

# Sequences of Primate Insulin Genes Support the Hypothesis of a Slower Rate of Molecular Evolution in Humans and Apes than in Monkeys<sup>1</sup>

Susumu Seino,\* Graeme I. Bell,\* † and Wen-Hsiung Li‡

\*Howard Hughes Medical Institute, Department of Medicine, and †Department of Biochemistry and Molecular Biology, The University of Chicago; and ‡Center for Demographic and Population Genetics, University of Texas at Houston

The chimpanzee and African green monkey insulin genes have been cloned and sequenced. These two sequences together with the previously reported sequences for the human and owl monkey insulin genes provide additional support for the hominoid-rate-slowdown hypothesis, i.e., a slower rate of nucleotide substitution in humans and apes than in monkeys. When these sequences and other primate sequences available for the relative-rate test were considered together, the substitution rate in the Old World monkey lineage was shown to be significantly higher than the rates in the human and chimpanzee lineages. This was true regardless of whether the  $\eta$ -globin pseudogene was included in the analysis. Therefore, in contrast to the claim by Easteal, the hominoid-rate-slowdown is not unique to the  $\eta$ -globin pseudogene but appears to be a rather general phenomenon. On average, the substitution rate at silent sites is about 1.5 times higher in the Old World monkey lineage than in the human and chimpanzee lineages.

## Introduction

There has been a longstanding controversy over the hominoid-rate-slowdown hypothesis, which postulates that the rate of molecular evolution has become slower in hominoids after their separation from the Old World (OW) monkeys. This hypothesis, first proposed by Goodman (1961) and Goodman et al. (1971), is based on rates estimated from immunological distance and protein sequence data. Sarich and Wilson (1967) and Wilson et al. (1977) contended that the slowdown was an artifact, owing to the use of an erroneous paleontological estimate of the ape-human divergence time. They conducted relative-rate tests using both immunological distance data and protein sequence data and concluded that there was no evidence for a hominoid slowdown. A similar conclusion was drawn from DNA hybridization studies (Kohne et al. 1972; Sibley and Ahlquist 1984); however, more recent DNA hybridization studies have produced conflicting conclusions (Sibley and Ahlquist 1987; Caccone and Powell 1989). On the other hand, comparative analyses of DNA sequence data by Britten (1986), Li and Tanimura (1987), and Li et al. (1987) provided strong support for the hominoid-slowdown hypothesis, and so the hypothesis has been accepted by many molecular evolutionists. However, recently Easteal (1991) has argued that the slowdown occurred only in the  $\eta$ -globin pseudogene, because, when this pseu-

1. Key words: nucleotide substitution, molecular clocks, hominoid slowdown, primate evolution.

Address for correspondence and reprints: Wen-Hsiung Li, Center for Demographic and Population Genetics, University of Texas, P.O. Box 20334, Houston, Texas 77225.

*Mol. Biol. Evol.* 9(2):193–203, 1992.

© 1992 by The University of Chicago. All rights reserved.  
0737-4038/92/0902-0001\$02.00

dogene was removed from comparison, the rate of nucleotide substitution in the OW monkey lineage was no longer significantly higher than that in the human lineage. To resolve this controversy, more DNA sequence data are needed.

In our studies of the molecular evolution of the insulin gene, we have determined the nucleotide sequence of the insulin gene from the owl monkey (*Aotus trivirgatus*), a New World (NW) monkey (Seino et al. 1987). Here, we report the sequences of the chimpanzee and African green monkey (an OW monkey) insulin genes. These three primate sequences together with the human insulin gene sequence (Bell et al. 1980; Ullrich et al. 1980) are used to test the hominoid-slowdown hypothesis. We show below that the new data support the slowdown hypothesis. Thus, in contrast to the claim by Easteal (1991), the phenomenon is not unique to the  $\eta$ -globin pseudogene.

## Material and Methods

### DNA

Chimpanzee (*Pan troglodytes*) DNA was isolated from fibroblast cultures (GMO3448A) obtained from the NIGMS Human Genetic Mutant Cell Repository, Coriell Institute for Medical Research, (Camden, N.J.). African green monkey (*Ceropithecus aethiops*) DNA was isolated from CV-1 cells (CCL70; American Type Culture Collection, Rockville, Md.).

### Southern Blotting

DNA was digested with *EcoRI*, electrophoresed, blotted, and hybridized with the <sup>32</sup>P-labeled human insulin gene probe phins96 (Bell et al. 1984), according to a method described elsewhere (Bell et al. 1981).

### Insulin Gene Isolation and Sequencing

Genomic DNA was digested with *EcoRI* and fragments of 10–15 kb were isolated by electrophoresis and cloned in  $\lambda$ Ch4A (Maniatis et al. 1982, pp. 270–294). Phage containing the chimpanzee and African green monkey insulin genes were identified by cross-hybridization with the human insulin gene (Seino et al. 1987).

The *EcoRI* fragments containing the chimpanzee and African green monkey insulin genes were subcloned into pUC9. Subfragments of these *EcoRI* fragments were cloned into M13mp19 and sequenced by the dideoxy chain-termination method (Sanger et al. 1980). More than 85% of the sequence was determined on both strands. The nucleotide sequence was compared with that of the human insulin gene (Bell et al. 1980, 1982; Ullrich et al. 1980). Sites at which the sequences differed were checked to confirm that the differences were not due to sequencing or clerical errors.

### Data Analysis

The sequences of the chimpanzee and African green monkey insulin genes are from the present study. The other insulin gene sequences were obtained from Bell et al. (1980, 1982), Ullrich et al. (1980), Wetekam et al. (1982), and Seino et al. (1987) (GenBank accession numbers J02989 and J00268). For the sources for the other sequences studied, see the legend to table 3.

For coding regions, we use the method of Li et al. (1985) to compute the number of substitutions per nonsynonymous site ( $K_A$ ) and the number of substitutions per synonymous site ( $K_S$ ) between two sequences. For noncoding regions we use Kimura's (1980) two-parameter method to estimate the number of substitutions per site ( $K$ ). For the relative-rate test (Sarich and Wilson 1967), we use the statistics developed by

Wu and Li (1985). Sequence alignment was simple and was done manually; the regions of tandem repeats in the 5' flanking region were excluded from comparison because they differ greatly in length and sequence among species.

## Results

### Sequences of Chimpanzee and African Green Monkey Insulin Genes

The human insulin gene hybridizes to a single *Eco*RI fragment in many primate species whose size is about 12–13 kb (S. Seino and G. I. Bell, unpublished observations). *Eco*RI fragments containing the chimpanzee and African green monkey insulin genes were cloned, and the sequences of regions of 2,483 and 1,909 bp, respectively, determined (fig. 1). The boundaries of the exons were assigned by comparison with the human insulin gene sequence (Bell et al. 1980; Ullrich et al. 1980). The exon-intron organization of the insulin genes of these two primates is the same as those previously reported for the human and owl monkey insulin genes (Bell et al. 1980; Ullrich et al. 1980; Seino et al. 1987) and corresponds to what is termed the "ancestral" insulin gene structure (Steiner et al. 1985), i.e., the coding region of the gene being interrupted by two introns. This is different from the rat I and mouse I insulin genes, which have only a single intron located in the region of the gene encoding the 5' untranslated region of the mRNA and which are believed to have been generated by a process of retrotransposition of a partially processed insulin transcript (Soares et al. 1985).

The identity between the sequences of human (chimpanzee) and African green monkey preproinsulin is 98% (97%); at the nucleotide level the corresponding identity between these protein-coding regions is 98% (95%). The amino acid sequence of human, chimpanzee, and African green monkey insulin is identical. Chimpanzee preproinsulin (fig. 1A) differs from the human protein at two sites in the signal peptide: amino acids -13 and -2 are Ala in the human protein. The sequence of African green monkey preproinsulin (fig. 1B) differs at amino acids -3 and -2 in the signal peptide and at residue 37 in the C-peptide portion of the precursor, which are Val, Pro, and Leu, respectively. In addition, there are several other notable differences between the human, chimpanzee, and African green monkey insulin gene sequences. The first is a deletion of 48 bp in the chimpanzee insulin gene at a site just after the TAG which specifies termination of translation (fig. 1). As a consequence, the predicted 3' untranslated region of chimpanzee mRNA is only 28 nucleotides, exclusive of the poly(A) tract and may be the shortest of any of the insulin mRNAs described to date if the site at which it is polyadenylated is the same as that for human insulin mRNA. This deletion is unlikely to be a cloning artifact, since it was a feature of eight independently isolated clones. The deletion of this 48-nucleotide segment in the chimpanzee insulin gene implies that this region is not required for translation of the mRNA.

The second notable difference was in the 5' flanking region of these genes. The chimpanzee insulin gene has a region of 206 bp that begins -365 bp from the start of transcription and that comprises 20 tandem repeats of sequences related to the 25-bp sequence ACAGGGGTCCTGGGG (fig. 1). A region of similar tandem repeats is present in the human gene (Bell et al. 1982), the most common sequence of which, ACAGGGGTGTGGGG, does not occur in the chimpanzee gene. However, the next two most common repeats in the human are ACAGGGTCCTGGGG and ACAGGGGTCTGGGG, which constitute most of the repeats in the chimpanzee. This region is characterized by extensive polymorphism in humans (Bell et al. 1981) and thus may be polymorphic in chimpanzees as well. Similar regions of multiple GC-rich tandem repeats are not present in the African green monkey or owl monkey



## B. African Green Monkey

-425 GGGCCATCCATGGGGGCATCTTGGGCAACCAGGGACTGAGAA**ACAGGGTCCAGGACAGGGGGTCTGGGG**ACTGTGGTGCAAAGAGCCCTGCCCTGCAGCCCCAGCTCTCTGGTGTAAATGT  
 -304 GGAAAGTGGGCCGGGGAGGGCTTTGCTCTGCTGGAGACATTTGCCCCCCAGCTGCGAGCAGGAACAGGCCTGGCCACCAGGCCCTGGTTAAGACTCTAATGACCCCGTGGCCCTGAGGAAG  
 -183 AGGTGCTGACGACCAAGATCTTCCACAGACCCAGCACCAGGAAATGATCTGGAAATTGCAGCCTCAGCCCCAGCCATCTGCTAACCCCCCAGCCAGGCCCTAATCGGGCAGCAGGG  
*Exon 1* *Intron A*  
 -62 GTTGGCAGGCTGGGGAGATGGGCTCTGAGACTATAAAGCCAGTGGGAGCCCCAGCAGCCCTC AACCTCGGACGCTGCATCAGAAGAGGCCAGCAAGCAG gtctgttccaagggccttcgcg  
 60 gtcaggtgggctcagggctgcccacttgggggttccaggggtggctggaccccaggccccagctctgcaacagggaggacatggctgggctcttgaagcgtttgaggtgaaccaggggcc  
 -24 *Exon 2* Met Ala Leu Trp Met Arg Leu Leu Pro Leu Leu Ala Leu  
 181 cagggcagggcactggcctcagctggcctcaggtgctctctcag GTCACTGTCCTTCCGCC ATG GCC CTG TGG ATG CGC CTC CTG CCC CTG CTG GCG CTG  
 -9 Leu Ala Leu Trp Gly Pro Asp Pro Val Pro Ala Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys  
 285 CTG GCC CTC TGG GGA CCT GAC CCG GTC CCG GCC TTT GTG AAC CAG CAC CTG TGC GGC TCC CAC CTG GTG GAA GCC CTC TAC CTG GTG TGC  
 20 Gly Glu Arg Gly Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Pro Gln V *Intron B*  
 375 GGG GAG CGA GGC TTC TTC TAC ACG CCC AAG ACC CGC CGG GAG GCA GAG GAC CCG CAG G gtgagcccagggcccagcgtgcccctggccgccccagcca  
 477 cctgtcctcctggcgtcccaccagcgtggacagaagggggcaggggctgccaccagcagggggtcagatgtacttttttaaaaaagaatgaagttctcttggtcacctcctaaaag  
 598 gaccggctcctgtggcccagtcagaagctcagcttgaggacgggtgttggttcagcagctgagtccaagatacatcagaggggtgggacgcttctctcctccactgcccgtcaaacac  
 719 atgccccacagcccgtttctccaccctcatttgatgaccacagattcaagtggtttgttgagtacaagctcctgggtgacctgggggtcaaggggtgccccacgctgctgctgagtgggcc  
 840 gaccctgtcgccagggcctcatggcagctctgagtcaggagatgggaagatgctggggacagggcctggggagaagaaccggggcactcattcaggtcccactatggcacactggg  
 961 gcgggggaaggagtaggacatgtgggcattggggcctgtgggtccacaccagctgtgggtgacctcctctaacctgggcccggcttggtggagacgggtgggagtgcgacctaggg  
 39 *Exon 3* al Gly Gln Val Glu  
 1080 tggtgggtaggtgggcaactgtctctcctgactgtgtcctcctgtgtccctctgcttgcgtgctgctccgggacactgctctgctggtggcccgcctggcag TG GGG CAG GTA GAG  
 44 Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys Cys Thr  
 1194 CTG GCG GGG GGC CCT GGC GCA GGC AGC CTG CAG CCC TTG GCG CTG GAG GGG TCC CTG CAG AAG CGC GGC ATC GTG GAG CAG TGC TGT ACC  
 74 Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn AM *Polyadenylation*  
 1284 AGC ATC TGC TCC CTC TAC CAG CTG GAG AAC TAC TGC AAC TAG ATCCCCGGTGAAGGCCGGCCACACCCTCCACCTCCTGCACCAAGAGAGATCGA**AATAAA**GCCTT  
*Site*  
 1391 GAACTAGC CCTGCTGTGCCATCTTTTTTTGGCCCTGGACAGGCCGCTTCCGGGCACCTGTTGTAAGCCCCCTCCAGCTCTCTCCACGCTCTCT - 1484

FIG. 1.—Sequences of chimpanzee and African green monkey insulin genes and flanking regions. The boundaries of the genes and the nucleotide which is polyadenylated in the mRNA were assigned by comparison with the human insulin gene. The predicted start of transcription has been designated as nucleotide 1. The deduced amino acid sequences of the chimpanzee and African green monkey preproinsulins are shown. The signal peptide is amino acids -24 to -1; the B chain, 1 to 30; C-peptide (including pairs of adjacent basic amino acids), 31 to 65; and A chain, 66 to 86. The sequences of the introns are shown in lowercase letters. The region of tandem repeats in the 5'-flanking region of the chimpanzee insulin gene is shown in boldface type. A similar region is not present in the African green monkey insulin gene, and the sequence which is in its place is shown in boldface type. The number of the nucleotide or amino acid is shown in italics at the beginning of each line. The arrow following the amber (TAG) codon of the chimpanzee insulin gene indicates the location of a 48-bp deletion relative to the human insulin gene. The polyadenylation signal, AATAAA, is underlined. The sequences have been submitted to the GenBank™/EMBL Data Bank under accession numbers X61089 and X61092.

insulin genes. Rather, the insulin gene of the African green monkey has in its place a sequence, ACAGGGGTCCCCAGGACAGGGGGTCTGGGG, which is homologous to two copies of the consensus repeats found in the human and chimpanzee genes. The owl monkey has the sequence GCAGGGGTCTGGGG in this region (Seino et al. 1987). Thus, these regions of tandem repeats may only be a feature of the insulin genes of humans and higher apes.

### Sequence Divergence between Species

Table 1 shows the number of nucleotide substitutions per 100 sites in noncoding regions between species; the 5' and 3' untranslated regions are not included because they are short. In the following comparisons, the data compiled by Li et al. (1987) and Easteal (1991) will be used as references. First, we consider the 5' flanking region. The divergence between human and chimpanzee is 2.1%, which is slightly higher than the average divergence ( $\sim 1.7\%$ ) in noncoding regions and at synonymous sites between human and chimpanzee sequences (see Li et al. 1987; Easteal 1991). The divergence (9.5%) between human (or chimpanzee) and the African green monkey is at the higher end of the variation observed among other genes. The divergence (10.7%) between human and the owl monkey is close to that (11.1%) observed for the  $\eta$ -globin pseudogene (between the same two species), and the divergence (14.3%) between the African green monkey and the owl monkey is also close to that (13.3%) observed in the  $\eta$ -globin pseudogene.

Next, we consider the two introns. The divergences (1.3% and 2.0%) in both introns between human and chimpanzee are within the variation observed among other genes. In intron B the divergence (8.6%) between human and the African green monkey is within the variation observed among the introns in other genes, whereas in intron A the divergence (13.7%) is higher than that observed in other genes. The divergences in introns A and B (15.3% and 16.1%, respectively) between human and the owl monkey are considerably higher than those observed in other genes and in the  $\eta$ -globin pseudogene. In the comparison between the African green monkey and the owl monkey the divergence is relatively low (11.7%) in intron A but relatively high (18.2%) in intron B, when compared with the divergence in the  $\eta$ -globin pseudogene. Finally, the 3' flanking region tends to show a higher degree of divergence

**Table 1**  
Mean  $\pm$  Standard Error of Number of Nucleotide Substitutions per 100 Sites, between Species, in Noncoding Regions of Insulin Gene

Species Pair	5' Flanking Region (395 bp)	Intron A (160 bp)	Intron B (699 bp)	3' Flanking Region (93 bp)
Human-Chimpanzee	2.1 $\pm$ 0.7	1.3 $\pm$ 0.9	2.0 $\pm$ 0.5	2.8 $\pm$ 1.0
Human-African green monkey	9.5 $\pm$ 1.6	13.7 $\pm$ 3.2	8.6 $\pm$ 1.2	13.0 $\pm$ 4.2
Human-owl monkey	10.7 $\pm$ 1.8	15.3 $\pm$ 3.4	16.1 $\pm$ 1.7	21.6 $\pm$ 5.3
Chimp-African green monkey	9.5 $\pm$ 1.6	13.7 $\pm$ 3.2	8.2 $\pm$ 1.2	13.0 $\pm$ 4.2
Chimp-owl monkey	11.9 $\pm$ 1.9	15.3 $\pm$ 3.4	15.4 $\pm$ 1.6	23.0 $\pm$ 5.5
African green monkey-owl monkey	14.3 $\pm$ 2.1	11.7 $\pm$ 2.9	18.2 $\pm$ 1.8	23.6 $\pm$ 6.0

than do other regions, though the higher values are probably partly due to large standard errors.

Table 2 shows  $K_S$  and  $K_A$ . The  $K_S$  value of 5.1% between human and chimpanzee is much higher than the values observed in other genes (see Li et al. 1987; Easteal 1991). This is also the case for the  $K_S$  values between human (or chimpanzee) and the three monkey species. The  $K_S$  value [17.5% (or 19.1%)] between the African green monkey (*Macaca fascicularis*) and the owl monkey is within the variation observed among other genes. The  $K_A$  values between human, chimpanzee, the African green monkey, and *M. fascicularis* are small, reflecting the fact that there are only a few nonsynonymous differences in both the signal peptide and the C peptide regions but no nonsynonymous differences in the regions coding for the A and B chains (see Seino et al. 1987). On the other hand, the  $K_A$  values between the owl monkey and the other four species are relatively large, because several amino acid differences have been found between the owl monkey insulin polypeptide and the insulin polypeptides from the other four species (see Seino et al. 1987).

#### Evidence for a Hominoid Slowdown

The insulin gene sequences are useful for examining the hominoid-slowdown hypothesis. We use the relative-rate test. For example, consider the African green monkey as species 1, human as species 2, and the owl monkey as species 3, i.e., a reference. Let  $K_{ij}$  denote the number of substitutions between species  $i$  and  $j$ . Then, a positive sign for  $K_{13} - K_{23}$  means that the African green monkey lineage evolved faster than the human lineage, and a negative sign means the opposite. It is seen from table 3 that in the introns and the flanking regions the African green monkey lineage has evolved faster than the human lineage. On the other hand, at synonymous sites the human lineage has evolved faster than the African green monkey lineage; this occurred because, as mentioned above, the human and chimpanzee sequences have evolved exceptionally fast at synonymous sites. [The  $K$  values at synonymous sites differ slightly from those reported by Li et al. (1987) because the African green monkey sequence was used instead of that of *M. fascicularis*.] The chimpanzee insulin gene sequence shows a pattern similar to that in the human sequence, i.e., compared with the African green monkey sequence, it has evolved faster at synonymous sites but slower in the introns and flanking regions. Overall, the insulin gene sequences provide some evidence for a rate slowdown in hominoids (humans and apes).

**Table 2**  
Mean  $\pm$  Standard Error of Number of Substitutions per 100 Synonymous Sites (above Diagonal) and per 100 Nonsynonymous Sites (below Diagonal)

	Human	Chimpanzee	African Green Monkey	Cynomolgus Monkey	Owl Monkey
Human . . . . .		5.1 $\pm$ 2.6	19.5 $\pm$ 5.6	17.7 $\pm$ 5.2	25.0 $\pm$ 6.6
Chimpanzee . . .	0.8 $\pm$ 0.6		16.6 $\pm$ 5.2	14.9 $\pm$ 4.7	21.7 $\pm$ 6.1
African green monkey . . . . .	1.3 $\pm$ 0.7	1.7 $\pm$ 0.9		9.2 $\pm$ 3.6	17.5 $\pm$ 5.6
Cynomolgus monkey . . . . .	0.8 $\pm$ 0.6	1.3 $\pm$ 0.7	0.4 $\pm$ 0.4		19.1 $\pm$ 5.8
Owl monkey . . .	6.6 $\pm$ 1.7	7.0 $\pm$ 1.8	7.0 $\pm$ 1.8	6.6 $\pm$ 1.7	

Table 3

Differences in Number of Nucleotide Substitutions per 100 Sites, between Old World Monkey (Species 1) Lineage and Human or Chimpanzee (Species 2) Lineage

SEQUENCE ( $N^a$ )	HUMAN		CHIMPANZEE	
	$K_{12}$	$K_{13} - K_{23}$	$K_{12}$	$K_{13} - K_{23}$
<b>Synonymous:</b>				
Insulin (28) .....	19.5	$-7.4 \pm 6.3$	16.6	$-4.2 \pm 5.8$
$\alpha$ -Globin (106) .....	6.1	$1.1 \pm 3.4$		
$\theta$ -Globin (114) .....	23.0	$-1.0 \pm 8.3$		
$\beta$ -Globin (74) .....	8.9	$2.8 \pm 5.6$	8.9	$2.8 \pm 5.6$
$\delta$ -Globin (105) .....	11.5	$-1.1 \pm 4.1$		
Transforming growth factor $\beta$ (276) .....	8.2	$-0.1 \pm 3.3$		
Alcohol dehydrogenase (260) .....	6.8	$2.0 \pm 1.8$		
Apolipoprotein AI (158) ...	7.9	$-5.3 \pm 4.8$		
Apolipoprotein E (228) ...	10.6	$5.1 \pm 3.2$		
Erythropoietin (145) .....	11.2	$5.1 \pm 5.9$		
Pepsinogen (263) .....	10.5	$0.0 \pm 3.3$		
$\alpha$ -Antitrypsin (140) .....	10.9	$6.7 \pm 6.8$		
Pro-opiomelanocortin (180)	15.6	$-6.2 \pm 4.2$		
<b>Introns:</b>				
Insulin (859) .....	9.6	$1.0 \pm 1.3$	9.2	$1.6 \pm 1.1$
$\alpha$ -Globin (237) .....	4.5	$-4.5 \pm 2.1^*$		
$\theta$ -Globin (190) .....	8.2	$3.2 \pm 3.5$		
$\beta$ -Globin (960/915) .....	7.4	$2.1 \pm 1.2$	5.4	$1.4 \pm 1.1$
$\delta$ -Globin (863) .....	9.8	$-0.4 \pm 1.1$		
<b>Flanking and untranslated regions:</b>				
Insulin (548) .....	9.8	$3.2 \pm 1.5^*$	9.8	$2.2 \pm 1.1$
$\alpha$ -Globin (195) .....	5.0	$2.4 \pm 2.3$		
$\theta$ -Globin (402) .....	10.2	$3.4 \pm 3.5$		
$\beta$ -Globin (146/179) .....	3.5	$1.7 \pm 1.8$	5.2	$0.8 \pm 2.1$
$\delta$ -Globin (338) .....	6.3	$1.0 \pm 1.6$		
Transforming growth factor $\beta$ (312) .....	8.1	$1.3 \pm 2.2$		
Pro-opiomelanocortin (230)	15.6	$4.7 \pm 3.9$		
Overall (7,409/2,655) ...	9.3	$1.1 \pm 0.4^{b,**}$	8.0	$1.5 \pm 0.7^{b,*}$

NOTE.—In all comparisons the first species is an Old World monkey and the second species is human (first case) or chimpanzee (second case). The insulin sequences are from the present study (the African green monkey and chimpanzee) and GenBank; the owl monkey is used as the third (reference) species. For chimpanzee  $\beta$ -globin, the OW monkey species used is *Macaca cynomolgus* and the reference species is brown lemur; all sequence data are from GenBank. All other results are taken from Li et al. (1987) and Easteal (1991).  $K_{ij}$  = no. of substitutions, per 100 sites, between species  $i$  and  $j$ .

<sup>a</sup> Approximate number of sites compared; since a site in the coding region can be partially synonymous and partially nonsynonymous,  $N$  is often not an integer, but for convenience we have used integers.

<sup>b</sup> Computed by using  $1/(K_{13} - K_{23})$  as weight.

\*  $P < 0.05$  (if the mean  $\geq 2$  times the standard error).

\*\*  $P < 0.01$  (if the mean  $\geq 2.7$  times standard error).

In the above test, we have used only the insulin gene sequences. We now also include other sequences. Since Easteal (1991) claims that a hominoid slowdown is observed only in the  $\eta$ -globin pseudogene sequence, we shall exclude this sequence from analysis. One way to compute the mean and variance of  $K_{13} - K_{23}$  when all available sequences are considered together is to combine the sequences in tandem



and treat them as a single sequence. In the present case, however, this approach is not applicable, because the reference species (i.e., species 3) used is not the same for all sequences. We have therefore computed the "overall" mean and variance by using the reciprocal of the variance of  $K_{13} - K_{23}$  as weight. This test is more conservative, since it tends to have a smaller chance of rejecting the null hypothesis of equal rate. Even so, it is clear from table 3 that, when all the sequences used are considered together, the substitution rate in the OW monkey lineage is significantly higher than the rates in the human and chimpanzee lineages (see bottom of table 3). Thus, even with the exclusion of the  $\eta$ -globin pseudogene sequence, there is evidence for the hominoid-slowdown hypothesis. Note that the sequences used are the same as those used by Easteal (1991), except that we have added the flanking sequences and the introns of the insulin gene.

## Discussion

The controversy over the molecular-clock hypothesis has been continuing unabated since the time of its proposal by Zuckerkandl and Pauling (1965). One major difficulty in resolving this controversy is that molecular changes are subject to strong stochastic effects, so that many data are needed to show a significant difference in rate. In Easteal's (1991) study, the molecular-clock hypothesis could not be rejected if the  $\eta$ -globin pseudogene was excluded from comparison. Now, with the addition of new data, the molecular-clock hypothesis can be rejected. Therefore, the hominoid slowdown is not unique to the  $\eta$ -globin pseudogene but seems to be a rather general phenomenon in the average sense. However, the fact that  $K_{13} - K_{23}$  is negative in many cases in table 3 suggests that some sequences have indeed evolved faster in the hominoid lineages than in the OW lineage, and so the hominoid slowdown probably does not occur in all genes.

In the present study a slower rate in the human lineage than in the OW monkey lineage has been interpreted as a slowdown in the former lineage rather than as a speedup in the latter lineage, because the rate in the human lineage also appears to be slower than those in other mammalian lineages such as rodents and artiodactyls (see Britten 1986; Li et al. 1987). Also, this interpretation fits better the generation-time-effect hypothesis (Kohne et al. 1972), which postulates that the rate of nucleotide substitution is higher in short-living organisms than in long-living ones because the number of DNA replications per unit time (and so also the rate of mutation) in the germ line would be higher in the former than in the latter.

As nucleotide substitutions are subject to strong stochastic effects, the rate difference between two lineages is likely to vary from sequence to sequence. Therefore, many data may be required for obtaining a reliable estimate of the rate difference between two lineages. For the sequences used to compare the rates in the OW monkey lineage and the human lineage, the  $K$  values for the two lineages are  $(9.3\% + 1.1\%) / 2 = 5.2\%$  and  $9.3\% - 5.2\% = 4.1\%$ , respectively; therefore, the OW monkey lineage has, on average, evolved 1.3 ( $=5.2/4.1$ ) times faster than the human lineages. For the sequences used to compare the rates in the OW monkey lineage and the chimpanzee lineage, the  $K$  values are 4.75% and 3.25%, respectively, and so the former lineage has evolved 1.5 times faster than the latter lineage. This ratio of 1.5 is actually the same as the ratio between the two lineages for the  $\eta$ -globin pseudogene ( $\sim 10,000$  bp), a ratio which can be computed from the  $K$  values in Easteal's (1991) study. Note that this ratio refers to the ratio of the average rates along each lineage from the common ancestor of the two lineages to the present time. Since at the early stage of divergence

the two lineages would have very similar rates of nucleotide substitution, the rate ratio at the present time would be considerably higher than the long-term average. Therefore, the ratio at the present time may be of the same order as that between the generation times in OW monkeys and humans, which is probably between two and three.

### Acknowledgments

This study was supported by the Howard Hughes Medical Institute and by NIH grants DK20595 (to G.I.B.) and GM30998 and GM45670 (to W.-H.L.).

### LITERATURE CITED

- BELL, G. I., S. HORITA, and J. H. KARAM. 1984. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* **33**:176–183.
- BELL, G. I., J. H. KARAM, and W. J. RUTTER. 1981. Polymorphic DNA region adjacent to the 5' end of the human insulin gene. *Proc. Natl. Acad. Sci. USA* **78**:5759–5763.
- BELL, G. I., R. L. PICTET, W. J. RUTTER, B. CORDELL, E. TISCHER, and H. M. GOODMAN. 1980. Sequence of the human insulin gene. *Nature* **284**:26–32.
- BELL, G. I., M. J. SELBY, and W. J. RUTTER. 1982. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequencing. *Nature* **295**:31–35.
- . 1982. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequencing. *Nature* **296**:585.
- BRITTEN, R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**:1393–1398.
- CACCONE, A., and J. R. POWELL. 1989. DNA divergence among hominoids. *Evolution* **43**:925–942.
- EASTEAL, S. 1991. The relative rate of cDNA evolution in primates. *Mol. Biol. Evol.* **8**:115–127.
- GOODMAN, M. 1961. The role of immunochemical differences in the phyletic development of human behavior. *Hum. Biol.* **33**:131–162.
- GOODMAN, M., J. BARNABAS, G. MATSUDA, and G. W. MOOR. 1971. Molecular evolution in the descent of man. *Nature* **233**:604–613.
- KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111–120.
- KOHNE, D. E., J. A. CHISCON, and B. H. HOYER. 1972. Evolution of primate DNA sequences. *J. Hum. Evol.* **1**:627–644.
- LI, W.-H., and M. TANIMURA. 1987. The molecular clock runs more slowly in man than in apes and monkeys. *Nature* **326**:93–96.
- 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. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**:150–174.
- MANIATIS, T., E. F. FRITSCH, and J. SAMBROOK. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- SANGER, F., A. R. COULSON, B. G. BARRELL, A. J. H. SMITH, and B. A. ROE. 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* **143**:161–178.
- SARICH, V. M., and A. C. WILSON. 1967. Immunological time scale for hominid evolution. *Science* **158**:1200–1203.
- SEINO, S., D. F. STEINER, and G. I. BELL. 1987. Sequence of a New World primate insulin having low biological potency and immunoreactivity. *Proc. Natl. Acad. Sci. USA* **84**:7423–7427.

- 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. Hybridization evidence of hominoid phylogeny: results from an expanded data set. *J. Mol. Evol.* **26**:99-121.
- SOARES, M. B., E. SCHON, A. HENDERSON, S. K. KARATHANASIS, R. CATE, S. ZEITLIN, J. CHIRGWIN, and A. EFSTRATIADIS. 1985. RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. *Mol. Cell. Biol.* **5**:2090-2103.
- STEINER, D. F., S. J. CHAN, J. M. WELSH, and S. C. M. KWOK. 1985. Structure and evolution of the insulin gene. *Annu. Rev. Genet.* **19**:463-484.
- ULLRICH, A., T. J. DULL, A. GRAY, J. BROSIUS, and I. SURES. 1980. Genetic variation in the human insulin gene. *Science* **209**:612-615.
- WETEKAM, W., J. GRONEBERG, M. LEINWEBER, F. WENGENMAYER, and E. L. WINNACKER. 1982. The nucleotide sequence of cDNA coding for preproinsulin from the primate *Macaca fascicularis*. *Gene* **19**:179-183.
- WILSON, A. C., S. S. CARLSON, and T. J. WHITE. 1977. Biochemical evolution. *Annu. Rev. Biochem.* **46**:573-639.
- WU, C.-I., and W.-H. LI. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA* **82**:1741-1745.
- ZUCKERKANDL, E., and L. PAULING. 1965. Evolutionary divergence and convergence in proteins. Pp. 97-166 in V. BRYSON and H. J. VOGEL, eds. *Evolving genes and proteins*. Academic Press, New York.

WESLEY M. BROWN, reviewing editor

Received June 27, 1991; revision received September 17, 1991

Accepted September 17, 1991