We have determined the nucleotide sequences of the linked γ'- and γ2- fetal globin genes from a single orangutan (Pongo pygmaeus) chromosome and compared them with the corresponding genes of other simian primates (γ1- and γ3-genes of human, chimpanzee, gorilla, and the single γ-gene of the spider monkey). Previous studies have indicated that the two γ-gene loci in catarrhine primates resulted from a duplication about 25–35 million years ago. However, comparisons of aligned γ-gene sequences show that these genes contain three regions with distinct histories of which only the 3' third clearly reflects the ancestral nature expected of the γ-gene duplication. To explain these different evolutionary histories and also hominid relationships we provide evidence for the occurrence of sequence conversions which affect region 1 (120 base pairs 5'-flanking through exon 2) in all hominid species and extend to varying degrees into region 2 (intron 2 through exon 3). Close examinations of the proposed conversions further suggest that 12 of the 13 conversions identified involved γ1 converting γ2. Polarity of these conversions may be a result of differential survival between these genes because during human fetal development the γ1-gene is preferentially expressed over the γ2-gene and it may be subjected to greater selection pressure to remain unaltered.

The β-type globin genes show a highly conserved chromosomal organization in the primate infraorder Catarrhini. The catarrhine species examined thus far include old world monkeys, Papio anubis (baboon) and Macaca mulatta (rhesus) and five hominoids, Hylobates lar (gibbon), Pongo pygmaeus (orangutan), Gorilla gorilla gorilla (lowland gorilla), Pan troglodytes (chimpanzee), and Homo sapiens (human) (Barrie et al., 1981; Zimmer, 1981). The β-globin gene cluster in each of these species spans about 50 kb of DNA and contains six globin-related genes. The β-type genes in these clusters are arranged (5' to 3') in the order of their developmental expression, 5'-ɛ-γ(embryonic), duplicated γ1-γ2 (fetal), inactive γ2 (a pseudogene), and δ and β (adult) (Efstratiadis et al., 1980; Barrie et al., 1981; Zimmer, 1981; Scott et al., 1984; Chang and Slightom, 1984; Goodman et al., 1984; Slightom et al., 1985). Evolutionary analyses indicate that the gene arrangement, 5'-ɛ-γ-δ-β-3', evolved by a series of tandem duplications which occurred between 200 and 100 million years ago in the lineage to basal eutherians (Czelusniak et al., 1982; Goodman et al., 1984). The γ-gene became a pseudogene in the early primates; however, the four other loci remained functional within Anthropoidea (Goodman et al., 1984; Menon and Lingrel, 1986). The new world monkeys Ateles geoffroyi (spider) and Aotus trivigatus (owl) show only a single γ-locus (Barrie et al., 1981; Harris et al., 1984; Giebel et al., 1985). The duplicated γ1- and γ2-genes of catarrhines arose by a recombinational event about 35–25 million years ago in basal Catarrhini after divergence of Platyrrhini, but prior to the divergence of superfamilies Cercopithecoidae (old world monkeys) and Hominioidea (Barrie et al., 1981; Shen et al., 1981). This recombinational event appears to have been an unequal crossing-over resulting in the duplication of about 5 kb of DNA (Shen et al., 1981).

After duplication of the ancestral γ-gene locus other recombinational events occurred in different species lineages between parts of the homologous duplicated sequences. Certainly of these events have resulted in either the expansion or contraction of the γ-gene locus by additional unequal crossing-over (occurring outside the structural gene regions) or by the exchange of structural gene sequences by way of gene conversion. Both unequal crossing-over and gene conversion share a common recombinational mechanism (Radding, 1982). In some cases, the outcome of recombinational events have survived in a species population thereby allowing its characterization. For example, individual human chromosomes containing triple or quadruple γ-genes as a result of unequal crossing-over events have been identified (Hill et al., 1986). Also in humans, nucleotide sequence comparison of paralogous γ1- and γ2-genes has shown that a gene conversion event resulted in a nonreciprocal transfer of γ1-sequences to the γ2-locus (Slightom et al., 1980).

In the case of γ-gene evolution, the effects of gene conversions can be evaluated by comparing the nucleotide sequences of orthologous and paralogous γ-genes in closely related species. Previously, we determined the nucleotide sequence for the linked fetal globin genes of human, chimpanzee, and gorilla and by using the parsimony method identified sequence regions within the γ-loci which appear to have been involved in gene conversion events. Comparing the human and gorilla γ-gene sequences enabled us to identify three possible human gene conversion events labeled C2-C4 (Scott et al., 1984) of which C4 represents the conversion originally described by Slightom et al. (1980). Similar analysis with the chimpanzee γ-gene sequences revealed three possible chimpanzee, C5-C7, and one gorilla, C8, conversion regions (Slightom et al., 1985). The finding that a stretch of simple-sequence DNA, (TG)n, is
located at the 3' boundary of conversion regions C2, C4, and C6 and at the 5' boundary of C5 supports the suggestion that this (TG), element may provide a "hot spot" for strand breakage and transfer between the paralogue y-genes (Slightom et al., 1985; Kilpatrick et al., 1984). However, conversion regions C7 and C8 in chimpanzee and gorilla, respectively, encompass exons 1 and 2, intron 1, and only the 5' boundary of intron 2 and thus do not appear to be associated with the hot spot element (Slightom et al., 1985). The polarity of these nucleotide sequence transfers appears not to be random because all regions except C5 have y-gene sequences (donor sequences) superimposed upon y'-gene sequences (acceptor sequences). Conversion C5 has this polarity reversed (Slightom et al., 1985). In addition to identifying species-specific conversions, our previous evidence (a less-than-expected divergence between y'- and y'-intron 2 sequences) suggested that a conversion event (C1) occurred in a late common ancestor of Homo, Pan, and Gorilla.

The present study extends our previous findings. We have now determined the nucleotide sequences of the orangutan y'- and y'-fetal globin genes and added these sequences to our evolutionary analysis. For our nearest outgroup, we have also determined the sequence of the single y-gene of spider monkey (Giebel et al., 1985). As in Homo, Pan, and Gorilla the 5'-600 bp of the two orangutan y-genes are more similar to each other than to orthologous positions in other species, whereas the reverse relationship is found for the 3'-flanking DNA. Thus, two different evolutionary histories are indicated for the 5' and 3' regions of these genes. In addition to this orangutan 5'-conversion region, labeled C11, two downstream conversion regions, C12 and C13, are indicated by shared sites between orangutan y'- and y'-loci. Moreover, perhaps because of having used a closer outgroup, our evidence now suggests that additional conversions, C9 and C10, may have occurred in gorilla. The present data continue to support the occurrence of a conversion, C1, in a common ancestor of Homininae (Homo, Pan, and Gorilla) and further suggest that a conversion, labeled C0, occurred in a late common ancestor of Ponginae (Pongo) and Homininae. In sorting out this history of primate gene conversion events, our evolutionary analysis suggests that conversion events in Homininae have occurred in these two subfamilies with much later divergence times for Homininae genera than for the Homininae subfamilies, 6-8 million years ago compared to 12-16 million years ago (Koop et al., 1986a, 1986b).

EXPERIMENTAL PROCEDURES

Materials—Restriction endonucleases; Apal, Accl, BamHI, SphII, EcoRI, HindIII, Sau3AI, Stul, Xbal, and XhoI were obtained from either Promega Biotec or New England Biolabs. Polynucleotide kinase was obtained from Pharmacia and bovine alkaline phosphatase was obtained from Boehringer Mannheim. Radioactive nucleotides (α-32P]ATP (2000-3000 Ci/mmol; 1 Ci = 3.7 x 1010 Eu) and T4 ligase were obtained from New England Nuclear. Chemicals used for DNA sequencing were obtained from vendors recommended by Maxam and Gilbert (1980). X-ray film rolls (20 cm x 25 m; XAR351) and sheet film (35 x 45 cm; XAR5) were obtained from Kodak. Intensifying screens (Quanta 111: 35 cm) were obtained from New England Nuclear. Chemicals used for DNA sequencing gel stands and sheet film (35 cm x 25 m; XAR351) and sheet film (35 cm x 45 cm; XAR5) were obtained from Kodak. Intensifying screens (Quanta 111: 35 cm) were obtained from Du Pont. Nitrocellulose paper (BA-85) was obtained from Schleicher and Schuell and 3MM paper from Whatman. DNA sequencing gel stands and safety cabinets were obtained from Fotodyne, Inc.

DNA Isolation and Gene Cloning—Total DNA was isolated from liver removed from two nonrelatated orangutans; orangutan NZO-1 (National Zoo, Washington, D. C.) DNA was isolated using the method described by Blin and Stafford (1976). Purified DNAs were partially digested with either EcoRI (YG-1) or Sau3AI (NZO-1) and DNA fragments of 15-20 kb were size selected on 5%-20% NaCl gradients (Slightom et al., 1980). Size-selected DNA fragments were cloned into the appropriate λ vector arms, EcoRI-cut fragments of YO-1 DNA were cloned into EcoRI-cut Charon 32 arms and Sau3AI-cut fragments of NZO-1 DNA were cloned into BamHI-cut Charon 35 arms (Loenen and Blattner, 1983). Recombinant phage DNAs were packaged into phage capsids using the in vitro phage-packaging procedure described by Hohn (1979). Both recombinant phage libraries, a total of 1-2 x 106 phage from each, were screened on the Reckard PR5000 (Bacteriological Instruments) host (Slightom et al., 1985) followed by nitrocellulose blotting as described by Benton and Davis (1977). Nitrocellulose filters were hybridized against a 32P-labeled Aval-EcoRI fragment (245 bp) cut from the human γ-globin clone pJW15) (Wilson et al., 1978). Clones containing regions of the orangutan y-gene were purified and phage DNA isolated previously (Slightom et al., 1980). Two recombinant phage clones were isolated from orangutan YO-1; PpyCh32-13.5 and PpyCh32-14.2, which contain the linked fetal globin and the linked y-y′ and δ-globin genes, respectively. From screening the Charon 35 recombinant phage library of orangutan NZO-1 we obtained clones PpyCh35-14.8 and PpyCh35-19.5. Both clones contain the linked fetal globin genes; however, clone PpyCh35-19.5, also contains the y′- and δ-globin gene linkage. The 7.25- and 6.24-kb EcoRI fragments containing the orangutan y′- and y'-fetal globin genes, respectively, were isolated from PpyCh32-13.5 and cloned into PBR322. BamHI fragments of PpyCh32-13.5 and 2.8-kb BamHI fragment of PpyCh32-19.5, respectively, were isolated previously (Slightom et al., 1980). The third of the y-gene (of PpyCh32-13.5 and 2.8 kb EcoRI fragment (y′-exon 3, and 3′-untranslated and 3′-flanking) of PpyCh32-19.5 were cloned into the respective sites of pUC8. Transformation of E. coli K12 strain HB101 was done using the extended calcium shock method described by Leder and Ebhlick (1977). These phage DNAs were designated pPpy13.5-R7.25, pPpy13.5-R2.64, pPpy13.5-B4.8, and pPpy13.5-R9.28. Plasmid DNAs were purified using the alkaline-extraction procedure described by Birnboim and Doly (1979), followed by two CsCl-ethidium bromide gradient bandings in type 70.1 Ti rotor (fixed angle rotor).

DNA Sequencing—Chemical sequencing was performed as described by Maxam and Gilbert (1980). Generally, 10-20 μl of phasmid or 50 μl of λ phage DNAs were digested with the appropriate restriction enzyme and end-labeled. The conditions and type of DNA sequencing gels used have been described previously (Chang and Slightom, 1984). Four percent agarose gels (0.2-mm thick) were used for routine sequencing. For sequence determination by the dideoxy method, a 20 cm-long gel is run until 600 bp can be read. The capacity of these gels can be increased by using a comb with 3-mm slots (International Biotechnologies Inc.) which allow 32 loads ranging from 0.2 (top) to 6.6 mm (bottom), a 0.2-mm spacer was used for the remaining (upper) 44 cm. These gels are referred to as "bellbottom gels." A typical 4% acrylamide bellbottom gel run until the xylene cyanol dye is 65 cm from the top, nucleotide sequence reads start at about 50 bp from the end-labeled site and are generally readable out to about 550 bp. Nucleotide sequencing reads greater than 600 bp can be obtained by using a 4% acrylamide gel filled with uniform 0.2-mm spacers and running the dye to the bottom. Nucleotide sequences near the labeled site can be obtained by running a 6% acrylamide gel (30 cm-long) until the xylene cyanol dye is at about 4000 bp can be read. The capacity of these gels can be increased by using a 3-mm slot (International Biotechnologies Inc.) which allows 32 loads to be made across a 20-cm-wide gel. Thus from a single 4% bellbottom gel loaded once with eight different sequenced fragments, as many as 4000 bp can be read.

Phylogenetic and Evolutionary Analysis—Using the principle of minimum change or parsimony as described in previous studies (Scott et al., 1984; Goodman et al., 1984; Slightom et al., 1985), we aligned the nucleotide sequences of 10 fetal globin genes (four y′-genes and five y-genes of the four hominids and the single y′-gene of the spider monkey Ateles geoffroyi) to the 5′ end of the cap site to 170 bp 3′ of the poly(A) site and then determined independently at each variable position the most parsimonious
RESULTS AND DISCUSSION

Restriction Maps of Orangutan γ-, δ- and ψ-Globin Gene Regions—From screening the orangutan Ch32 library, YO-1, with 32P-labeled human γ-globin probe prepared from the cDNA clone pW151 (Wilson et al., 1978) we isolated two γ-gene hybridizing λ clones, PyCh32-14.2 and PyCh32-13.5. Detailed restriction enzyme site mapping showed that clone PyCh32-14.2 contains the orangutan ψ- to δ-globin gene linkage and that clone PyCh32-13.5 contains the linked γ- and γ'-fetal globin genes (Fig. 1). The nucleotide sequence of this orangutan ψ-γ-globin gene has already been reported (Koop et al., 1986a). The linkage between the γ- to ψ-γ-globin genes was not obtained from this screening. We subsequently screened the second orangutan library, NZO-1, and isolated clones PyCh35-16.8 and PyCh35-19.5. Both of these clones contain linked γ- and γ'-fetal globin genes and clone PyCh35-19.5 contains the linkage region between the γ-2- and ψ-γ-genes (see Fig. 1).

The EcoRI fragments containing the γ-1- and γ'-fetal globin genes, from λ clone PyCh32-13.5, were isolated and subcloned into pBR322 yielding clones pPy32-13.5-7.25 (γ-1-gene) and pPy32-13.5-R2.64 (γ'-gene) (Fig. 2). Restriction enzyme site mapping reveals that most of the sites found in and around the human fetal globin genes are also present in these orangutan genes. Notable exceptions include the absence of EcoRI sites in the 3'-flanking DNAs of each γ-gene (the EcoRI fragments 3' of γ- and γ'-genes are 2.2 and 2.8 kb, respectively). The XhoI site in each intron 2 is also absent, as is the case in chimpanzee fetal globin genes (Sligh et al., 1985). One additional HindIII site is located in the 3'-untranslated region of the orangutan γ'-gene (see Fig. 1).

Nucleotide Sequence of Orangutan Fetal Globin Genes—The strategies used to sequence the orangutan fetal globin genes are shown in Fig. 2. The nucleotide sequences of both orangutan γ- and δ-globin genes were determined starting 154-bp 5' of the expected capped nucleotide and extending 171-bp 3' of the expected poly(A) addition nucleotide. Additional 5'-flanking
### Nucleotide Sequence of Orangutan Fetal Globin Genes

![Nucleotide Sequence](image)

**FIG. 3.** Nucleotide sequence of orangutan γ1- and γ2-fetal globin genes and their alignment with other primate fetal globin gene sequences. Human γ1- and γ2-globin gene sequences are from Slightom et al. (1980) and Shen et al. (1981) and are denoted as Hsa αδ and Hsa αδ (γ1- and γ2-genes; respectively) from chromosome A and Hsa αA (γ2-gene) is from chromosome B of a single individual. Gorilla (Ggo) and chimpanzee (Pmc) γ1- and γ2-globin gene sequences are from Scott et al. (1984) and Slightom et al. (1985), respectively. The single γ-globin gene sequence of the spider monkey is from Giebel et al. (1985). The numbering system was set by the overall alignment of the 10 primate γ-gene sequences; asterisks indicate gaps which were used to increase sequence identities among the genes. The complete nucleotide sequence of the gorilla γ2-globin gene (which is the longest gene) is presented on the top line (above the counting line). For any position where one sequence differs from another, the nucleotide for each gene is shown. Nucleotides that may have biological importance are noted: single overlined promoter sequences CCAAT and CAAAT, elements and double overlined poly(A) addition signal, AATAAA. The fetal globin amino acid sequences are shown below the second dashes a counting line, with amino acid replacements being printed below the appropriate codon. The initiator codon is the first Met and the terminator codon is designated TIR. Vertical arrows indicate exon-intron boundaries that all conform to the GT/AG rule (Breathnach et al., 1978).
Nucleotide Sequence of Orangutan Fetal Globin Genes

HOT SPOT SEQUENCES

FIG. 3—continued
nucleotide sequence information was obtained from the gorilla (Scott et al., 1984) and chimpanzee (Slightom et al., 1985) fetal globin genes so that the CCAAT-element could be added to our analysis. Although most of the orangutan \( \gamma \) and \( \gamma' \)-gene sequences presented in Fig. 3 were obtained from the same chromosome of orangutan YO-1, sequences extending 3' from the EcoRI site in exon 3 of the \( \gamma \)-gene were obtained from a different orangutan chromosome (NZO-1). Because these two \( \gamma \)-gene region clones differ in their intron 2 nucleotide sequences (data not shown) they certainly must represent different orangutan chromosome alleles. Nevertheless, our previous studies of this 3' noncoding and flanking DNA region indicate that the use of a "hybrid" \( \gamma \)-gene nucleotide sequence will not distort our analysis because this region is not active in gene conversions and has retained the expected paralogous relationship (see discussion below).

As evident from the nucleotide sequence shown in Fig. 3, the fetal globin genes from the five simian species share a high degree of identity not only in exons, but also in noncoding regions. Tables I and II summarize regional divergence results from noncoding and coding DNAs, respectively. A high degree of sequence conservation exist in and around the 5'-flanking promoter elements (CCAAT and AATATA), 5'-untranslated, exons 1 and 2, and intron 1 sequence regions. In contrast intron 2, 3'-untranslated, and flanking DNA regions have accumulated many more substitutions. The 3' noncoding regions all share the same poly(A) addition signal (AATATA), with the spider monkey containing two signals; however, only the shared poly(A) signal is functional (Giebel et al., 1985). The intron 1 nucleotide sequence of the orangutan \( \gamma \)-gene is unusual because it shows a one base pair insertion (T at position 256, Fig. 3) which is identical to that found in the single \( \gamma \)-gene of spider monkey and may have been present in the simian ancestor. Lengths of intron 2 sequence of the orangutan \( \gamma \)- and \( \gamma' \)-genes (874 and 878 bp, respectively) are shorter than intron 2 sequences of the other primate \( \gamma \)-genes due to the presence of shorter tracts of the simple sequence DNA (TG)\( _n \); values of \( n \) are 11 for the \( \gamma \)-gene and 13 for the \( \gamma' \)-gene. The value of \( n \) for the orangutan \( \gamma' \)-gene is half that found in the orthologous gene of the other hominids shown in Fig. 3 (positions 1076-1137), which have \( n \) values ranging between 22 and 25. The shortness of the (TG), repeats in the orangutan \( \gamma \)- and \( \gamma' \)-genes may be important because there is a lack of any clear gene conversion pattern in intron 2 which has been initiated by their (TG), sequence elements (see Fig. 4 and discussion below).

**Evolution of the Orangutan \( \gamma \)-Globin Genes—Comparison of the coding nucleotide sequences of both orangutan \( \gamma \)-gene reveals that they differ by a total of five nucleotide substitutions, two nonsynonymous and three synonymous. The two nonsynonymous substitutions occurred in descent of the \( \gamma \)-gene; the first is at position 276 (G/C; changing amino acid position 33, Val to Leu) and the second is at position 402 (A/G; changing amino acid 75, Ile to Val). These positions in the orangutan \( \gamma' \)-gene are identical to those found in all the other hominid \( \gamma \)-genes (see Fig. 3 and Table III). The nonsynonymous substitution at amino acid position 75 is predicted by amino acid sequence data (Huisman et al., 1973; Schroeder et al., 1978), but amino acid analysis at position 33 is not available. Amino acid sequence heterogeneity is predicted for amino acid position 135, Thr/Ala (Huisman et al., 1973); however, our nucleotide sequences show that for position 135 both \( \gamma \)-genes encode Ala. This difference could be indicative of allelic polymorphism in the orangutan population.

Synonymous substitutions occur in both orangutan \( \gamma \)-genes; the first is located at nucleotide position 137 of the \( \gamma' \)-gene (C/T), the second is position 377 of the \( \gamma' \)-gene (G/A), and the last is located at position 449 of the \( \gamma' \)-gene (A/G). All of these nonsynonymous and synonymous substitutions are unique to orangutan \( \gamma \)-genes except for the substitution at position 449 in which orangutan \( \gamma' \) shares a G-nucleotide with the single \( \gamma \)-gene of the spider monkey. This latter substitution provides evidence indicating that a conversion occurred in the conserved 5' region of the duplicated hominine \( \gamma \)-genes,
TABLE I
Percent divergence table for noncoding sequences

| Species Comparisons | 5' | 5' intron | 3' intron | 3' | Expected Orthologue | Expected Paralogue |
|---------------------|----|----------|----------|----|----------------------|-------------------|
| H1 vs H2a, H2b     | 0.5, 0.0 | 0.0, 2.0 | 3.0, 3.8 | 13.2, 13.6 |                       | 8.0 - 11.6       |
| H2a vs H2b         | 0.5 | 2.3      | 0.8      | 0.4 |                       |                   |
| H1 vs C1, C2       | 1.0, 0.5 | 3.3, 4.2 | 2.3, 2.6 | 2.3, 13.6 | 1.7                  | 8.0 - 11.6       |
| H2a vs C1, C2      | 1.4, 1.0 | 3.3, 4.2 | 2.3, 1.1 | 14.0, 0.4 |                       |                   |
| H1 vs G1, G2       | 1.0, 1.0 | 4.4, 4.8 | 1.1, 2.6 | 2.7, 14.4 |                       |                   |
| H2a vs G1, G2      | 1.4, 1.4 | 4.4, 4.8 | 2.6, 2.6 | 12.8, 1.2 |                       |                   |
| H2b vs G1, G2      | 1.0, 1.0 | 4.8, 4.3 | 3.4, 2.6 | 13.2, 1.6 |                       |                   |
| H1 vs O1, O2       | 1.9, 2.9 | 4.5, 4.9 | 4.2, 4.9 | 5.4, 13.6 | 3.4                 | 8.0 - 11.6       |
| H2a vs O1, O2      | 2.4, 2.9 | 4.5, 4.9 | 3.8, 3.4 | 14.8, 5.9 |                       |                   |
| H2b vs O1, O2      | 1.9, 2.9 | 4.2, 4.4 | 4.5, 3.4 | 15.2, 6.3 |                       |                   |
| H1 vs S             | 7.7 | 13.2     | 10.9     | 15.7 | 11.6                |                   |
| H2a vs S            | 7.1 | 13.2     | 10.2     | 18.3 |                     |                   |
| H2b vs S            | 7.7 | 13.1     | 10.2     | 18.3 |                     |                   |
| C1 vs C2            | 0.5 | 3.6      | 1.1      | 14.4 |                     | 8.0 - 11.6       |
| C1 vs G1, G2        | 1.0, 1.0 | 4.5, 5.0 | 1.9, 1.1 | 12.1, 15.2 | 2.2                |                   |
| C2 vs G1, G2        | 0.5, 0.5 | 4.8, 2.8 | 2.3, 1.5 | 13.3, 0.8 |                       |                   |
| C1 vs O1, O2        | 1.9, 2.9 | 3.8, 6.1 | 4.5, 3.4 | 4.7, 12.9 | 4.0                 |                   |
| C2 vs O1, O2        | 1.0, 2.4 | 4.7, 4.0 | 3.4, 2.3 | 15.2, 6.3 |                       |                   |
| C1 vs S             | 7.7 | 12.8     | 10.2     | 16.3 | 11.6                |                   |
| C2 vs S             | 7.1 | 12.2     | 9.1      | 19.0 |                     |                   |
| G1 vs G2            | 0.0 | 5.2      | 2.3      | 14.0 |                     | 8.0 - 11.6       |
| G1 vs O1, O2        | 1.9, 2.8 | 5.7, 6.2 | 3.4, 4.2 | 5.1, 12.5 | 3.9                |                   |
| G2 vs O1, O2        | 1.9, 2.9 | 5.0, 4.5 | 4.9, 3.8 | 15.2, 7.1 |                       |                   |
| G1 vs S             | 7.7 | 14.2     | 9.8      | 17.6 | 11.6                |                   |
| G2 vs S             | 7.7 | 12.4     | 10.6     | 20.4 |                     |                   |
| O1 vs O2            | 2.9 | 4.8      | 4.5      | 12.9 |                     | 8.0 - 11.6       |
| O1 vs S             | 7.7 | 13.0     | 10.2     | 17.0 | 11.6                |                   |
| O2 vs S             | 8.2 | 12.9     | 9.8      | 19.6 |                     |                   |

* From 4n-gene comparison (Koop et al., 1986a).

* Values of 8% are for old world and hominoid comparisons values of 11.6% are for new world and hominoid comparisons.

with γ² accepting sequence transfer from the γ⁻-gene (see discussion below).

In addition to the above mentioned substitutions, the human and orangutan γ-gene coding sequences differ by six substitutions, four of these substitutions are nonsynonymous and are located at nucleotide positions 108 (stem to both orangutan γ-genes), 408 (stem to all hominine γ-genes), 490 (stem to hominine γ-genes including converted human and chimpanzee γ²-genes), and 1500 (stem to hominine +genes). A summary of primate γ-gene nonsynonymous substitutions (most of which are at the root of the split between catarhinine and platyrrhine species) are presented in Table III. The orangutan γ⁻- and γ- genes share the two synonymous substitutions (see positions 356 and 1447 in Fig. 3), the substitution at position 356 is also shared with both chimpanzee γ- genes (Fig. 3). Two more synonymous substitutions are specific for the chimpanzee genes, one at position 419 (γ²) and another at position 437 (γ²). In addition, inspection of the aligned sequences (Fig. 3) shows that a minimum of 21 mutations (at positions 69, 250, 408, 490, 937, 1138, 1240, 1241, 1383, 1500, 1537, 1539, 1540, 1551, 1612, 1712, 1724, 1768, 1776) unequivocally favor the genealogical grouping of Homo, Pan, and Gorilla into the subfamily Homininae, separate from the subfamily Ponginae (Pongo).

The percent divergence values tabulated for these coding sequences (Table II) reveal that the silent site divergence between Pongo and Homininae, 3.5% is similar to the divergence found from orangutan ε-, ψγ-, α⁺-, and α²-globin gene nucleotide sequence and from total unique sequence genomic DNA hybridization data (Koop et al., 1986a, 1986b; Marks et al., 1986; Sibley and Ahlquist, 1984). Thus, silent sites in the coding sequences of these γ-genes are not evolving at a slower
## TABLE II
Percent divergence for coding regions

| Species Comparisons | Exons 1 & 2 | Exons 3 | All Coding | Nontcoding Expected Orthologues | Nontcoding Expected Paralogues |
|---------------------|-------------|---------|------------|--------------------------------|-------------------------------|
| Silent              | AAChg       | Silent  | AAChg      | Silent                         | AAChg                        |                             |
| H1 vs H2A, H2B     | 0.0.0.0     | 0.0.0.0 | 1.1.1.1    | 0.0.0.0                        | 1.2.0.3                       |                             |
| H1 vs H2B          | 0.0.0.0     | 0.0.0.0 | 0.0.0.0    | 0.0.0.0                        | 0.0.0.0                       | 1.7                         |
| H1 vs C1, C2       | 2.8.2.8     | 0.0.0.0 | 0.0.0.0.1  | 1.9.19                         | 0.0.3.0                       | 8.0-11.6                   |
| H1 vs C1, C2       | 2.8.2.8     | 0.0.0.0 | 1.1.0.0    | 1.9.19                         | 0.3.0.0                       |                             |
| H1 vs C1, C2       | 2.8.2.8     | 0.0.0.0 | 1.1.0.0    | 1.9.19                         | 0.3.0.0                       |                             |
| H1 vs G1, G2       | 0.0.0.0     | 0.0.0.8 | 0.0.0.0    | 0.0.0.0                        | 0.0.0.0                       |                             |
| H1 vs G1, G2       | 0.0.0.0     | 0.0.0.8 | 0.0.0.0    | 0.0.0.0                        | 0.0.0.0                       |                             |
| H1 vs O1, O2       | 2.8.4.2     | 1.2.2.1 | 3.2.3.2    | 0.0.0.0                        | 2.9.3.9                       | 1.4                         |
| H1 vs O1, O2       | 2.8.4.2     | 1.2.2.1 | 3.2.3.2    | 1.1.1.1                        | 2.9.3.9                       | 1.2.18                      |
| H1 vs O1, O2       | 2.8.4.2     | 1.2.2.1 | 3.2.3.2    | 1.1.1.1                        | 2.9.3.9                       |                             |
| H1 vs S            | 13.3.5.6    | 6.4    | 5.3        | 11.2                           | 5.5                           | 11.6                        |
| H2 vs S            | 13.3.5.6    | 6.4    | 3.2        | 11.2                           | 5.8                           |                             |
| H2 vs S            | 13.3.5.6    | 6.4    | 6.3        | 11.2                           | 5.8                           |                             |
| C1 vs C2           | 2.8        | 0.0.4  | 0.0        | 1.1                            | 1.9                           | 0.3                         |
| C1 vs G1, G2       | 2.8.2.8     | 0.0.0.8 | 0.0.0.0    | 0.0.1.1                        | 1.9.19                         | 2.2                         |
| C1 vs O1, O2       | 2.8.4.2     | 1.2.2.1 | 3.2.3.2    | 0.0.0.0                        | 2.9.3.9                       | 4.0                         |
| C1 vs O1, O2       | 2.8.4.2     | 1.2.2.1 | 3.2.3.2    | 1.1.1.1                        | 2.9.3.9                       | 1.2.18                      |
| C1 vs S            | 16.0.5.6    | 6.4    | 5.3        | 13.1                           | 5.5                           | 11.6                        |
| C1 vs S            | 16.0.5.6    | 6.4    | 6.3        | 13.1                           | 5.8                           |                             |
| G1 vs G2           | 0.0        | 0.0.8  | 0.0        | 1.1                            | 0.0                           | 0.9                         |
| G1 vs O1, O2       | 2.8.4.2     | 1.2.2.1 | 3.2.3.2    | 0.0.0.0                        | 2.9.3.9                       | 3.9                         |
| G1 vs O1, O2       | 2.8.4.2     | 1.2.2.1 | 3.2.3.2    | 1.1.1.1                        | 2.9.3.9                       |                             |
| G1 vs S            | 13.3.5.6    | 6.4    | 5.3        | 11.2                           | 5.5                           | 11.6                        |
| G2 vs S            | 13.3.5.6    | 6.4    | 6.3        | 11.2                           | 5.7                           |                             |
| O1 vs O2           | 4.2        | 0.0.8  | 0.0        | 1.0                            | 2.9                           | 0.6                         |
| O1 vs S            | 16.0.5.2    | 9.6    | 5.3        | 14.1                           | 5.2                           | 11.6                        |
| O2 vs S            | 14.6.6.0    | 9.6    | 5.3        | 13.1                           | 5.8                           |                             |
| Positional Orthologues |          |        |            |                                |                               |                             |
| H, C, G            | 1.9        | 0.0.3  | 0.0        | 1.2                            | 0.2                           | 1.9                         |
| O vs H, C, G       | 3.6        | 1.7    | 3.2        | 0.6                            | 3.5                           | 1.4                         |
| S vs H, C, G       | 14.3       | 5.6    | 6.4        | 5.4                            | 12.2                          | 5.6                         |
| Positional Paralogues |          |        |            |                                |                               |                             |
| H, C, G            | 1.8        | 0.2    | 0.0        | 1.1                            | 1.2                           | 0.5                         |
| O vs H, C, G       | 3.4        | 1.6    | 3.2        | 0.6                            | 3.3                           | 1.3                         |

* Calculations are in accordance with the method described by Nei and Gojobori (1986).
* From Xn-gene comparison (Koop et al., 1986a).
* Values of 8% are for old world and hominoid comparisons and values of 11.6% are for new world and hominoid comparisons.

The divergence between coding sequences is less within each hominid species than the divergence of positional orthologues among species (γ1* to γ1, γ2 to γ2), a finding which strongly suggests the occurrence of gene conversions. Divergence at silent sites (Table II) between positional paralogues (γ1 and γ2) for any rate than average noncoding DNA of hominin genomes, although in previous studies (Scott et al., 1984; Slightom et al., 1985) this was thought to be the case. From nucleotide sequence analyses of hominin genomic DNAs, we and others calculate the overall neutral drift rate to be about 1.3 × 10⁻⁸ substitutions/site/year (Goodman et al., 1984, Slightom et al., 1985; Koop et al., 1986a, 1986b; Britten, 1986) which is about one-fourth the average reported mammalian rate of 5 × 10⁻⁹ substitutions/site/year (Li et al., 1985).
Nucleotide Sequence of Orangutan Fetal Globin Genes

FIG. 4. A parsimony analysis of aligned γ1- and γ2-genes of human (H), chimpanzee (C), gorilla (G), and orangutan (O), plus the hypothetical ancestor of human, chimpanzee, and gorilla (An). Directly above the map of the duplicated γ1- and γ2-genes are the results of a position by position parsimony analysis using the spider monkey to indicate the primitive condition. Dots indicate positions where intraspecific similarity of γ1- and γ2-genes (paralogous positions) is greater than the interspecific similarity of γ1- or γ2-genes (orthologous positions). Asterisks indicate positions where interspecific similarity between γ1- (or γ2-) genes is greater than intraspecific similarity between γ1- and γ2-genes. Two consecutive dots indicate greater intraspecific similarity and suggests a conversion event (labeled C1–C13). Below the map of γ1- and γ2-genes we have indicated three regions and using regional parsimony analysis we have determined the overall phylogenetic order for these three distinct γ-gene regions. The complex nature of region 2 necessitated breaking up sequences into smaller subsequences that reflected patchy conversion regions (see text).

TABLE III
Amino acid replacements encoded in primate γ-globin gene nucleotide sequences amino acid residue number (nucleotide position)

| Species  | 1 (37) | 2 (60) | 5 (70.71) | 10 (84) | 18 (108) | 33 (278) | 47 (319) | 51 (330) | 65 (372, 373) | 73 (397, 398) | 75 (402) | 77 (408) | 86 (436) | 104 (490) | 117 (1443) | 118 (1445, 1446) | 135 (1498) | 136 (1500) | 142 (1517) |
|----------|--------|--------|-----------|---------|----------|----------|----------|----------|----------------|----------------|---------|----------|----------|----------|----------------|----------------|------------|--------------|-------------|
| Hsa Gy   | Gly    | His    | Glu       | Thr     | Val      | Val      | Asn      | Ala       | Lys             | Asp             | Ile      | His       | Ala       | Lys       | His           | Phe           | Thr        | Gly          | Ser          |
| A γ      | "      | "      | "         | "       | "        | "        | "        | "        | "               | "               | "       | "        | "        | "        | "             | "             | "         | "            | "            |
| Ptr Gy   | "      | "      | "         | "       | "        | "        | "        | "        | "               | "               | "       | "        | "        | "        | "             | "             | "         | "            | "            |
| Ggo γ    | "      | "      | "         | "       | "        | "        | "        | "        | "               | "               | "       | "        | "        | "        | "             | "             | "         | "            | "            |
| A γ      | "      | "      | "         | "       | "        | "        | "        | "        | "               | "               | "       | "        | "        | "        | "             | "             | "         | "            | "            |
| Ppy γ1   | "      | "      | "         | "       | "        | Leu      | "        | "        | Asp             | Asn             | "       | "        | "        | "        | "             | "             | "         | "            | "            |
| γ2       | "      | "      | "         | "       | "        | Leu      | "        | "        | Asp              | Val             | Asn     | "        | "        | "        | "             | "             | "         | "            | "            |
| Age γ    | Ser    | Asn    | Ala       | Ala      | Val      | Val      | Ser      | Pro       | Val              | Glu             | Ile     | Asn       | Gly       | Arg       | Leu           | His           | Ala        | Gly          | Ala          |

pair of hominid species (e.g. human and orangutan) is much less than that expected in the absence of gene conversion for duplicated loci which arose in the stem of the Catarrhini about 35 million years ago, long before the separation of Homininae from Ponginae. In addition, evidence supporting the occurrence of conversions comes from noncoding 5' flanking and intron 2 sequences. For these regions, we calculate divergence values for paralogous genes that are much less than that expected for these duplicated loci (see calculation results listed in Table I).

Parsimony Analysis Reveals Gene Conversions Between γ1 and γ2 Genes—Our initial examination of the aligned γ-gene
sequences in Fig. 3 reveals that these \( \gamma \)-gene sequences consist of three distinct regions which exhibit quite different evolutionary histories. These three regions are represented by nucleotide positions \(-154 \) to \( 495 \) (5'-flanking and untranslated, exons 1 and 2, and intron 1), \( 496-1535 \) (intron 2 and exon 3), and \( 1506-1792 \) (3'-untranslated and flanking), respectively, see Fig. 4. Of these regions, only region 3 shows evidence that the two ancient paralogous \( \gamma \)-gene lineages are unaffected by gene conversions. The orthologous splittings within these lineages show divergence of an orangutan clade from a human, chimpanzee, and gorilla clade (Fig. 4). Although an identical speciation pattern is evident in region 1, the sequence relationships suggest that latter parallel duplications produced the duplicated \( \gamma \)-genes of human, chimpanzee, gorilla, and orangutan. However, this suggestion is clearly inconsistent with the pattern shown in region 3 and the fact that two linked \( \gamma \)-loci are present in all extant catarrhine primates. The pattern of sequence relationships found in region 1 can be explained if we hypothesize, in each species lineage, a transfer of genetic information via gene conversion events between paralogues of \( \gamma \)-genes. The complex mosiac pattern (see below) found in region 2 could then be explained by the initiation and/or termination of patchy type gene conversion events extending from or into region 1. In order to test the gene conversion hypothesis in more detail, and also to delineate additional conversion regions, we examined each variable position in the aligned \( \gamma \)-gene sequences for evidence of species-specific conversion event.

Fig. 4 presents results from position-by-position parsimony analysis over the 1946 aligned nucleotide positions of the 10 simian fetal-globin genes shown in Fig. 3. A "dot" at a variable site indicates that the positional paralogues are closer genealogically to each other than to positional orthologues, therefore supporting an hypothesized conversion, whereas an "asterisk" indicates the reverse and supports no conversion. For example, position 108 (Fig. 3) is represented by a dot in the orangutan row (O\( \circ \)) of Fig. 4 because at this position the two orangutan \( \gamma \)-genes share the same T-nucleotide whereas the other eight \( \gamma \)-genes have a G-nucleotide. In contrast, position 583 is represented by an asterisk in the O\( \circ \) row because the orangutan, chimpanzee, and gorilla \( \gamma \)-orthologues share an A-nucleotide whereas the \( \gamma \)-orthologues of these three species share a G-nucleotide. A 13-bp deletion over positions 572-584 in the three human \( \gamma \)-genes (H\( \mathrm{I} \), H\( \mathrm{II} \), H\( \mathrm{H} \)) includes position 583, and this deletion is represented by a dot at position 572 in each of the two human rows (H\( \mathrm{H} \)\( \mathrm{II} \) and H\( \mathrm{H} \)\( \mathrm{H} \)). Positions 69, 210, 250, 256, 408, and 449, each viewed independently of the others, have dots in the ancestral hominine (An'An') row because in each case all hominine \( \gamma \)-1 and \( \gamma \)-2 genes share a nucleotide or deletion which is different from the deduced ancestral nucleotide (rooted with the spider monkey \( \gamma \)-gene sequence). Positions 256 and 449 are especially informative because they provide evidence that in conversion C1 (ancestral Hominine) \( \gamma \)-sequences accepted \( \gamma \)-sequences. It should be noted that on putting together these results at different positions, the most parsimonious solution does not require conversion of \( \gamma \)-2 by \( \gamma \)-1 in the Hominine ancestor, but it does require that \( \gamma \)-2 sequence accept \( \gamma \)-1 sequences in the 5' conversion region of each species (C7, C8, and C11). A highly nonrandom distribution of dots and asterisks are found for the \( \gamma \)-1 and \( \gamma \)-2 gene conversions compared in Fig. 4 with dots being clustered among sites in region 1 and asterisks clustered in region 3. This distribution of characters indicates that the \( \gamma \)-gene 5' regions are prone to gene conversions, whereas the 3' regions resist conversions (Fig. 4). Between these two regions (from the end of exon 2 to the end of exon 3) is region 2 which contains approximately 1000 bases that show a complex, mosaic arrangement of dots and asterisks. To account for the distribution of dots by the minimum number of gene conversion events, we treat two or more consecutive dots as part of the same conversion region. In fact, because of the scarcity of dots in the 3' region and the presence of two widely separated dots which are not interrupted by asterisks appears to be a nonrandom occurrence. Analyses of mutational patterns in the 3' region indicate that parallel mutations resulting in the same nucleotide being present at an homologous position in \( \gamma \)-1 and \( \gamma \)-2-gene pairs are unlikely. As a result of these observations our minimal criterion for assigning a hypothetical gene conversion, using the data in Fig. 4, is the presence of two consecutive dots (not necessarily adjacent or even close) which are not interrupted by an asterisk. Once having identified a gene conversion region by this criterion the conversion is allowed to spread until it meets the first of two consecutive asterisks or enters a sequence region of relatively frequent mismatches between positional paralogues. However, a gene conversion does not have to be continuous, we can postulate a conversion event in which converted sequences are interspersed among stretches of unconverted sequences, i.e. a patchy or bubble conversion (Kourilsky, 1983; Michelson and Orkin, 1983; Stoeckert et al., 1984, Powers and Stoeckert, 1986). By using these criteria, we have determined for each row of paired \( \gamma \)-1 \( \gamma \)-2-genes in Fig. 4, the minimum number of gene conversion events needed to account for the pattern of dots and these results are delineated (as bars) in the upper part of Fig. 4. Even though the six sets of \( \gamma \)-1 \( \gamma \)-2-gene pairs show 13 conversion regions (C1-C13) we need hypothesize no more than a single conversion event per \( \gamma \)-gene pair if we assume that the single conversion occurred via a patchy conversion mechanism.

Our present parsimony analysis confirms the presence of C2, C3, and C4 (human) suggested by Scott et al. (1984) and C5, C6, C7 (chimpanzee), and C8 (gorilla) suggested by Sligh-tom et al. (1985). By the criterion of two consecutive dots, we further hypothesize two additional conversion regions in gorilla, C9 (extending approximately 120 bp 5' of the hot spot (TG), and C10 for positions 1240 and 1241, downstream of the hot spot (Fig. 4). From our analysis of the orangutan \( \gamma \)-genes, we hypothesize three conversion regions; C11, extending from the CCAAT-element through position 250 and possibly through exon 2, C12, extending from about position 620 to 720 in the upstream region of intron 2, and C13 which encompasses exon 3 and extends through position 1600 (see Fig. 4). In contrast to the hominine results, the hot spot does not border any of the orangutan conversion regions. The hypothesized conversion, C1, which occurred in a late common ancestor of Homininae is neither supported or contradicted by our parsimony analysis, but is suggested by the lack of divergence in intron 2 (Table I). For instance, 3' of the hot spot in intron 2, there are no two linked dots which directly support C1 (An'An') and no linked dots which support a human conversion (H\( \mathrm{H} \)\( \mathrm{H} \)\( \mathrm{H} \); H\( \mathrm{H} \)\( \mathrm{H} \)); yet the human \( \gamma \)-gene has diverged only less than expected (8-12%) for sequences which descended from a gene duplication that occurred in the basal catarrhines about 35 MYA (Shen et al., 1981; Scott et al., 1984). Only region 3' of exon 3 shows the expected degree of divergence (see Table I). Even in the 5' region of intron 2, where asterisks at 11 positions oppose conversion C1, the hominine \( \gamma \)-1 and \( \gamma \)-2-genes still differ by less than 7%; this suggests that another ancient gene conversion (C0, not shown in Fig. 4) occurred in the basal Homininae prior to the separation of subfamilies Ponginae (Pongo) from Homininae (see conversion summary.
in Fig. 5). Alternatively, the 5' intron 2 region of each hominid species may have been involved in species-specific conversions in which the $\gamma^2$-gene always accepted $\gamma^1$-gene sequences. One further observation is that the sequence stretches which immediately flank many of the proposed conversions have accumulated mutations at a rate greater than that found in comparable parts of these genes. This is particularly evident in orthologous comparisons of 5' intron 2 and 3'-flanking regions (see calculation results listed in Table I). Presently, it is difficult to determine whether these higher mutation rates are the cause or are the result of conversion terminations.

To further test whether the proposed conversions shown in Fig. 4 yield a meaningful picture we separated region 2 into sequence subregions with presumed independent histories and then determined their parsimonious branching arrangement. This branching arrangement is shown below region 2 in Fig. 4. Regions 1 and 3 were left intact and their most parsimonious branching arrangements are also shown below the gene map in Fig. 4. Note that when conversions are accounted for in region 1 and 2, sequence and species relationships for all three regions are compatible. Also, when conversions are accounted for the region 2 tree depicts distinct paralogous, $\gamma^1$- and $\gamma^2$-lineages which predate the separation of Ponginae from Homininae. Thus, the three trees shown in Fig. 4 support our conversion hypotheses and divide Hominidae into subfamilies Ponginae (Pongo) and Homininae (Homo, Pan, Gorilla).

With regard to the directions of conversions, the branching pattern of the region 2 tree shows all converted sequences joining the $\gamma^1$ branch with one exception. The exception, chimpanzee conversion C5, is clearly a member of the $\gamma^2$ branch as supported by the parsimony solutions at positions 1138, 1142, 1160, and 1172. The parsimony solutions at other positions show that, $\gamma^2$ sequences accepted $\gamma^1$ sequences. Several examples of these positions are as follows. Position 626 provides parsimony evidence that the $\gamma^1$-gene converted $\gamma^1$-gene sequences in the C12 region of orangutan and C3 and C4 regions of human chromosomes A and B, respectively. However, gorilla and chimpanzee $\gamma^2$-genes are not converted here and contain a G-nucleotide, whereas the ancestral sequence (which contains an A-nucleotide, as deduced from spindle monkey $\gamma$-gene) was retained in the hominid $\gamma^1$ lineage and transferred by conversions to orangutan and human $\gamma^2$-genes. Position 1023 provides evidence that either in an ancestral hominine conversion (C1) or in later conversions of gorilla (C9), chimpanzee (C6), and human (C2 and C4), $\gamma^2$-genes accepted $\gamma^1$-gene sequences.

**Fig. 5.** A summary of 35 million years of $\gamma$-gene lineage evolution under the constraints imposed by species evolution. The thicker outlines reflect species evolution and the thinner lines reflect gene sequence (open or filled box) evolution. An $\times$ at the end of a line indicates the disappearance of a gene lineage. Regions 1 and 3 are the same as those in Fig. 4 and clearly show different evolutionary histories with region 1 being involved in many gene conversions, whereby $\gamma^2$-gene sequences have accepted $\gamma^1$-gene sequences. Region 3 shows no history of gene conversions, whereas region 2 shows an intermediate and more complex history (see Fig. 4).
discontinuous and patchy, as well as infrequent.

Perhaps, the orangutan (TG), hot spot has physical features which reduce its ability to initiate genetic recombination. Two important physical parameters are length of the (TG), element (Stringer, 1985) and the degree of nucleotide sequence identity immediately adjacent to the sites of strand transfer and invasion (Nicolas and Rossignol, 1983; Radding, 1982). Tracts of (CG), with values of n as small as five can adopt the Z-DNA conformation (Singleton et al., 1983); however, (CG), repeats from a Z-DNA helix more readily than (TG), repeats. The importance of length may rest with the increased efficiency in pairing permitted by longer tracts of a repeated dinucleotide (Stringer, 1985). The degree of sequence identity immediately adjacent to the (TG), elements may be the more important physical parameter affecting the frequency of γ-gene conversions. We note that in the gene conversions (C2, C4, C5; Fig. 4) which extend from (TG), elements in the orangutan γ-genes and the gene sequences immediately 5′ of the (TG), tract do not share complete identity as they differ by a 3-bp deletion in the γ1-gene and a 6-bp insertion in the γ2-gene. The orangutan γ-gene nucleotide sequences immediately 3′ of their hot spots also differ as a result of a 2-bp deletion in the γ2-gene.

Concluding Remarks—Our analyses of the closely linked γ- and γ-fetal globin genes in these four hominid species provide us with a great deal of information concerning the effect that recombinational events such as gene conversions have on the evolution of a pair of duplicated genes. The results presented here clearly show that in all four species the evolution of the paralogous γ- and γ2-genes are affected because these gene pairs are evolving more like orthologous genes, and in addition the rate at which these gene pairs are, diverging is reduced. Thus, this fetal globin gene conversion mechanism is probably playing a role in the evolution of the γ-gene pairs in all catarrhine species, although the degree and frequency of conversion appears to be quite different for the different species.

Why should there be more examples of the γ-gene sequences (5′ of the hot spot element) accepting γ-gene sequences than the reverse? A possible clue comes from human clinical studies (Bunn and Forget, 1986) which show that the ratio of γ- to α-chains in fetal hemoglobin is about three to one. We suggest, therefore, that harmful mutations in regulatory and coding positions are more likely to have functional effects and be selected against if they occur in the γ1 locus rather than if they occur in the γ2 locus. That is, selection should act more stringently on the γ1-gene than the γ2-gene, and the transfer of γ1-sequences to γ2-sequences would be more favorable than the reverse.

Acknowledgments—We are grateful to Dr. E. A. Zimmer for the gift of orangutan, YO-1, DNA. We thank Drs. Frederick R. Blattner and Oliver Smithies for the use of shared laboratory space and equipment in the Department of Genetics, University of Wisconsin, Madison, WI, where this work was initiated. We thank Julie Nelson for her excellent help in organizing the arrangement of sequences shown in Fig. 3 and Laurie Tuinstra for typing the manuscript.

REFERENCES
Barrie, P. A., Jeffreys, A. J., and Scott, A. F. (1981) J. Mol. Biol. 149, 319-336
Benton, W. D., and Davis, R. W. (1977) Science 196, 180-182
Blumenthal, T. H., and Stamatatos, D. W. (1976) Nucleic Acids Res. 3, 2203-2308
Britten, R. J. (1980) Science 213, 1393-1398
Bunn, H. F., and Forget, B. G. (1986) Hemoglobin; Molecular, Genetics and Clinical Aspects, S. B. Saunders Co., Philadelphia
Chang, L. Y.-E., and Tachtitides, K. J. (1984) J. Mol. Biol. 180, 767-794
Collins, F. S., and Weissman, S. M. (1984) Proc. Nucleic Acid Res. Mol. Biol. 31, 313-319
Czelusniak, J., Goodman, M., Hewett-Emmett, D., Weiss, M. L., Venta, P. J., Olsson, A., Moks, T., Uhlen, M., and Gaal, A. B. Proc. Natl. Acad. Sci. U.S.A. 75, 4853-4857
Dagert, M., and Ehrlich, S. D. (1979) Gene 6, 23-28
Efratiadis, A., Posakony, J. W., Maniatis, T., Lown, R. M., O'Connell, C., Suul, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Stightom, J. L., Blelch, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C., and Proudfoot, N. J. (1980) Cell 21, 652-668
Giebel, L. B., vanOosten, V. L., Slighstom, J. L., and Spritz, R. A. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 6985-6989
Goodman, M., Koop, B. F., Casazza, J., Weiss, M. L., and Stightom, J. L. (1984) J. Mol. Biol. 180, 903-923
Hill, A. V. S., Bowden, D. K., Weatherall, D. J., and Clegg, J. B. (1986) Blood 67, 1011-1016
Hobn, B. (1979) Methods Enzymol. 68, 299-309
Huisman, T. H. J., Schroeder, W. A., Keeling, M. E., Gengozian, N., Miller, A., Brodie, A. R., Shetton, J. R., Shetton, J. B., and Apell, J. (1973) Biochem. Genet. 10, 309-318
Jeffreys, A. J. (1979) Cell 18, 1-10
Kourilsky, P., Li, W.-H., Luo, C.-C., and Wu, C.-I. (1983) Nature 308, 43-52
Koop, B. F., Goodman, M., Xu, P., Chan, K., and Stightom, J. L. (1986a) Proc. Natl. Acad. Sci. U.S.A. 83, 1413-1417
Koop, B. F., Miyamoto, M. M., Embury, J. E., Goodman, M., Czelusniak, J., and Stightom, J. L. (1986b) J. Mol. Evol. 24, 94-102
Koop, B. F., and Forget, B. G. (1986) J. Mol. Evol. 29, 171-179
Marks, J. W., Shaw, J.-P., and Chen, K.-J. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 1413-1417
Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499-560
McKusick, A. A., and Ingelfinger, J. R. (1986) Gene 42, 141-150
Michelson, A. M., and Orkin, S. H. (1983) J. Biol. Chem. 258, 1524-1525
Murray, N. E., Branner, W. J., and Murray, R. K. (1977) Mol. Genet. 150, 53-61
Nei, M., and Gojobori, T. (1986) Mol. Biol. Evol. 3, 418-426
Olsson, A., and Rossingol, J. L. (1985) EMBO J. 4, 2265-2270
Oslo, H., Moks, T., Uhlen, M., and Gaal, A. B. (1984) J. Biochem. Biophys. Methods 10, 93-101
Powers, P. A., and Smithies, O. (1986) Genetics 112, 243-358
Radding, C. M. (1982) Annu. Rev. Genet. 16, 405-437
Radding, C. M., Slighstom, J. R., Shelton, J. H., and Huisman, T. H. J. (1978) Biochem. Genet. 16, 1293-1295
Scott, A. F., Health, P., Trusko, S., Boyer, S. H., Prass, W., Goodman, M., Czelusniak, J., Chang, L.-Y. E., and Slighstom, J. L. (1984) Mol. Biol. Evol. 109, 571-586
Shen, S.-H. S., Slighstom, J. L., and Smithies, O. (1981) Cell 26, 191-203
Shiley, C. G., and Aulestia, J. E. (1984) J. Mol. Evol. 20, 2-15
Singleton, C. K., Klysik, J., Stibrivaint, S. M., and Wells, R. D. (1982) Nature 299, 312-316
Singleton, C. K., Klysik, J., and Wells, R. D. (1980) Proc. Natl. Acad. Sci. U.S.A. 80, 2447-2451
Slighstom, J. L., Blelch, A. E., and Smithies, O. (1980) Cell 21, 627-638
Slighstom, J. L., Chang, L.-Y. E., Koop, B. F., and Goodman, M. (1985) Mol. Biol. Evol. 2, 370-389
Schooert, C. J., Jr., Collins, J. B., and Weissman, S. M. (1984) Nucleic Acids Res. 12, 4489-4497
Stringer, C. R. (1986) Mol. Cell. Biol. 5, 1247-1259
Wilson, J. T., Wilson, L. B., DeRiel, J. K., Villa-Komaroff, L., Efratiadis, A., Forget, B. G., and Weismann, S. M. (1979) Nucleic Acid Res. 5, 583-585
Zimmer, E. A. (1981) Ph.D. Thesis, University of California, Berkeley

Nucleotide Sequence of Orangutan Fetal Globin Genes

7483