixed in species-lineage 3 but that the speciation events first separated lineage 1 from the stem of lineages 2 and 3 and then separated lineages 2 and 3. If so, the branching pattern of the alleles would not be congruent with the branching pattern of the species. Such lack of congruence has probably occurred, although perhaps infrequently, in the case of the gene branching patterns and species branching patterns of the gorilla-chimpanzee-human clade. It seems more plausible to us that it happened for the involucrin gene than for two pseudogenes (the $\psi\eta$-globin and the immunoglobulin-epsilon gene), for the 28S ribosomal genes, for the mitochondrial cytochrome oxidase II locus and, as indicated by the DNA-DNA hybridization results, for the majority of orthologous DNA segments in the genomes representing the gorilla-chimpanzee-human clade. The degree of error that polymorphisms introduce into phylogenetic analysis of species relationships can more accurately be assessed when additional nucleotide sequences are available from different individuals of the same species and from unlinked regions of the genome.

Although the inclusion of several human alleles in our $\psi\eta$-globin study does not conclusively eliminate the possibility that branching error in the species phylogeny may be due to intraspecific variation (Miyamoto et al. 1987), inclusion of the gibbon and galago sequences in the phylogenetic analysis corroborates our previous findings that chimpanzee and human form a monophyletic clade that narrowly excludes gorilla (Miyamoto et al. 1987; Goodman et al. 1989; Williams and Goodman 1989). This conclusion is supported by 12 synapomorphic character traits, as opposed to three characters supporting a human-gorilla grouping and four characters shared by chimpanzee and gorilla.

Our study also suggests that the rate of accumulation of nucleotide changes in noncoding DNA was relatively high in early primate history but has since decreased in the descent of hominoids. Moreover, the pattern of evolutionary rates among pri-

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

mates is graded, in that a gradual decrease is observed from galago (4.5×10^{-9}) to tarsier (3.4×10^{-9}) to spider monkey (1.9×10^{-9}), to gibbon (1.7×10^{-9}) and finally to great apes (1.2×10^{-9}) and humans (1.1×10^{-9}), which show the slowest rate of all. This slowdown may be attributed to increases in the generation times of higher primates (i.e., to a decrease in the number of germ-line replications per unit of time) and perhaps to mechanisms that have increased DNA repair fidelity, which would decrease the rate of occurrence of spontaneous mutations (Goodman 1985; Wu and Li 1985; Britten 1986).

This finding of a gradual rate slowdown in the Anthropoidea contradicts the premise of a global molecular clock (Zuckerkandl and Pauling 1965). A more accurate measurement of divergence may be obtained by using a local molecular clock calibrated by well-established branching dates for a specific group of related organisms. Local clock calculations in this gibbon study used 25 Mya as the branching point of Old World monkeys and hominoids and for branch points among hominoids obtained clock dates (table 3) that are consistent with the hominoid fossil record.

Sequence Availability

The sequences for the gibbon and galago have been deposited in GenBank under accession numbers M54985 and M54984, respectively.

Acknowledgments

We thank Ben Koop for advice on both laboratory techniques and computer algorithms used in this work, and we thank Nancy Moncrief, Doug Wisniewski, Michael Stanhope, and Asha Kamat for helpful suggestions and stimulating discussions. We also thank Walter Fitch and two anonymous reviewers for their useful comments and critical evaluation of the original version of this paper.

APPENDIX A

Sequence Positions Supporting African Ape–Human Clade

Transition	Transversion	Gap
548	1622	292–293 (deletion)
779	3266	1729–1730 (insertion)
841	3833	4962 (deletion)
1032	4923	5544 (insertion)
1483	5298	6675–6685 (insertion)
1622	5372	6686–6692 (deletion)
1624	6022	8714 (deletion)
1634	6095	8919–8923 (insertion)
1852	7128	12663–12667 (deletion)
2181	7287	
2286	7537	
2387	8644	
2483	8747	
3243	9419	
3362	9807	
3804	10072	
3817	12248	
3883	12258	

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

APPENDIX A (*Continued*)

Transition	Transversion	Gap
3929	12277	
4003	12309	
4089	12600	
4566		
4674		
4759		
5508		
6101		
6106		
6740		
8625		
9112		
9346		
9436		
9447		
9707		
10030		
10923		
10957		
11095		
11503		
11566		
11630		
12027		
12087		
12349		

APPENDIX B
Sequence Positions Supporting Human-Ape Clades

A. Hsa/Ptr/Ggo/Ppy

Transition	Transversion	Gap
3467	118	6033–6050 (deletion)
3502	4044	6278 (deletion)
3544	5032	6305–6308 (insertion)
3815	5277	9136 (deletion)
3956	5558	**10634** (deletion)
4586	6563	**10772** (insertion)
4770	6699	**11610–11618** (deletion)
5245	6774	**12328–12329** (deletion)
5390	7090	
5617	7587	
6194	9310	
6277	9316	
7270	9458	
7339	**11109**	
7682	**11528**	

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

A. Hsa/Ptr/Ggo/Ppy

Transition	Transversion	Gap
8048	**12149**	
8565	**12261**	
8976		
9298		
9351		
9370		
9577		
9578		
9719		
10649		
10743		
10866		
10919		
10998		
11008		
11024		
11036		
11053		
11404		
11413		
11560		
11671		
12078		
12114		
12264		
12641		

B. Hsa/Ptr/Ggo/Ppy/Hla

Transition	Transversion	Gap
46		2223 (deletion)
461		2324–2362 (insertion)
518		3587–3590 (deletion)
941		4692–4719 (deletion)
990	103	5997 (deletion)
1961	109	7175–7179 (deletion)
2059	154	7632–7634 (deletion)
2125	419	7790 (deletion)
2566	506	8605–8617 (insertion)
3185	861	8702–8712 (insertion)
3392	870	9431 (insertion)
3728	2009	
3889	2050	
4749	2064	
4942	3612	

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

APPENDIX B (*Continued*)

B. Hsa/Ptr/Ggo/Ppy/Hla

Transition	Transversion	Gap
5177	3803	
5411	5055	
5424	5085	
5834	5272	
5956	5360	
6151	6004	
6438	7123	
7027	7207	
7042	7781	
7060	8091	
7298	8641	
7373	9125	
7498	9139	
7647	9432	
8540	9433	
9158	9533	
9195	9651	
9435	10128	
9514		
9559		
9652		
9769		
9814		

NOTE.—Base positions shown underlined and in boldface occur in the 3′ portion of the nucleotide alignment, where gibbon sequence is not present.

LITERATURE CITED

ANDREWS, P. 1985. Improved timing of hominoid evolution with a DNA clock. Nature **314**: 498–499.

———. 1986. Fossil evidence on human origins and dispersal. Cold Spring Harbor Symp. Quant. Biol. **51**:419–428.

BEARD, K. C., M. DAGOSTO, D. L. GEBO, and M. GODINOT. 1988. Interrelationships among primate higher taxa. Nature **331**:712–714.

BENVENISTE, R. E., and G. F. TODARO. 1976. Evolution of type c viral genes: evidence for an Asian origin of man. Nature **261**:101–108.

BONNER, T. I., R. HEINEMANN, and G. J. TODARO. 1980. Evolution of DNA sequences has been retarded in Malagasy primates. Nature **286**:420–423.

BRITTEN, R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science **231**:1393–1398.

BROWN, W. M., E. M. PRAGER, A. WANG, and A. C. WILSON. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. **18**:225–239.

CACCONE, A., and J. R. POWELL. 1989. DNA divergence among hominoids. Evolution **43**:925–942.

CHANG, L. Y. E., and J. L. SLIGHTOM. 1984. Isolation and nucleotide sequence analysis of the β-type globin pseudogene from human, gorilla, and chimpanzee. J. Mol. Biol. **180**:767–784.

CHIARELLI, B. 1966. Caryology and taxonomy of the catarrhine monkeys. Am. J. Phys. Anthropol. **24**:155–170.

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134148 by guest on 21 August 2022

CHU, E. H. Y., and M. A. BENDER. 1962. Cytogenetics and evolution of primates. Ann. NY Acad. Sci. **102**:253–266.

COLLINS, F. S., and S. WEISSMAN. 1984. The molecular genetics of human hemoglobin. Prog. Nucleic Acid Res. Mol. Biol. **31**:315–462.

CZELUSNIAK, J., M. GOODMAN, B. F. KOOP, D. A. TAGLE, J. SHOSHANI, G. BRAUNITZER, T. K. KLEINSCHMIDT, W. W. DE JONG, and G. MATSUDA. 1990. Perspectives from amino acid and nucleotide sequences on cladistic relationships among higher taxa of eutheria. Pp. 545–572 *in* H. H. GENOWAYS, ed. Current mammalogy. Vol. **2**. Plenum, New York.

DENE, H., M. GOODMAN, W. PRYCHODKO, 1976. Immunodiffusion evidence on the anthropology of the primates. Pp. 171–196 *in* M. GOODMAN and R. E. TASHIAN, eds. Molecular anthropology. Plenum, New York.

DJIAN, P., and H. GREEN. 1990. The involucrin gene of the gibbon: the middle region shared by the hominoids. Mol. Biol. Evol. **7**:220–227.

DUNN, I. S., and F. R. BLATTNER. 1987. Charons 36 to 40: multienzyme, high capacity, recombination deficient replacement vectors with polylinkers and polystuffers. Nucleic Acids Res. **15**:2677–2698.

FIEDLER, W. 1956. Übersicht über das system der primaten. Pp. 1–126 *in* H. HOFER, A. H. SCHULTZ, and D. STARK, eds. Primatologia. I. Systematik phylogenie, ontogenie. Karger, Basel.

FITCH, D. H. A., C. MAINONE, J. L. SLIGHTOM, and M. GOODMAN. 1988. The spider monkey $\psi\eta$-globin gene and surrounding sequences: recent or ancient insertions of LINEs and SINEs? Genomics **3**:237–255.

FITCH, D. H. A., C. MAINONE, M. GOODMAN, and J. L. SLIGHTOM. 1990. Molecular history of gene conversions in the primate fetal γ-globin genes. J. Biol. Chem. **265**:781–793.

FLEAGLE, J. G. 1988. Primate adaptation and evolution. Academic Press, San Diego.

FRITSCH, E. F., R. M. LAWN, and T. MANIATIS. 1980. Molecular cloning and characterization of the human β-like globin gene cluster. Cell **19**:959–972.

GINGERICH, P. D. 1984. Primate evolution: evidence from the fossil record, comparative morphology, and molecular biology. Yearbook Phys. Anthropol. **27**:57–72.

GOJOBORI, T., W. H. LI, and D. GRAUR. 1982. Patterns of nucleotide substitutions in pseudogenes and functional genes. J. Mol. Evol. **18**:360–369.

GONZALEZ, I. L., J. E. SYLVESTER, T. F. SMITH, D. STAMBOLIAN, and R. D. SCHMICKEL. 1990. Ribosomal RNA gene sequences and hominoid phylogeny. Mol. Biol. Evol. **7**:203–219.

GOODMAN, M. 1963. Man's place in the phylogeny of the primates as reflected in serum proteins. Pp. 204–234 *in* L. S. WASHBURN, ed. Classification and human evolution. Aldine, Chicago.

———. 1981. Decoding the pattern of protein evolution. Prog. Biophys. Mol. Biol. **38**:105–164.

———. 1985. Rates of molecular evolution: the hominoid slowdown. BioEssays **3**:9–14.

———. 1986. Molecular evidence on the ape subfamily Homininae. Pp. 121–132 *in* H. GERSHOWITZ, D. L. RUCKNAGEL, and R. E. TASHIAN, eds. Evolutionary perspectives and the new genetics. A. R. Liss, New York.

———. 1989. Update to "Evolution of the immunologic species specificity of human serum proteins". Hum. Biol. **61**:925–934.

GOODMAN, M., J. CZELUSNIAK, B. F. KOOP, D. A. TAGLE, and J. L. SLIGHTOM. 1987. Globins: a case study in molecular phylogeny. Cold Spring Harbor Symp. Quant. Biol. **52**:875–900.

GOODMAN, M., J. CZELUSNIAK, G. W. MOORE, A. J. ROMERO-HERRARA, and G. MATSUDA. 1979. Fitting the gene lineage into its species lineages: a parsimony strategy illustrated by cladograms constructed from globin sequences. Syst. Zool. **28**:132–163.

GOODMAN, M., B. F. KOOP, J. CZELUSNIAK, D. H. A. FITCH, D. A. TAGLE, and J. L. SLIGHTOM. 1989. Molecular phylogeny of the family of apes and humans. Genome **31**:316–335.

GOODMAN, M., B. F. KOOP, J. CZELUSNIAK, M. L. WEISS, and J. L. SLIGHTOM. 1984. The η-globin gene: its long evolutionary history in the β-globin gene family of mammals. J. Mol. Biol. **180**:803–823.

GOODMAN, M., and G. W. MOORE. 1971. Immunodiffusion systematics of the primates. I. The catarrhini. Syst. Zool. **20**:19–62.

Downloaded from https://academic.oup.com/mbe/article/8/2/155/130... by guest on 21 August 2022

Goodman, M., G. W. Moore, and J. Barnabas. 1974. The phylogeny of human globin genes investigated by the maximum parsimony method. J. Mol. Evol. **3**:1–48.

Goodman, M., D. A. Tagle, D. H. A. Fitch, W. Bailey, J. Czelusniak, B. F. Koop, P. Benson, and J. L. Slightom. 1990. Primate evolution at the DNA level and a classification of hominoids. J. Mol. Evol. **30**:260–266.

Hafleigh, A. S., and C. A. Williams, Jr. 1966. Antigenic correspondence of serum albumins among primates. Science **151**:1530–1535.

Hamerton, J. L. 1963. Primate chromosomes. Symp. Zool. Soc. (Lond.) **10**:211–219.

Hamerton, J. L., H. P. Klinger, D. E. Mutton, and E. M. Lang. 1963. The somatic chromosomes of hominoidea. Cytogenetics **2**:250–263.

Harris, S., P. A. Barrie, M. L. Weiss, and A. J. Jeffreys. 1984. The primate ψβ1 gene: an ancient β-globin pseudogene. J. Mol. Biol. **180**:785–801.

Hayasaka, K., T. Gojobori, and S. Horai. 1988. Molecular phylogeny and evolution of primate mitochondrial DNA. Mol. Biol. Evol. **5**:626–644.

Hayashida, H., and T. Miyata. 1983. Unusual evolutionary conservation and frequent DNA segment exchange in class I genes of the major histocompatibility complex. Proc. Natl. Acad. Sci. USA **80**:2671–2675.

Hoyer, B. H., N. W. van de Velde, M. Goodman, and R. B. Roberts. 1972. Examination of hominoid phylogeny by DNA sequence homology. J. Hum. Evol. **1**:645–649.

Jagadeeswaran, P., J. Pan, B. Forget, and S. M. Weissman. 1982. Sequences of non-alpha-globin genes in man. Cold Spring Harbor Symp. Quant. Biol. **47**:1081–1082.

Jeffreys, A. J., P. A. Barrie, S. Harris, D. H. Fawcett, Z. J. Nugent, and A. C. Boyd. 1982. Isolation and sequence analysis of a hybrid δ-globin pseudogene from the brown lemur. J. Mol. Biol. **156**:487–503.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 21–123 *in* H. N. Munro, ed. Mammalian protein metabolism. Academic Press, New York.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge.

Klinger, H. P. 1963. The somatic chromosomes of some primates: *Tupaia glis, Nycticebus caucang, Tarsius bancanus, Cercocebus aterimus, Symphalangus syndactylus.* Cytogenetics **2**:140–151.

Klinger, H. P., J. L. Hamerton, D. Mutton, and E. M. Lang. 1963. The chromosomes of the hominoidea. Pp. 235–242 *in* L. S. Washburn, ed. Classification and human evolution. Aldine, Chicago.

Kluge, A. G. 1983. Cladistics and the classification of the great apes. Pp. 151–177 *in* R. L. Ciochon and R. S. Corruccini, eds. New interpretations of ape and human ancestry. Plenum, New York.

Kohne, D. E. 1975. DNA evolution data and its relevance to mammalian phylogeny. Pp. 249–261 *in* W. P. Luckett and F. S. Szalay, eds. Phylogeny of the primates. Plenum, New York.

Koop, B. F., and M. Goodman. 1988. Evolutionary and developmental aspects of two β-hemoglobin genes (ε^M and β^M) of opossum. Proc. Natl. Acad. Sci. USA **85**:3893–3897.

Koop, B. F., M. Goodman, P. Xu, K. Chan, and J. L. Slightom. 1986. Primate η-globin DNA sequences and man's place among the great apes. Nature **319**:234–238.

Koop, B. F., D. Siemìeniak, J. L. Slightom, M. Goodman, J. Dunbar, P. C. Wright, and E. L. Simons. 1989*a*. Tarsius δ and β globin genes: conversions, evolution, and systematic implications. J. Biol. Chem. **264**:68–79.

Koop, B. F., D. A. Tagle, M. Goodman, and J. L. Slightom. 1989*b*. A molecular view of primate phylogeny and important systematic and evolutionary questions. Mol. Biol. Evol. **6**:580–612.

Li, W. H., C. C. Luo, and C. I. Wu. 1985. Evolution of DNA sequences. Pp. 1–84 *in* R. J. MacIntyre, ed. Molecular evolutionary genetics. Plenum, New York.

Li, W. H., and M. Tanimura. 1987. The molecular clock runs more slowly in man than in apes and monkeys. Nature **326**:93–96.

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022

LI, W. H., M. TANIMURA, and P. M. SHARP. 1987. An evaluation of the molecular clock hypothesis using mammalian DNA sequences. J. Mol. Evol. **25**:330–342.

LI, W. H., C. I. WU, and C. C. LUO. 1984. Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J. Mol. Evol. **21**: 58–71.

MAEDA, N., C.-I. WU, J. BLISKA, and J. RENEKE. 1988. Molecular evolution of intergenic DNA in higher primates: pattern of DNA changes, molecular clock, and evolution of repetitive sequences. Mol. Biol. Evol. **5**:1–20.

MARTIN, R. D. 1988. Several steps forward for Eocene primates. Nature **331**:660–661.

———. 1990. Primate origins and evolution: a phylogenetic reconstruction. Chapman & Hall, London.

MAXAM, A. M., and W. GILBERT. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **68**:499–560.

MIYAMOTO, M. M., B. F. KOOP, J. L. SLIGHTOM, and M. GOODMAN. 1988. Molecular systematics of higher primates: genealogical relations and classification. Proc. Natl. Acad. Sci. USA **85**:7627–7631.

MIYAMOTO, M. M., J. L. SLIGHTOM, and M. GOODMAN. 1987. Phylogenetic relations of humans and African apes from DNA sequences in the ψη-globin region. Science **238**:369–373.

MOORE, G. W., J. BARNABAS, and M. GOODMAN. 1973. A method for constructing maximum parsimony ancestral amino acid sequences on a given network. J. Theor. Biol. **38**:459–485.

MURRAY, N. E., W. J. BRAMMAR, and K. MURRAY. 1977. Lambdoid phages that simplify the recovery of in vitro recombinants. Mol. Gen. Genet. **150**:53–61.

PILBEAM, D. 1984. The descent of hominoids and hominids. Sci. Am. **250**:84–96.

ROSENBERGER, A. L. 1984. Fossil new world monkeys dispute the molecular clock. J. Hum. Evol. **13**:737–742.

RUVOLO, M. T., T. DISOTELL, M. W. ALLARD, W. M. BROWN, and R. C. HONEYCUTT. Resolution of the African hominoid trichotomy using a mitochondrial gene sequence. Proc. Natl. Acad. Sci. USA (accepted).

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SANGER, F., S. NICHLEN, and A. R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5468.

SARICH, V. W., and J. E. CRONIN. 1976. Molecular systematics of the primates. Pp. 141–170 *in* M. GOODMAN and R. E. TASHIAN, eds. Molecular anthropology. Plenum, New York.

SARICH, V. M., and A. C. WILSON. 1967. Immunological time scale for hominoid evolution. Science **158**:1200–1203.

SCHWARTZ, J. H. 1984. The evolutionary relationships of man and orangutans. Nature **308**: 501–505.

———. 1986. Primate systematics and a classification of the order. Pp. 1–41 *in* D. R. SWINDLER and J. ERVIN, eds. Comparative primate biology. Vol. **1**: Systematics, evolution and anatomy. A. R. Liss, New York.

SHAPIRO, S. G., E. A. SCHON, T. M. TOWNES, and J. B. LINGREL. 1983. Sequence and linkage of the goat ε^{I} and ε^{II} β-globin genes. J. Mol. Biol. **180**:803–823.

SIBLEY, C. G., and J. E. AHLQUIST. 1984. The phylogeny of the hominoid primates, as indicated by DNA-DNA hybridization. J. Mol. Evol. **20**:2–15.

———. 1987. DNA hybridization evidence of hominoid phylogeny: results from an expanded data set. J. Mol. Evol. **26**:99–121.

SIBLEY, C. G., J. A. COMSTOCK, and J. E. AHLQUIST. 1990. DNA hybridization evidence of hominoid phylogeny: a reanalysis of the data. J. Mol. Evol. **30**:202–236.

SIMPSON, G. G. 1945. The principles of classification and a classification of mammals. Bull. Am. Museum Nat. Hist. **85**:1–350.

———. 1961. Principles of animal taxonomy. Columbia University Press, New York.

———. 1963. The meaning of taxonomic statements. Pp. 1–31 *in* S. L. WASHBURN, ed. Classification and human evolution. Aldine, Chicago.

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1341108 by guest on 21 August 2022

SLIGHTOM, J. L., and R. F. DRONG. 1988. Procedures for constructing genomic clone banks. Pp. 1–42 *in* S. B. GELVIN and R. A. SCHILPEROORT, eds. Introduction, expression, and analysis of genes and gene products in plants. Martinus Nijhoff, Dordecht, The Netherlands.

SMITH, F., and M. S. WATERMAN. 1981. Identification of common molecular subsequences. J. Mol. Biol. **147**:195–197.

SNEATH, P. H. A., and R. R. SOKAL. 1973. Numerical taxonomy: the principles and practice of numerical classification. W. H. Freeman, San Francisco.

TAGLE, D. A. 1990. Molecular evolutionary genetics of the β-globin cluster of the prosimian primate *Galago crassicaudatus:* nucleotide sequence determination of the 41 kb cluster, developmental expression patterns, and comparative analyses. Ph.D. diss., Wayne State University, Detroit.

TAUTZ, D., M. TRICK, and G. A. DOVER. 1986. Cryptic simplicity in DNA is a major source of genetic variation. Nature **322**:652–656.

UEDA, S., Y. WATANABE, N. SAITOU, K. OMOTO, H. HAYASHIDA, T. MIYATA, H. HISAJIMA, and T. HONJO. 1989. Nucleotide sequences of immunoglobulin-epsilon pseudogenes in man and apes and their phylogenetic relationships. J. Mol. Biol. **205**:85–90.

VON KOENIGSWALD, G. H. R. 1968. The phylogenetical position of the hylobatinae. Pp. 271–276 *in* B. CHIARELLE, ed. Taxonomy and phylogeny of old world primates with references to the origin of man. Rosenberg & Sellier, Turin.

WILLIAMS, S. A., and M. GOODMAN. 1989. A statistical test that supports a human/chimpanzee clade based on noncoding DNA sequence data. Mol. Biol. Evol. **6**:325–330.

WILSON, J. T., L. B. WILSON, J. K. DE RIEL, L. K. VILLA-KOMAROFF, A. EFSTRATIADIS, B. G. FORGET, and S. M. WEISSMAN. 1978. Insertion of synthetic copies of human globin genes into bacterial plasmids. Nucleic Acids Res. **5**:563–581.

WU, C. I., and W. H. LI. 1985. Evidence for higher rates of nucleotide substitutions in rodents than in man. Proc. Natl. Acad. Sci. USA **82**:1741–1745.

ZAGURSKI, R. J., K. BAUMEISTER, N. LOMAX, and M. L. BERMAN. 1985. Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Tech. **2**:89–94.

ZUCKERKANDL, E., and L. PAULING. 1965. Evolutionary divergence and convergence in proteins. Pp. 97–116 *in* V. BRYSON and H. J. VOGEL, eds. Evolving genes and proteins. Academic Press, New York.

WALTER M. FITCH, reviewing editor

Received July 2, 1990; revision received September 24, 1990

Accepted October 19, 1990

Downloaded from https://academic.oup.com/mbe/article/8/2/155/1134108 by guest on 21 August 2022