First publ. in: Journal of Molecular Evolution 53 (2001), pp. 89-103

The Cytochrome *b* Gene as a Phylogenetic Marker: The Limits of Resolution for Analyzing Relationships Among Cichlid Fishes

Izeni P. Farias[1,4], Guillermo Ortí[2], Iracilda Sampaio[3], Horacio Schneider[3], Axel Meyer[4]

[1] Departamento de Biologia, Instituto de Ciências Biológicas, Universidade do Amazonas, 69077-000, Estrada do Contorno 3000, Manaus, Amazonas, Brazil
[2] School of Biological Sciences, 314 Manter Hall University of Nebraska, Lincoln, NE 68588-0118, USA
[3] Laboratório de Biologia Molecular, Campus Universitário de Bragança Universidade Federal do Pará, Bragança, Pará, Brazil
[4] Department of Biology, PO Box M617, University of Konstanz, D-78457 Konstanz, Germany

Received: 11 October 2000 / Accepted: 26 February 2001

Abstract. The mitochondrial cytochrome *b* (cyt-*b*) gene is widely used in systematic studies to resolve divergences at many taxonomic levels. The present study focuses mainly on the utility of cyt-*b* as a molecular marker for inferring phylogenetic relationship at various levels within the fish family Cichlidae. A total of 78 taxa were used in the present analysis, representing all the major groups in the family Cichlidae (72 taxa) and other families from the suborders Labroidei and Percoidei. Gene trees obtained from cyt-*b* are compared to a published total evidence tree derived from previous studies. Minimum evolution trees based on cyt-*b* data resulted in topologies congruent with all previous analyses. Parsimony analyses downweighting transitions relative to transversions (ts1:tv4) or excluding transitions at third codon positions resulted in more robust bootstrap support for recognized clades than unweighted parsimony. Relative rate tests detected significantly long branches for some taxa (LB taxa) which were composed mainly by dwarf Neotropical cichlids. An improvement of the phylogenetic signal, as shown by the four-cluster likelihood mapping analysis, and higher bootstrap values were obtained by excluding LB taxa. Despite some limitations of cyt-*b* as a phylogenetic marker, this gene either alone or in combination with other data sets yields a tree that is in agreement with the well-established phylogeny of cichlid fish.

Key words: Cichlid fishes — Phylogeny — Cytochrome *b* — Systematics — Long branches — DNA substitution rate

Correspondence to: Axel Meyer; *email:* axel.meyer@uni-konstanz.de

Introduction

While nuclear ribosomal rRNA gene and complete mitochondrial DNA (mtDNA) have been used to assess phylogenetic relationships involving ancient divergences (Mindell and Honeycutt, 1990; Zardoya and Meyer, 1996; Van de Peer and De Wachter, 1997; Abouheif et al., 1998; Naylor and Brown, 1998; Zardoya et al., 1998), the majority of research on animals has used single mitochondrial DNA genes to assess population or low-level taxonomic relationships (reviewed in Meyer, 1993; Rocha-Olivares et al., 1999; Lovejoy and de Araújo, 2000; Tsigenopoulos and Berrebi, 2000). Cytochrome *b* (cyt-*b*) has been considered one of the most useful genes for phylogenetic work, and is probably the best-known mitochondrial gene with respect to structure and function of its protein product (Esposti et al., 1993). Cyt-*b* gene contains both slowly and rapidly evolving codon positions, as well as more conservative and more variable regions or domains overall. Therefore, this gene

has been used for a diversity of systematic questions, from "deep" phylogeny (e.g., Meyer and Wilson, 1990; Irwin et al., 1991; Normark et al., 1991; Cantatore et al., 1994; Lydeard and Roe, 1997; Kumazawa and Nishida, 2000) to the population and recent divergence levels (Sturmbauer and Meyer, 1992; Rocha-Olivares et al., 1999; Kirchman et al., 2000; Lovejoy and de Araújo, 2000). However, many problems have been encountered when using cyt-*b*, including base compositional biases, rate variation between lineages, saturation at third codon positions, and limited variation in first and second codon position, resulting in little phylogenetic information for "deep" evolutionary questions, or few informative sites for the third codon position at the population levels (Meyer, 1994). Therefore, the rather naïve notion that this gene would be useful as phylogenetic marker at all phylogenetic levels has been called into question (Hillis and Huelsenbeck, 1992; Meyer, 1994; Honeycutt et al., 1995).

The phylogenetic utility of the cyt-*b* gene has been studied at several taxonomic levels among vertebrate taxa (Irwin et al., 1991; Moritz et al., 1992; da Silva and Paton, 1993; Graybeal, 1993; Lamb and Lydeard, 1994; Moore and DeFilippis, 1997; Nunn and Stanley, 1998) and particularly in fish taxa (Ortí and Meyer, 1996, 1997; Lydeard and Roe, 1997; Martin and Bermingham, 1998; Zardoya and Doadrio, 1999; Lovejoy and de Araújo, 2000). Molecular phylogenetic studies using cyt-*b* in cichlid fishes included only few selected representatives from the lineages of the family Cichlidae. Some of them used partial cyt-*b* sequence to address phylogenetic relationships among closely related cichlids from African Lakes (Meyer et al., 1990, 1991; Sturmbauer and Meyer, 1993; Sturmbauer et al., 1994), while others more recently carried out extensive taxonomic sampling of Central American cichlids using the complete and nearly complete sequences of cyt-*b* (Lydeard and Roe, 1997; Roe et al., 1997; Martin and Bermingham, 1998). Considering that cichlids are found throughout Africa, the Neotropics, Madagascar, and India, in all of previous studies, most lineages of cichlids at the family level remained poorly represented.

Here we are interested in investigating the utility of the cyt-*b* gene as a molecular marker for inferring phylogenetic relationships in the monophyletic family Cichlidae, thought to have originated early in the Cretaceous (Stiassny, 1991; Zardoya et al., 1996) about ~130–150 million years ago. Our previous work (Farias et al., 2000) based on a portion of mitochondrial gene (16S rRNA, 556 bp), microsatellite flanking regions (*Tmo-M27,* 302 bp), and a single-copy nuclear (scn) DNA locus (*Tmo-4C4,* 511 bp), and published morphological data (91 characters, Kullander, 1998), combined in a total evidence analysis (1460 characters for 34 taxa), provided increased resolution and higher bootstrap support compared to any of the individual data partitions, supporting a robust phylogenetic hypothesis for the family Cichlidae (Fig. 1). The total evidence analyses confirmed the placement of Malagasy/Indian cichlids as the most basal lineages, with a sister group relationship to the monophyletic African and Neotropical assemblages, and were concordant with previous morphological and molecular studies (Stiassny, 1991; Zardoya et al., 1996; Streelman and Karl, 1997; Farias et al., 1999, 2000). Total evidence suggested that the enigmatic cichlid genus *Heterochromis* is at the base of the African radiation. Among the Neotropical assemblage, *Retroculus, Cichla,* and *Astronotus* were found to be the basal lineages, followed by a monophyletic group of the geophagines, and the sister and monophyletic groups of cichlasomines and heroines.

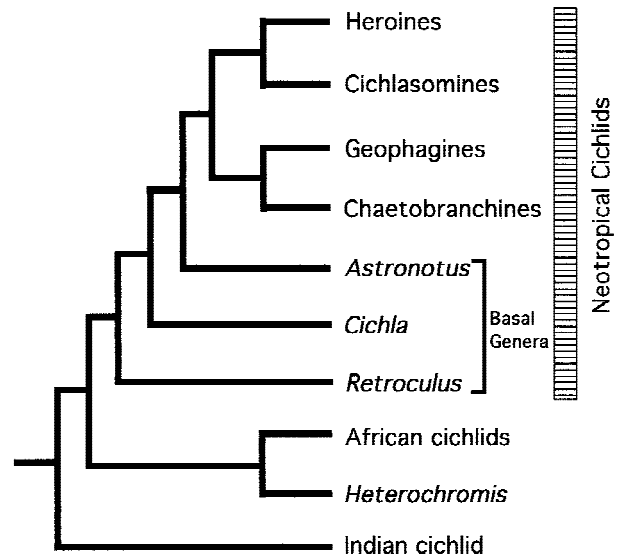

Fig. 1. Phylogenetic relationships of the major groups of cichlid fish based on a total evidence analysis (rRNA 16S + nuclear loci + morphological data, 1460 characters; Farias et al., 2000).

One way to evaluate the phylogenetic efficiency of a gene tree is by assessing congruence with a well-supported phylogeny based on independently derived data sets. Frequently, this has involved comparisons among molecular phylogenies with phylogenies based on morphological data (e.g., Lydeard and Roe, 1997; Ortí and Meyer, 1997; Farias et al., 1998, 1999, 2000). However, tests of congruence have been increasingly used to compare several molecular phylogenetic estimates with one another (Yoder et al., 1996). Therefore, our main goal in the present study is to explore the utility of cyt-*b* to infer phylogenetic relationships at various taxonomic levels in the family Cichlidae by assessing congruence among data sets. Gene trees from cyt-*b* data are compared to our previous results, based on total evidence tree (Farias et al., 2000), and all data available were subsequently combined for a total evidence analysis. The degree of robustness of particular relationships of the main cichlid lineages and the phylogenetic signal (phylogenetic information in each data partition of cyt-*b*) were

analyzed. We also examined rates and patterns of evolution in cyt-*b* sequences among the cichlid lineages.

Materials and Methods

Samples and DNA Sequencing. All main cichlid lineages are represented in our data set, from the most closely related taxa to the most distantly related groups in the family Cichlidae. A total of 78 taxa was used in the present analysis, including cichlids (72 taxa) and outgroup taxa from the suborders Labroidei and Percoidei. The complete list of taxa and the sources of sequences is presented in Table 1. For the new taxa sequenced in the present study, total DNA was extracted by standard proteinase K/phenol/chloroform extraction (Sambrook et al., 1989).

Polymerase chain reaction (PCR) was used to amplify the complete or partial sequence of the cyt-*b* mitochondrial gene. PCR reactions were carried out in 25 μl reaction containing 5 μl dNTP (1 mM), 2.5 μl (10×) reaction buffer, 1 μl (25 mM) $MgCl_2$, 1 μl (10 mM) each primer, 1 μl total DNA (50 ng/μl), and 1 U Taq DNA polymerase. PCR cycles for amplications were performed using the following conditions: 25 cycles of denaturation at 93°C for 1 min, annealing at 50°C for 1 min, and extension at 72°C for 2 min. A final extension at 72°C for 10 min was performed to completely extend the amplified product. The sequencing primers used in this study are listed in Table 2. The PCR product was purified with the QIAquick (QIAGEN) kit. Each PCR product was sequenced using the BigDye Terminator cycle sequencing ready reaction kit (Applied Biosystems Inc.) on an automated DNA sequencer (Applied Biosystems 377), following manufacturer's instructions. The nucleotide sequence data determined for the present paper was deposited in GenBank under accession numbers AF370623-AF370679.

Sequence Alignment and Data Analysis. The present data set is composed by complete and nearly complete cyt-*b* gene sequences, which did not show any insertions or deletions (indels). The final alignment comprised 1138 base pairs.

Phylogenetic analyses were assessed by maximun-parsimony (MP) using PAUP* version 4.0 b2a (Swofford, 2000), using heuristic searches with 50 repetitions using random stepwise addition of taxa. For all analyses, tree length (L), consistency indices (CI, excluding uninformative characters, Kluge and Ferris, 1969), and retention indices (RI, Farris, 1989) are reported. The internal stability of the inferred MP tree was measured by bootstrapping using 1000 replications (Felsenstein, 1985).

Minimum evolution (ME) methods were also applied on the data using maximum likelihood distances (Kidd and Sgaramella-Zonta, 1971; Rzhetsky and Nei, 1992). The MODELTEST program (Posada and Crandall, 1998) was applied to estimate the model that fits the cyt-*b* data best. The model selected was general time-reversible (GTR; Yang, 1994: assuming a different rate for all six classes of substitutions), with some sites assumed to be invariable, and variable sites assumed to follow a gamma distribution (GTR + I + Γ). Parameters for this model were estimated by optimizing the data on the MP trees. The average number of changes per site for each codon position was computed using MacClade (Maddison and Maddison, 1992). The level of confidence in each node of the ME trees was assessed using non-parametric bootstrapping with 100 pseudoreplicates (each with 10 random addition replicates).

The four-cluster likelihood mapping analysis (Strimmer and Haeseler, 1996) was performed to test for phylogenetic signal present in the data. Sequences were grouped according to the total evidence tree (Fig. 1) in four clusters: cluster a = African cichlids, cluster b = Neotropical cichlids, cluster c = Malagasy/Indian cichlids, and cluster d = outgroup. The result is shown by a triangle with values at the corners indicating the percentages of well-resolved phylogenies (resolved quartets) for all possible quartets. Values at the central and lateral regions are percentages of unresolved phylogenies (unresolved, "bad" quartets). A high percentage of bad quartets usually indicates that the data set is not very well suited for a phylogenetic analysis. If more than 10%–15% of the quartets are bad, then the quartet-puzzling tree is, in general, not completely resolved. The frequencies of bad quartets indicate a measure of "noise" obscuring phylogenetic signal present in the data (Strimmer and Haeseler, 1996).

Our previous analyses of 16S rRNA sequences from cichlid fishes (Farias et al., 1999, 2000) suggested heterogeneous rates of nucleotide change between Neotropical and African cichlids. Therefore, we tested the constancy of nucleotide substitution rates in the cyt-*b* data set using relative rate tests. The two-cluster test implemented in LINTRE (Takezaki et al., 1995) assumes a tree topology (neighbor-joining tree) to test for differences in root-to-tip distance of each sequence from the average of all sequences and identifies taxa that show significantly different substitution rates at the 5% level. Stability in base composition was tested using the chi-square test for unequal base composition implemented in PAUP* version 4.0 b2a (Swofford, 2000) and PUZZLE version 4.0.1 (Strimmer and Haesler, 1996).

A variety of analyses were conducted in order to test for phylogenetic effects of different weighting schemes. In the first step, heuristic searches under MP and ME criteria were performed with all data equally weighted. This was done to assess the results of posteriori weighting and to verify if weighted parsimony would filter the noise out of the data (Griffiths, 1997). The level of saturation (multiple substitutions) was assessed by plotting genetic distance (*P* distance) against the number of nucleotide substitution for each codon position. Then, by examining the phylogenetic results from the first step and the patterns of nucleotide substitution for each codon position, the following weighting schemes were applied: (i) the third codon positions were excluded from the data set; (ii) different weight schemes were applied according to saturation level and transition/transversion ratios (ts/tv) from the data set. An additional test was also performed by exclusion of fast evolving taxa from the cyt-*b* data set.

A simple test, the incongruence length difference (ILD) test, described by Farris et al. (1994, 1995), measures the significance of incongruence among data sets. This test, also implemented in PAUP* 4.0 b2a (Swofford, 2000) as homogeneity test, was applied to access the congruency among cyt-*b* and the total evidence tree (Farias et al., 2000). Furthermore, considering the homogeneity test results, we combined all data available in a total evidence analysis.

Results

Pattern of Sequence Variation

We compiled a data set of 78 nearly complete and complete cyt-*b* gene sequences, including 72 representatives of the family Cichlidae (of which 61 are new sequences determined for the present study). Considering a total of 1138 bp for the analysis, 683 characters were variable, and 581 were phylogenetically informative under parsimony. Considering the phylogenetically informative sites, 142 (24%) were at first codon positions, 62 (11%) were at second positions, and 377 (65%) were at third codon positions (see Table 3).

The three codon positions differed greatly in their base composition (Table 3). First codon positions showed no bias, but second (G = 13%) and third positions (G = 4%) exhibited a strong bias anti-guanine. Considering all codon positions, the mean base composition for cyt-*b* was 25% adenine, 32% cytosine, 14%

Table 1. Fish included in the present study. The taxonomy for the family Cichlidae follows Kullander (1998). Genera between quotation marks indicate species on taxonomic revision (e.g., "*Cichlasoma*"). Numbers identify the source of the sequence data analyzed in the present paper. 1: new data determined in the present study; 2: data from Whitmore et al. (1994); 3: data from Song (1994); 4: data from Cantatore et al. (1994); 5: data from Roe et al. (1997); 6: data from Martin and Bermingham (1998). Voucher specimen and locality information from South American cichlids is available from IPF

Taxa	Cytochrome *b*
Percomorpha	
Percoidei	
Moronidae	
Morone mississippiensis	3
Dicentrarchus labrax	4
Centrarchidae	
Micropterus salmoides	2
Percidae	
Etheostoma kennicotti	3
Labroidei	
Embiotocidae	
Cymatogaster aggregata	1
Labridae	
Halichoeres maculipinna	1
Cichlidae	
Etroplinae (Malagasy/India)	
Etroplus maculatus	1
Oxylapia polleni	1
Paratilapia sp.	1
Paretroplus polyactis	1
Ptychochromoides betsileanus	1
Ptychochromis oligocantus	1
Pseudocrenilabrinae (Africa)	
Astatotilapia calliptera	1
Boulengerochromis microlepis	1
Chalinochromis brichardi	1
Chromidotilapia sp.	1
Cyrtocara sp.	1
Hemichromis bimaculatus	1
Labidochromis caeruleus	1
Oreochromis mossambicus	4
Serranochromis robustus	1
Thysochromis sp.	1
Tylochromis polylepis	1
Heterochromidinae (Africa)	
Heterochromis multidens	1
Retroculinae (South America)	
Retroculus sp.	1
Retroculus xinguensis	1
Cichlinae (South America)	
Cichlini	
Cichla monoculus	1
Cichla orinocensis	1
Cichla temensis	1
Crenicichlini	
Crenicichla sp.	1
Crencichla regani	1
Teleocichla centrarchus	1
Teleocichla monogramma	1
Teleocichla proselicts	1

Table 1. Continued

Taxa	Cytochrome *b*
Astronotinae (South America)	
Astronotini	
Astronotus crassipinnis	1
Astronotus ocellatus	1
Chaetobranchini	
Chaetobranchopsis orbicularis	1
Chaetobranchus flavescens	1
Geophaginae (South America)	
Acarichthyini	
Acarichthys heckelli	1
Guianacara sp.	1
Crenicaritini	
Crenicara sp.	1
Geophagini	
Apistogramma sp.	1
Biotodoma wavrini	1
Geophagus sp.	1
"Geophagus" brasiliensis	1
"Geophagus" steindachneri	1
Gymnogeophagus gymnogenys	1
Gymnogeophagus labiatus	1
Satanoperca acuticeps	1
Satanoperca jurupari	1
Taeniacara candidi	1
Cichlasomatinae (South America)	
Acaroniini (South America)	
Acaronia sp.	1
Cichlasomatini (South America)	
"Aequidens" sp.	1
Bujurquina sp.	1
Cichlasoma boliviensis	6
Cichlasoma portalegrense	5
Cichlasoma amazonarum	1
Nannacara anomala	1
Heroini (South America)	
Caquetaia spectabilis	1
Heros sp.	1
Hoplarchus psittacus	1
Hypselecara coryphaenoides	1
Mesonauta insignis	1
Pterophyllum scalare	1
Symphysodon aequifasciatus	1
Uaru amphiacanthoides	1
Heroini (Central America)	
Amphilophus labiatus	5
Archocentrus spilurus	5
Astatheros bussingi	5
Heros appendiculatus	6
Herotilapia multispinosa	6
Nandopsis dovii	2
Parachromis loiselli	6
Paraneetroplus sieboldii	6
Petenia splendida	1
Theraps maculicauda	5
Thorichthys ellioti	5
Tomocichla tuba	6

Table 2. Sequencing primers used in this study

Primer sequence	Reference
L14725 5′-CGAAGCTTGATATGAAAAACCATCGTTG-3′	Pääbo (1990)
H15915 5′-AACTGCAGTCATCTCCGGTTTACAAGAC-3′	Irwin et al. (1991)
L14981 5′-GGCGCATCCTTCTTCTTYATCTGT-3′	Present study
L15162 5′-GCAAGCTTCTACCATGAGGACAA-3′	Taberlet et al. (1992)
L15379 5′-GCAGCCATAACAATAATTCA-3′	Lydeard and Roe (1997)
H15149 5′-AAACTGCAGCCCCTCAGAATGATA-3′	Kocher et al. (1989)
H15573 5′-AATAGGAAGTATCATTCGGGTTT-3′	Taberlet et al. (1992)

Table 3. Parameters obtained in the analysis of cyt *b* data set at each codon position

Parameters	First position	Second position	Third position	All taxa and all codon positions	Excluding third codon position	Excluding LB taxa
Parsimony informative sites	142	62	377	579	203	525
Ts/Tv Ratio (κ)	4.21	2.86	6.62	3.93	4.23	4.04
Proportion of invariable sites	0.41	0.52	0.003	0.37	0.46	0.47
Gamma shape (α)	0.69	0.56	1.80	0.73	0.47	0.99
Average number of changes per site	3.4	0.9	17.0	7.2	2.2	5.3
Proportion of Adenine	25%	20%	30%	25%	22%	25%
Proportion of Cytosine	27%	25%	44%	32%	26%	32%
Proportion of Guanine	24%	13%	4%	14%	19%	14%
Proportion of Tymine	24%	42%	22%	29%	33%	29%
Chi-square tests of base frequencies[a]	$X^2 = 57.3$ df = 234 $P > 0.05$	$X^2 = 17.9$ df = 234 $p > 0.05$	$X^2 = 568.5$ df = 234 $p < 0.05$	$X^2 = 205.6$ df = 231 $p > 0.05$	$X^2 = 42.1$ df = 231 $p > 0.05$	$X^2 = 120.5$ df = 192 $p > 0.05$

[a] The chi-square tests of homogeneity compares the base frequencies across taxa, performed with PAUP* (Swofford 2000).

guanine, and 29% tymine. These values are close to observations reported in other studies of cyt-*b* gene in different taxa (for review see Meyer, 1993, and Johns and Avise, 1998). Base composition was slightly variable among taxa, with seven taxa (*Bujurquina, Crenicichla regani, Chaetobranchopsis, Oxylapia, Cymatogaster, Etheostoma,* and *Dicentrarchus*) rejecting the chi-square test of homogeneity (assessed by PUZZLE 4.0), however, it was homogeneous in the overall variable sites ($\chi^2 = 208.51$; df = 234, $p > 0.05$). Variation in base composition among taxa was homogeneous at first and second positions, but not at third positions which showed high heterogeneity with 22 taxa failing the chi-square test of homogeneity ($\chi^2 = 568.48$; df = 234, $p < 0.05$). Detailed analyses of the contribution of phylogenetic signal at different codon positions were assessed by likelihood mapping analysis (PUZZLE 4.0), and are shown in Table 4.

Considering all cyt-*b* data, analysis of substitution patterns shows that the number of transitions increases linearly with genetic distance, but beyond 0.15 (uncorrected *p* distance), transversions begin to increase substantially, with a rapid decrease in the number of transitions indicating some leves of transition saturation (Fig. 2D). This effect is not seen at first and second positions (Figs. 2A, B), but is most remarkable at third positions (Fig. 2C). The maximum pairwise divergence value observed among all taxa was 0.31 (uncorrected *p* distance) between *Dicentrarchus* and *Crenicara.* The minimum pairwise divergence value observed among all taxa was 0.02, between *Labidochromis* and *Cyrtocara.* The pairwise sequence divergences within the family Cichlidae ranged from 0.02 to 0.29 (Table 5). Except for comparisons among heroines and among East African cichlids, most pairwise comparisons have genetic distance values greater than 0.15. Based on our previous work (Farias et al., 1999, 2000), major lineages of cichlid fishes were defined (Fig. 1) and are used here to gauge the level of divergence for cyt-*b*. Divergence values between cichlid lineages averaged higher (0.19–0.29) than within-lineage comparisons (0.15–0.19). The lowest nucleotide divergence values were observed among recently diverged species such as the East African cichlids and the Neotropical heroines. Among African cichlids low values were obtained among the genera *Boulengerochromis, Astatotilapia, Cyrtocara, Labidochromis,* and *Serranochromis,* which are all endemic to the East African lakes. Higher divergence values were observed among heroines from South America than from Central America, suggesting that the latter form a more recent lineage. Maximum divergence values were noted among geophagine cichlids (up to 0.29), suggesting elevated substitution

Table 4. Results of likelihood mapping analysis (PUZZLE 4.0.1) showing the contributing of phylogenetic signal at each codon position of cyt-*b* data. Sequences were grouped in the following clusters: cluster a = African cichlids, cluster b = Neotropical cichlids, cluster c = Malagasy/Indian cichlids, cluster d = outgroup

Percentage of quartets	Sequences are grouped in clusters	First position	Second position	Third position	All taxa and all codon positions	Excluding third codon position	Excluding LB taxa
Resolved quartets	(a,b)–(c,d)[a]	36.2%	50.9%	21.0%	46.8%	53.5%	60.5%
	(a,c)–(b,d)	8.2%	7.9%	12.9%	17.4%	6.2%	17.5%
	(a,d)–(b,c)	14.8%	4.1%	12.4%	23.1%	7.9%	11.2%
Unresolved quartets	...	40.8%	37.1%	53.7%	12.7%	32.4%	10.8%

[a] Topology found in total evidence tree (Fig. 1).

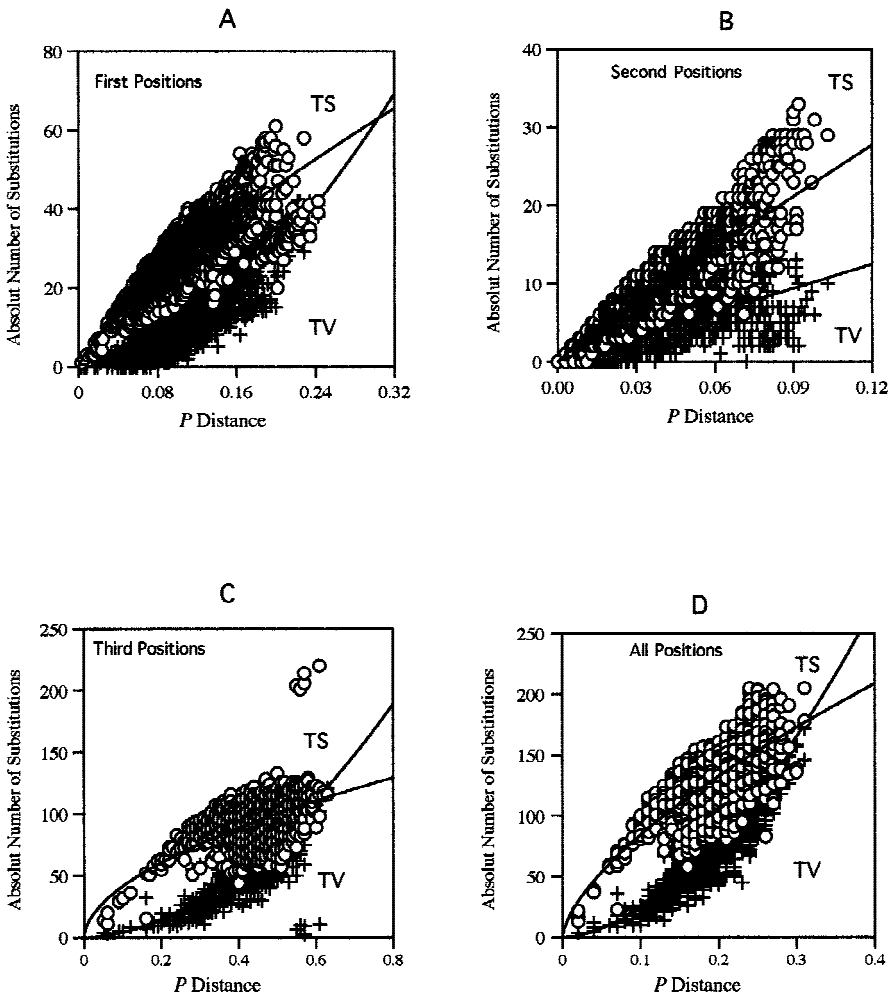

Fig. 2. Substitution pattern of transitions (TS = circles) and transversions (TV = crosses) for each codon position (A, B, C), and all cyt-*b* data (D). The number of transitions and transversions is plotted against total sequence divergence (uncorrected *p* distance) for all pairwise comparisons of taxa.

rates in this group (see below). Other genetic distances within the family Cichlidae are shown in Table 5.

Phylogenetic Analysis

Equally weighted (i.e., unweighted) parsimony analyses considering all codon positions resulted in a single MP tree (L = 7964; CI = 0.160; RI = 0.402). The topology of the MP tree (not shown) was strongly incongruent with previous results (Fig. 1). The major discrepancies included paraphyly of the geophagine (Neotropical group) and paraphyly of the African and Neotropical assemblage, since *Heterochromis,* an African genus, was placed among the Neotropical lineage. The bootstrap values were very low, and high bootstrap support was only observed for a clade of East African cichlids. These results demonstrate limited resolution of the cyt-*b* data in the unweighted MP analysis.

Table 5. Genetic distance (uncorrected *p* distance) of cyt-*b* for the main lineages of cichlid fishes

Genetic distance (uncorrected *p* distance) between the main cichlid lineages		
Africans × Neotropical Average = 0.20 ± 0.02 Maximum = 0.29 Minimum = 0.15	Africans × Malagasy/Indian Average = 0.19 ± 0.02 Maximum = 0.23 Minimum = 0.16	Neotropical × Malagasy/Indian Average = 0.21 ± 0.02 Maximum = 0.29 Minimum = 0.16
Genetic distance (uncorrected *p* distance between genera) within representative cichlid lineages		
Africans Average = 0.15 ± 0.04 Maximum = 0.20 Minimum = 0.02	Malagasy/Indian Average = 0.18 ± 0.04 Maximum = 0.23 Minimum = 0.07	Neotropical Average = 0.19 ± 0.04 Maximum = 0.29 Minimum = 0.04
Heroines Average = 0.13 ± 0.02 Maximum = 0.17 Minimum = 0.04	Cichlasomines Average = 0.18 ± 0.03 Maximum = 0.22 Minimum = 0.08	Geophagines Average = 0.21 ± 0.04 Maximum = 0.29 Minimum = 0.12

The ME tree (score = 12.44) showed improved bootstrap support for the relationships (Fig. 3). The monophyly of the family Cichlidae was supported by a bootstrap value of 67% and high bootstrap support was observed for most groups of African cichlids and for a monophyletic grouping of the Neotropical cichlasomine and heroines. The only discrepancy, compared to the total evidence analysis (Farias et al., 2000, Fig. 1), was the paraphyly of the South America geophagines and the paraphyly of the African group, since *Heterochromis* was placed as the basal taxon in the Neotropical lineage, albeit, with very low bootstrap support. Both MP and ME trees were in agreement regarding the paraphyly of Malagasy/Indian cichlids, placing them as the basal lineages of the family Cichlidae.

Excluding Third Positions. Analyses of the data excluding the third position were unable to find strong support for intergeneric relationships for most cichlid lineages. The removal of all third codon positions from the data set resulted in 64 equally parsimonious trees (L = 1567; CI = 0.288; RI = 0.533) with a mostly unresolved consensus tree. This topology showed inconsistent results compared to the topology of total evidence analysis and the parsimony bootstrap values were relatively low (results not shown). The major disagreement included the paraphyly of the family Cichlidae, and the paraphyly of the African and Neotropical lineages. The same resolution was observed in the ME tree (score = 2.73).

Excluding Transitions at Third Positions. Only one most parsimonious tree (L = 8082; CI = 0.157; RI = 0.391) resulted from MP analysis when we considered first and second positions and excluded transitions at the third codon position. The MP tree (not shown) presented the same internal arrangements for the heroines-cichlasomines groups (Neotropical lineages) and African cichlids obtained using all the cyt-*b* data, without the exclusion of transitions at the third position. Despite the low bootstrap support, the geophagines form a paraphyletic group followed by *Retroculus* and *Cichla* as the basal lineages of the Neotropical assemblage. A major difference concerns the position of the African genus *Heterochromis,* now placed among the Malagasy/Indian cichlids. The ME tree (score = 12.52) shows similar topology and bootstrap values as presented in Fig. 3.

Transition/Transversion Weighting Schemes. Since the cyt-*b* data set showed a transition/transversion ratio = 4.0 (Table 3), we applied the weighting scheme where transversions were weighted 4× over transitions (ts1:tv4) under parsimony analysis for all codon positions. This weighting scheme yielded a single MP tree (L = 15094; CI = 0.170; RI = 0.472) which is shown in Fig. 4. This tree agrees with the total evidence analysis (Fig. 1) by showing Neotropical and African cichlids as monophyletic sister clades (since *Heterochromis* is now placed as the basal African genus), but differs from the total evidence tree by placing *Retroculus* (a basal Neotropical genus) within the geophagine group.

Exclusion of Fast Evolving Taxa. In our previous analyses, based on the 16S rRNA mitochondrial gene, high rates of molecular evolution in some Neotropical cichlid lineages have been discovered (Farias et al., 1999; Farias et al., 2000). Likewise, in the present study, the cyt-*b* data evidenced longer branches in the ME tree for some of Neotropical taxa (shown with an asterisk in Fig. 2). Therefore, we performed the two-cluster test, as implemented in the LINTRE program (Takezaki et al., 1995) to identify the taxa that show significantly higher rates of nucleotide substitution. Simulation results have shown that long branches can be problematic for parsimony and likelihood, and will cause both methods to fail under some conditions (Wiens and Hollingsworth, 2000). Given this, the taxa placed on significantly long branches (see Figs. 3 and 4) were excluded from subsequent analyses. Parsimony analysis without the long-branch taxa (LB taxa) resulted in two equally parsimo-

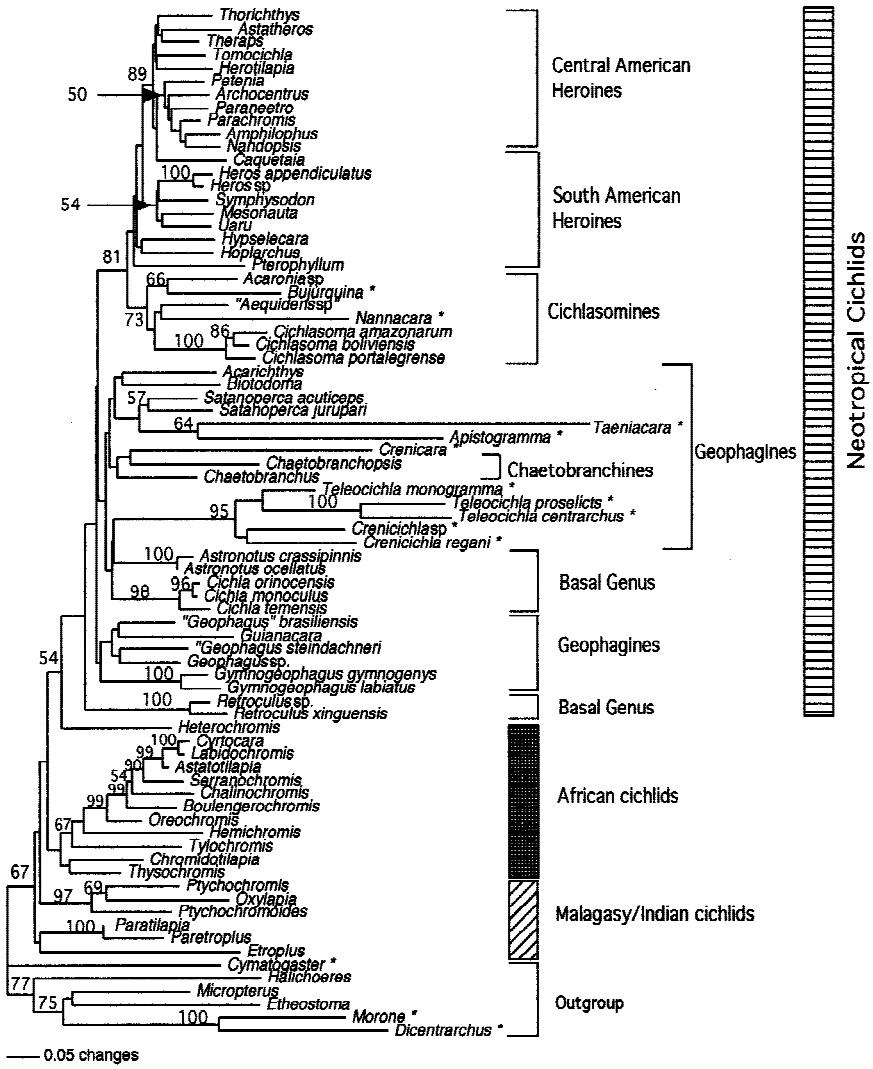

Fig. 3. ME tree from all cyt-*b* data set obtained using maximum likelihood distance (score = 12.44, Tratio = 3.78, Pinvar = 0.37, gamma shape = 0.71). The numbers above branches are bootstrap values (only values above 50% are shown). The asterisks identify the long branch taxa (LB taxa).

nious trees (L = 5937; CI = 0.180; RI = 0.425). The topologies of those trees were similar to the other MP trees where no weighting schemes were applied. The ME tree (score = 8.59) shows high bootstrap support for the monophyly of the family Cichlidae, the Neotropical lineage, and for the sister clades cichlasomine-heroines (Fig. 5). The geophagines appear as a monophyletic group, albeit, not strongly supported by bootstrap values. In agreement with the total evidence tree, the genera *Cichla, Astronotus,* and *Retroculus* were placed basal in the Neotropical assemblage. In agreement with the all ME analyses (e.g., Fig. 3), *Heterochromis* forms the sister lineage to the Neotropical, rather than the African cichlids. However, this is not convincingly supported by bootstrap support, and disagrees with the total evidence analysis.

Nine trees resulted from MP analysis excluding transitions at third positions (L = 2505; CI = 0.217; RI = 0.548), a strict consensus of which is shown in Fig. 6. These trees were very similar to the MP trees using ts1:tv4 at first and second position and excluded transitions at the third position (50 trees; L = 3114; CI = 0.273; RI = 0.587). MP resulted in high bootstrap support for the monophyly of the family Cichlidae, monophyly of Neotropical lineage, monophyly of the sister clades cichlasomine-heroines, and high bootstrap support for the monophyly of East African cichlids.

Tables 3 and 4 show an overview of the results obtained with different analyses applied to cyt-*b*. In general, the results indicate an improvement of the phylogenetic signal contained in the cyt-*b* data set when the LB taxa were excluded from the analysis.

As expected from the overall taxonomic congruence among cyt-*b* data and the total evidence tree (Fig. 1) the partition-homogeneity test indicated no significant heterogeneity among the data partitions (Table 6), and thus all data available (mitochondrial genes, nuclear genes, and morphological data) were combined in a total evidence analysis, in which transitions were excluded for the third positions of cyt-*b* data. The combined data set comprised a total of 2595 characters for 32 taxa. Of 1225 variable characters, 902 were phylogenetically informative under parsimony. The total evidence analysis resulted in only one parsimonious tree (L = 3379; CI =

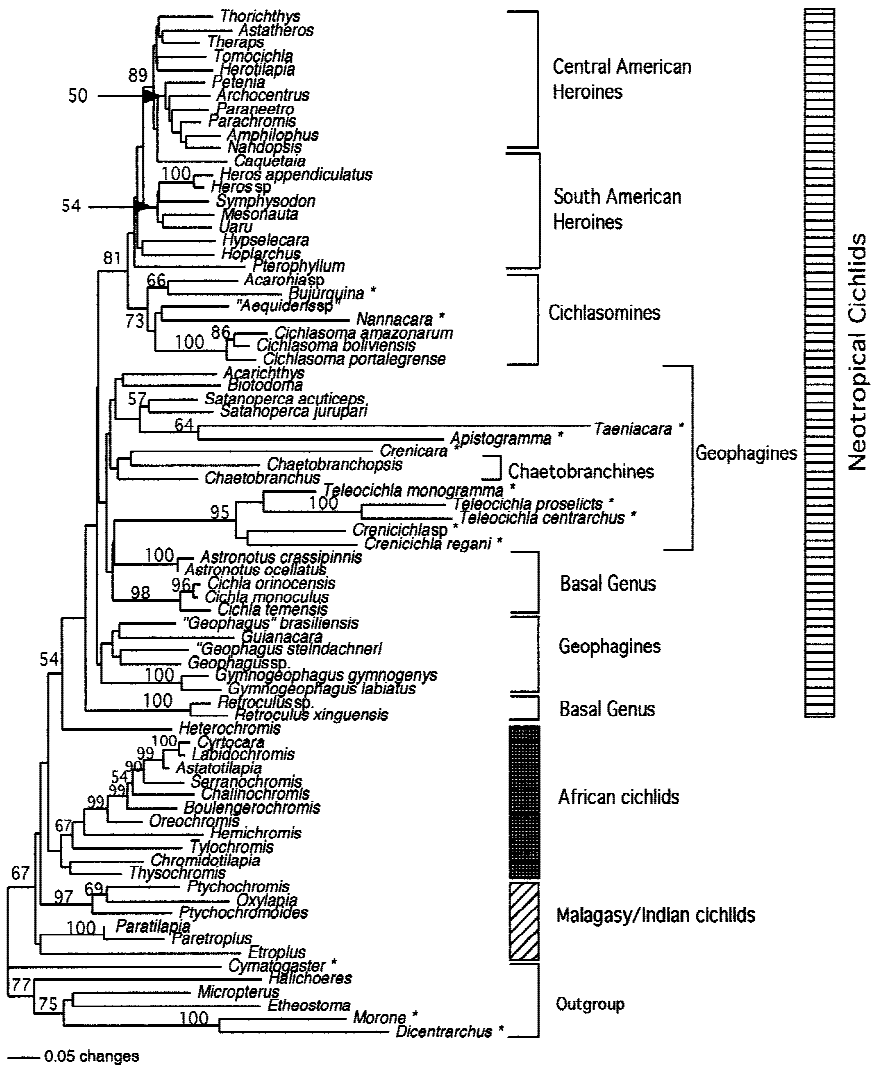

Fig. 4. Single MP tree (L = 15094; CI = 0.170; RI = 0.472) from all cyt-*b* data using the weighting schemes ts1:tv4 for all codon positions. The asterisks identify the long branch taxa (LB taxa).

0.390; RI = 0.434) that is shown in Fig. 7. This tree is highly congruent to the previous total evidence tree (Fig. 1). The only main difference involved the placement of Neotropical genus *Astronotus.* While the previous total evidence tree placed *Astronotus* as one of the basal lineage of Neotropical assemblage (Fig. 1) the present total evidence tree placed *Astronotus* as closed related to cichlisomines-heorines groups (Fig. 7) despite of low bootstrap values.

Amino Acid Data. Cyt-*b* in cichlid fishes, as in other vertebrates, suffers effects of functional constraints on amino acid evolution, resulting in transmembrane regions more variable (54.2%) than either outer membrane (20.5%) or inner membrane (25.3%). The transmembrane regions are relatively unconstrained by specific protein functions (Naylor et al., 1995), and the evolutionary rates in these parts are expected to be much higher. On the other hand, the portions of the outer membrane and inner membrane segments perform essential biological roles (Esposti et al., 1993), and the rates of substitutions of amino acids in these parts are expected to be much lower.

In addition, we also performed phylogenetic analysis based on amino acid sequences from the cyt-*b* data set. As expected, the topologies of the MP and ME trees (not shown) were strongly incongruent with previous results (Fig. 1), and presented very low bootstrap supports. The major discrepancies included paraphyly of the family Cichlidae, since the outgroup taxa were placed among the Malagasy/Indian cichlids, paraphyly of the African and Neotropical assemblage, since *Heterochromis,* was placed among the Neotropical lineage, and the paraphyly of the geophagine (Neotropical group). Probably, the major problem responsible for the low resolution in cyt-*b* protein is the limited number of parasimony-informative characters to construct a well-supported hypothesis about the inter-relationships of the cichlids.

Discussion

Codon Position Analysis

The cyt-*b* gene contains discrete character classes (i.e., the three codon positions), which exhibit mutation rates

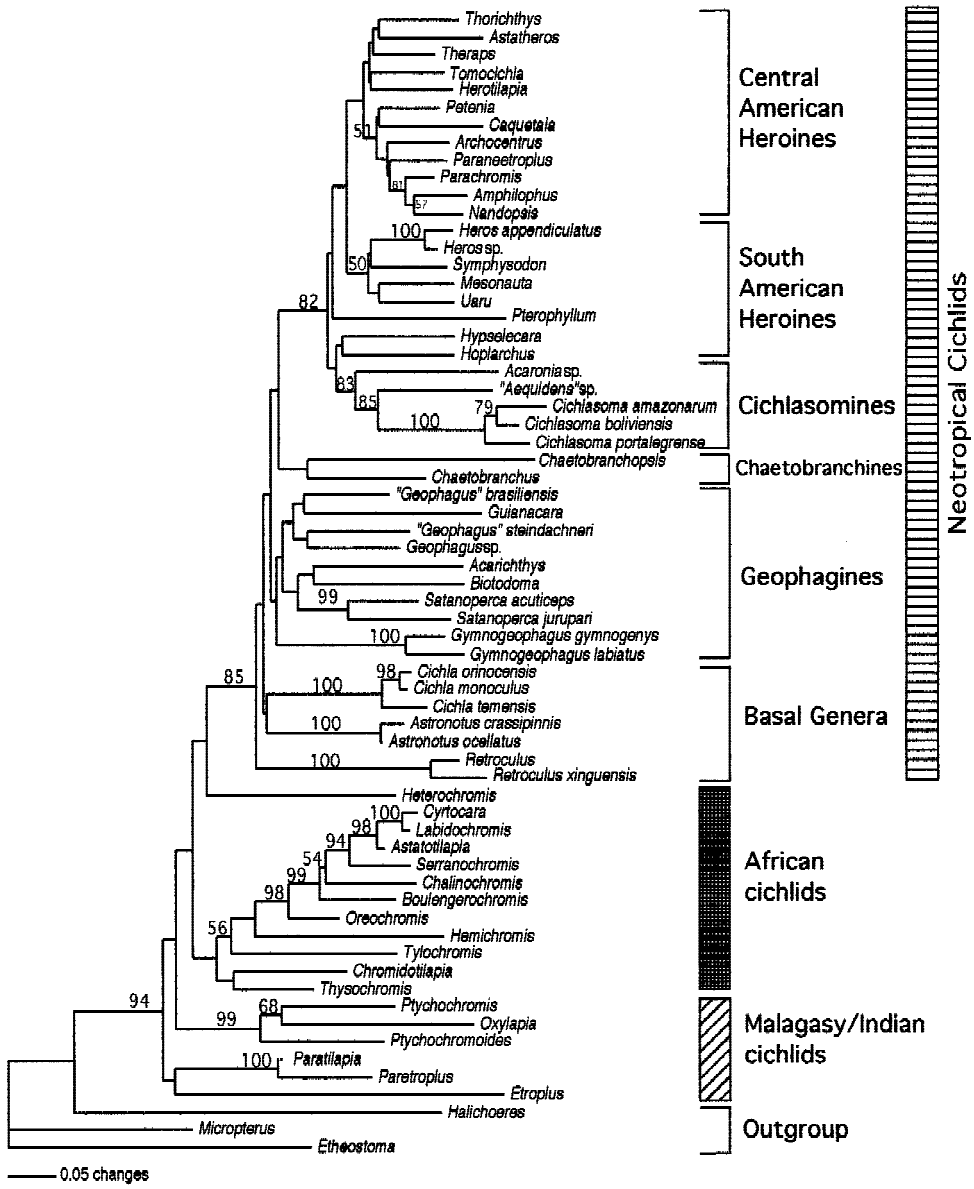

Fig. 5. ME tree from cyt-*b* data set obtained using maximum likelihood distance (score = 12.44; Tratio = 3.78; Pinvar = 0.37; gamma shape = 0.71) in which the LB taxa were excluded from the data. The numbers above branches are bootstrap values (only values above 50% are shown).

ranging from rapid to conservative (Irvin et al., 1991). Several studies have assessed the variability, signal to noise ratio, and utility for phylogenetic inferences of cyt-*b* (e.g., Griffiths, 1997; Groth, 1998; Björklund, 1999; Prychitko and Moore, 2000). Overall, the present results show that values of *k* (ts/tv ratio), the proportion of invariable sites, and α (gamma-shape distribution parameter) varied substantially across codon positions (Table 3). Relatively low α values for first and second positions suggest that most sites have either very low substitution rates or are invariable, whereas a few sites exist with very high rates (Sullivan et al., 1995; Yang, 1996; Voelker and Edwards, 1998). Third codon position, however, showed a higher value of α, suggesting a more uniform distribution, in which most sites have intermediate rates and few sites have very low or very high rates (Yang, 1996). Third position also had much higher *k* value relative to first and second positions (Table 3). These patterns also were observed in the saturation plots for each codon position (Fig. 2), in which only the curve of genetic distance versus transition at third codon position shows saturation (Fig. 2B, C, D). Meyer (1993) found the same pattern of saturation in cyt-*b* sequences for a few representative taxa from the family Cichlidae, suggesting that transitions at third position would not be reliable indicators of evolutionary relationships for the whole family, although they could be informative among closely related groups of species.

When the third positions were excluded from the parsimony analysis, the topology changed to an unreasonable grouping of Neotropical cichlids and included several polytomies, suggesting that these sites contain

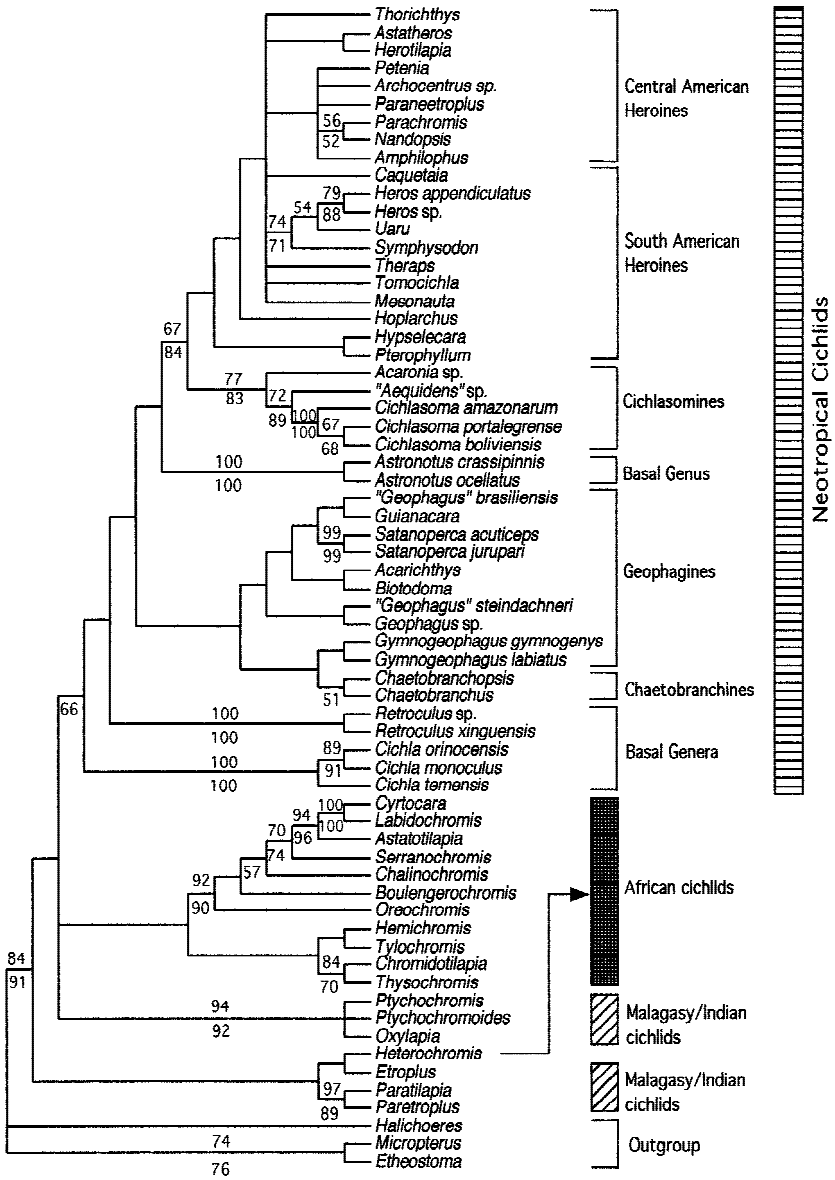

Fig. 6. Strict consensus of 9 MP trees from cyt-*b* data (L = 2505; CI = 0.217; RI = 0.548) without LB taxa and excluding transitions at third codon positions. The numbers above branches are bootstrap values (only values above 50% are shown). The numbers below internodes correspond to bootstrap values to the MP analysis without transitions at third positions and applying the weighting scheme ts1:tv4 to the first and second positions (50 trees; L = 3114; CI = 0.273; RI = 0.587).

Table 6. *P* values for partition-homogeneity tests (100 random replicates)

Comparisons	*P* value
Cytochrome *b* × 16S rRNA	0.27
Cytochrome *b* × the rest (16S, nuclear loci and morphology)	0.10
Mitocondrial genes × nuclear loci	0.99
Molecular data × morphological data	0.02

Note. Significant heterogeneity at the 99% confidence level.

important signal and that transitions were informative at least for closely related species. Indeed, several studies have shown that removal of third positions resulted in a problem of finding the true tree, especially at low levels of variation (Edward et al., 1991; Ortí et al., 1994; Håstad and Björklund, 1998; Broughton et al., 2000). Lydeard and Roe (1997) studied the phylogenetic utility of the cyt-*b* gene for inferring high level relationships among the Actinopterygian fishes. They found a strong phylogenetic signal when the third codon position was excluded for studying deep relationships. Although strong support was shown for the monophyly of certain groups within the Perciformes, no robust pattern of relationship among groups was found. The authors found some support for the monophyly of family Cichlidae and strong support for monophyly of Neotropical cichlids. However, the main cichlid lineages were poorly represented in their phylogeny of cyt-*b*.

Studies of the cyt-*b* gene, have shown that some categories of base substitution become saturated with change after 10 million years of divergence time (Irwin et al., 1991; Kornegay et al., 1993; Moore et al., 1999) due to "multiple hits" (Brown et al., 1982). Given that divergence of many cichlid groups probably occurred during the late Cretaceous (Stiassny, 1991) the saturation

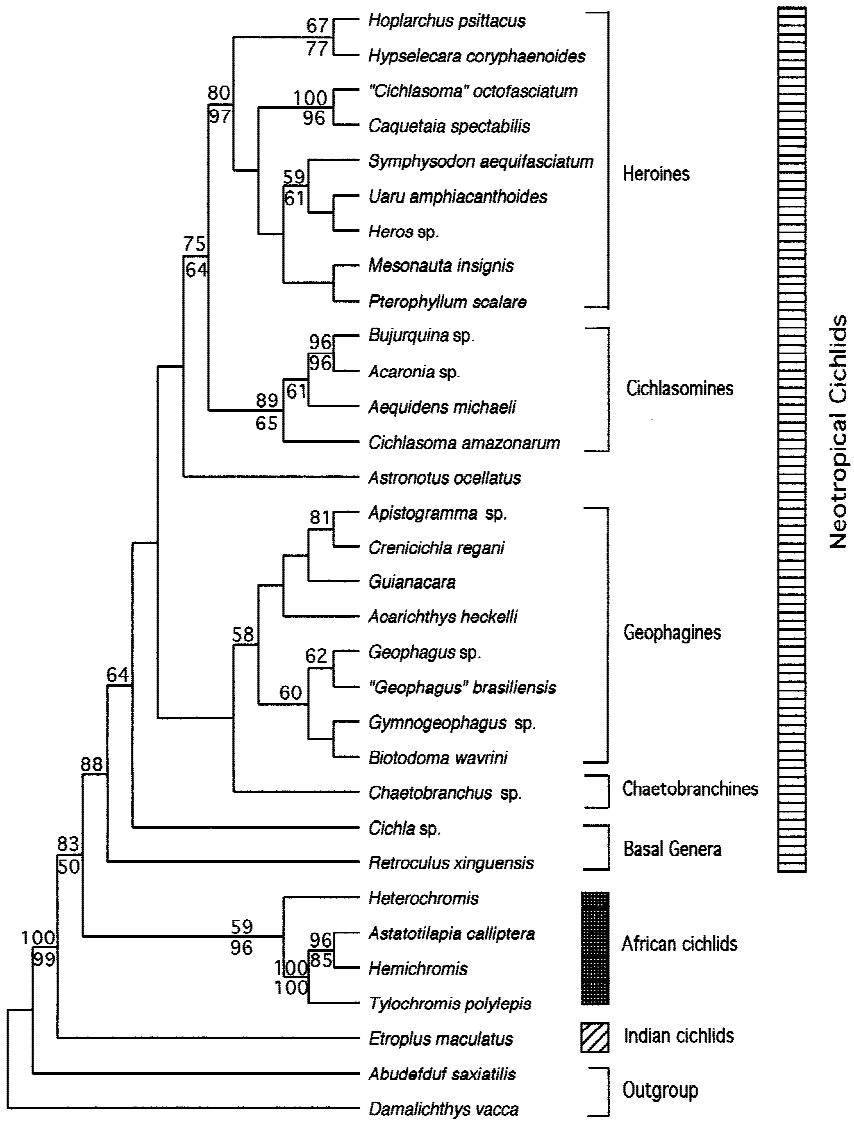

Fig. 7. Single MP tree (L = 3379; CI = 0.390; RI = 0.434) from combined mitochondrial data (rRNA 16S, Cytochrome *b*), nuclear data (Tmo-M27, Tmo-4C4) and morphological data (91 characters) in a total evidence analysis (2595 characters, 32 taxa). The numbers above branches are bootstrap values from MP analysis (only values above 50% are shown). The numbers below internodes correspond to bootstrap values from the ME analysis (only values above 50% are shown). For the cyt-*b* data partition, transitions at third positions were excluded for the analysis.

effect was expected in this study. The results of saturation analysis (Fig. 2A) show that for genetic distance close to 0.15 (uncorrected *p* distance) the number of transitions decreased compared to transversions, indicating saturation. As can be seen in Table 5, the lowest nucleotide divergence values were observed within groups with recent divergence, such as among East African cichlids and among the Neotropical heroines. However, in other groups the genetic distance values exceeded 0.15, especially in the Neotropical geophagine group. This indicates that, in fact, saturation may be a problem at the intergeneric level for some of the Neotropical taxa.

*Limits of Resolution of Cyt-*b *for Cichlid Fishes*

In the present study the phylogenetic utility of cyt-*b* was evaluated using taxonomic congruence (Miyamoto and Fitch, 1995; Kluge, 1998) by comparing the results obtained with cyt-*b* to a total evidence analysis (Fig. 1, Farias et al., 2000).

Alternative weighting schemes and taxon-exclusion sets applied in MP analysis resulted in different tree topologies. Compared to the total evidence tree (Fig. 1) the major discrepancies were due to some Neotropical groups (e.g., geophagines) and the placement of the African genus *Heterochromis* in the Neotropical lineage (no weighting schemes applied), or either as basal lineage of the African clade (weighting schemes applied) or within Malagasy/Indian lineage (weighting schemes applied). The exclusion of LB taxa in the present study resulted in better phylogenetic signal (see Table 4 and Fig. 6).

Methods such ME and maximum likelihood, under models that consider multiple substitutions that occur along long branches, rate heterogeneity among sites, dif-

ferent equilibrium nucleotide frequencies, and a ts:tv rate, have been shown to be consistent under more experimental conditions (Lyons-Weiler and Takahashi, 1999; Sanderson et al., 2000; Wiens and Hollingsworth, 2000). Indeed, all ME trees obtained here had similar topologies, regardless of taxon sampling schemes.

Overall, regardless of the phylogenetic methods or weighting strategy employed, some clades were always stronly supported: cichlasomines-heroines (Neotropical groups), African cichlids (except riverine taxa), and the basal position of the Malagasy/Indian cichlids. Although we found similar topologies for the relationship among heroines and cichlasomines with high bootstrap support for their monophyly, the internal arrangements for heroines, believed to be the most recent lineage in Neotropical radiation (Farias et al., 1998), was not well resolved. Roe et al. (1997) and Martin and Bermingham (1998) using cyt-*b* as a molecular marker to infer the relationship among Central American cichlids found saturation at third codon position of this gene, and they also were unable to find strong support for intergeneric relationships. The lack of resolution among closely related genera may reflect rapid speciation events among the Central America heroines and insufficient time for the accumulation of many shared derived substitutions to reliably resolve relationships among these rapidly diverging groups. Indeed, Martin and Bermingham (1998) suggested that the diversification within genera for Central American cichlids pointed to an adaptive radiation. This scenario was also supported by ecomorphological studies (Winemiller et al., 1995). However, neither cyt-*b* data nor the total evidence analyses support the hypothesized monophyly of Central American heroines, since *Caquetaia spectabilis* (only distributed in South America) is placed within the Central American group (Figs. 3–6).

In our previous findings, relative-rate tests suggested that Neotropical cichlids have experienced accelerated rates of molecular evolution (Farias et al., 1999, 2000) especially among geophagines. In the present study the two-cluster analysis (Takazaki et al., 1995) showed that at least five genera (eight species) from the geophagine groups had significantly higher substitution rates. Interestingly, almost all of them are dwarf or smaller sized cichlids (See Fig. 3). This observation, in agreement with our previous findings, suggests that the rate of mitochondrial evolution is faster for dwarf and smaller-sized Neotropical cichlids. Increased substitution rates may have many possible explanations such as generation time (Kohne, 1970), population size effects (Ohta, 1972, 1992), and metabolic rate (Martin and Palumbi, 1993; Rand, 1994). It is improbable that rates of molecular evolution are determined by a single factor (Avise, 1994), since hypotheses related to body size are not independent of generation time effects, population size, and metabolic rates (Martin and Palumbi, 1993; Nunn and Stanley, 1998). Unfortunately, there is not sufficient physiological and natural history information available for those species. However, we believe that fast rates of molecular evolution in these taxa are possibly affected by body size, since the smaller species tend to be associated with longer branches in ME trees.

According to Hillis and Huelsenbeck (1992) a particular data set may contain historical information about one clade and none on another. Then, despite some limitations of cyt-*b* to infer relationships among the major clades of cichlid fishes the partition-homogeneity test (Table 6) did not show significant heterogeneity levels in the comparison between cyt-*b* data and the total evidence tree (Fig. 1). When data partitions are congruent, it is generally accepted that all data should be combined to obtain the best estimate of phylogeny (Chippindale and Wiens, 1994). Thus, the cyt-*b* data was combined with the total evidence tree. The new total evidence tree (Fig. 7), resulting from the analysis, is highly congruent to the previous total evidence tree (Fig. 1). The only main difference involved the placement of *Astronotus.*

In conclusion, the Cyt-*b* gene, either alone, or in combination with total evidence tree confirms: (i) Malagasy/Indian cichlids are paraphyletic and the most basal group in the family, suggesting two distinct groups: *Etroplus* + (*Paretroplus* + *Paratilapia*), and (*Ptychochromis, Ptychochromoides, Oxylapia*); (ii) the West African cichlids are placed as the most basal African taxa; (iii) the African genus *Heterochromis* is an independent lineage of the African assemblage; (iv) among the Neotropical cichlids, *Retroculus* and *Cichla* are placed as the most basal genera; and (v) *Astronotus* is closely related to the cichlasomines-heroines which are monophyletic and sister group. Although it is apparent that cyt-*b* fails to resolve deeper relationships in the major lineage of the family Cichlidae, i.e., geophagine group (from Neotropical lineages) and the placement of *Heterochromis* in African lineage, it was shown to be useful for assessing more recent relationships (e.g., cichlasomines, heroines, and East African cichlids).

Acknowledgments. We thank the many colleagues who provided tissue samples and species identification: A. Basolo, C. Silva, E. Feldberg, E. Ferreira, M. Jegú, I. Koslowski, J. Alves-Gomes, J. Porto, J. Zuanon, M. Garcia, R. Reis, S. Kullander, T. Ploesch, W. Staeck, and P. Wimberger. We thank Tom Titus for technical support and Y. Van de Peer for providing useful comments on earlier versions of the manuscript. This work was supported in part by University of Nebraska-Lincoln (USA), Universidade Federal do Pará (Brazil), and Universidade do Amazonas (Brazil). We thank R. Zardoya and C. Sturmbauer for insightful comments on earlier versions of this manuscript. We also acknowledge financial support from the US National Science Foundation (BSR-912367, BSR-9119867, DEB-9615178), the Fond der Chemischen Industrie, the Deutsche Forschungsgemeinschaft, and University of Konstanz to A.M.

References

Avise JC (1994) Molecular markers, natural history, and evolution. Chapman and Hall, New York

Abouheif E, Zardoya R, Meyer A (1998) Limitations of metazoan 18S rRNA sequence data: implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the cambrian explosion. J Mol Evol 47:394–405

Björklund M (1999) Are third position really that bad? a test using vertebrate cytochrome b. Cladistics 15:191–197

Broughton RE, Stanley SE, Durrett RT (2000) Quantification of homoplasy for nucleotide transitions and transversions and a reexamination of assumptions in weighted parsimony analysis. Syst Biol 4:617–627

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial of DNA sequences of primates: tempo and mode of evolution. J Mol Evol 18:225–239

Cantatore P, Roberti M, Pesole G, Ludovico A, Milella F, Gadaleta MN, Saccone C (1994) Evolutionary analysis of cytochrome *b* sequences in some Perciformes: evidence for a slower rate of evolution than in mammals. J Mol Evol 39:589–597

da Silva MNF, Patton JL (1993) Patterns of evolutionary divergence in arboreal Amazonian echimyid rodents: the perspective of mtDNA sequence variation. Mol Phyl Evol 2:243–255

Edward SV, Arctander P, Wilson AC (1991) Mitochondrial resolution of a deep branch in the genealogical tree for perching birds. Proc R Soc London B Biol Sci 243:99–107

Esposti DM, De Vries S, Crimi M, Ghelli A, Patarnello T, Meyer A (1993) Mitochondrial cytochrome *b*: evolution and structure of the protein. Biochim Biophys Acta 1143:243–271

Farias IP, Schneider H, Sampaio I (1998) Molecular phylogeny of neotropical cichlids: the relationships of cichlasomines and heroines. In: Malabarba LR, Reis RE, Vari RP, Lucena ZM, Lucena CAS (eds) Phylogeny and classification of neotropical fishes. porto alegre, edipucrs, pp. 499–508

Farias IP, Ortí G, Sampaio I, Schneider H, Meyer A (1999) Mitochondrial DNA phylogeny of the family Cichlidae: monophyly and fast molecular evolution of the neotropical assemblage. J Mol Evol 48:703–711

Farias IP, Meyer A, Ortí G (2000) Total evidence: molecules, morphology, and the phylogenetics of cichlids fishes. J Exp Zool (Mol Dev Evol) 288:76–92

Farris JS, Källersjö M, Kluge AG, Bult C (1994) Testing significance of incongruence. Cladistics 10:315–391

Farris JS, Källersjö M, Kluge AG, Bult C (1995) Constructing a significance test for incongruence. Syst Biol 44:570–572

Farris JS (1989) The retention index and rescaled consistency index. Cladistics 5:417–419

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1–15

Graybeal A (1993) The phylogenetic utility of cytochrome *b*: lessons from bufonid frogs. Mol Phy Evol 2:256–269

Griffiths CS (1997) Correlation of functional domains and rates of nucleotide substitution in cytochrome b Mol Phy Evol 7(3):352–365

Groth JG (1998) Molecular phylogenetics of finches and sparrows: consequences of character state removal in cytochrome *b* sequences. Mol Phy Evol 10(3):337–390

Håstad O, Björklund M (1998) Nucleotide substitution models and estimation of phylogeny. Mol Biol Evol 15:1381–1389

Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular evolutionary analyses. J Hered 83:189–195

Honeycutt RL, Nedbal MA, Adkins RM, Janecek LL (1995) Mammalian mitochondrial DNA evolution: a comparison of the cytochrome *b* and cytochrome *c* oxidase II genes. J Mol Evol 40:260–272

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of cytochrome *b* gene in mammals. J Mol Biol Evol 2:13–34

Johns GC, Avise J (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. Mol Biol Evol 15(11):1481–1490

Kidd KK, Sgaramella-Zonta LA (1971) Phylogenetic analysis: concepts and methods. Am J Human Genetics 23:235–252

Kirchman JJ, Whittingham LA, Sheldon FH (2000) Relationships among cave swallow populations (*Petrochelidon fulva*) determined by comparisons of microsatellite and cytochrome *b* data. Mol Phy Evol 14(1):107–121

Kluge AG, Farris JS (1969) Quantitative phyletics and the evolution of anurans. Syst Zool 18:1–32

Kluge AG (1998) Total evidence or taxonomic congruence: cladistics or consensus classification. Cladistics 14:151–158

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo SF, Villablanca FX, Wilson AC (1989) Dynamics of mtDNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci USA 86:6196–6200

Kohne DE (1970) Evolution of higher-organism DNA. Q Rev Biophys 33:327–375

Kornegay JR, Kocher TD, Williams LA, Wilson AC (1993) Pathways of lysozyme evolution inferred from the sequences of cytochrome *b* in birds. J Mol Evol 37:367–379

Kullander SO (1998) A phylogeny and classification of the South American Cichlidae (Teleoster: Perciformes). In: Malabarba LR, Reis RE, Vari RP, Lucena ZM, and Lucena CAS (eds) Phylogeny and classification of Neotropical fishes. Edipucrs, Porto Alegre, Brasil, pp. 461–498

Kumazawa Y, Nishida M (2000) Molecular phylogeny of Osteoglossoids: a new model for Gondwanian origin and plate tectonic transportation of the Asian arowana. Mol Biol Evol 17(2):1869–1878

Lamb T, Lydeard C (1994) A molecular phylogeny of the gopher tortoises, with comments on familial relationships within the Testudinoidea. Mol Phyl Evol 3:283–291

Lovejoy NR, de Araújo MLG (2000) Molecular systematics, biogeography, and population structure of Neotropical freshwater needlefishes of the genus Potamorrhaphis. Mol Ecol 9(3):259–268

Lydeard C, Roe KJ (1997) The phylogenetic utility of the mitochondrial cytochrome *b* gene for inferring relationships among Actinopterygian fishes. In: Kocher TD, Stepien CA (eds) Molecular systematics of fish. Academic Press, New York, pp. 285–303

Lyons-Weiler J, Takahashi K (1999) Branch length heterogeneity leads to nonindependent branch length estimates and can decrease the efficiency of methods of phylogenetic inference. J Mol Evol 49:392–405

Maddison WP, Maddison DR (1992) MacClade: analysis of phylogeny and character evolution, version 3.0. Sinauer, Sunderland, MA

Martin AP, Palumbi SR (1993) Body size, metabolic rate, generation time, and the molecular clock. Proc Natl Acad Sci USA 90:4087–4091

Martin AP, Bermingham E (1998) Systematics and evolution of lower Central American cichlids inferred from analysis of cytochrome *b* gene sequence. Mol Phyl Evol 9:192–203

Meyer A (1993) Evolution of mitochondrial DNA in fishes. In: Hochachka, Mommsen (eds) Biochemistry and molecular biology of fishes, vol. 2. Elsevier, Amsterdam, pp. 1–38

Meyer A (1994) Shortcomings of the cytochrome *b* gene as a molecular marker. Trend Ecol Evol 9(8):278–280

Meyer A, Wilson C (1990) Origin of tetrapods inferred from their mitochondrial DNA affiliation to lungfish. J Mol Evol 31:359–364

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature Lond 347:550–553

Meyer A, Kocher TD, Wilson AC (1991) African fishes—a replay. Nature 351:467–468

Mindell DP, Honeycutt RL (1990) Ribosomal RNA in vertebrates: evolution and phylogenetic applications. Annu Rev Ecol Syst 21:541–566

Miyamoto MM, Fitch WM (1995) Testing species phylogenies and phylogenetic methods with congruence. Syst Biol 44:64–76

Moore WS, DeFilippis VR (1997) The window of taxonomic resolution for phylogenies based on mitochondrial cytochrome *b*. In: Mindell DP (ed) Avian molecular evolution and systematics. Academic Press, New York, pp. 83–120

Moore WS, Smith SM, Prychitko, T (1999) Nuclear gene introns versus mitochondrial genes as molecular clocks. In: Adams N, Slotow R (eds) Proceedings of the 22nd International Ornithology Congress University of Natal, Durban, South Africa

Moritz C, Schneider CL, Wake DB (1992) Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. Syst Biol 41:273–291

Naylor GJP, Collins TM, Brown WM (1995) Hydrophobicity and phylogeny. Nature 373:565–566

Naylor GJP, Brown WM (1998) Amphioxus mitochondrial DNA, chordate phylogeny, and the limits of inference based on comparisons of sequences. Syst Biol 47(1):61–76

Normark BB, McCune AR, Harrison RG (1991) Phylogenetic relationships of neopterygian fishes, inferred from mitochondrial DNA sequences. Mol Biol Evol 8:19–834

Nunn GB, Stanley SE (1998) Body size effects and rates of cytochrome *b* evolution in tube-nosed seabirds. Mol Biol Evol 15(10):1360–1371

Ohta T (1972) Population size and rate of evolution. J Mol Evol 1:305–314

Ohta T (1992) The nearly neutral theory of molecular evolution. Ann Rev Ecol Syst 23:263–286

Ortí G, Bell MA, Reimchen TE, Meyer A (1994) Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migration. Evolution 48:608–622

Ortí G, Meyer A (1996) Molecular evolution of ependymin and the phylogenetic resolution of early divergences among teleost fishes. Mol Biol Evol 13(4):556–573

Ortí G, Meyer A (1997) The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. Syst Biol 46(1):75–100

Pääbo S (1990) Amplifying ancient DNA. In: Innes MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, San Diego, pp. 159–166

Posada D, Crandall KA (1998) MODEL TEST: testing the model of DNA substitution. Bioinformatics 14(9):817–818

Prychitko TM, Moore WS (2000) Comparative evolution of the mitochondrial cytochrome *b* gene and nuclear β-fibrinogen intron 7 in woodpeckers

Rand DM (1994) Thermal habit, metabolic rate, and evolution of mitochondrial DNA. Trends Ecol Evol 9:125–131

Roe KL, Conkel D, Lydeard C (1997) Molecular systematic of Middle American cichlid fishes and evolution of trophic-types in '*Cichlasoma* (*Amphilophus*)' and *C.* (*Thorichthys*). Mol Phyl Evol 7:366–376

Rocha-Olivares A, Rosenblatt RH, Vetter RD (1999) Molecular evolution, systematics, and zoogeography of the rockfish subgenus *Sebastomus* (*Sebastes,* Scorpaenidae) based on mitochondrial cytochrome *b* and control region sequences. Mol Phy Evol 11(3):441–458

Rzhetsky A, Nei M (1992) A simple method for estimating and testing minimum-evolution trees. Mol Biol Evol 9(5):945–967

Sanderson ML, Wojciechowski MF, Hu J-M, Khan TS, Brady SG (2000) Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. Mol Biol Evol 17(5):782–797

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

Song C (1994) Molecular evolution of the cytochrome *b* gene among Percid fishes. Unpublished dissertation thesis, University of Illinois at Urbana-Champaign, IL

Stiassny MLJ (1991) Phylogenetic intrarelationships of the family Cichlidae: an overview. In: Keenleyside MHA (ed) Cichlid fishes: behaviour, ecology, and evolution. Chapman and Hall, London, pp. 1–35

Streelman JT, Karl SA (1997) Reconstructing labroid evolution with single-copy nuclear DNA. Proc R Soc Lond B 264:1011–1020

Strimmer K, Haeseler A (1996) Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol Biol Evol 13:964–969

Sturmbauer C, Meyer A (1992) Genetic divergence, speciation, and morphological stasis in a lineage of African cichlid fishes. Nature Lond 359:578–581

Sturmbauer C, Meyer A (1993) Mitochondrial phylogeny of the endemic mouthbrooding lineages of cichlid fishes of Lake Tanganyika, East Africa. Mol Biol Evol 10:751–768

Sturmbauer C, Verheyen E, Meyer A (1994) Mitochondrial phylogeny of the Lamprologini, the major substrate spawning lineage of cichlid fishes from Lake Tanganyika in Eastern Africa. Mol Biol Evol 11:691–703

Sullivan J, Holsinger KE, Simon C (1995) Among-site rate variation and phylogenetic analysis of 12S rRNA in sigmodontine rodents. Mol Biol Evol 12:988–1001

Swofford DL (2001) PAUP* Phylogenetic analysis using parsimony and other methods, Version 4d64. Sinauer, Sunderland, MA (in press)

Taberlet P, Meyer A, Bouvet J (1992) Unusual mitochondrial DNA polymorphism in two local populations of blue tit (*Parus caerulens*). Mol Ecol 1:27–36

Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. Mol Biol Evol 12:823–833

Tsigenopoulos CS, Berrebi P (2000) Molecular phylogeny of North Mediterranean freshwater barbs (genus *Barbus*: Cyprinidae) inferred from cytochrome *b* sequences: biogeographic and systematic implications. Mol Phy Evol 14(2):165–179

Van de Peer Y, De Wachter R (1997) Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site rate variation in 18S rRNA. J Mol Evol 45:619–630

Voelker G, Edwards SV (1998) Can weighting improve bushy trees? models of cytochrome *b* evolution and the molecular systematics of pipits and wagtails (Aves: Motacillidae). Syst Biol 47(4):589–603

Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39:306–314

Yang Z (1996) Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol Evol 11:367–372

Yoder AD, Vilgalys R, Ruvolo M (1996) Molecular evolutionary dynamics of cytochrome *b* in Strepsirrhine primates: the phylogenetic significance of third-position transversions. Mol Biol Evol 13:1339–1350

Whitmore DH, Thai TH, Craft CM (1994) The largemouth bass cytochrome *b* gene. J Fish Biol 44:637–645

Wiens JJ, Hollingsworth BD (2000) War of the iguanas: conflicting molecular and morphological phylogenies and long-branch attractions in Iguanid lizards. Syst Biol 49(1):143–159

Winemiller KO, Kelso-Winemiller LC, Brenkert AL (1995) Ecomorphological diversification and convergence in fluvial cichlid fishes. Environ Biol.Fish 44:235–261

Zardoya R, Vollmer DM, Craddock C, Streelman JT, Karl SA, Meyer A (1996) Evolutionary conservation of microsatellite flanking regions and their use in resolving the phylogeny of cichlid fishes (Pisces: Perciformes). Proc R Soc Lond B 263:1589–1598

Zardoya R, Cao Y, Hasegawa M, Meyer A (1998) Searching for the closest living relative(s) of tetrapods through evolutionary analyses of mitochondrial and nuclear data. Mol Biol Evol 15(5):506–517

Zardoya R, Meyer A (1996) Evolutionary relationships of the coelacanth, lungfishes, and tetrapods based on the 28S ribosomal RNA gene. Proc Natl Acad Sci USA 93:5449–5454

Zardoya R, Doadrio I (1999) Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. J Mol Evol 49:227–237