J Mol Evol (1998) 47:249–257

JOURNAL OF MOLECULAR EVOLUTION

Accelerated Evolution of Cytochrome *b* in Simian Primates: Adaptive Evolution in Concert with Other Mitochondrial Proteins?

T. Daniel Andrews, Lars S. Jermiin, Simon Easteal

Human Genetics Group, John Curtin School of Medical Research, Australian National University, Canberra ACT 0200, Australia

Received: 15 December 1997 / Accepted: 24 February 1998

Abstract: We have sequenced the cytochrome *b* gene of Horsfield's tarsier, *Tarsius bancanus*, to complete a data set of sequences for this gene from representatives of each primate infraorder. These primate cytochrome *b* sequences were combined with those from representatives of three other mammalian orders (cat, whale, and rat) in an analysis of relative evolutionary rates. The nonsynonymous nucleotide substitution rate of the cytochrome *b* gene has increased approximately twofold along lineages leading to simian primates compared to that of the tarsier and other primate and nonprimate mammalian species. However, the rate of transversional substitutions at fourfold degenerate sites has remained uniform among all lineages. This increase in the evolutionary rate of cytochrome *b* is similar in character and magnitude to that described previously for the cytochrome *c* oxidase subunit II gene. We propose that the evolutionary rate increase observed for cytochrome *b* and cytochrome *c* oxidase subunit II may underlie an episode of coadaptive evolution of these two proteins in the mitochondria of simian primates.

Key words: Cytochrome *b* — Primate — Coevolution — Adaptive — Relative rate — Nonsynonymous — Cytochrome *c* oxidase — Mitochondrial DNA — Fourfold degenerate — Transversion

Correspondence to: T.D. Andrews; *e-mail:* d.andrews@anu.edu.au

Introduction

Analysis of evolutionary rates can be a useful approach for detecting genes and proteins that have undergone an episode of adaptive evolution or have been subjected to positive natural selection. While rates of protein evolution may be increased by elevated mutation rates, this can be distinguished from adaptive evolution through analysis of the relative frequencies of synonymous and nonsynonymous nucleotide changes. The hallmark of adaptive evolution is an increased rate of nonsynonymous nucleotide change relative to synonymous change, while increased mutation rates result in proportional increases to both synonymous and nonsynonymous nucleotide changes (Kreitman and Akashi 1995).

Analysis of the evolutionary rates of seven complete mammalian mitochondrial genomes (rat, mouse, human, seal, cow, whale, and opossum as outgroup) has shown significant variation among these lineages for the cytochrome *c* oxidase subunit I and subunit II genes and the cytochrome *b* gene (Janke et al. 1994). More detailed comparative analysis of the cytochrome *c* oxidase subunit II gene from mammals has confirmed an evolutionary rate acceleration along lineages leading to simian primates (apes including humans, Old and New World monkeys) (Adkins and Honeycutt 1994; Adkins et al. 1996; Ramharack and Deeley 1987). It appears that cytochrome *c* oxidase has undergone a nearly twofold increase in the rate of amino acid substitution relative to other primates (Adkins and Honeycutt 1994), implying an elevated rate of nonsynonymous nucleotide substitution. Related to this, the nuclear-encoded cytochrome *c* gene has undergone a period of rapid evolution in the

lineage leading to humans (Evans and Scarpulla 1988) and cytochrome *c* amino acid sequences demonstrate a higher number of replacements in lineages leading to simian primates than they do for other mammalian species (Baba et al. 1981). The nuclear-encoded cytochrome *c* oxidase subunit IV gene may have also undergone a brief elevation of its nonsynonymous substitution rate in ancestral lineages leading to apes and Old World monkeys (Wu et al. 1997).

Cytochrome *c* functions by shuttling electrons between the cytochrome *c* reductase complex (cytochrome *b* complex) and the cytochrome *c* oxidase complex and, in doing so, binds directly to subunit II of cytochrome *c* oxidase (Hatefi 1985). This intimate functional relationship between cytochrome *c* and cytochrome *c* oxidase subunit II has led to the suggestion that their correlated increase in evolutionary rates could be due to coevolution (Cann et al. 1984). This suggestion is supported by kinetic studies of *in vitro* reconstituted reactions of cytochrome *c* with cytochrome *c* oxidase, which show that the nature of the reaction in simian primates is different from that of other mammals, including strepsirhine primates (Osheroff et al. 1983).

If adaptive coevolution has occurred in the mitochondrial electron transport chain of simian primates, the effects of this may not be restricted to cytochrome *c* and cytochrome *c* oxidase subunit II. The genes encoding other components of the electron transport chain may also exhibit unusual evolutionary rates as a result of adaptive evolution. While not coming into direct contact with either cytochrome *c* or cytochrome *c* oxidase, cytochrome *b* is a near neighbor, forming an important part of the cytochrome *c* reductase complex (Hatefi 1985). Cytochrome *b* has been shown to have an accelerated evolutionary rate in lineages leading to humans (Irwin et al. 1991; Ma et al. 1993). We have sequenced the cytochrome *b* gene of *Tarsius bancanus*, thus completing a data set representing cytochrome *b* sequences from each extant primate infraorder. The separation of tarsiers from simian primates is the earliest divergence among extant haplorhine primates (Groves 1989, p105), and hence, the tarsier sequence is important in the phylogenetic mapping of any evolutionary rate change that may have occurred among the primates. Here we present an analysis of the relative rates of evolution of cytochrome *b* across the primate order in comparison to other mammals.

Materials and Methods

DNA and Data Sources. A lung sample from Horsfield's tarsier (*Tarsius bancanus*) and a liver sample from the Philippine tarsier (*Tarsius syrichta*) were provided by the Duke University Primate Center. Isolation of pure mitochondrial DNA was precluded by the limited quantity of primary tissue and therefore DNA was phenol/chloroform extracted from homogenates of whole samples (Sambrook et al. 1989). Other cytochrome *b* nucleotide sequences were obtained from the following published sources: *Homo sapiens* (human; J01415) (Anderson *et al.* 1981), *Colobus guereza* (black-and-white colobus monkey; U38264), *Saimiri sciureus* (squirrel monkey; U38273), *Lemur catta* (ring-tail lemur; U38271) (Collura and Stewart 1995), *Galago crassicaudatus* (thick-tailed bushbaby; U53579), *Nycticebus coucang* (slow loris; U53580) (Yoder et al. 1996), *Felis catus* (domestic cat; U20753) (Lopez et al. 1996), *Balaenoptera physalus* (fin whale; X61145) (Arnason et al. 1991), *Rattus norvegicus* (rat; X14848) (Gadaleta et al. 1989), *Giraffa camelopardalis* (giraffe; X56287), *Dama dama* (fallow deer; X56290), *Camelus dromedarius* (dromedary camel; X56281), *Sus scrofa* (pig; X56295), *Equus grevyi* (Grevy's zebra; X56282), *Stenella longirostris* (spinner dolphin; X56292), *Diceros bicornis* (black rhinoceros; X56283) (Irwin et al. 1991), *Ursus maritimus* (polar bear; X82309), *Phoca groenlandica* (harp seal; X82303) (Arnason et al. 1995), *Dugong dugong* (dugong; U07564), *Oryctolagus cuniculus* (rabbit; U07566) (Irwin and Arnason 1994), and *Bos taurus* (cattle) (unpublished DDBJ/EMBL/GenBank entry d34635).

Table 1. Oligonucleotide primer sequences

Primer[a]	Nucleotide sequence
H14581	5′-CACTAAGGATCCATAAATAGGIGAAGG-3′
H15132	5′-TAGGCTAIGTICTCCCATGAGG-3′
H15359	5′-TCCAITAACCCATCAGGAAT-3′
H15542	5′-CCCCATATTAAICCAGAATGA-3′
L16016	5′-TAGTTTAAAGTAGAAGCTTAGCTTTGGG-3′
L15562	5′-TCATTCTGGITTAATATGGGG-3′
L15379	5′-GATTCCTGATGGGTTAITIGA-3′
L15186	5′-TGGCGCCTCAIAATGATATTTG-3′

[a] Primer numbers refer to the nucleotide position at the 5′ end of the oligonucleotide and are numbered after Anderson et al. (1981).

PCR and Sequencing. The tarsier cytochrome *b* gene was isolated by the polymerase chain reaction (PCR) using primers that anneal to conserved regions upstream in the ND6 gene and downstream in the tRNA-Pro gene (Table 1 and Fig. 1). Previously published primers (Irwin et al. 1991) proved to have a very short life span after synthesis, even when stored at −70°C, possibly because they are homologous to tRNA regions and can easily self-anneal. PCR products were run on 0.8% low-melting temperature agarose (FMC Bioproducts) gels and the band of the correct size was excised and cleaned using Wizard Preps (Promega). The sequences of both strands of products prepared in this way were determined using dye-terminator cycle sequencing (Perkin–Elmer) with the original PCR primers and a series of internal sequencing primers (Table 1) on an ABI377 automated sequencer (Applied Biosystems).

Data Analysis. Cytochrome *b* nucleotide sequences were aligned using CLUSTAL W (Thompson et al. 1994), without the need to insert gaps. Two subsets of the aligned sequences were used. The first 10 sequences (*Homo sapiens*, *Colobus guereza*, *Saimiri scuireus*, *Tarsius bancanus*, *Lemur catta*, *Galago crassicaudatus*, *Nycticebus coucang*, *Felis catus*, *Balaenoptera physalus*, and *Rattus norvegicus*) were used to perform an exhaustive search of tree space, generate maximum-likelihood trees, and conduct relative rate tests (see below). The second subset of sequences was used to analyze the distribution of variation along the length of an alignment that consisted of 19 mammalian species, these being representatives of mammalian orders and suborders for which complete cytochrome *b* nucleotide sequences are present in the DDBJ/EMBL/GenBank database. The sequences are those listed in the *DNA and Data Sources* section, including the three strepsirhine primate species and *Tarsius bancanus* but excluding *Homo sapiens*, *Colobus guereza*, and *Saimiri scuireus.*

The number of synonymous (K_s) and nonsynonymous (K_a) substitutions per site were calculated using the method of Wu and Li (1985),

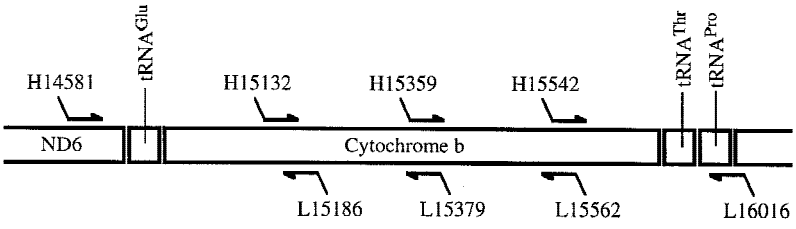

Fig. 1. Binding sites of primers (see Table 1) used in the amplification and sequencing of the *Tarsius bancanus* cytochrome *b* gene.

as modified by Li (1993) and Pamilo and Bianchi (1993) and implemented by the NewDiverge program from the GCG package (Genetics Computer Group, WI, USA).

Potential base compositional heterogeneity among the nucleotide sequences was assessed using the Distance program by L.S.J., whereby pairwise comparisons of nucleotide sequences were conducted and χ^2 values calculated to test the independence of each pair of sequences with respect to base composition. Such calculations were conducted for data sets consisting of first, second, and third codon positions and *Z* scores were calculated from the resultant χ^2 values (Smith 1986). The *Z* scores calculated for each codon position were then graphed as a frequency distribution and used to identify the most compositionally homogeneous sites in the data.

The exhaustive search of tree space for the maximum-likelihood tree of the alignment of 10 nucleotide sequences was made using the TrExML program (Wolf et al. 1998), and a standardized, exponentially weighted majority-rule consensus tree was generated using the TreeCons program (Jermiin et al. 1997). Relative-likelihood support for internal edges in the consensus tree was compared with that for internal edges in the maximum-likelihood tree and the tree representing a conventional primate phylogeny.

Relative rate tests were conducted using the method of Wu and Li (1985), with variances and *Z* scores calculated as by Muse and Weir (1992) and implemented by the K2WuLi program by L.S.J. K_{01}/K_{02} values were calculated by the method described by Li (1997, p. 217).

The spatial distribution of variability along the length of the mammalian cytochrome *b* protein was determined with the Sliding Window program by T.D.A. A sliding window of 20 residues was moved along the length of the alignment of 19 mammalian sequences and variability was scored as the number of variable positions of the possible 20.

Ancestral amino acid sequences were predicted with the maximum-likelihood method of Yang et al. (1995) and implemented through the PAML package (Yang 1997).

Results and Discussion

Tarsier Cytochrome b

The nucleotide sequence of cytochrome *b* from Horsfield's tarsier (*Tarsius bancanus*) is 1140 base pairs in length and translates to a protein of 379 residues. In this respect it is like all other mammalian cytochrome *b* genes sequenced to date, with the exception of the gene from the elephant (Irwin et al. 1991).

We sequenced parts of a number of independent *Tarsius bancanus* cytochrome *b* PCR products, and while the sequence obtained was identical in each case, varying regions were deleted. To investigate the possibility that we had preferentially amplified a nuclear pseudogene instead of the functional cytochrome *b* gene, we sequenced a short, 350-base pair region from the product of an identical PCR from *Tarsius syrichta.* No deletions were found in the short region of the *Tarsius syrichta* gene we sequenced, and a pairwise comparison between the *Tarsius bancanus* and the *Tarsius syrichta* sequences revealed K_a and K_s values of 0.048 and 0.630 substitutions per site, respectively. These values are not characteristic of comparisons between diverged pseudogenes or between a pseudogene and a functional gene. Additionally, calculation of K_s/K_a ratios for comparisons of these sequences with those of other nonsimian primates and other mammals resulted in values with an average of 13.19 and a standard deviation of 1.30. Hence, we concluded that the *Tarsius bancanus* sequence we obtained was from a functional copy of cytochrome *b.* Due to the fact that the mitochondrial genome becomes progressively damaged with age, usually due to extensive deletion of semi-random segments (Linnane et al. 1992), we attributed our finding to being the result of deletions from the ''functional'' gene of an aged animal. This is a plausible explanation, as our source of DNA was the Duke University Primate Center, where tissue samples are obtained predominantly from aged animals that have died of natural causes.

Analysis of DNA

Heterogeneity of nucleotide composition has been observed among the cytochrome *b* genes of mammals (Jermiin et al. 1994), and this may affect both evolutionary rates and their estimation. Therefore an analysis of evolutionary rates among these genes must take steps to ensure that any results obtained are not compromised by this factor. Table 2 shows the base composition at each codon site of each gene analyzed in this study. At first and second codon positions the base composition is similar among the sequences, but at third codon positions the variation in base composition is more pronounced. The distribution of base compositional heterogeneity among codon positions can be visualized by plotting *Z* scores (Smith 1986) for pairwise sequence comparisons of each codon position over the length of the gene (Fig. 2). The magnitude and range of the *Z* scores are lower and tighter for first and second codon positions than for third codon positions. This implies a more homogeneous base composition among sequences at the first and second codon positions than among those at the third codon

Table 2. Base compositions of cytochrome *b* genes calculated for each codon position

	First codon				Second codon				Third codon			
Species	% A	% T	% G	% C	% A	% T	% G	% C	% A	% T	% G	% C
Homo sapiens	29.6	23.5	19.3	27.7	20.5	40.1	12.9	26.9	36.4	12.1	3.7	47.8
Colobus guereza	31.7	23.2	17.7	27.4	19.8	40.1	12.7	27.4	39.3	18.0	3.7	39.0
Saimiri sciureus	32.7	21.6	18.7	26.9	20.1	39.6	12.7	27.7	38.3	21.6	4.7	35.4
Tarsius bancanus	30.6	23.2	20.8	25.3	20.3	42.0	13.5	24.2	41.2	17.2	1.6	40.1
Lemur catta	30.2	24.9	21.2	23.8	19.8	41.3	14.0	24.7	38.1	23.5	2.9	35.5
Galago crassicaudatus	28.3	23.0	20.9	27.8	20.6	40.2	13.2	26.0	34.9	16.1	4.0	45.0
Nycticebus coucang	28.0	23.2	22.4	26.4	20.1	41.2	13.5	25.3	39.1	19.5	4.5	36.9
Felis cattus	27.7	23.7	22.2	26.4	20.1	40.9	14.2	24.8	40.1	15.8	4.5	39.6
Balaenoptera physalus	28.8	21.4	22.7	27.2	20.1	41.7	13.7	24.5	41.4	13.7	2.6	42.2
Rattus norvegicus	28.2	24.0	21.4	26.4	20.1	42.2	14.0	23.7	42.2	15.6	2.6	39.6

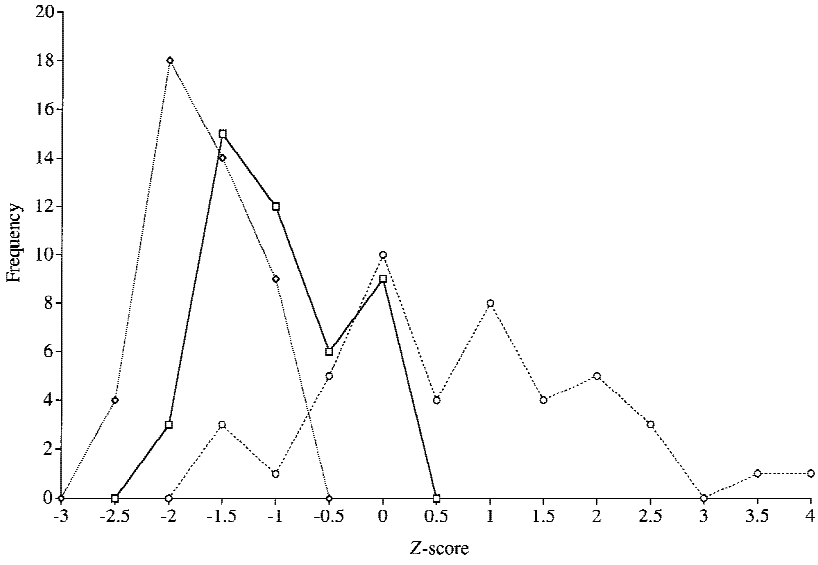

Fig. 2. Frequency distribution of *Z* scores for pairwise comparisons between cytochrome *b* sequences. *Squares* show first codon sites; *diamonds*, second codon sites; and *circles*, third codon sites.

position. Hence, the first and second codon positions provide a better data set for phylogenetic analysis and estimation of evolutionary rates than does the third codon position.

Tree Generation

An exhaustive search of tree space (Wolf et al. 1998) by maximum-likelihood analysis was conducted using the alignment of the first and second codon positions of the 10 mammalian cytochrome *b* nucleotide sequences as described in the *Materials and Methods.* This analysis produced 1731 trees that were not significantly ($\alpha = 0.05$) different from the maximum-likelihood tree shown in Fig. 3A. A consensus of the 1731 trees and the maximum-likelihood tree is shown in Fig. 3B. The trees shown in Figs. 3A and B are congruent neither with each other nor with any generally accepted phylogeny of primates and other mammals. The tree that represents an acceptable phylogeny is shown in Fig. 3C. This tree was among the 1731 trees that are not significantly different from the maximum-likelihood tree. In addition, we found that trees generated from the same data set using the neighbor-joining (Saitou and Nei 1987) and maximum-parsimony (Fitch 1977) methods were also inconsistent with currently accepted primate phylogeny (data not shown). When the cytochrome *c* oxidase subunit II gene (Adkins and Honeycutt 1994) is subjected to the same phylogenetic analysis as described above for cytochrome *b*, it is similarly unable to resolve the currently accepted phylogeny (data not shown).

Only two groupings of species are clearly resolved by the cytochrome *b* data. These are the simian primates (*Homo sapiens*, *Colobus guereza* and *Saimiri sciureus*) and the lorisoid strepsirhine primates (*Nycticebus coucang* and *Galago crassicaudatus*). The branching order of the other taxa, including *Tarsius bancanus* and *Lemur catta*, cannot be determined from the data. However, it is evident from Fig. 3 that the simian primate lineage has undergone more evolutionary change (has longer branch lengths) than all other lineages, including that of the tarsier, irrespective of which tree is correct. This phenomenon is the same in character as that observed for the

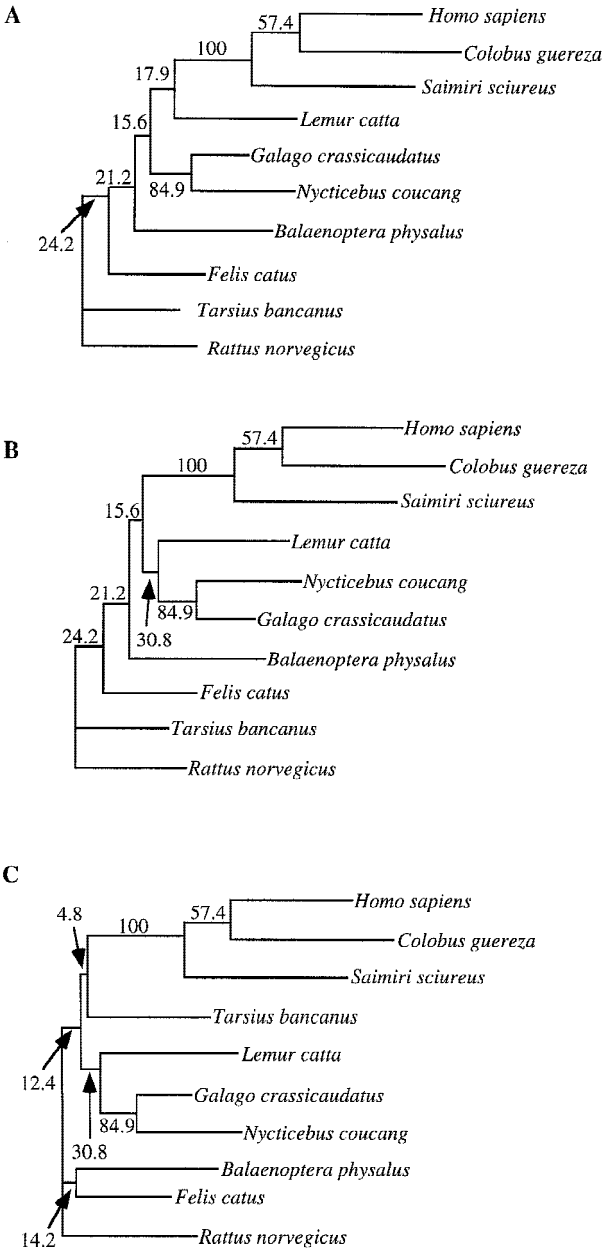

Fig. 3. **A** Maximum-likelihood tree found by exhaustive search of tree space (log likelihood score = −3420.10652). **B** Consensus of the maximum-likelihood tree and the 1731 trees not significantly different from it (log likelihood score = −3420.25773). **C** Tree most compatible with currently accepted mammalian phylogenetic relationships (log likelihood score = −3442.13304). Relative-likelihood support for each branch node was determined as described by Jermiin et al. (1997).

cytochrome *c* oxidase subunit II gene (Adkins and Honeycutt 1994).

Rates of Evolution

Relative rate tests using a data set of nondegenerate substitutions is presented in Table 3. Using *Rattus norvegicus* as an outgroup, sizable differences were detected in the rate of nondegenerate substitutions between simian primates and other mammals. Focusing on this phenomenon, we performed successive tests using outgroups separated from simian primates by progressively shorter periods of evolution (*Felis catus*, *Lemur catta*, and *Tarsius bancanus*). In addition, rate differences among simian primates were analyzed using *Saimiri sciureus* as the outgroup. For all comparisons between simians and other nonsimian mammals substantial increases in evolutionary rate were found for the simian species. Comparisons between simian primates showed that little rate heterogeneity exists among these species. From these findings we conclude that the cytochrome *b* gene of simian primates has an elevated evolutionary rate compared to the same gene from the other mammals we tested. K_{01}/K_{02} ratios (Table 3) show that the magnitude of the rate acceleration is about twofold compared to other, nonsimian species. These findings are similar to those presented for the cytochrome *c* oxidase subunit II gene (Adkins and Honeycutt 1994; Ramharack and Deeley 1987).

From these relative rate tests either of two conclusions can be made about the evolution of simian primate cytochrome *b*. First, the rate acceleration could be the result of an increased mutation rate restricted to the simian lineage, or second, an episode of adaptive evolution could have taken place. To differentiate between these possibilities, additional rate tests were undertaken using a data set consisting of nucleotide substitutions at fourfold degenerate sites (Table 3). However, to do this without obtaining a result that was confounded by the AT/GC compositional heterogeneity at third codon sites (Table 2), we compared only transversional substitution rates. Following the same method of outgrouping as before, we determined that little rate heterogeneity exists between any of the nucleotide sequences we tested. When this lack of rate heterogeneity at fourfold degenerate sites is viewed in relation to the substantial rate heterogeneity observed at nondegenerate sites between simian cytochrome *b* sequences and those from other mammals, it is evident that the evolutionary forces that have accelerated the rate of substitution for simian cytochrome *b* have resulted in an increase in nonsynonymous changes only. This is strong evidence for positive natural selection having acted on cytochrome *b* in lineages leading to simian primates and, hence, that simian primate cytochrome *b* has undergone an episode of adaptive evolution.

The Nature of Amino Acid Replacements in Simian Cytochrome b

Perutz (1983) hypothesized that large modification or refinement of the function of a protein can evolve through changes to only a small number of key residues. This has been found in a number of cases, such as crocodile hemoglobin (Komiyama et al. 1995), stomach lysozymes (Kornegay et al. 1994; Stewart et al. 1987), and visual pigments (Asenjo et al. 1994; Yokoyama and Yokoyama 1990; Yokoyama 1995). If simian primate cy-

Table 3. Relative rates of substitutions (*K*) and transversional substitutions (*B*) per site for the cytochrome *b* gene from primates and other mammalian species

			Nondegenerate substitutions[b]				Fourfold degenerate transversions			
Species 1	Species 2	Species 3 (outgroup)	K_{12} ± SE	K_{13}–K_{23} ± SE	*Z* score	K_{01}: K_{02}^{a}	*L*	B_{12} ± SE	B_{13}–B_{23} ± SE	*Z* score
Felis catus	*Homo sapiens*	*Rattus norvegicus*	0.139 ± 0.015	−0.061 ± 0.015	−3.99*	2.574	75.0	0.720±0.146	−0.123±0.258	−0.479
	Colobus guereza		0.162 ± 0.017	−0.075 ± 0.017	−4.52*	2.731	85.0	1.002±0.262	−0.005±0.239	−0.022
	Saimiri scuireus		0.161 ± 0.016	−0.070 ± 0.017	−4.21*	2.535	79.5	0.822±0.181	−0.069±0.248	−0.281
	Tarsius bancanus		0.083 ± 0.012	0.005 ± 0.011	0.409	1.121	75.5	0.772±0.165	−0.174±0.282	−0.615
	Lemur catta		0.099 ± 0.013	−0.024 ± 0.013	−1.91	1.651	80.0	0.916±0.223	−0.609±0.619	−0.984
	Balaenoptera physalus		0.104 ± 0.013	−0.025 ± 0.013	−1.94	1.647	68.0	0.591±0.111	−0.114±0.243	−0.468
Lemur catta	*Homo sapiens*	*Felis catus*	0.137 ± 0.015	−0.040 ± 0.015	−2.59*	1.815	77.5	0.846±0.193	0.195±0.248	0.789
	Colobus guereza		0.161 ± 0.017	−0.063 ± 0.017	−3.74*	2.285	67.0	0.611±0.118	−0.086±0.293	−0.293
	Saimiri scuireus		0.153 ± 0.016	−0.062 ± 0.016	−3.77*	2.364	68.0	0.625±0.121	0.093±0.248	0.376
	Tarsius bancanus		0.091 ± 0.013	−0.016 ± 0.012	1.34	1.436	64.5	0.594±0.115	0.144±0.238	0.606
	Nycticebus coucang		0.014 ± 0.014	0.001 ± 0.013	0.105	1.027	75.5	0.787±0.171	−0.280±0.408	−0.687
	Galago crassicaudatus		0.088 ± 0.012	0.016 ± 0.012	1.31	1.433	66.5	0.596±0.114	0.124±0.240	0.516
Tarsius bancanus	*Homo sapiens*	*Lemur catta*	0.154 ± 0.017	−0.046 ± 0.016	−2.87*	1.856	77.0	0.815±0.180	−0.251±0.210	−1.198
	Colobus guereza		0.159 ± 0.017	−0.070 ± 0.017	−4.19*	2.558	78.5	0.857±0.197	−0.164±0.156	−0.105
	Saimiri scuireus		0.157 ± 0.017	−0.061 ± 0.016	−3.73*	2.284	64.0	0.549±0.102	−0.031±0.143	−0.219
Saimiri scuireus	*Homo sapiens*	*Tarsius bancanus*	0.134 ± 0.015	0.003 ± 0.015	−0.183	1.044	64.0	0.527±0.096	−0.266±0.178	−1.491
	Colobus guereza		0.142 ± 0.016	0.002 ± 0.016	−0.152	1.036	70.0	0.623±0.119	−0.308±0.199	−1.549
Colobus guereza	*Homo sapiens*	*Saimiri scuireus*	0.130 ± 0.015	0.007 ± 0.015	0.464	1.116	50.5	0.362± 0.64	0.096±0.116	0.827

[a] $K_{01} = (K_{13} + K_{12} - K_{23})/2$ and $K_{02} = (K_{12} + K_{23} - K_{13})/2$.
[b] Average number of sites nondegenerate substitutions was $L = 713.9$.
* Denotes *Z* scores >1.96 or <−1.96, this being a significance level of $\alpha = 0.05$ if the comparisons were independent.

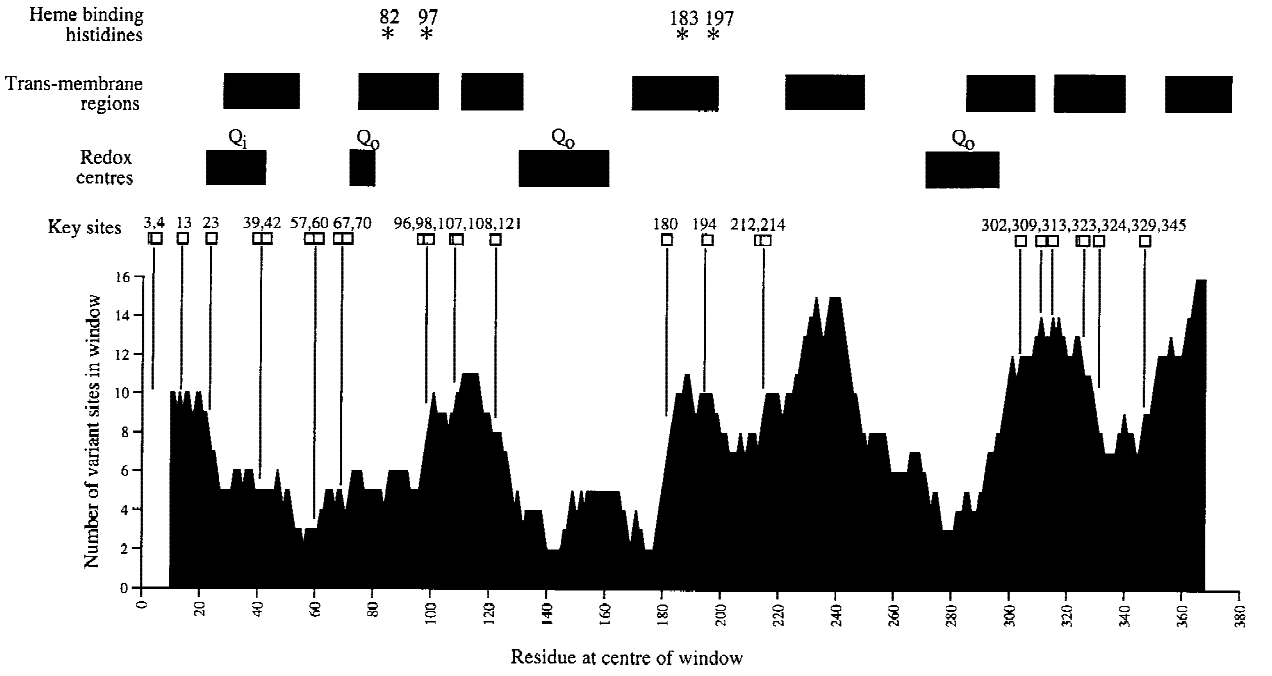

Fig. 4. Variability along the length of an alignment of 19 nonsimian-primate mammalian cytochrome *b* amino acid sequences, as counted in a moving window of 20 residues. Also shown are the transmembrane regions from the crystal structure of beef-heart cytochrome *b*, key residue changes between ancestral sequences (see text), redox centers, and heme-binding histidine residues.

tochrome *b* has acquired a new or altered biochemical function of character similar to that of simian cytochrome *c* and cytochrome *c* oxidase subunit II (Osheroff et al. 1983), then it should be possible to identify the key residue changes that facilitated this. To this end, we reconstructed ancestral cytochrome *b* sequences for the tree shown in Fig. 3C using the method of Yang et al. (1995) and identified amino acid changes that occurred between the most recent common ancestral sequence of all haplorhine primates and the most recent common ancestral sequence of simian primates. Presuming that the altered function of simian cytochrome *b* arose between these two ancestral sequences, the amino acid changes that occurred along this lineage would include the key changes that caused any functional differentiation. Twenty-six amino acid changes between the ancestral sequences were detected (Fig. 4).

In an effort to determine which of these amino acid changes were functionally silent or, alternatively, which occurred in important conserved regions of the cytochrome *b* protein, a variability plot (Fig. 4) along the length of cytochrome *b* was constructed using a sliding-window approach (after that of Irwin et al. 1991). For this analysis a larger alignment of 19 mammalian species (see *Materials and Methods*) was used, including tarsier and strepsirhine primates but excluding simian species, and the absolute number of variable sites in a window 20 residues in length was counted. Using this approach, generally conserved and variable regions of the mammalian cytochrome *b* protein were identified and these showed correlations with postulated structural regions, such as the Q_i and Q_o redox centers and the transmembrane domains (Fig. 4).

Twenty of the twenty-six sites that changed along the ancestral lineage immediately preceding the simian primates are situated in regions of the protein that were classified as nonconserved (defined as containing more than 6 variant sites per window of 20 sites). The remaining six amino acid changes, at positions 39, 42, 57, 60, 67, and 70, are in conserved regions of the protein (defined as containing 6 or fewer variant sites per window of 20 sites). While some of the 20 residue changes that occur in the nonconserved region of the protein could be important for a modified function of simian cytochrome *b*, it is more likely that the majority of the changes are neutral or compensating changes. The six residue changes that occur in the conserved regions of the protein are more likely to be residues responsible for a change in the function of simian cytochrome *b*, as they are located in the region proposed to contain the Q_i and part of the Q_o redox sites. As yet, the coordinates of the crystal structure of the cytochrome bc_1 complex (Xia et al. 1997), of which cytochrome *b* forms a subunit, have not been released, and hence it is not possible to model the structural consequences of the change of this combination of six residues. However, the residues at positions 39 and 42 lie at the end of the Q_i redox site at the N terminus of the protein, in a region of the protein that has been found to lend resistance to inhibitors in bacterial respiratory chains (Brasseur et al. 1996). The other residues, at positions 57, 60, 67, and 70, are not directly within regions associated with either of the redox sites,

but they are on a loop region that is proposed to lie close to the extramembranous α-helix that forms the central portion of the Q_o redox site. These six amino acid changes that arose in the protein ancestral to extant simian primate cytochrome *b* may be some of the key residues important in an alteration of cytochrome *b* function that may have occurred in simian primates but not in other mammals.

Acknowledgments. Thanks go to A. Yoder for advice and tips on sequencing the tarsier cytochrome b gene, L. Croft for technical advice, G. Chelvanayagam for encouragement and comment, D. Kaufmann and M. Smith for IT support, and the Duke University Primate Center for the tarsier tissues. This work was partially supported by an Australian Postgraduate Award to T.D.A. The tarsier cytochrome *b* nucleotide sequence has been submitted to DDBJ/EMBL/GenBank and has the accession number AB011077. All alignments, ancestral sequences, and programs written by the authors can be obtained directly from them or via the World Wide Web at <http://jcsmr.anu.edu.au.> and pages therein.

References

Adkins RM, Honeycutt RL (1994) Evolution of the primate cytochrome c oxidase subunit II gene. J Mol Evol 38:215–231

Adkins RM, Honeycutt RL, Disotell TR (1996) Evolution of eutherian cytochrome c oxidase subunit II: heterogeneous rates of protein evolution and altered interaction with cytochrome c. Mol Biol Evol 13:1393–1404

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organisation of the human mitochondrial genome. Nature 290:457–465

Arnason U, Gullberg A, Widegren B (1991) The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus.* J Mol Evol 33:556–568

Arnason U, Bodin K, Gullberg A, Ledje C, Mouchaty S (1995) A molecular view of pinniped relationships with particular emphasis on the true seals. J Mol Evol 40:78–85

Asenjo AB, Rim J, Oprian DD (1994) Molecular determinants of human red/green colour discrimination. Neuron 12:1131–1138

Baba ML, Darga LL, Goodman M, Czelusniak J (1981) Evolution of cytochrome c investigated by the maximum parsimony method. J Mol Evol 17:197–213

Brasseur G, Saribas AS, Daldal F (1996) A compilation of mutations located in the cytochrome b subunit of the bacterial and mitochondrial *bcl* complex. Biochim Biophys Acta 1275:61–69

Cann RL, Brown WM, Wilson AC (1984) Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics 106:479–499

Collura RV, Stewart CB (1995) Insertion and duplications of mtDNA in the nuclear genomes of old world monkeys and hominoids. Nature 378:485–489

Evans MJ, Scarpulla RC (1988) The human somatic cytochrome c gene: two classes of processed pseudogenes demarcate a period of rapid molecular evolution. Proc Natl Acad Sci USA 85:9625–9629

Fitch W (1977) On the problem of discovering the most parsimonious tree. Am Nat 111:223–257

Gadaleta G, Pepe G, De Candia G, Quagliariello C, Sbisa E, Saccone C (1989) The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J Mol Evol 28:497–516

Groves CP (1989) A theory of human and primate evolution. Oxford University Press, Oxford

Hatefi Y (1985) The mitochondrial electron transport and oxidative phosphorylation system. Annu Rev Biochem 54:1015–1069

Irwin DM, Arnason U (1994) Cytochrome b gene of maritime mammals: phylogeny and evolution. J Mammal Evol 2:37–55

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. J Mol Evol 32:128–144

Janke A, Feldmaier-Fuchs G, Thomas WK, von Haessler A, Pääbo A (1994) The marsupial mitochondrial genome and the evolution of placental mammals. Genetics 137:243–256

Jermiin LS, Graur D, Lowe RM, Crozier RH (1994) Analysis of directional mutation pressure and nucleotide content in mitochondrial cytochrome b genes. J Mol Evol 39:160–173

Jermiin LS, Olsen GJ, Mengersen KL, Easteal S (1997) Majority-rule consensus of phylogenetic trees obtained by maximum-likelihood analysis. Mol Biol Evol 14:1296–1302

Komiyama NH, Miyazaki G, Tame J, Nagai K (1995) Transplanting a unique allosteric effect from crocodile into human haemoglobin. Nature 373:244–246

Kornegay JR, Schilling JW, Wilson AC (1994) Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. Mol Biol Evol 11:921–928

Kreitman M, Akashi H (1995) Molecular evidence for natural selection. Annu Rev Ecol Syst 26:403–422

Li W-H (1993) Unbiased estimation of the rates of synonymous and non-synonymous substitution. J Mol Evol 36:96–99

Li W-H (1997) Fundamentals of molecular evolution. Sinauer, Sunderland, MA

Linnane AW, Zhang C, Baumer A, Nagley P (1992) Mitochondrial DNA mutation and the ageing process: bioenergy and pharmacological intervention. Mutat Res 275:195–208

Lopez JV, Cevario S, O'Brien SJ (1996) Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and transposed mtDNA repeat, *Numt*, in the nuclear genome. Genomics 33:229–246

Ma D-P, Zharkikh A, Graur D, VandeBerg JL, Li W-H (1993) Structure and evolution of opossum, guinea pig and porcupine cytochrome b genes. J Mol Evol 36:327–334

Muse SV, Weir BS (1992) Testing for equality of evolutionary rates. Genetics 132:269–276

Osheroff N, Speck SH, Margoliash E, Veerman ECI, Wilms J, Konig BW, Muijsers AO (1983) The reaction of primate cytochromes c with cytochrome c oxidase. J Biol Chem 258:5731–5738

Pamilo P, Bianchi NO (1993) Evolution of the *Zfx* and *Zfy* genes: rates and interdependence between genes. Mol Biol Evol 10:271–281

Perutz MF (1983) Species adaptation in a protein molecule. Mol Biol Evol 1:1–28

Ramharack R, Deeley RG (1987) Structure and evolution of primate cytochrome c oxidase subunit II gene. J Biol Chem 262:14014–14021

Saitou N, Nei M (1987) The neighbour joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbour, NY.

Smith CAB (1986) Chi-squared tests with small numbers. Ann Hum Genet 50:163–167

Stewart CB, Schilling JW, Wilson AC (1987) Adaptive evolution in the stomach lysozymes of foregut fermenters. Nature 330:401–404

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Wolf MJ, Jermiin LS, Easteal S, Kahn S, McKay BD (1998) TrExML—a maximum likelihood program for exhaustive tree-space exploration (submitted)

Wu C-I, Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents that in man. Proc Natl Acad Sci USA 82:1741–1745

Wu W, Goodman M, Lomax MI, Grossman LI (1997) Molecular evolution of cytochrome c oxidase subunit IV: evidence for positive selection in simian primates. J Mol Evol 44:477–491

Xia D, Yu C-A, Kim H, Xia J-Z, Kachurin AM, Zhang L, Yu L, Deisenhofer J (1997) Crystal structure of the cytochrome bcl complex from bovine heart mitochondria. Science 277:60–66

Yang Z (1997) Phylogenetic analysis by maximum likelihood (PAML), Version 1.3. Department of Integrative Biology, University of California at Berkeley, Berkeley

Yang Z, Kumar S, Nei M (1995) A new method of inference of ancestral nucleotide and amino acid sequences. Genetics 141:1641–1650

Yoder AD, Cartmill M, Ruvolo M, Smith K, Vilgalys R (1996) Ancient single origin for Malagasy primates. Proc Natl Acad Sci USA 93: 5122–5126

Yokoyama R, Yokoyama S (1990) Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus*, and human. Proc Natl Acad Sci USA 87:9315–9318

Yokoyama S (1995) Amino acid replacements and wavelength absorption of visual pigments in vertebrates. Mol Biol Evol 12:53–61