Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 7622–7626, October 1988
Evolution

Rate constancy of globin gene evolution in placental mammals

(molecular clock/neutral theory/molecular phylogeny)

SIMON EASTEAL

Human Genetics Group, Division of Clinical Sciences, John Curtin School of Medical Research, The Australian National University, G.P.O. Box 334, Canberra City, Australian Capital Territory 2601, Australia

Communicated by Douglas E. Yen, June 27, 1988 (received for review March 28, 1988)

ABSTRACT **The molecular clock hypothesis is investigated by comparison of the rates of nucleotide substitution in globin genes of mice, cows and goats, humans, and rabbits, using the relative rate test. These comparisons are based on a branching order of genes and species established by cladistic analysis of nucleotide sequences. The species branching order is shown to be mouse, cow/goat, human, and rabbit. Relative rate tests involving paralogous and orthologous genes provide no evidence of heterogeneity, among species, in the rate of evolution of the genes. This result is discrepant with the conclusions of most other recent, similar studies. By comparison with previous studies, the present study is based on a sound phylogeny and involves a larger sample of species, genes, and genic regions. The result provides strong support for the neutral theory of molecular evolution and demonstrates that molecular evolutionary rate does not depend on generation time.**

The question of rate constancy in molecular evolution, the molecular clock hypothesis, has been at issue for more than two decades (1–3). The question is of interest for two main reasons.

(*i*) It is the basis of a critical test of the importance of positive natural selection in molecular evolution. If positive selection is a major force affecting the rate of molecular evolution, then variation in evolutionary rate will be the general rule, reflecting the differential effects of selection in time and among lineages; on the other hand, a general pattern of rate constancy could only result if positive selection is of little or no importance in molecular evolution (3).

(*ii*) If the rate of molecular evolution were constant and could be determined by calibration with the fossil record, then the divergence times of taxa, whose evolutionary history is not well documented in the fossil record, could be estimated from the degree of difference between their genes and proteins (2, 4).

The general conclusion, particularly from recent studies of mammalian genes and proteins, using four different approaches to resolving the issue, has been that the rate of molecular evolution differs among mammalian lineages (5–15), although the validity of this conclusion has been questioned (4, 16–18). A slowdown of the rate of molecular evolution in the human lineage has been particularly emphasized (7, 10, 11, 13–15).

The method of choice in evaluating the molecular clock hypothesis is the relative rate test (10, 14, 15). This test involves comparison of the rates of nucleotide or amino acid change between two sequences relative to a third, reference sequence. Unlike other tests, it does not depend on fossil-record-derived divergence times, which may not be accurate, and it can be applied to paralogous genes. Use of the test does, however, require a knowledge of the relative divergence order of the compared sequences to ensure that the reference sequence is the most distantly related.

In this paper, relative rate tests of orthologous and paralogous genes are applied to functional globin genes in placental mammals. The size of the data set is increased beyond that used in previous studies to include the genes of rabbits, cows and goats, as well as humans and mice. The analysis thus involves genes from species belonging to four mammalian orders: Rodentia (mice), Artiodactyla (cows and goats), Primates (humans), Lagomorpha (rabbits). The relative rate tests are based on branching orders of the species and genes determined by cladistic analysis of aligned nucleotide sequences.

MATERIALS AND METHODS

All nucleotide sequences were obtained from either the GENBANK (1987, release 48) or EMBL (1987, release 13) data base. Sequences were aligned by using an iterative, multiway method, contained in a computer package, ALIGN, written by D. Smith (The Australian National University). In this method, which is similar to those described by Hogeweg and Hesper (19) and Feng and Doolittle (20), all pairwise alignments are made of a group of sequences. A phenetic tree of the sequences, based on their pairwise similarities, is then constructed. The alignments are examined for inconsistencies in the positioning of gaps between sequence pairs. The alignments are adjusted to remove any such inconsistencies, based on the criterion that the positioning of gaps is most reliably determined by the alignment of the most similar sequences. After one complete round of such adjustment, another tree is constructed and adjustments are made to remove any further gap inconsistencies. This process of tree construction and gap position adjustment is repeated until all inconsistencies have been removed and a completely stable tree results. This usually occurs after two or three iterations. The advantages of using iterative, multiway alignment methods over the intuitive integration of pairwise alignments have been discussed by Feng and Doolittle (20).

Genic cladograms were constructed by using the DNAMETRO and DNAPENNY algorithms from J. Felsenstein's (University of Washington, Seattle) phylogenetic reconstruction package PHYLIP (version 3.0). These algorithms use Metropolis annealing (21) and branch and bound (22) methods, respectively, to identify the most parsimonious trees. Maximum parsimony is most reliably determined by using DNAPENNY; however, constraints of computer time made its use prohibitive in comparisons of more than nine sequences. The reliability of branching order estimates was determined by the bootstrap method (23), using the DNABOOT algorithm of PHYLIP.

The rates of synonymous and nonsynonymous substitutions between the coding regions of genes were estimated by the method of Li, Wu, and Luo (24). The rates of nucleotide substitution in noncoding regions were corrected for multiple

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

```
                    10        20        30        40        50        60        70        80        90       100       110
MOUSE    βMAJ  CATATAAGGTGAGGTAGGATCAGTTGCT-CCTACATTT-GCTTCTGACATAGTTGTGTTGA----------CTCACAA-CCCCAGAAACAGACATC GTTGGTATCCAGGTTACA
COW      β     CATAAAAGGAAGAGCTGGGCCAGCTGCTGCTTACACTT-GCTTCTGACACAACCGTGTTCA----------CTAGCAA-CTACACAAACAGACACC GTAGGTATCCCACTTACA
HUMAN    β     CATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTT-GCTTCTGACACAACTGTGTTCA----------CTAGCAA-C--CTCAAACAGACACC GTTGGTATCAAGGTTACA
RABBIT   β1    CATAAAAGGCAGAGCAGGG-CAGCTGCTGCTTACACTT-GCTTTTGACACAACTGTGTTTA----------CTTGCAATCCCCCAAAACAGACAGA GTTGGTATCCTTTTTACA
MOUSE    γ2    ATAAAAGGCCACCA-CTTCTAGCAGCAGTACGTACTTCGCTTCTGACACTCCTGTGATCA----------CCAGCAACCTCCCAGACTTGCCATC GTAAG----AAT------
GOAT     ε1    AATAAAAGGCCACAG-CATCCAGCAGCAGCACAGACTT-GCTTCTGATGCTTCTGTGATCA----------CCTGTAAGCTCCACGACTTGACATC GTAGA----AAGTGG---
HUMAN    ε     AATAAAAGGCCAGAC-AGAGAGGCAGCAGCACATATCT-GCTTCCGACACAGCTGCAATCA----------CTAGCAAGCTCTCAGGCCTGGCATC GTAAG----CATTGGTTC
RABBIT   β4    AATAAAAGGCCAAGC-CTTGAAGCAGCAGCACAAAGCT-GCTTCTGACACATTTGTGATCGATCA------CCAGCAAGCTCCCAGACGTGACACC GTAAG----TATTGGTTC
MOUSE    βH1   AATAAAAGGACAGGT-CTTCAGCCTCTTGAACATTCTG-GCTTTTG-CACACTTGAGATCA----------TCTCCAAGCTTCTAGACCTCACACC GTAAG-------------
GOAT     ε2    AATAAAAGGCCATGC-AGTGAAGCAGCGGCACAGACTT-GCTTCTGGCCCATTATGGATCA----------CCAGTAAGCTCCCA-----GACACC GTAAG----CAG------
HUMAN    Gγ    AATAAAAGGAAGCAC-CCTTCAGCAGTTCCACACACTC-GCTTCTGGAACGTCTGAGGTTA----------TCAATAAGCTCCTAGTCCAGACGCC GTAGG--CTCTGGTGACC
RABBIT   β3    AATAAAAGGACGAGC-CTTAGAGCAGTTTCACATACTT-GCTTCTGAGACATCTGAGACTA-TCAGCAAGCTCAGCGAGCTCCTAGACCAGACATC GTAGG--CCCTGGGGTCC

                   120       130       140       150       160       170       180       190       200       210       220
MOUSE    βMAJ  AGGCAGCT--------CACAAGAAGAAGTTG-GGTGCTTGGAGACAG----------AG--GTCT-----GCTTTCCAGCAGACACTAACTTTCAGTGT-CCCCT-GTCTATGT
COW      β     AGGCAGGTTTAAGGAGAGTGAAATGCACCTG-GGCGTGTGAGGACAG----------AGCCGTC-C-CTGAGATTCAGAGAGCTGCTGGCTTCCCTCTGA-C-CTTGTGCTGTTT
HUMAN    β     AGACAGGTTTAAGGAGA-CCAATAGAAACTG-GGCATGTGGAGACAG----------AGAAGACTC-TTGGGTTTCTGATAGGCACTGACTCTCTCTGC-CTATTGGTCTATTT
RABBIT   β1    GCACAAC-TTAATGAGA-CAGATAGAAACTG-GTCTTGTAGAAACAG----------AGTAGTCGC-CTGCTTTTCTGCCAGGTGCTGACT---TCTCT-CCCCTGGGCTGTTT
MOUSE    γ2    ----------------------TGTTTTTTAG---TGTACAAAGAGCT--------AGGGAAA---TCAAGAATTCTGAGGCTCCCTTATATTCT-ACTTGTCTTTCTC-----
GOAT     ε1    --ACTTCATGGGGGAGG--ATGGTGAATATGAGCCTGGCA-AATCGGCC--------AGAAAAATTCTTCAAAA--ATCTGAGTTGCTGATTTTCCATCTGCTATGTTTC-----
HUMAN    ε     TCAATGCATGGGAATGA--AGGGTGAATATTACCCTAGCA-AGTTGATT--------GG-GAAAGTCCTCAAGATTTTTTGCATCTCTAATTTTGTATCTGATATGGTGT-----
RABBIT   β4    TCAATACTTGGTAGAGA--AAGGCAAATATGACCCTGGTA-GACTGACC--------AG-GAAATTGCTCAA-AGTTTTTGTATCTCTGATTTTCTATTTGCTGTTGTCC-----
MOUSE    βH1   -----GAATGGAGGGAA--ATTATCCTTAT--GCATGGCAGAAATTTCC--------AGGG--TTTCTATAGGGTTTTGTGGCACACTGATTTTTTAACTGCTAATGCAC-----
GOAT     ε2    TGGACACAGGTAGGAGA--GGAGTG-----TACAAAGGCTGAAAGTGTTCCAGAAAAGAGGGACTGGTTAGGTTTCTTACATACTCTGACTTCTTATCTGTTCTGTGACTATGA
HUMAN    Gγ    AGGACA-AGGGAGGGAA--GGAAGGACCCTGTGCCTGGCAAAA----GTCC-----AGGTCGCTTCTCAGGATTTGTGGCACCTTCTGACTGTCAAACTGTTCTTG-TCAAT--
RABBIT   β3    AGGACA-AGGCAGAGGA--GGAATGAAACTGAGCCTGGCAGGA----CTCC-----AGGCCCTTTCTCAGGACTTGTGAAGCTCTCTGATACTCCCACCACTATTGTTCTGT--

                  230       240       250       260       270       280       290       300       310       320       330       340
MOUSE    βMAJ  TTCCCTTTTTAG ACCCCCTTTCCTGCTCTTGCCTGT-GAACAA-TGGTTAATTGTTCCCAAG-AGAGCATCTGTCAGTTGTTGGCAAAATGAT--AGACATTTGAAAATCTGTC
COW      β     TCTC-CCCCTAG GCTCCCTTTCCTGC-------TTTCCAGGAA-AGGTTTTTTCATCCTCAG-AGCCCAAAGATTGAAT-ATGGAAAAATTATGAAGTGTTTTGAGCATCTGGC
HUMAN    β     TCCCACCCTTAG GCTCGCTTTCTTGCTGTCCAATTTCTATTAA-AGGTTCCTTTGTTCCCTA-AGTCCAACTACTAAAC-TGGGGGATATTATGAAGGGCCTTGAGCATCTGGA
RABBIT   β1    TCATTTTCTCAG GATCTTTTTCCCTCTGCCAA----------------------------------------AAATTATGGGGACATCATGAAGCCCCTTGAGCATCTGAC
MOUSE    γ2    --CACTTACCAG GCCCTCTCTCTAGCTGTCCAGCAATCCTGT--GTGTCCGCTATGCCTCCTCTCTGCA---CATGAATACTG---CTGTTCCTGA-A-----------GCACA
GOAT     ε1    --CATCTCATAG ATTCTCT-TTACA-ATTCACCATTTTGT------GTCCCCAGTGCCTTCCTTCTGCC--CCTTGGGACTGGGGTTTGGCCTTGTGA-----------ACCCA
HUMAN    ε     --CATTTCATAG GTTCTCT-TCCAGTTTGCAGGTGTTCCTGT----GACCCTGACACCCTCCTTCTGCA---CATGGGGACTGGGCTTGGCCTTGAGA-----------GAAAG
RABBIT   β4    --CATCATATAG ACTCTCT-TCCCAGTTTCTTTAGTGCCCCTACTCATCCCCAGCGTCCTCCTTCTGTA--CCTTGGGGACTAGGCTCAGCCTTGATG-----------GCACA
MOUSE    βH1   --TATGTCTTAG ACTCCAT-TCCAGTACACTGGCAATCCCAT--GTGTCTATGATGCCCT-CTTTGACT--CCATGGGGACTGAATTAGGCATTGAGA-----------GCACA
GOAT     ε2    T-CATCCCATAG TTGCCTGGCCTACCATGCTGG--TGCCTATCTGAAGGCCCAGTGTCCCAGAAGTTCATCTCCTGAAGACAGAG--------GGAGA-----------GAGCT
HUMAN    Gγ    ----CTC-ACAG GCTCACTGCCCATGATGCAGAGCTTTCAAGGATAGGCTTTATTCTGCAAGCAAT--A---CA--AATAATAAATCTATTCTGCTAA-----------GAGAT
RABBIT   β3    ----CTCTATAG ACCCCTTCACTGTAGGACAGAGCTTCTAGGAAGAAGCTTTATCCCTCAAATAAT--AATGAA--AATAATAAAACTACTCTAAGAA-----------ATTAT
```

FIG. 1. Multiway alignment of 5′ (nucleotides 1–96) and 3′ (nucleotides 241–343) noncoding and intron 1 (nucleotides 97–240) sequences for β-like globin genes used in the cladistic analysis. A gap penalty of 2.5 was used in the alignment. Gaps are indicated by dashes.

substitution by Kimura's (3) method, with transition-transversion ratios determined directly from the aligned sequences. Rates of substitution in different lineages were compared by the relative rate test method of Wu and Li (10).

RESULTS

All available sequences for functional and pseudo β-like globin genes in mice, cows, goats, rabbits, and humans were aligned. Intron 2 was omitted as it was too long to permit multiway alignment using the available computer facilities. The phenograms constructed during the sequence alignment confirmed the results of previous studies in showing that the following pairs of genes diverged from each other after the divergence of the different mammalian orders: mouse β^{maj}, β^{min}; mouse β_H0, β_H1; cow β, γ; cow ε_1, ε_3; human Gγ, Aγ; human β, δ. Only one gene from each of these pairs [mouse β^{maj}, mouse β_H1, cow β, goat (in place of cow) ε_1, human Gγ, human β] was included in the cladistic and relative rate analyses. In all, these analyses involved 12 genes, 3 from each of the four mammalian orders. The alignments of noncoding sequences used in the cladistic analysis are shown in Fig. 1.

It was not possible to construct a reliable α-globin-rooted cladogram of the 12 genes, probably because of the low degree of similarity between the α- and β-like globin genes. Different α-globin genes selected from each of the four orders, when used as out-groups, indicated different root positions, and in all cases a low percentage of bootstraps gave the consensus tree.

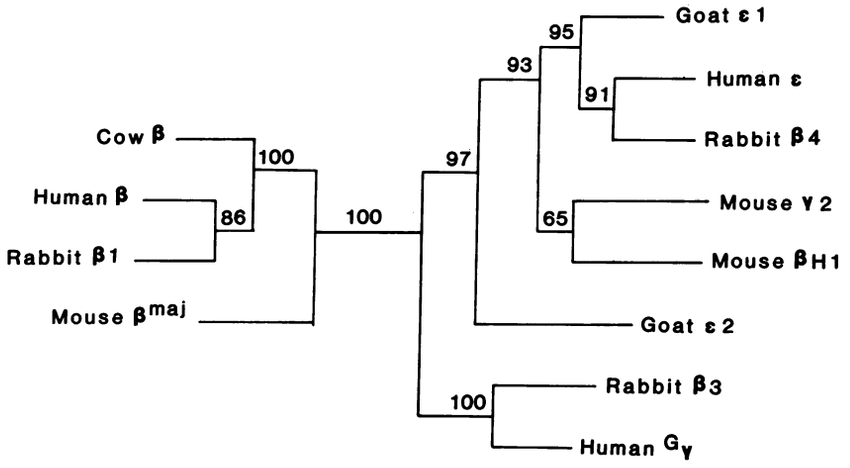

FIG. 2. Single most parsimonious cladogram of representative functional β-globin-like genes from species belonging to four orders of placental mammals. The tree is unrooted. Horizontal branch lengths are drawn in proportion to the numbers of substitutions occurring along the branches. The compared sequences extend from the start of the TATAA box to a position ≈10 bases 5′ to the poly(A) site, excluding intron 2. The numbers shown at each branch are the percentage of bootstraps for which the clustering of the genes above that branch in the tree occurred, based on 100 bootstraps. The sources of the sequences are as follows: mouse β^{maj} (25), cow β (26), human β (27), rabbit β_1 (28), mouse γ_2 (29), mouse β_H1 (30), goat ε_1 and ε_2 (31), human γ (32), human ε (33), rabbit β_3 (34), rabbit β_4 (35).

The unrooted cladogram of the 12 genes is shown in Fig. 2, with the horizontal branch lengths drawn in proportion to the numbers of substitutions occurring along the branches. The percentage of bootstraps (based on 100 bootstraps) that result in the various pairings and clusterings of genes is shown at the branches of the tree. With the exception of the mouse γ_2–β_H1 branch, all branches occur in >85% of the bootstraps. Although an accurate root position was not obtained for the cladogram, the α-globin out-groups indicated that the root lies between a monophyletic group comprising the adult-expressed genes, shown branching to the left of the tree, and another monophyletic group comprising goat ε_1, human ε, rabbit β_4, mouse γ_2, and mouse β_H1.

These two monophyletic groups of genes indicate, with high probabilities, a branching order of mouse, goat/cow, rabbit, and human. This is also the single most parsimonious branching order found by a similar analysis of the coding regions of the α-globin genes (36–39), using the mouse ζ-globin (40) as an out-group. The percentage of bootstraps showing this species' branching order for these three sets of genes combined, end-to-end (using a combination of mouse γ_2, human β, and mouse ζ genes as out-groups for the adult β-like, embryonic β-like, and α-globin genes, respectively), is 100% for the mouse divergence and 89% for the cow/goat divergence (Fig. 3).

The branch lengths in Fig. 3, drawn in proportion to numbers of substitutions, show that the times between divergences, particularly between the mouse and goat/cow divergences, are long, relative to the total length of the tree. This demonstrates the invalidity of the previously assumed "star phylogeny" of the orders to which these species belong. The branching order of the species, particularly the early divergence of rodents, is different from that assumed in previous relative rate tests (10, 15).

In the previous relative rate test involving paralogous β-like globin genes (10) it was assumed that mouse β_H1 was orthologous to human $G\gamma$. Fig. 2 shows this assumption to be invalid.

Thirty-seven relative rate tests were performed, involving regions of orthologous goat/cow, rabbit, and human genes, with reference genes selected on the basis of the relationships shown in Fig. 2 (Table 1). Rates of nucleotide substitution in intron 2 are based on pairwise alignments, since multiway alignment was not possible in this region. In only two of these tests are there significant differences (at the 5% level) in substitution rate between genic regions. Among all tests, no one lineage has a predominantly faster rate of evolution, and, when all genic regions are combined, there are no significant differences in evolutionary rate between the lineages (Table 1).

The number of nucleotide substitutions between adult and non-adult β-like globin genes is shown in Tables 2 and 3. For the adult genes, there is little difference, among species, for any of the genic regions in the average values for comparisons

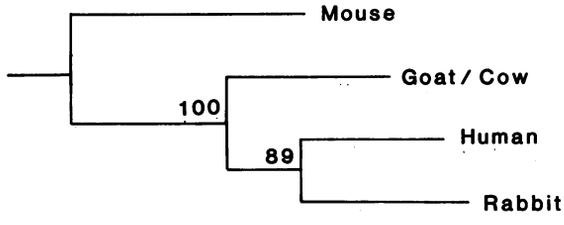

FIG. 3. Single most parsimonious cladogram of combined α, adult-expressed β-like, and four of the five monophyletic non-adult-expressed β-like (mouse γ_2, goat ε_1, human ε, rabbit β_4) genes. The sequences compared were a combination of α-, β- and ε-globin genes, aligned end-to-end; the combined out-group consisted of mouse ζ, mouse γ_2, and human β.

Table 1. Comparisons of the numbers of substitutions per 100 nucleotides between orthologous, mammalian globin genes using the relative rate test (9, 13)

	R (1), H (2) vs. C/G (3)		C/G (1), H (2) vs. M (3)		R (1), C/G (2) vs. M (3)	
Gene	K_{12}	K_{13}–K_{23}	K_{12}	K_{13}–K_{23}	K_{12}	K_{13}–K_{23}
			Exon replacement sites (n = 310–347)			
β	5	1 ± 2	8	3 ± 2	9	−3 ± 2
α	11	2 ± 2	9	0 ± 2	10	2 ± 2
ε	8	3 ± 2	5	−2 ± 2	8	3 ± 2
γ	12	2 ± 2				
			Exon silent sites (n = 93–106)			
β	36	1 ± 8	37	8 ± 10	38	3 ± 11
α	31	−5 ± 8	38	−18 ± 13	35	7 ± 12
ε	55	10 ± 13	46	20 ± 12	56	−10 ± 13
γ	57	38 ± 16*				
			Intron 1 and flanking sequences (n = 262–309)			
β	36	6 ± 5	33	2 ± 6	39	1 ± 6
ε	41	−8 ± 6	51	0 ± 8	43	6 ± 8
γ	39	0 ± 14				
			Intron 2 (n = 512–757)			
β	32	0 ± 3	40	−2 ± 4	51	2 ± 4
ε	34	7 ± 3*	39	4 ± 4	48	6 ± 4
γ	44	0 ± 5				
			Introns 1 and 2 (n = 149–204)			
α	41	11 ± 6	39	5 ± 7	32	6 ± 8
Total	29	2 ± 3	30	1 ± 3	30	1 ± 3

K_{ij} is the number of substitutions between species i and j. The values of i and j identifying particular species are shown in parentheses after the species name. n, Number of nucleotides compared. The relative rate test involves a comparison (K_{13}–K_{23}) of the number of substitutions between species 1 and the reference species (species 3) and between species 2 and the reference species. The β genes are mouse (M) β^{maj}, cow (C) β, human (H) β, and rabbit (R) β_1. The ε genes are mouse γ_2, goat (G) ε_1, human ε, and rabbit β_4. The γ genes are human $G\gamma$ and rabbit β_3 (compared with goat ε_1).
*Significant at the 5% level.

with all eight non-adult genes. Furthermore, the order of evolutionary rate among species is not consistent across genic regions, and there is no indication that the human gene is evolving more slowly than the genes of the other species. In fact, in none of the genic regions is the human adult gene the slowest evolving of the adult genes.

For the non-adult genes, with the exception of the exon silent sites, there is, again, little difference among genes in the mean numbers of substitutions for comparisons with the adult genes. As in the case of the adult genes, there is no consistent pattern in the order of evolutionary rate among genes and no evidence for a slower rate of evolution in human genes. There is also no consistency in evolutionary rate between the genes of the same species. Thus, for example, for the exon replacement sites, goat ε_1 is the slowest evolving gene and goat ε_2 is one of the three fastest evolving genes. Similarly, for intron 2, rabbit β_3 is the fastest evolving gene and rabbit β_4 is the slowest evolving gene. This suggests that any differences between the genes are not the result of a species-specific factor such as generation time or global mutation rate.

In the exon silent sites there is more variation among genes than there is in the other regions, with the human genes appearing to be the slowest evolving. Some of this extra variation may be due to sampling error as fewer nucleotides are compared in this region than in the other regions.

Table 2. Numbers of substitutions per 100 nucleotides in exons of adult- and non-adult-expressed genes of the β-globin gene family in different mammalian orders

Non-adult gene	Adult gene				Mean
	M β^{maj}	C β	H β	R β	
	Replacement sites (n = 310–347)				
M β_H1	20	20	20	17	19
M γ_2	18	18	17	15	17
G ε_1	13	16	14	12	14
G ε_2	19	19	20	19	19
H ε	16	17	15	15	16
H γ	18	16	17	17	17
R β_3	19	20	18	17	19
R β_4	18	19	16	16	17
Mean	18	18	17	16	
	Silent sites (n = 93–106)				
M β_H1	89	84	107	97	94
M γ_2	99	88	91	80	90
G ε_1	80	68	77	55	70
G ε_2	116	102	102	130	113
H ε	67	62	64	56	62
H γ	69	60	75	52	64
R β_3	89	100	112	101	101
R β_4	92	77	85	83	84
Mean	87	80	89	82	

n, Number of nucleotides compared. Species are designated by the letters M for mouse, G for goat, C for cow, H for human, and R for rabbit.

The number of possible relative rate tests between adult genes, using non-adult genes as reference genes, and vice versa, is prohibitively large. Since there has been particular interest in the relative evolutionary rates of human and mouse genes (10), only tests involving the genes of these species were performed. The rates of goat/cow, human, and rabbit genes have already been compared. Of 32 tests conducted (Table 4), only 2 showed a significant difference (at the 5% level) in evolutionary rate between human and mouse genic regions. When all genic regions that were compared using a particular reference species are combined, there are no significant differences between the species.

In all, there are four significant differences in 69 tests. There is one case of a human genic region evolving faster than a mouse region and three cases of human regions evolving more slowly than either mouse or rabbit regions. Two of these significant differences were found in comparisons with only one of the four reference species used. This number of significant differences can be explained as resulting by chance. There is thus no evidence in the rate of nucleotide change among the four orders of placental mammals.

Table 3. Numbers of substitutions per 100 nucleotides in introns of adult- and non-adult-expressed genes of the β-globin gene family in different mammalian orders

Non-adult gene	Adult gene				Mean
	M β^{maj}	C β	H β	R β_1	
	Intron 1 and flanking sequences (n = 227–308)				
M β_H1	159	131	140	124	138
M γ_2	141	113	106	98	114
G ε_1	143	131	110	110	124
G ε_2	129	110	99	107	111
H ε	151	153	135	112	138
H γ	133	102	108	90	108
R β_3	143	93	106	94	109
R β_4	165	153	135	107	140
Mean	146	123	117	105	
	Intron 2 (n = 463–757)				
M β_H1	47	62	55	53	54
M γ_2	63	51	55	47	54
G ε_1	51	57	57	57	56
G ε_2	53	55	59	53	55
H ε	51	53	62	51	54
H γ	51	57	51	55	54
R β_3	53	64	57	57	58
R β_4	57	49	55	47	52
Mean	53	56	56	53	

n, Number of nucleotides being compared. Species are designated by the letters M for mouse, G for goat, C for cow, H for human, and R for rabbit.

DISCUSSION

The conclusion drawn from the results presented here is that genes are evolving at the same rate in the different mammalian orders. This conclusion is in sharp contrast to those drawn from most other recent studies (5–15). There are a number of reasons for this.

Previous orthologous and paralogous relative rate tests have assumed branching orders of genes or species that appear to be incorrect from the cladistic analysis presented here. In the case of orthologous genes, the branching order and relative divergence times, indicated by the cladistic analysis, are those previously proposed (17) to account for the differences in evolutionary distance between the species (10), under the assumption of rate constancy.

The species branching order obtained from the cladistic analysis presented here might be incorrect if there were large differences in evolutionary rate among branches (41), and this could affect the conclusions drawn from the orthologous sequence comparisons. Thus, for instance, the data might be consistent with a branching order of artiodactylids, primates, rodents, and lagomorphs (42), with a greater rate of evolution in the rodent lineage. However such a scheme seems unlikely since no difference in rate between rodents and the other

Table 4. Comparisons of the numbers of substitutions per 100 nucleotides between paralogous β-globin-like genes in humans and mice using the relative rate test

Rate-tested genes	K_{12}	K_{13}–K_{23}			
		M (3)	C/G (3)	H (3)	R (3)
	Exon replacement sites				
M β, H β	12	0 ± 2	−1 ± 2	1 ± 2	2 ± 2
M γ_2, H ε	9	2 ± 2	1 ± 2	2 ± 2	0 ± 2
	Exon silent sites				
M β, H β	50	−18 ± 19	3 ± 14	3 ± 12	7 ± 15
M γ_2, H ε	57	32 ± 17	26 ± 15	27 ± 14	24 ± 13
	Intron 1 and flanking sequences				
M β, H β	51	19 ± 40	33 ± 27	16 ± 34	30 ± 42
M γ_2, H ε	37	−10 ± 37	−41 ± 33	−29 ± 9	−15 ± 18
	Intron 2				
M β, H β	51	−8 ± 5	−6 ± 5	−11 ± 5*	2 ± 5
M γ_2, H ε	48	12 ± 5*	−2 ± 5	−7 ± 5	−4 ± 5
Total	40	1 ± 3	0 ± 3	0 ± 3	1 ± 3

K_{ij} is the number of substitutions between species i and j. For the rate-tested genes, mouse (M) is always species 1 and human (H) is species 2; (3) denotes the reference species. The following reference species were used: mouse β_H1, goat (G) ε_1, human ε, and rabbit (R) β_4 for mouse β–human β comparisons; mouse β, cow (C) β, human β, and rabbit β_1 for the mouse γ_2–human ε comparisons.
*Significant at the 5% level.

orders is apparent in comparisons with paralogous genes (Tables 2–4) or in comparisons with marsupial sequences (unpublished data).

The apparently incorrect assumption that mouse $\beta_H l$ and human $G\gamma$ are orthologues may account for some of the variation in evolutionary rate suggested by the previous paralogous relative rate tests involving human and mouse genes (10). In those tests, only exon silent sites were considered, and no attempt was made to evaluate whether the observed differences were significant. In only 2 of the 32 relative rate tests described here between humans and mice is there a significant difference between the two species, and in only 1 of these has the mouse genic region evolved faster than the human genic region.

The cladistic analysis suggests that the star phylogeny, assumed in comparisons of the observed and expected variance in substitution rate among lineages (3, 5, 6, 8, 9, 12), is not valid. The heterogeneity in the degree of sequence difference between species appears to be due not to variation in substitution rate but to variation in the species' divergence times. The fossil record over the period during which the main orders of placental mammals diverged is too inadequate to determine definitively the pattern of ordinal branching. The assumption of a star phylogeny is more a reflection of ignorance about branching orders than a reflection of any pattern in the data (43).

The discrepancy between the present results and those from studies in which species' divergence times are derived from the fossil record (7, 11, 13–15) can be explained in a similar way. The species' divergence times proposed in such studies are presumably based on an incorrect interpretation of the poor fossil record.

The results provide strong support for the neutral theory of molecular evolution, as it relates to nucleotides and, since exon replacement sites appear to evolve at a constant rate, proteins. The results also show that the rate of accumulation of new mutants is not affected by generation time and that the rate of molecular evolution has not slowed down in the primate lineage leading to humans. This has important implications for mammalian divergence times. The nucleotide sequences of human and ape genes are more similar than expected under the current assumptions about the divergence times of the main mammalian orders and of humans from the African apes (7, 11, 13–15). Since there has been no slowdown in the primate lineage to explain this discrepancy, either the mammalian orders diverged earlier than has been supposed or humans diverged from the African apes later than the currently accepted five to eight million years ago.

I thank J. Felsenstein, W.-H. Li, and D. Smith for making available their computer programs for phylogenetic reconstruction, estimation of silent and replacement site differences, and sequence alignment, respectively; and P. W. Easteal, D. Penny, S. W. Serjeantson, S. R. Wilson, and two anonymous reviewers for helpful comments on the manuscript.

1. Zuckerkandl, E. & Pauling, L. (1965) in *Evolving Genes and Proteins*, eds. Bryson, V. & Vogel, H. J. (Academic, New York), pp. 97–116.
2. Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639.
3. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, London).
4. Wilson, A. C., Ochman, H. & Prager, E. M. (1987) *Trends Genet.* **3**, 241–247.
5. Ohta, T. & Kimura, M. (1971) *J. Mol. Evol.* **1**, 18–25.
6. Langley, C. H. & Fitch, W. M. (1974) *J. Mol. Evol.* **3**, 161–177.
7. Goodman, M., Moore, G. W. & Matsuda, G. (1975) *Nature (London)* **253**, 603–608.
8. Gillespie, J. H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 8009–8013.
9. Gillespie, J. H. (1984) *Evolution* **38**, 1116–1129.
10. Wu, C.-I. & Li, W.-H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1741–1745.
11. Britten, R. J. (1986) *Science* **231**, 1393–1398.
12. Gillespie, J. H. (1986) *Mol. Biol. Evol.* **3**, 138–155.
13. Koop, B. F., Goodman, M., Xu, P., Chan, K. & Slightom, J. L. (1986) *Nature (London)* **319**, 234–237.
14. Li, W.-H. & Tanimura, M. (1987) *Nature (London)* **326**, 93–96.
15. Li, W.-H., Tanimura, M. & Sharp, P. M. (1987) *J. Mol. Evol.* **25**, 330–342.
16. Takahata, N. (1987) *Genetics* **116**, 169–179.
17. Easteal, S. (1985) *Mol. Biol. Evol.* **2**, 450–453.
18. Easteal, S. (1987) *Mol. Biol. Evol.* **4**, 78–80.
19. Hogeweg, P. & Hesper, B. (1984) *J. Mol. Evol.* **20**, 175–186.
20. Feng, D.-F. & Doolittle, R. F. (1987) *J. Mol. Evol.* **25**, 351–360.
21. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953) *J. Chem. Phys.* **21**, 1087–1092.
22. Hendy, M. D. & Penny, D. (1982) *Math. Biosci.* **59**, 277–290.
23. Felsenstein, J. (1985) *Evolution* **39**, 783–791.
24. Li, W.-H., Wu, C.-I. & Luo, C.-C. (1985) *Mol. Biol. Evol.* **2**, 150–174.
25. Konkel, D. A., Tilghman, S. M. & Leder, P. (1978) *Cell* **15**, 1125–1132.
26. Schimenti, J. C. & Duncan, C. H. (1984) *Nucleic Acids Res.* **12**, 1641–1655.
27. Lawn, R. M., Efstratiadis, A., O'Connell, C. & Maniatis, T. (1980) *Cell* **21**, 647–651.
28. Rohrbaugh, M. L., Johnson, J. E., James, M. D. & Hardison, R. C. (1985) *Mol. Cell. Biol.* **5**, 147–160.
29. Hansen, J. N., Konkel, D. A. & Leder, P. (1982) *J. Biol. Chem.* **257**, 1048–1052.
30. Hill, A., Hardies, S. C., Phillips, S. J., David, M. G., Hutchison, C. A. & Edgell, M. H. (1980) *J. Biol. Chem.* **259**, 3739–3747.
31. Shapiro, S. G., Schon, E. A., Townes, T. M. & Lingrel, J. B. (1983) *J. Mol. Biol.* **169**, 31–52.
32. Slightom, J. L., Blechl, A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
33. Baralle, F. E., Shoulders, C. C. & Proudfoot, N. J. (1980) *Nucleic Acids Res.* **8**, 4393–4404.
34. Hardison, R. C. (1981) *J. Biol. Chem.* **256**, 11780–11786.
35. Hardison, R. C. (1983) *J. Biol. Chem.* **258**, 8739–8744.
36. Nishioka, Y. & Leder, P. (1979) *Cell* **18**, 875–882.
37. Schon, E. A., Wernke, S. M. & Lingrel, J. B. (1982) *J. Biol. Chem.* **257**, 6825–6835.
38. Michelson, A. M. & Orkin, S. H. (1980) *Cell* **22**, 371–377.
39. Heindell, H. C., Liu, A., Paddock, G. V., Studnicka, G. M. & Salser, W. A. (1978) *Cell* **15**, 43–54.
40. Leder, A., Weir, L. & Leder, P. (1985) *Mol. Cell. Biol.* **5**, 1025–1033.
41. Felsenstein, J. (1978) *Syst. Zool.* **27**, 401–410.
42. Penny, D. & Hendy, M. (1986) *Mol. Biol. Evol.* **3**, 403–417.
43. Wyss, A. R., Novacek, M. J. & McKenna, M. C. (1987) *Mol. Biol. Evol.* **4**, 99–116.