Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 34 (2005) 227–244
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Molecular phylogeny and evidence for an adaptive radiation of geophagine cichlids from South America (Perciformes: Labroidei)

Hernán López-Fernández, Rodney L. Honeycutt*, Kirk O. Winemiller

Section of Ecology and Evolutionary Biology, Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 77843-2258, United States

Received 3 September 2004

Abstract

Nucleotide sequences from the mitochondrial ND4 gene and the nuclear RAG2 gene were used to derive the most extensive molecular phylogeny to date for the Neotropical cichlid subfamily Geophaginae. Previous hypotheses of relationships were tested in light of these new data and a synthesis of all existing molecular information was provided. Novel phylogenetic findings included support for : (1) a 'Big Clade' containing the genera *Geophagus sensu lato*, *Gymnogeophagus*, *Mikrogeophagus*, *Biotodoma*, *Crenicara*, and *Dicrossus*; (2) a clade including the genera *Satanoperca*, *Apistogramma*, *Apistogrammoides*, and *Taeniacara*; and (3) corroboration for Kullander's clade Acarichthyini. ND4 demonstrated saturation effects at the third code position and lineage-specific rate heterogeneity, both of which influenced phylogeny reconstruction when only equal weighted parsimony was employed. Both branch lengths and internal branch tests revealed extremely short basal nodes that add support to the idea that geophagine cichlids have experienced an adaptive radiation *sensu* Schluter that involved ecomorphological specializations and life history diversification.

Keywords: Geophaginae; Molecular phylogeny; Neotropical cichlids; Adaptive radiation; Rate heterogeneity

1. Introduction

The Neotropical cichlid subfamily Geophaginae encompasses 18 genera and over 180 described species (Kullander, 2003), with many more in need of description (e.g., Kullander, 2003; López-Fernández and Taphorn, 2004). Although our knowledge of geophagine biology is limited, this group of fishes displays diverse ecology, morphology, and reproductive behavior. Their overall morphological and behavioral diversity suggests ecomorphological specialization for feeding and habitat use (e.g., Winemiller et al., 1995; López-Fernández, unpublished). For instance, some taxa share a common feeding mode based on sifting of benthic invertebrates (e.g., Lowe-McConnell, 1991; Winemiller et al., 1995), while others are strict piscivores. Geophagines also exhibit a variety of reproductive modes, from typical substrate spawners to mouth-brooding, and are the only riverine cichlids approaching the reproductive versatility of lacustrine cichlids (e.g., Barlow, 2000; Weidner, 2000; Wimberger et al., 1998). Several genera and species of geophagines are syntopic in South American rivers (e.g., Arrington and Winemiller, 2003; Winemiller et al., 1995), thus ecomorphological and behavioral specialization may facilitate niche partitioning within species-rich ecological communities. Although this Neotropical fish assemblage offers many opportunities for those interested in evolutionary ecology and the processes responsible for ecomorphological diversification, such studies require an interpretive framework based on knowledge of phylogenetic relationships and the timing and duration of speciation events. As can be seen in the following paragraphs, considerable controversy still

* Corresponding author. Fax: +1 979 845 4096.
E-mail address: rhoneycutt@tamu.edu (R.L. Honeycutt).

1055-7903/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2004.09.004

surrounds the phylogenetics and classification of the Geophaginae.

Recent phylogenetic analyses of the Cichlidae (Farias et al., 1999, 2000, 2001; Kullander, 1998) have improved understanding of higher-level relationships within the Neotropical clade (e.g., establishment of subfamilies), yet even here there are disagreements. Using a morphology-based phylogeny, Kullander (1998) subdivided the Neotropical Cichlidae and the African genus *Heterochromis* into six subfamilies. The Retroculinae (genus *Retroculus*) and Cichlinae (*Cichla*, *Crenicichla*, and *Teleocichla*) constituted the basal clades of the American assemblage. The African Heterochromidinae (*Heterochromis*) was nested between the latter two and Astronotinae (*Astronotus* and *Chateobranchus*), which were sister to the rest of the Neotropical assemblage, thus the Neotropical cichlids were rendered paraphyletic. The more derived subfamilies Cichlasomatinae and Geophaginae included all the remaining genera within the American cichlids. Cichlasomatinae included over 25 genera placed in the tribes Cichlasomini, Heroini, and Acaroniini. Geophaginae included 16 genera and was divided into three tribes: Acarichthyini (*Acarichthys* and *Guianacara*), Crenicaratini (*Biotoecus*, *Crenicara*, *Dicrossus*, and *Mazarunia*), and Geophagini (*Geophagus*, *Mikrogeophagus*, '*Geophagus*' *brasiliensis*, '*Geophagus*'*steindachneri*, *Gymnogeophagus*, *Satanoperca*, *Biotodoma*, *Apistogramma*, *Apistogrammoides* and *Taeniacara*).

In disagreement with the definition of Kullander, molecular studies (Farias et al., 1998, 1999) and total evidence analyses (Farias et al., 2000, 2001), including Kullander's morphological data, supported a monophyletic Neotropical Cichlidae and the placement of *Heterochromis* as basal to the African clade. Farias et al. (1999, 2000, 2001) also found the genera *Crenicichla* and *Teleocichla* nested within the Geophaginae, expanding the subfamily to 18 genera, and challenging the previously proposed relationship between *Crenicichla*, *Teleocichla* and the basal genus *Cichla* (Stiassny, 1987, 1991; Kullander, 1998). Despite the contribution of these studies to the clarification of higher-level relationships, the lack of relevant taxa limits their phylogenetic resolution and leaves many questions of geophagine relationships unanswered. Although geophagine monophyly seems indisputable, there is considerable disagreement between morphological and molecular evidence when analyzed separately, and the relationships within the Geophaginae are not clear. Kullander's study included an extensive taxon sampling of cichlids, and his proposed geophagine relationships were based on the analysis of 13 genera of geophagines (*sensu* Kullander) plus *Crenicichla* and *Teleocichla*. The studies of Farias et al. (1999, 2000, 2001) are not suited for testing Kullander's hypothesis, because taxon sampling is insufficient in their combined analyses. Farias et al. (2000) included only 11 genera in their molecular total evidence analysis and 9 in the combined analysis of molecular and morphological data. Their total molecular evidence analyses lacked the genera *Satanoperca*, *Biotoecus*, *Crenicara*, *Dicrossus*, and the '*Geophagus*' *steindachneri* group (Farias et al., 2000), and several additional genera were absent from their analyses of molecular and morphological data combined (Farias et al., 2000, 2001). Clearly, exclusion of these taxa makes it impossible to test the monophyly of Kullander's (1998) tribes Crenicaratini and Geophagini, and impedes further resolution of internal relationships within the subfamily. Better taxon sampling and incorporation of new data are needed to clarify relationships within the Geophaginae.

In this paper, we used newly obtained sequences from the mitochondrial ND4 gene and the nuclear RAG2 gene to derive a molecular phylogeny of the Geophaginae. We also performed a combined analysis of the new data with previously published sequences from Neotropical cichlids, thus integrating all available molecular evidence into the resolution of geophagine relationships. In addition, taxon sampling was largely expanded with respect to previous studies to include 16 of the 18 genera and 30 species of geophagines. We used these data to: (1) evaluate relationships among genera of Geophaginae, comparing our results to those from previous studies; (2) determine the extent of substitutional saturation and heterogeneity of molecular evolutionary rates within the subfamily and their effect on phylogenetic reconstruction; and (3) evaluate the phylogenetic evidence supporting an adaptive radiation of the group, as a necessary step in studying patterns of evolution of morphology, ecology, and behavior within the Geophaginae.

2. Materials and methods

2.1. Taxon sampling

DNA sequence data were collected for both the mitochondrial ND4 (NADH dehydrogenase subunit 4) gene and the nuclear RAG 2 gene (Recombination Activating Gene 2). Specimens examined included 21 genera and 38 species of Neotropical cichlids, and when possible, sequences were obtained from two individuals of each species. Ingroup samples included 16 of 18 genera and 30 species (Table 1) of Geophaginae *sensu* Farias et al. (1999, 2000, 2001), excluding only the genera *Teleocichla* and *Mazarunia*, for which tissue samples could not be obtained. The absence of these taxa from the dataset should not affect the resolution of the phylogeny, because *Teleocichla* has been clearly established as the sister group of *Crenicichla* (Farias et al., 1999, 2000, 2001; Stiassny, 1987), and *Mazarunia* is known to be related to *Crenicara* and *Dicrossus* (Kullander, 1990). The genus *Geophagus sensu lato* includes three distinct genera, of

Table 1
List of taxa for which ND4 and RAG2 were sequenced in this study with collection localities and accession numbers to GenBank; more detailed locality data are available from HLF on request

	# Fish	Collection locality	Accession numbers					
			ND4	16S	Cyt *b*	RAG2	Tmo-4C4	Tmo-M27
Outgroup taxa								
Astronotus sp.	2	Aquarium trade	AY566776	AF048998[b] *A. crassipinnis*	AB018987[e] *A. ocellatus*	AY566740	AOU70345[f] *A. ocellatus*	AOU63668[d] *A. ocellatus*
Cichla intermedia	1	Río Cinaruco, Venezuela	AY566788			AY566752		
Cichla orinocensis	1	Río Cinaruco, Venezuela	AY566786	AF049018[b]	AF370643[a]	AY566751	AF113064[c]	AF112602[c]
Cichla temensis	2	Río Cinaruco, Venezuela	AY566793	AF049019[b]	AF370644[a]	AY566755		
Retroculus sp.	1	Macapá, Brazil	AY566774	AF112591[b]	AF370640[a]	AY566737	AF113061[c]	AF112599[c]
Cichlasomatinae								
Cichlasoma orinocense	2	Apure, Venezuela	AY566778	AF045845[i]	AF145128[c] *C. bimaculatum*	AY566747	AF113075[c] *C. amazonarum*	AF112613[c] *C. amazonarum*
Hoplarchus psittacus	1	Río Cinaruco, Venezuela	AY566789	AF045855[i]	AF370673[a]	AY566760	AF113074[c]	AF112612[c]
Mesonauta egregius	2	Caño Maporal, Venezuela	AY566782	AF045859[i] *M. insignis*	AF370675[a] *M. insignis*	AY566748	AF113066[c] *M. insignis*	AF112604[c] *M. insignis*
Geophaginae								
Acarichthys heckelii	2	Aquarium trade	AY566768	AF049004[b]	AF370653[a]	AY566733	AF113083[c]	AF112621[c]
Apistogrammoides pucallpaensis	2	Río Orosa, Perú	AY566770			AY566735		
Apistogramma hoignei	1	Caño Maporal, Venezuela	AY566781	AF049006[b] *Apisto*. sp.2	AF370656[a] *Apisto*. sp.	AY566746	AF113095[c] *Apisto*. sp.2	AF112633[c] *Apisto*. sp.2
Apistogramma agassizi	2	Río Orosa, Perú	AY566787	AF049005[b] *Apisto*. sp.1		AY566749		
Biotodoma wavrini	2	Río Cinaruco, Venezuela	AY566784	AF049007[b]	AF370657[a]	AY566726	AF113082[c]	AF112620[c]
Biotodoma cupido	2	Río Orosa, Perú	AY566772			AY566723		
Biotoecus dicentrarchus	2	Río Cinaruco, Venezuela	AY566792	AF112641[g] *Biotoecus* sp.		AY566754		
Crenicara punctulatum	2	Río Nanay, Perú	—	AF049008[b] *Crenicara* sp.	AF370655[a] *Crenicara* sp.	AY566742	AF113090[c] *Crenicara* sp.	AF112628[c] *Crenicara* sp.
Crenicichla geayi	2	Río Las Marí as, Venezuela	AY566771	AF045848[i] *Creni*. sp.	AF370645[a] *Creni*. sp.	AY566736		
Crenicichla af. *lugubris*	2	Río Cinaruco, Venezuela	AY566785	AF049002[b] *C. lugubris*	AF370646[a] *C. regani*	AY566750	AF113087[c] *C. regani*	AF112625[c] *C. regani*
Crenicichla sveni	2	Apure, Venezuela	AY566779	AF285939[h] *C. lepidota*		AY566743	U70335[f] *C. saxatilis*	CSU63667[d] *C. saxatilis*
Crenicichla af. *wallacii*	2	Río Cinaruco, Venezuela	AY566790			AY566753		
Dicrossus sp.	2	Aquarium trade	AY566767			AY566731		
Geophagus brachybranchus	2	Río Cuyuní, Venezuela	AY566763			AY566727		
Geophagus grammepareius	1	Río Claro, Venezuela	AY566796	AF112642[g] *G. argyrostictus*		AY566724	AF113092[c] *G. argyrostictus*	AF112630[c] *G. argyrostictus*
Geophagus abalios	2	Río Cinaruco, Venezuela	AY566795	AF045850[i] *G. altifrons*		AY566757	AF113091[c] *G. altifrons*	AF112629[c] *G . altifrons*
Geophagus dicrozoster	2	Río Cinaruco, Venezuela	AY566794	AF049009[b] *G*. cf. *proximus*		AY566756		
Geophagus surinamensis	2	Haut Maroni, French Guiana	AY566777	AF112597[b] *Geophagus* sp.	AF370658[a] *Geophagus* sp.	AY566741	AF113093[c] *Geophagus*	AF112631[c] *Geophagus* sp.

(*continued on next page*)

Table 1 (*continued*)

	# Fish	Collection locality	Accession numbers					
			ND4	16S	Cyt *b*	RAG2	Tmo-4C4	Tmo-M27
'*Geophagus*' *brasiliensis*	2	Aquarium trade	AY566766	AF049016[b]	AF370659[a]	AY566732	AF113088	AF112626
'*Geophagus*' *steindachneri*	2	Aquarium trade, origin not known	AY566765		AF370660[a]	AY566730		
Guianacara n. sp. 'Caroní'	2	Río Claro, Venezuela	AY566762	AF049010[b] *Guianacara* sp.	AF370654[a] *Guianacara* sp.	AY566725	AF113084[c] *Guianacara* sp.	AF112622[c] *Guianacara* sp.
Gymnogeophagus balzanii	1	Aquarium trade, probably from Uruguay	—	AF112594[g] *G. gymnogenys*	AF370661[a] *G. gymnogenys*	AY566739	AF113085[c] *G. gymnogenys*	AF112623[c] *G. gymnogenys*
Gymnogeophagus rhabdotus	2	Aquarium trade, probably from Uruguay	AY566775	AF049011[b] *G. labiatus*	AF370662[a] *G. labiatus*	AY566738		
Mikrogeophagus altispinosus	2	Aquarium trade	AY566764	AF045857[b]		AY566729	AF113089[c]	AF112627[c]
Mikrogeophagus ramirezi	2	Caño Maporal, Venezuela	AY566780			AY566744		
Satanoperca daemon	2	Río Cinaruco, Venezuela	AY566791	AF049013[b] *S. acuticeps*	AF370663[a] *S. acuticeps*	AY566758		
Satanoperca jurupari	2	Río Orosa and R. Nanay, Perú	AY566783	AF049014[b]	AF370664[a]	AY566745		
Satanoperca mapiritensis	2	Río Pao and R. Morichal Largo, Venezuela	AY566761			AY566728		
Satanoperca pappaterra	2	Río Paraná, Brazil	AY566773			AY566759		
Taeniacara candidi	2	Aquarium trade	AY566769	AF112592[b]	AF370665[a]	AY566734	AF113094[c]	AF112632[c]

Taxa and accession number for sequences of 16S, Cyt *b*, *Tmo-M27*, and *Tmo-4C4* used to build the supermatrix are also given. Superscript numbers indicate the original publication of non-original sequences. References are as follows: [a]Farias et al. (2001); [b]Farias et al. (1999); [c]Farias et al. (2000); [d]Zardoya et al. (1996); [e]Kumazawa et al. (1999); [f]Streelman and Karl (1997); [g]Farias et al., unpublished; [h]Tang (2001); and [i]Farias et al. (1998).

which two are in need of description (e.g., Kullander, 1986; Kullander and Nijssen, 1989). Each of these undescribed genera was represented by one species in our analysis, and we have referred to them as '*Geophagus*' *brasiliensis* and '*Geophagus*' *steindachneri* to distinguish them from *Geophagus sensu stricto*. One species of each of the genera *Cichlasoma*, *Mesonauta*, and *Hoplarchus* were added to the ingroup to further test geophagine monophyly against the closely related Cichlasomatinae (Farias et al., 2000, 2001; Kullander, 1998). Based on previous knowledge of cichlid relationships (Farias et al., 1999, 2000, 2001; Kullander, 1998; Oliver, 1984; Stiassny, 1991), three species of *Cichla* and one of *Astronotus* and *Retroculus* were used as outgroups. Whenever possible, we sampled more than one species per genus to test genus-level monophyly, and to improve robustness and resolution of the analysis (e.g., Graybeal, 1998; Zwickl and Hillis, 2002). All references to ingroup and outgroup throughout the paper correspond to the above sets of taxa.

2.2. Molecular methods

The DNeasy kit (Qiagen) was used to extract total genomic DNA from muscle tissues stored in 95% ethanol. The mitochondrial ND4 gene (648 bp) was amplified using standard PCR protocols (94 °C denaturation, 60 s, 48 °C annealing 60 s, 77 °C extension 45 s for 35 cycles) and directly sequenced using primers Nap 2 (Arévalo et al., 1994) and ND4LB (Bielawsky et al., 2002) and the internal sequencing primer Geo-ND4F (5′ TCCTCC CCCTRATAATTCTKGC 3′), specifically designed for this study. The nuclear RAG 2 gene (~1000 bp) was amplified using a touchdown PCR protocol (94 °C, 30 s denaturation, 62, 60, and 58 °C, 2 cycles each; then 56 °C for 25 cycles, 60 s annealing, and 72 °C 90 s extension). PCR products were gel extracted using the Quiaquick kit (Qiagen) before sequencing. Amplification and external sequencing primers (F2 & R7) were from Lovejoy and Collette (2001) and the internal sequencing primers Geo IF (5′ AGGTCCTACATGCCTACATGC) and GeoIR (5′ GGGGCTGCCTTGCARAAGC) were developed specifically for this study. Forward and reverse DNA strands were sequenced with fluorescent-labeled dideoxynucleotide terminators (BigDye, PE Biosystems) following the protocol of Sanger et al. (1977) and using an automated ABI Prism 377 or 3100 automated sequencer (PE Biosystems).

2.3. Alignment

NCBI's BLAST search was used to confirm the identity of all new sequences. Since cichlid sequences of ND4 and RAG2 were not available before this study, sequences significantly matching the expected nucleotide regions in any teleosts were considered accurate and in-

cluded in the analyses. Forward and reverse sequences were edited and aligned in Sequencher 4.0 (Genecodes) and a consensus sequence was constructed for each specimen of each taxon. Unambiguous sequences of the ND4 fragment could not be obtained for the geophagine taxa *Crenicara punctulatum* and *Gymnogeophagus balzanii*, thus they were removed from the ND4 matrix and treated as missing data in all analyses (Table 1). Preliminary multiple alignments of all sequences were determined with Clustal X (Thompson et al., 1994), using default gap penalties. Several gap penalties were used to account for any indels involving codons. The same procedure was followed when combining our data with previously published sequences from mitochondrial genes (16S and cytochrome *b*), and two microsatellite flanking regions (*Tmo-M27* and *Tmo-4C4*). The 16S fragment was aligned using the secondary structural model of *Xenopus laevis* predicted by the Gutell Lab at the University of Texas at Austin (Cannone et al., 2002: http://www.rna.icmb.utexas.edu). Twenty-nine base pairs in regions of ambiguous structural alignment were excluded because positional homology could not be established. The alignments of microsatellite flanking region *Tmo-M27* and nuclear locus *Tmo-4C4* were checked for unnecessary gaps or obvious alignment mistakes, but otherwise were used unmodified. GenBank accession numbers for new sequences used in this study are given in Table 1; a Nexus file with the alignments for all loci is available from RLH on request.

2.4. Tests of saturation effects and rate heterogeneity

Potential substitutional saturation in ND4 and RAG2 was evaluated using saturation plots and chi square tests. We plotted uncorrected pairwise sequence divergence distances against maximum likelihood (ML) distances. In the absence of substitutional saturation, these plots should reveal a linear increment of uncorrected distance in relation to ML distance. If sequences are saturated (i.e., multiple substitutions occur at a given nucleotide position) the plots should plateau beyond a certain amount of uncorrected sequence divergence. Incorrect assumptions about the number of substitutions because of saturation can lead to incorrect phylogenetic inference (e.g., Swofford et al., 1996). We used saturation plots to separately evaluate transition and transversion substitutions at each codon position. ML model parameters were derived using nested likelihood ratio tests as implemented in ModelTest (Posada and Crandall, 1998). χ^2 tests were used to determine if frequency and type (i.e., transitions, transversions) of substitutions at a given position were significantly more likely to occur than at other positions (see Larson, 1994). The number of changes at each position was determined from either the most parsimonious (MP) tree (ND4) or the strict consensus of the MP trees (RAG2) derived from each data set (see below). Only parsimony-informative characters were included in the counts.

The branch length test (BLT) of Takezaki et al. (1995) was performed using the LINTRE program (1995, http://www.bio.psu.edu/People/Faculty/Nei/Lab). The BLT tests whether the branch length of a lineage is significantly longer or shorter than the average branch length across the tree (Takezaki et al., 1995), thus detecting whether a lineage evolves significantly faster or slower. Accuracy and power of the test increases when the tree is rooted with the nearest outgroup (Robinson et al., 1998), and when taxon sampling is increased (Robinson et al., 1998; Sorhannus and Van Bell, 1999). With this in mind, we used *Astronotus* as an outgroup following the results of Farias et al. (2000), and performed the test including all species in the phylogeny.

2.5. Phylogenetic analyses

Parsimony and Bayesian analyses were performed separately on the ND4 and RAG2 datasets; additionally, a simultaneous analysis (SA) was performed on both genes combined. Under parsimony, both equally and differentially weighted analyses were performed to assess whether possible saturation effects or lineage-specific rate heterogeneity affect phylogenetic resolution. We used two-parameter step matrices to differentially weight transitions and transversions for each of the aligned gene fragments. We estimated transition–transversion ratios using ML on a Neighbor-Joining tree constructed using the HKY85 model of nucleotide substitution (Hasegawa et al., 1985) in PAUP*. All parsimony phylogenetic analyses were performed in PAUP* v4.0b10 (Swofford, 2002) using 100 replicates of heuristic search with random addition sequence and tree bisection and reconnection (TBR) branch swapping. Support for the parsimony-derived topologies was estimated with non-parametric bootstrap (Felsenstein, 1985) and Bremer support indices (Bremer, 1988, 1994) with searches performed in PAUP*. Bootstrap included 100 pseudoreplicates and 10 heuristic search replicates under the same conditions of the original search, for both the equally and differentially weighted analyses. Bremer support was estimated for the equally weighted trees using topological constraints implemented in MacClade 4.0 (Maddison and Maddison, 2000) and 100 replicates of heuristic search using random addition sequences and TBR.

We used Bayesian phylogenetic analyses of each data set (ND4, RAG2, and SA) to determine whether the use of a model-based approach improved phylogenetic resolution by explicitly accounting for the effect of rate heterogeneity and sequence saturation. Analyses were run in MrBayes 3.0b4 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). This version of

MrBayes implements a modified Metropolis coupled Markov Chain Monte Carlo (MC^3) algorithm that independently samples trees and the parameters of the model of evolution from each data partition (Ronquist and Huelsenbeck, 2003). The program produces topologies of the combined data, but uses separate models of evolution for each partition. Initial models of molecular evolution were selected using nested Likelihood Ratio Tests as implemented in ModelTest (Posada and Crandall, 1998). Once the general model was obtained for each partition, specific parameters of nucleotide substitution were left as unknown using default priors, thus each model could accommodate several possible rate models (Huelsenbeck and Imennov, 2002). For each data set, phylogenetic analyses were run for 2×10^6 generations, sampling every 100 generations for a total of 20,000 samples per run (Leaché and Reeder, 2002). We plotted the log-likelihood values of each sample against the number of generations, and considered the Markov chain attained stationarity when log-likelihood estimates reached a stable value (Huelsenbeck and Ronquist, 2001; Leaché and Reeder, 2002). All samples with likelihood values below the stationarity level were discarded as burn-in. Three methods were applied to each data set to avoid estimating phylogenies corresponding to local optima. (1) The MC^3 algorithm implemented in MrBayes was employed. This approach facilitates an efficient search of tree space by using three incrementally heated chains along with the cold chain from which parameters are derived. Heated chains reduce the height of suboptimal peaks and fill valleys between peaks; by randomly swapping with the heated chains, the cold chain can more effectively explore tree space, reducing the chance of being trapped on local suboptimal peaks (Geyer and Thompson, 1995; Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). (2) For each data set, we repeated the MC^3 analyses with different starting trees until not less than four runs converged on the same mean stationarity value. Samples were then used from the four runs that converged on the highest likelihood. (3) For each of the four convergent analyses, we separately estimated a 50% majority rule consensus tree, and calculated mean values for the parameters of nucleotide substitution.

2.6. Internal branch tests

An internal branch test (IBT) was employed to determine whether short basal branches in the phylogeny represent credible relationships or a polytomy. IBTs are designed to establish the reliability of a tree by determining whether the length of its internal branches is significantly different from zero (Nei and Kumar, 2000). A bootstrap-based IBT, in which a distribution of internal branch lengths for the topology being tested is created, was used to determine whether branches were significantly positive (Dopazo, 1994; Sitnikova, 1996; Sitnikova et al., 1995). The IBT was performed in MEGA2 version 2.1 (Kumar et al., 2001) using each data set to build distance-based trees under neighbor-joining (NJ) and minimum evolution (ME). Trees were produced under the Kimura-2-parameter (1980) and the Tamura and Nei (1993) models of nucleotide substitution, both with and without gamma distributions to account for among-site rate heterogeneity (Felsenstein, 2004; Swofford et al., 1996).

3. Results

3.1. Sequence divergence and saturation effects

The aligned ND4 dataset included 648 bp. A single codon deletion was found in *Retroculus* sp. at position 126, and at position 130 in all species of *Cichla* as well as in *Biotodoma cupido*. Homogeneity of nucleotide composition was not rejected by the χ^2 test ($\chi^2 = 90.111$, $df = 105$, $p > 0.8$) as implemented in PAUP* 4.0b10 (Swofford, 2002). The HKY85 model of substitution revealed an overall transition–transversion ratio of 2.23, which we used for weighted parsimony analysis. The largest overall genetic distance (uncorrected sequence divergence = 41.52%) occurred among the geophagine taxa *Crenicichla wallacii* and *Dicrossus* sp., and the smallest divergence (4.78%) between *Geophagus brachybranchus* and *G. surinamensis*. Sequence divergence between geophagine cichlids and outgroup taxa ranged between 16.2% (*Guianacara* n. sp. 'Caroní' × *Cichla orinocense*) and 36.8% (*Crenicichla wallacii* × *Cichla intermedia*). Divergence between geophagine and cichlasomine taxa varied between 18.3% (*Cichlasoma orinocense* × *G. brachybranchus*) and 37.9% (*Mesonauta egregius* × *Crenicichla wallacii*). Saturation plots revealed a linear pattern of substitution for first and second positions and for third position transversions. However, beyond approximately 35% uncorrected divergence, a clear saturation effect was observed for third position transitions (Fig. 1). χ^2 Tests showed that third-position transitions occur significantly more frequently than any other type of substitution ($\chi^2 = 57.71$, $df = 9$, $p < 0.001$).

The RAG2 dataset included 976 bp, with no differences in length among the 38 taxa included. Homogeneity of base composition was not rejected by the chi square test ($\chi^2 = 7.498$, $df = 111$, $p = 1.0$), and the transition–transversion ratio was estimated at 2.41. The largest overall genetic distance (7.2%) occurred among the geophagine taxa *Apistogramma hoignei* and *B. cupido*, and the smallest between *G. brachybranchus* and *G. surinamensis*, which showed no difference in their sequences for the RAG2 fragment. Sequence divergence between geophagine cichlids and outgroup taxa ranged

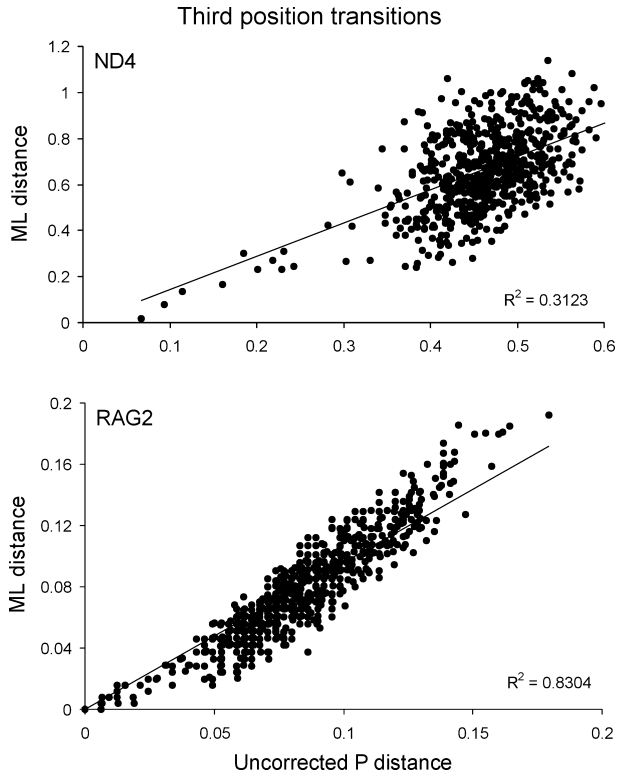

Fig. 1. Saturation plots for third position transversions in ND4 and RAG2. The graphs represent the increase in maximum likelihood distance correcting for multiple nucleotide substitutions versus the increase in uncorrected genetic distance between pairs of sequences; R^2 values show the fit of the relationship to a linear regression model. In the absence of saturation, both distances should increase linearly.

between 6.7% (*Acarichthys heckelii* × *Retroculus* sp.) and 2.6% (*Acarichthys heckelii* × *Cichla temensis*). Divergence between geophagine and cichlasomine taxa varied between 6.4% (*A. hoignei* × *Mesonauta egregius*) and 3.0% (*Acarichthys heckelii* × *Hoplarchus psittacus*). Saturation plots for the aligned 976 bp from 38 taxa revealed no apparent saturation of nucleotide substitution at any codon position, even at third position transitions (Fig. 1). These results were corroborated by the χ^2 test, which indicated that third position transitional substitutions do not occur with higher frequency than other types of substitutions ($\chi^2 = 3.56$, df = 3, $p = 0.313$).

The BLT indicated that rate heterogeneity is present in both ND4 and RAG2, but is more extensive in the former. In ND4, the genera *Biotodoma*, *Dicrossus*, *Crenicichla*, *Apistogramma*, and *Taeniacara* showed at least one species with significantly longer branches than average (Table 2), indicating an accelerated rate of molecular evolution. In contrast, *Guianacara*, *Acarichthys*, *Mikrogeophagus*, and *Geophagus sensu lato* showed significantly shorter branches than average. With the exception of *Retroculus* ($p > 0.05$), all non-geophagine taxa showed lower rates of evolution in the mitochondrial gene. RAG2 showed a less heterogeneous pattern

Table 2
Branch length test of rate heterogeneity (Takezaki et al., 1995)

Taxon name	BLT, *P* value, rate of evolution			
	df = 35 ND4		df = 37 RAG2	
Guianacara n.sp. 'Caroní'	<0.001	−	<0.05	−
Acarichthys heckelii	<0.01	−	<0.001	−
Biotodoma wavrini	NS		NS	
Biotodoma cupido	<0.001	+	<0.05	+
Mikrogeophagus altispinosus	NS		NS	
Mikrogeophagus ramirezi	<0.05	−	NS	
Biotoecus dicentrarchus	NS		NS	
Crenicara punctulatum	N/A		NS	
Dicrossus sp.	<0.01	+	NS	
Geophagus surinamensis	<0.001	−	NS	
Geophagus brachybranchus	<0.001	−	NS	
Geophagus abalios	<0.001	−	NS	
Geophagus dicrozoster	<0.001	−	<0.05	−
Geophagus grammepareius	<0.001	−	NS	
'*Geophagus*' *brasiliensis*	<0.001	−	NS	
'*Geophagus*' *steindachneri*	<0.001	−	NS	
Satanoperca pappaterra	NS		NS	
Satanoperca jurupari	NS		NS	
Satanoperca mapiritensis	NS		NS	
Satanoperca daemon	NS		NS	
Gymnogeophagus rhabdotus	NS		NS	
Gymnogeophagus balzanii	N/A		NS	
Crenicichla af. *wallacii*	<0.001	+	NS	
Crenicichla sveni	<0.001	+	NS	
Crenicichla af. *lugubris*	<0.001	+	NS	
Crenicichla geayi	<0.001	+	NS	
Apistogrammoides pucallpaensis	NS		<0.001	+
Apistogramma agassizi	<0.001	+	<0.001	+
Apistogramma hoignei	NS		<0.01	+
Taeniacara candidi	<0.001	+	NS	
Hoplarchus psittacus	<0.001	−	NS	
Mesonauta egregius	<0.001	−	NS	
Cichlasoma orinocense	<0.001	−	NS	
Retroculus sp.	NS		NS	
Cichla orinocensis	<0.001	−	<0.001	−
Cichla intermedia	<0.001	−	<0.05	−
Cichla temensis	<0.001	−	<0.001	−

P values are reported followed by a sign indicating rate increase (+) or rate decrease (−) in comparison with the average rate for all taxa. Model used for BLT was Tamura Nei + Γ as implemented in LINTRE.

of evolution, and only *Biotodoma* and *Apistogramma* (including *Apistogrammoides*) showed accelerated rates of molecular evolution, whereas *Guianacara*, *Acarichthys*, and one species of *Geophagus sensu stricto* showed significantly lower rates. Of the non-geophagine taxa, only *Cichla* had significantly shorter branches, whereas branches in other taxa were not significantly different from average.

3.2. Phylogenetic relationships

3.2.1. ND4

Parsimony analysis of ND4 included 395 informative sites for 36 taxa. Unweighted analysis produced 2 MP trees of 2770 steps with a global consistency index (CI)

of 0.30, retention index (RI) of 0.39, and rescaled consistency index (RC) of 0.12. Transition–transversion weighted parsimony resulted in a single MP tree of 3918.91 steps and CI = 0.32, RI = 0.41 and RC = 0.13. Both analyses produced essentially the same topology (Fig. 2A), with a weakly supported monophyletic Geophaginae (denoted by an asterisk in Fig. 2). Support for all basal nodes within the Geophaginae (i.e., intergeneric relationships) was weak. Additionally, only monophyly of the genera *Cichla*, *Crenicichla*, *Biotodoma*, and, to a lesser extent, *Geophagus sensu stricto* was strongly supported. The unweighted analysis rendered the genera *Satanoperca* and *Mikrogeophagus* paraphyletic, grouping *Satanoperca pappaterra* with *Biotoecus*, *Mikrogeophagus altispinosus* with *Crenicichla*, and *M. ramirezi* with a clade including the remainder *Satanoperca*, *Apistogramma* and *Taeniacara*. Weighted parsimony recovered *Mikrogeophagus* monophyly, but *Satanoperca pappaterra* remained weakly grouped with *Biotoecus*. The slight improved resolution observed in the weighted analysis suggested that sequence saturation and/or rate heterogeneity may affect the performance of parsimony analysis of ND4, and that a model-based analysis that corrects for these factors may further improve results. We used a general time reversible model of molecular evolution with invariants and gamma-approximated site-specific rate heterogeneity (GTR + I + Γ) for the Bayesian analysis. Four independent runs converged in the same likelihood range after approximately 25,000 generations, and we discarded the first 50,000 (500 trees) as burn in. Fifty percent majority rule topologies and model parameters for each run (19,500 trees/samples per run) were identical, thus all trees were combined into the final topology (Fig. 2B). The Bayesian analysis produced a monophyletic Geophaginae, albeit weakly supported and with visibly short basal branches. Once again, geophagine basal nodes were weakly supported. The position of the genera *Gymnogeophagus*, '*Geophagus*' *steindachneri*, '*Geophagus*' *brasiliensis*, and *Dicrossus* was largely inconsistent between the parsimony and Bayesian trees. *Mikrogeophagus* was monophyletic and strongly supported, but *Satanop-*

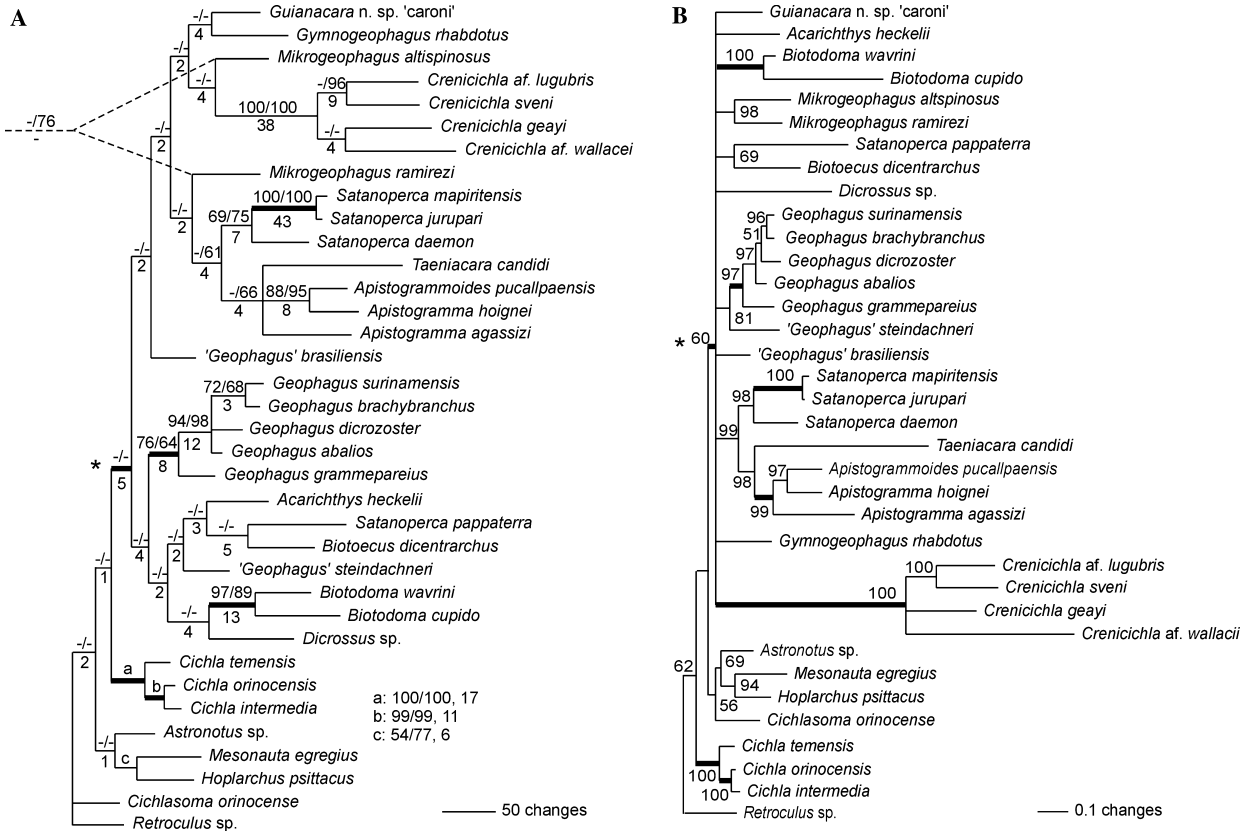

Fig. 2. (A) Consensus of most parsimonious topologies derived from equally weighted (2 MP trees) and transition–transversion weighted (1 MP tree) analysis of a 648 bp of the mitochondrial NADH dehydrogenase subunit 4 (ND4). Boostrap support, based on 100 pseudoreplicates, for unweighted/weighted analyses is given above branches (only scores >50% are shown); Bremer decay indices for the equally weighted analysis are given below branches. Support values for nodes a–c are given to the right of the tree. Dashed lines indicate how a transition–transversion weighted parsimony analysis recovers a monophyletic *Mikrogeophagus*. See text for tree statistics. (B) 50% majority rule Bayesian topology derived from ND4 sequences with a GTR + I + Γ general model of evolution. The topology resulted from combining 78,000 trees from four independent runs of 2×10^6 generations sampling every 100 trees with burn in of 50,000 generations/500 trees for each analysis. Posterior probabilities are given above branches or near nodes. Node marked with an asterisk denotes the basal node of the subfamily Geophaginae.

erca pappaterra was weakly grouped with *Biotoecus* as in the MP analyses. Interestingly, parsimony analysis after removal of third position transitions (not shown), does recover a monophyletic *Satanoperca*, suggesting that the models used for analysis (i.e., t_i/t_v weighting, GTR + I + Γ) do not completely correct the effect of sequence saturation.

3.2.2. RAG2

Parsimony analysis of RAG2 included 155 parsimony informative sites for 38 taxa. Unweighted analysis resulted in 152 MP trees of 483 steps and CI = 0.70, RI = 0.71 and RC = 0.50. The transition–transversion weighted analyses produced 36 MP trees of 681.99 steps and CI = 0.71, RI = 0.73 and RC = 0.52. Differential weighting did not produce a different topology from that obtained by equally weighted characters (Fig. 3A). Geophaginae were monophyletic and moderately supported, and all genera were monophyletic. MP analyses of RAG2 could not resolve the more basal relationships within the Geophaginae, but recovered three multi-genus clades: (1) the tribe Acarichthyini (*Acarichthys* and *Guianacara*, Kullander, 1998), (2) a clade formed by *Satanoperca*, *Apistogramma*, *and Taeniacara* (''*Satanoperca* clade,'' from here on), and (3) a large clade including the sister *Crenicara* and *Dicrossus* (crenicarine clade, from here on), *Biotodoma*, *Mikrogeophagus*, *Geophagus sensu lato*, *Gymnogeophagus*, referred to as the ''Big clade'' from here on. RAG2 Bayesian analyses were run using a GTR + I + Γ model of nucleotide substitution. Four analyses with convergent likelihood values were used for tree construction as described for ND4, but in this case, we discarded the first 1000 trees as burn in. As before, 50% majority rule consensus trees and parameters of sequence evolution for each of the four analyses were identical, thus topologies were combined into a single tree (Fig. 3B). The Bayesian topology was identical to the parsimony tree, but *Crenicichla* and *Biotoecus* were weakly placed at the base of the ''Big clade'' (Fig. 3B).

3.2.3. Simultaneous analysis

The combined datasets of ND4 and RAG2 produced a matrix of 1624 bp with 550 parsimony infor-

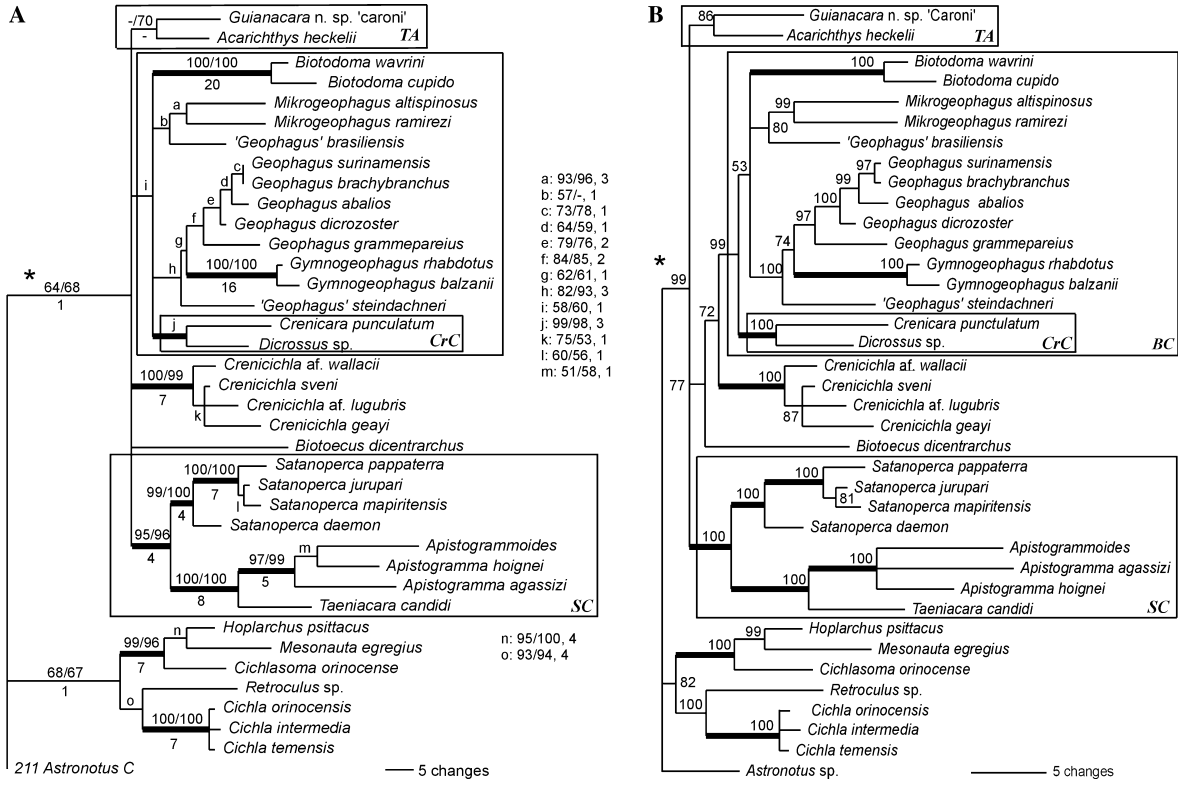

Fig. 3. (A) Consensus of most parsimonious topologies derived from equally weighted (152 MP trees) and transition–transversion weighted (36 MP trees) analysis of a 976 bp of the nuclear Recombination Activation Gene 2 (RAG2). Boostrap support, based on 100 pseudoreplicates, for unweighted/weighted analyses is given above branches (only scores >50% are shown); Bremer decay indices for the equally weighted analysis are given below branches. Support values for nodes a–o are given to the right of the tree. See text for tree statistics. (B) 50% majority rule Bayesian topology using RAG2 sequences with a GTR + I + Γ general model of evolution. The topology resulted from combining 76,000 trees from four independent runs of 2×10^6 generations sampling every 100 trees with burn in of 100,000 generations/1000 trees for each analysis. Posterior probabilities are given above branches or nodes. Node marked with an asterisk denotes the basal node of the subfamily Geophaginae. Highlighted clades: BC = Big clade; CrC = Crenicarine clade; SC = *Satanoperca* clade; TA = Tribe Acarichthyini.

mative positions for 38 taxa. Since ND4 sequences were not available for *C. punctulatum* and *G. balzanii* (see Section 2), these were treated as missing data. The unweighted parsimony analysis produced 1 MP tree of length 3304 with CI = 0.36, RI = 0.42 and RC = 0.15, and the transition–transversions weighted analysis resulted in 1 MP tree of 4667.94 steps with CI = 0.37, RI 0.44 and RC = 0.16 (Fig. 4A). The Geophaginae were monophyletic, but with short, unresolved basal branches. *Satanoperca pappaterra* was rendered paraphyletic, clearly because of the saturation effect from ND4; once again, removal of third position transitions from the mitochondrial data made the genus strongly monophyletic (not shown). Despite the exclusion of *S. pappaterra*, parsimony SA analysis recovered a mildly supported "*Satanoperca* clade," especially after differential weighting. Bayesian SA analyses were performed under unlinked GTR + I + Γ models of nucleotide substitution for each partition. The final topology (Fig. 4B) was derived exactly as described for ND4; parameters of sequence evolution for ND4 and RAG2 were not different from those estimated during the individual analyses. The Bayesian topology was similar to the RAG2 trees (Fig. 3), with short and marginally supported basal branches. All genera were strongly monophyletic, including *Satanoperca*, and *S. pappaterra* revealed a longer branch than other species in the genus. Clearly, the combination of the two genes under a model-based analysis correctly recovered this relationship, which was not resolved by parsimony analyses including ND4. Well-supported intergeneric relationships included the tribe Acarichthyini, the "*Satanoperca* clade," the crenicarine clade, a clade uniting *Geophagus sensu stricto* with '*Geophagus*' *steindachneri* and *Gymnogeophagus*, and a clade of '*Geophagus*' *brasiliensis* and *Mikrogeophagus*.

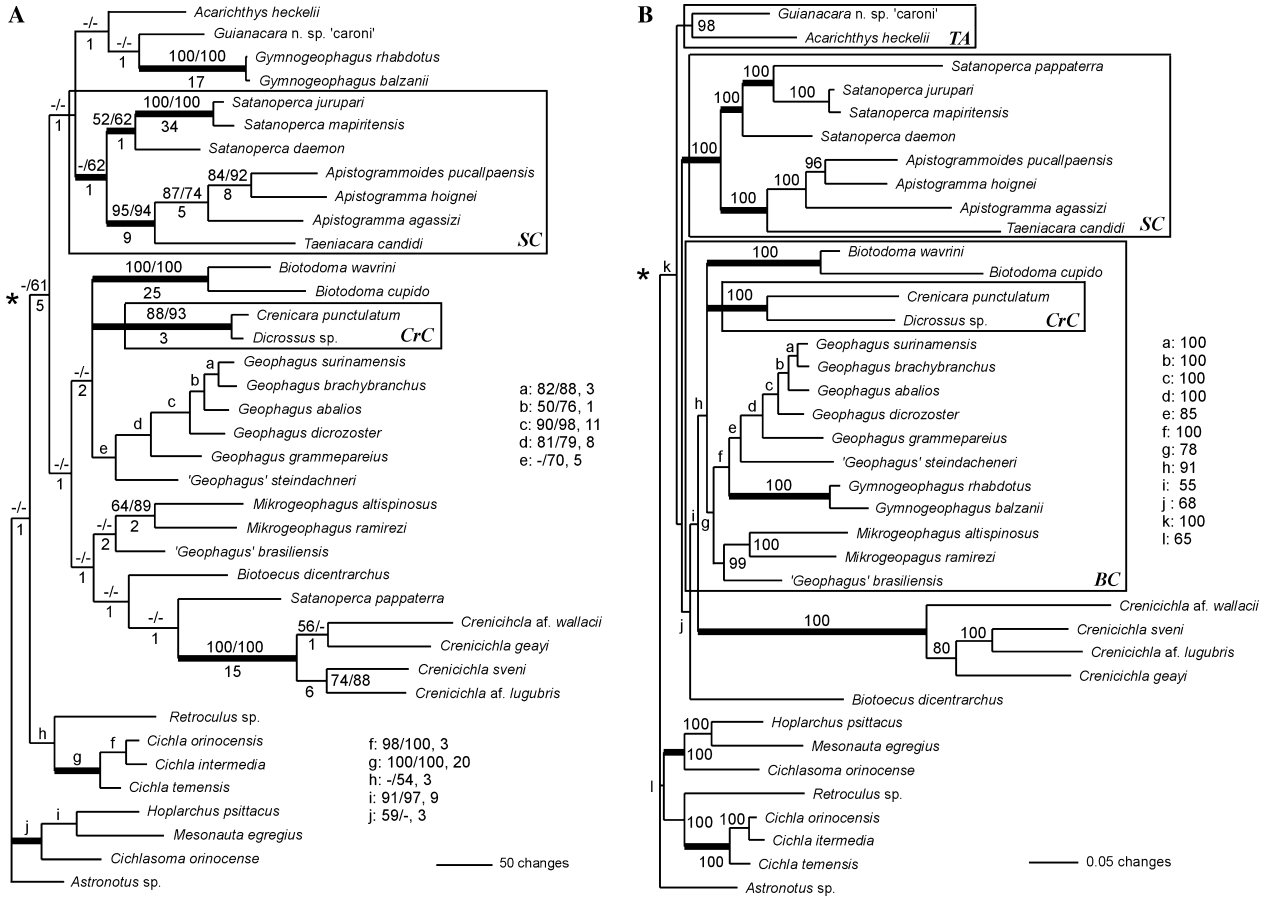

Fig. 4. (A) Consensus of most parsimonious topologies derived from equally weighted (1 MP tree) and transition–transversion weighted (1 MP tree) analysis of the combined 1624 bp of the mitochondrial ND4 and the nuclear RAG2. Boostrap support, based on 100 pseudoreplicates, for unweighted/weighted analyses is given above branches (only scores >50% are shown); Bremer decay indices for the equally weighted analysis are given below branches. Support values for nodes a–j are given to the right of the tree. See text for tree statistics. (B) 50% majority rule Bayesian topology using ND4 and RAG2 sequences with unlinked GTR + I + Γ general models of evolution. The topology resulted from combining 78,000 trees from four independent runs of 2×10^6 generations sampling every 100 trees with burn in of 50,000 generations/500 trees for each analysis. Posterior probabilities are given above branches or near nodes. Node marked with an asterisk denotes the basal node of the subfamily Geophaginae. Highlighted clades: BC = Big clade; CrC = Crenicarine clade; SC = *Satanoperca* clade; TA = Tribe Acarichthyini.

The "Big clade" (Fig. 4B, node h) was not as well supported by posterior probabilities as it was in the RAG2 analyses (Fig. 3).

3.3. Internal branch tests

Under all models of nucleotide substitution in all datasets, with and without a gamma distribution, only a few branches were significantly different from zero (Figs. 2 and 3). Most intergeneric relationships showed branches whose length confidence probabilities were much lower than 95% (Nei and Kumar, 2000; Kumar et al., 2001). The only exceptions were the branches at the base of the "*Satanoperca* clade" and the *Crenicara* + *Dicrossus* clade. These results reveal that basal branches are extremely short, and suggest a scarcity of characters associated with the early stages of geophagine diversification.

4. Discussion

4.1. Saturation, rate heterogeneity, and phylogenetic resolution

Saturation plots and χ^2 analysis revealed some saturation at third position transitions in the mitochondrial gene ND4, but no saturation was observed for the nuclear locus RAG2 (Fig. 1). Our analyses also revealed substantial lineage-specific rate heterogeneity, especially in ND4 (Table 2). Biases introduced by saturation and/or rate heterogeneity appear to have a significant effect on the resolution of geophagine phylogenetic relationships when mitochondrial data are used. Particularly, parsimony analysis of ND4 failed to recover the monophyletic genera *Mikrogeophagus* and *Satanoperca* (Fig. 2A). Differential weighting of transitions and transversions partially corrected the problem and identified a monophyletic *Mikrogeophagus*, but only total removal of third position transitions recovered a monophyletic *Satanoperca* under parsimony. Bayesian analysis of ND4 alone was partially effective, indicating that implementation of a model-based phylogenetic approach improves results, but the selected model of molecular evolution for ND4 (i.e., GTR + I + Γ) was not able to completely correct the bias in the sequence of *S. pappaterra*. Parsimony analysis of RAG2 recovered a strongly monophyletic *Satanoperca* (Fig. 3) and the monophyly of the genus is well established based on independent data (e.g., Kullander, 1986). Lack of resolution in ND4 is probably due to the evolutionary patterns observed for this particular locus, and the use of a specific model of molecular evolution compensates for the effect of saturation and rate heterogeneity, yielding improved phylogenetic results. This interpretation is further supported by the simultaneous analyses, which under parsimony failed to recover a monophyletic *Satanoperca* (Fig. 4A) but strongly supported monophyly of the genus when the model-based Bayesian analysis was used (Fig. 4B). It is well known that Bayesian posterior probabilities tend to inflate the actual support offered by the data to a particular node (Douady et al., 2003; Erixon et al., 2003; Simmons et al., 2004; Suzuki et al., 2002). However, due to the biases introduced by the use of ND4, we believe that the SA Bayesian tree represents the most reliable inference of geophagine relationships that can be obtained from our data. It is particularly interesting that most relationships with highest posterior probabilities coincide with those few for which branch lengths were significantly positive (Fig. 4B, e.g., "*Satanoperca* clade", crenicarine clade, *Biotodoma*, *Gymnogeophagus*), or that were repeatedly recovered in various analyses (e.g., Figs. 3 and 4B, "Big Clade"). Even considering the caveats associated with posterior probabilities, the consistent recovery of these relationships in various analyses suggests that they probably represent monophyletic clades. The Bayesian results are also interesting in that they show low support for most basal nodes, as do parsimony results and branch length tests. Given the explicit correction for saturation and rate heterogeneity effects, low posterior probabilities should not result from biases in molecular evolution of a locus, and an additional cause should be found for these weakly supported nodes.

4.2. Phylogenetic relationships of the Geophaginae

Phylogenetic relationships of geophagine cichlids found with our ND4 + RAG2 molecular dataset were significantly more likely than those proposed by Kullander (1998) based on morphology (Shimodaira–Hasegawa test, $p < 0.05$, $-\ln L$ Kullander tree = 21590.92, $-\ln L$ this study = 21566.08). Our results are also qualitatively different from Farias et al.'s (1999, 2000, 2001) topologies based on separate and total-evidence analyses of molecular and morphological characters. A quantitative comparison of our results to those of Farias et al.'s (1999, 2000, 2001) was not possible due to the reduced taxon sampling in their datasets. For instance, the absence of the genera *Dicrossus*, *Crenicara*, *Biotoecus*, *Satanoperca*, and '*Geophagus*' *steindachneri* from some or all of their analyses did not allow for meaningful direct comparison of their trees and ours. We discuss our results (summarized in Fig. 5A) in the context of previously proposed hypotheses of geophagine relationships, particularly the morphology-based phylogeny of Kullander (1998, Fig. 5B), and the total molecular evidence tree of Farias et al. (2000, Fig. 5C). Of several trees proposed by Farias et al. (1999, 2000, 2001), their total molecular evidence tree simultaneously offers the largest dataset (1371 bp from three genes) and the most extensive taxon

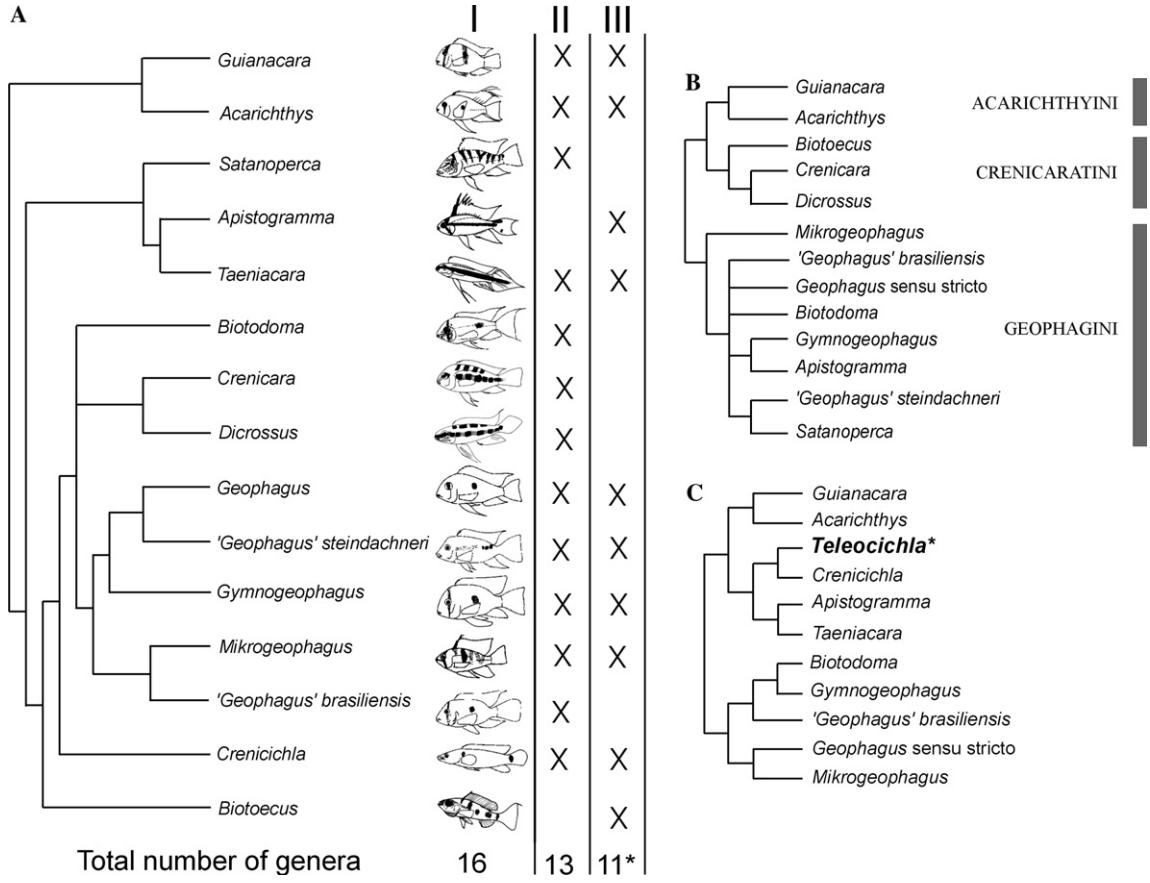

Fig. 5. Geophagine phylogenetic relationships from molecular and morphological data. (A) Bayesian tree from simultaneous analysis of ND4 and RAG2. (B) Kullander (1998) topology derived from analysis of 91 morphological characters showing the internal classification proposed based on that tree. (C) Farias et al. (2000) topology derived from the combined analysis of molecular data from three loci (1 mitochondrial, 2 nuclear). The table in the center compares the taxon sampling in each of three studies. Genera illustrated in column I were included in our study; shared genera between our study and those of Kullander and Farias et al. are marked with an "×" in columns II and III, respectively. The asterisk highlights the genus *Teleocichla*, which was included in Farias et al.'s (2000) study, but was absent from our dataset (see text). Tree comparisons clearly show that molecular evidence does not support Kullander's classification; taxon sampling in Farias et al.'s studies was not sufficient to test Kullander's hypothesis.

sampling of relevance for our study (11 geophagine genera). In other analyses, they included either more data for less taxa (e.g., Farias et al., 2001; 4 genes plus Kullander's morphology for 8 geophagine genera) or more taxa with less data (e.g., Farias et al., 2001; 1 gene for 13 geophagine genera), but these trees were less supported than the one we selected for comparison.

To further evaluate the contribution of our data to resolving geophagine relationships, we constructed a "supermatrix" (Gatesy et al., 2002) combining our dataset with published sequences of the mitochondrial genes 16S and cytochrome *b*, the microsatellite flanking region *Tmo-M27*, and the nuclear locus *Tmo-4C4* (Table 1). The obtained six-gene supermatrix allowed for a total-evidence analysis of all molecular data available for the Geophaginae, and provided additional data to test support for intergeneric relationships at the base of the tree. The large diversity (over 180 described species, Kullander (2003)) and extended geographic distribution of geophagine cichlids (southern Panama to northern Argentina) makes it largely impossible to obtain samples of exactly the same taxa used in other studies. As our objective was to resolve intergeneric relationships, we circumvented species-level taxonomic mismatches between our new data and published sequences by creating composites of species within each genus. Given that our SA tree provided strong support for genus-level monophyly within the Geophaginae, this strategy should increase the amount of data to resolve the weakly supported basal nodes of the phylogeny. Whenever possible, we based species combinations on previous knowledge of intra-generic phylogeny. For example, *Satanoperca daemon* was combined with *S. acuticeps* following Kullander and Ferreira (1988), and *Gymnogeophagus rhabdotus* with *G. labiatus* based on Wimberger et al. (1998). However, a lack of explicit phylogenetic information often forced us to base our decisions on rather arbitrary criteria that helped reduce

the amount of missing data. The result of combining the six gene data was a matrix of 38 OTUs with concatenated species within most genera and missing data whenever a locus was not available for a particular genus. When published sequences were lacking for a genus or locus, we coded that part of the matrix as missing data (?). The final structure of the supermatrix is illustrated in Table 1, which shows the species that were concatenated within each genus, and the accession numbers for all sequences employed. We analyzed the supermatrix under parsimony, but did not use Bayesian methods because, to our knowledge, the effect of missing data on this type of analysis has not been evaluated. On the other hand, parsimony is known for being robust to the effect of relatively large amounts of missing data (e.g., Wiens and Reeder, 1995). To partially reduce the effect of saturation and rate heterogeneity observed in ND4, we decided to remove third position transitions from the dataset. Although this strategy is not as effective as using an explicit model of sequence evolution, it is compatible with parsimony analysis. After removal of ND4 third position transitions, the six-gene dataset included 3748 nucleotides, of which 954 were parsimony informative. Phylogenetic analysis produced a single MP tree of 4389 steps with CI = 0.47, RI = 0.42, and RC = 0.20. The tree derived from the supermatrix (Fig. 6) is of interest when comparing the results from our newly derived data and previously published sequence analyses. Additionally, the supermatrix provides further evidence suggesting that weak support and lack of resolution at the base of the geophagine phylogeny is not due to lack of characters or the effect of biased substitution patterns in any particular locus. In the following paragraphs, we discuss the impact of our ND4 and RAG2 analyses on currently available hypotheses of geophagine relationships, and evaluate how these new data fit within published datasets.

Our results coincide with those of Farias et al. (2001) in including *Crenicichla* within Geophaginae (Fig. 5A and C), whereas Kullander (1998) found *Crenicichla* to group with *Cichla* at the base of the Neotropical Cichlidae (see also Stiassny, 1987, 1991). The widely corroborated sister-group relationship of *Crenicichla* and *Teleocichla* (e.g., Farias et al., 1999, 2000, 2001; Stiassny, 1987) and the established relationship of *Crenicara* and *Dicrossus* with the rare genus *Mazarunia* (Kullander, 1990) indicate that these genera also belong within the Geophaginae. Although monophyly of the subfamily was strongly supported, relationships among most geophagine genera remain difficult to resolve. Despite the large amount of data analyzed and nearly complete genus-level taxon sampling, support for suprageneric relationships was generally poor. Nonetheless, several clades were recovered and each will be discussed separately.

4.2.1. The tribe acarichthyini

A sister-group relationship between *Acarichthys* and *Guianacara* (representing the formally recognized Acarichthyini) was recovered by the RAG2 data and by the Bayesian SA, with increased posterior probabilities in the latter (Figs. 4 and 5A). Nevertheless, parsimony analysis of both the combined data (Fig. 4A) and the supermatrix (Fig. 6) failed to recover the Acarichthyini, even after either differential weighting or removal of saturated positions from ND4. Repeated recovery of the clade suggests that it is likely monophyletic, but it is also evident that support for the clade is reduced by the effect of saturation and/or rate heterogeneity, as the ND4 and uncorrected datasets were not able to recover the relationship. The tribe Acarichthyini was repeatedly recovered by morphological (Kullander, 1998) and some, but not all, molecular analyses (Farias et al., 1999, 2000, 2001). Nevertheless, support for the clade was always low with respect to other relationships (e.g., Farias et al. 1999, Fig. 3; Farias et al. 2001, Fig. 4). It is interesting that the addition of a significant amount of new data (i.e., ND4 and RAG2) does not provide stronger support for either the Acarichthyini or for its possible relationships to the remainder of the Geophaginae.

4.2.2. The ''Satanoperca clade''

This clade was consistently retrieved in all our analyses under parsimony and Bayesian approaches (Figs. 2–6). The ''*Satanoperca* clade,'' including *Satanoperca*, *Apistogramma*, *Apistogrammoides*, and *Taeniacara*, was not supported by previous morphological analyses (Kullander, 1998), in which *Satanoperca* grouped with '*Geophagus*' *steindachneri* and *Apistogramma* with *Gymnogeophagus* as part of Kullander's (1998) tribe Geophagini (Fig. 5B). Farias et al. (1999, 2000, 2001), using mitochondrial 16S and cytochrome *b* sequences, recovered the clade in independent analyses, but their total evidence analysis grouped *Apistogramma* with *Crenicichla*, probably because *Satanoperca* was not included in the dataset (Farias et al., 2000, 2001; Fig. 5C). In our analyses, the ''*Satanoperca* clade'' was generally well supported, and also was one of two suprageneric groupings with significant branch lengths (internal branch test, Figs. 2–4 and 6), supporting monophyly of this clade. Within the genus *Apistogramma*, most analyses grouped *A. hoignei* with the monotypic *Apistogrammoides* (Figs. 2B, 3 and 4), suggesting that this genus consists of an autapomorphic species of *Apistogramma* and does not warrant separate status.

4.2.3. The ''Big clade''

A clade including the genera *Biotodoma*, *Mikrogeophagus*, *Geophagus sensu lato*, *Gymnogeophagus*, *Crenicara*, and *Dicrossus* (Figs. 3, 4B, and 6) was moderately supported and consistently recovered by RAG2, the Bayesian SA, and the supermatrix. Despite its limited

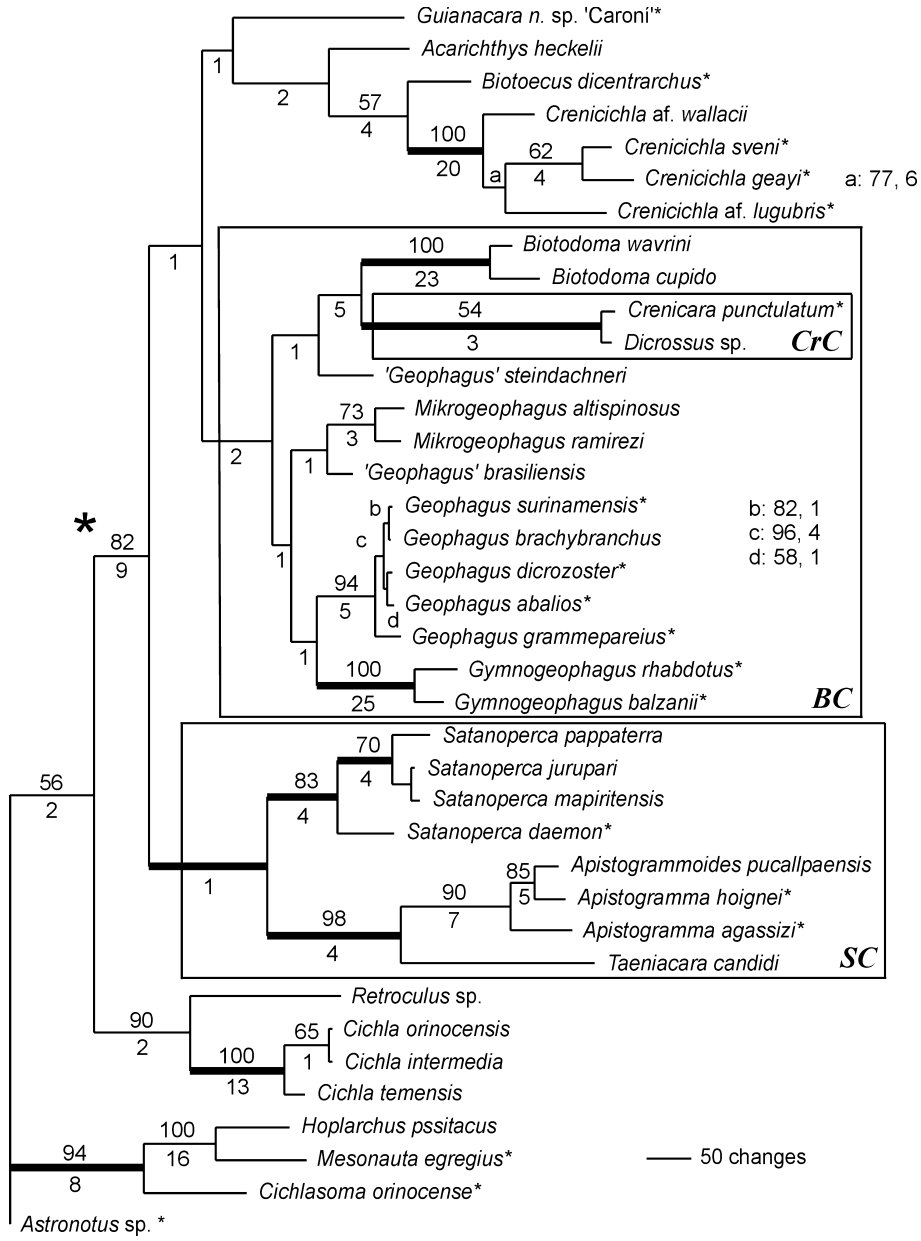

Fig. 6. Single MP tree derived from analysis of 3748 nucleotides of the combined RAG2, ND4 (after removal of saturated third position transitions), cytochrome *b*, 16S, *Tmo-M27*, and *Tmo-4C4*. Boostrap support, based on 100 pseudoreplicates, is given above branches (only scores >50% are shown); Bremer decay indices are given below branches. Support values for nodes a–d are given to the right of the tree. Names of taxa marked with an asterisk represent concatenated sequences from different species within the same genus, see Table 1 for species used in concatenation and for GenBank accession numbers. Branches highlighted in bold correspond to those with lengths significantly higher than zero in the internal branch tests. Node marked with an asterisk denotes the basal node of the subfamily Geophaginae. The "*Satanoperca* clade" (SC), the "Big clade" (BC), and the crenicarine clade (CrC) are highlighted within the box.

statistical support, the "Big clade" is of interest when compared to previous findings. With the exception of *Satanoperca* and *Apistogramma*, the "Big clade" includes all the genera formerly placed in Kullander's (1998) tribe Geophagini, plus *Crenicara* and *Dicrossus*, which Kullander placed in the tribe Crenicaratini along with *Biotoecus* (Fig. 5B). Farias et al. (2000) did not examine *Crenicara* and *Biotoecus* in their molecular analysis, and *Dicrossus* was absent from all of their datasets. Although their results are compatible with the "Big clade" (Fig. 5C), their data did not provide insight on whether the relationships of Kullander's Crenicaratini were supported by molecular evidence. By inclusion of the relevant taxa in our analysis, all available molecular evidence indicates that Kullander's tribes Geophagini and Crenicaratini are paraphyletic (Figs. 5A, B and 6). Interestingly, our results suggest that *Crenicara*, *Dicrossus*, and *Biotoecus*, all small-bodied taxa known as

"dwarf cichlids" in the aquarium trade, probably grouped together in Kullander's morphological analysis due to shared homoplastic morphological characters associated with body-size reduction (see Buckup, 1993; López-Fernández, 2004).

4.2.4. The crenicarine clade, Crenicichla and Biotoecus

In his morphological analysis, Kullander (1998) proposed the tribe Crenicaratini to include the genera *Crenicara*, *Dicrossus*, and *Biotoecus*. Our results consistently indicated that *Crenicara* and *Dicrossus* do form a monophyletic clade (Figs. 3–6), but no evidence suggested that *Biotoecus* is part of the same group. Furthermore, it seems clear that *Crenicara* and *Dicrossus* are nested within the "Big clade," and that *Biotoecus* is perhaps more closely related to *Crenicichla* than to other geophagines, as was weakly suggested by the supermatrix analysis (Fig. 6). Bayesian analyses tended to place *Crenicichla* and *Biotoecus* at the base of the "Big clade" (Figs. 3B and 4B), but with marginal support; parsimony analyses were not consistent across datasets. *Biotoecus* was weakly related to *Mikrogeophagus* in the only dataset of Farias et al. (2000) that included the genus; their analysis grouped the *Biotoecus-Mikrogeophagus* clade with the *Crenicichla-Teleocichla* clade, but with no clear support. The phylogenetic position of *Crenicichla* remains uncertain. For instance, morphological analyses tend to group *Crenicichla* with the basal genus *Cichla* (Kullander, 1998; Stiassny, 1987, 1991; subfamily Cichlinae), but all molecular evidence groups it with the Geophaginae (Farias et al., 1999, 2000, 2001; this study). Unfortunately, *Crenicichla's* relationships within geophagines are not clear, although our analysis of the supermatrix weakly supported a previously unidentified clade grouping *Crenicichla* with *Biotoecus* (Fig. 6). In previous studies, *Crenicichla* grouped with *Gymnogeophagus* (Farias et al., 1999) or with *Apistogramma* (e.g., Farias et al., 2000, 2001; total evidence), but support for these relationships was never high. No evidence in our data supports an association to *Gymnogeophagus*, but the topology from the supermatrix does not allow discarding a relationship with *Acarichthys* and *Guianacara*. The relative support for the "*Satanoperca* clade" found in our study suggests that *Apistogramma* is probably related to *Satanoperca*, and not to *Crenicichla*.

Our study confirms previous findings that include *Crenicichla* within the Geophaginae. Increased amounts of data and improved taxon sampling with respect to earlier studies suggest some previously unreported relationships, mainly the "*Satanoperca* clade," grouping *Satanoperca*, *Apistogramma*, and *Taeniacara*, and the possible relationship of *Crenicichla* and *Biotoecus*. Some support also was found for a "Big clade" that includes the genera *Biotodoma*, *Crenicara*, *Dicrossus*, *Geophagus sensu lato*, and *Gymnogeophagus*. Because of its expanded taxon sampling, this study allowed the first detailed comparison of molecular and morphological phylogenies of geophagine cichlids. We found that molecular evidence is incongruent with published morphological data, suggesting that fundamental changes in geophagine classification are necessary. The most remarkable result, however, is the generalized lack of well-supported relationships within the geophagine phylogeny. Support for basal branches, as measured by Bremer support, bootstrap, and Bayesian posterior probabilities is lower than support for branches near the tips, and especially lower than support for clades outside the subfamily. The internal branch tests corroborate the weakness of the basal relationships, and indicate that basal branches are extremely short. A tree with short basal branches may result from inadequate data, but also can result from an episode of rapid lineage diversification that leaves little opportunity for character fixation.

4.3. Implications of the molecular phylogeny: are geophagine cichlids an adaptive radiation?

According to Schluter (2000), adaptive radiations are characterized by four features: common ancestry, rapid speciation, phenotype-environment correlation, and trait utility. Based on this definition, our results fulfill the phylogenetic requisites to regard geophagines as an adaptive radiation of Neotropical cichlids. The subfamily Geophaginae is clearly a monophyletic clade, both according to the results presented in this paper and to previous research (Farias et al., 1999, 2000, 2001). More revealing is the fact that the phylogeny recovered in this study is characterized by short basal branches (Figs. 2–4 and 6). Relationships at the base of the geophagine clade were poorly supported by both bootstrap and Bremer values, and internal branch tests confirmed that most basal branches are not significantly different from zero. Furthermore, lack of resolution and support seems restricted to the base of the geophagine clade, as relationships of outgroup taxa and cichlasomatine lineages are much better supported by our dataset (Figs. 2–4 and 6). Within geophagines, support for monophyly of genera (i.e., recent divergence) is much stronger than observed for basal, intergeneric nodes (i.e., older divergence, Figs. 2–4 and 6). Poor resolution and weak support at the base of a phylogeny can be due to either lack of information necessary to resolve basal relationships or fast differentiation of lineages at the base of the tree (e.g., Hodges, 1997; Jackman et al., 1999; Kontula et al., 2003; Poe and Chubb, 2004). Our datasets included a large amount of phylogenetically informative characters, suggesting that lack of basal resolution within the geophagine clade is not due to scarcity of data. In fact, the ND4 + RAG2 dataset included 550 informative characters, providing nearly fifteen times as many characters than taxa (38), and the supermatrix dataset

included 954 informative sites and was roughly 25 times larger than the number of taxa. Low resolution is, likewise, unlikely to be caused by sequence substitutional saturation or rate heterogeneity, as the analysis in which these biases were corrected did not show better support or longer branches within the Geophaginae (Figs. 2B, 3B, and 4B). Addition of previously published sequences from four markers into a total molecular evidence analysis improved resolution at the base of the tree, but did not increase support for short branches, and only support for geophagine monophyly improved significantly by addition of the supplemental data (e.g., Figs. 4A and 6). Finally, the dataset included both mitochondrial and nuclear loci, thus encompassing a variety of evolutionary rates (see Table 2) that should help resolve relationships at different depths of the tree. Considering all these elements together, it is improbable that lack of resolution and support at the base of the geophagine tree is due to either inadequate or insufficient data. Instead, short branches at the base of the tree suggest that the different geophagine genera may have originated rapidly and/or over a short period, with subsequent reduced fixation of character states associated with the radiation. Whether differentiation of geophagine genera occurred simultaneously (e.g., Hoelzer and Melnick, 1994) or through a series of rapid sequential events (e.g., Jackman et al., 1999; Poe and Chubb, 2004) cannot be resolved with any certainty. In the first case, the geophagine tree would represent a "hard polytomy" (e.g., Maddison, 1989), and because this represents the null hypothesis of phylogenetic reconstruction, an effective test is difficult to apply (Walsh et al., 1999). If geophagine differentiation occurred sequentially and rapidly, an inordinate amount of data will be required to resolve this "soft polytomy" (Hoelzer and Melnick, 1994). In either case, the results suggest a fast radiation of geophagine cichlid genera over a relatively short period of time. Geophagine monophyly and fast diversification of lineages thus fulfill the phylogenetic requirements of a geophagine cichlid adaptive radiation *sensu* Schluter (2000).

Further study of geophagine cichlids should reveal whether the clade fulfills the remaining requisites of phenotype-environment correlation and trait utility. A compelling observation is the coherence between molecular and morphological patterns of variation within and among genera. Patterns of molecular evolution in the group are characterized by strong lineage-specific rate heterogeneity, suggesting a different molecular evolutionary trajectory for each genus after its differentiation. Likewise, overall morphological differentiation among genera is quite remarkable, suggesting that each genus followed a unique pattern of morphological diversification (see Winemiller et al., 1995; HLF, unpublished). Interestingly, species within a genus tend to be morphologically indistinct and may form relatively large complexes of highly similar taxa (see López-Fernández, 2004; López-Fernández and Taphorn, 2004), indicating that most morphological differentiation occurred during the origin of the clade when genera differentiated from each other. Morphological variation in geophagines is clearly associated with the diverse ecological and behavioral repertoire of the group (e.g., Winemiller et al., 1995). Geophagine reproductive strategies range from large-bodied, mouth-brooding, monogamous taxa with relatively large generation times (e.g., *Geophagus*) to small-bodied, polygynous substrate spawners that reach sexual maturity much earlier (e.g., *Apistogramma*). The importance of ecological specialization in the evolution of the group also is evident in the association between form and function in relation to trophic biology and habitat use. Ecomorphological patterns range from deep bodied fishes with ventrally oriented mouths that ingest and sift sandy substrates for invertebrates (e.g., *Geophagus*, *Satanoperca*), to elongate piscivores with terminal mouths (*Crenicichla*) and small species that inhabit highly structured habitats and feed mostly on epibenthic invertebrates (e.g., *Taeniacara*, *Biotoecus*, *Apistogramma*). Ecomorphology may determine both lineage-specific evolutionary trajectories and patterns of community assembly. Although inconclusive, these patterns strongly suggest a correlation between phenotype and environment, which is Schluter's third requirement for an adaptive radiation. At present, evidence for trait utility is almost completely lacking. Explicit analyses of functional morphology and ecological performance of geophagines are needed. In conclusion, geophagine cichlids provide an outstanding group for the investigation of adaptive radiation of fishes in fluviatile habitats.

Acknowledgments

We are indebted to the following people for providing tissue samples of different taxa: Nathan Lovejoy and Stuart Willis (University of Manitoba, Canada), Izeni Farias (Universidade do Amazonas, Brazil), Eloisa Revaldaves (Universidade Estadual de Maringá, Brazil), and Yves Fermon (Museum National d'Histoire Naturelle, France). April Harlin, Jim Woolley, Anthony Cognato, and John McEachran read earlier versions of this manuscript and offered comments that significantly improved it. Field work in Venezuela was largely possible due to support from Donald Taphorn (Museo de Ciencias Naturales de Guanare, UNELLEZ, Venezuela) and Carol Marzuola and Edgar Peláez from the Cinaruco Fishing Club. Field collections in Venezuela were partially funded by grants from the National Geographic Society to K.O.W. Fishing permits to H.L.F. and K.O.W. were provided by the Servicio Autónomo de los Recursos Pesqueros y Acuí colas from the Venezuelan Ministerio de Agricultura y Crí a. We thank Stuart Willis, Albrey

Arrington, Craig Layman, Tom Turner, Horacio Julio, José Vicente Montoya, Carmen Montaña, Carol Marzuola and Jeremy Walther for help in the field. Laboratory work was partially funded by a Jordan Endowment Fund from the American Cichlid Association to H.L.F. and grants from the National Science Foundation grants DEB 9615163 to R.L.H. and DEB 0089834 to K.O.W. H.L.F. thanks April Harlin and Diane Rowe for their help in the lab. Joe Gillespie helped align the 16S ribosomal sequences. Comments from an anonymous reviewer greatly improved the original manuscript.

References

Arévalo, E., Davis, S., Sites, J., 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of *Sceloporus grammicus* complex (Phrynosomatidae) in Central Mexico. Syst. Biol. 43, 387–418.

Arrington, D.A., Winemiller, K.O., 2003. Diel changeover in sandbank fish assemblages in a Neotropical floodplain river. J. Fish Biol. 63, 442–459.

Barlow, G.W., 2000. The Cichlid Fishes: Nature's Grand Experiment in Evolution. Perseus Publishing, Cambridge.

Bielawsky, J., Brault, A., Gold, J., 2002. Phylogenetic relationships within the genus *Pimephales* as inferred from ND4 and ND4L nucleotide sequences. J. Fish Biol. 61, 293–297.

Bremer, K., 1988. The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42, 795–803.

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.

Buckup, P., 1993. Phylogenetic interrelationships and reductive evolution in Neotropical characidiin fishes (Characiformes, Ostariophysi). Cladistics 9, 305–341.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D''Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Müller, K.M., Pande, N., Shang, Z., Yu, N., Guttell, R.R., 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics 3 (1), 2.

Dopazo, J., 1994. Estimating errors and confidence intervals for branch lengths in phylogenetic trees by a bootstrap approach. J. Mol. Evol. 38, 300–304.

Douady, C., Delsuc, F., Boucher, Y., Doolittle, W., Douzery, E., 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Mol. Biol. Evol. 20, 248–254.

Erixon, P., Svennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52, 665–673.

Farias, I.P., Schneider, H., Sampaio, I., 1998. Molecular phylogeny of Neotropical cichlids: the relationships of Cichlasomines and Heroines. In: Malabraba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M., Lucena, C.A.S. (Eds.), Phylogeny and Classification of Neotropical Fishes. Editora Universitaria, Pontificia Universidad Catolica do Rio Grande do Sul, Porto Alegre, pp. 499–508.

Farias, I.P., 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, I.P., Ortí, G., Meyer, A., 2000. Total evidence: molecules, morphology, and the phylogenetics of cichlid fishes. J. Exp. Zool. 288, 76–92.

Farias, I.P., Ortí, G., Sampaio, I., Schneider, H., Meyer, A., 2001. The cyrochrome *b* gene as a phylogenetic marker: the limits of resolution for analyzing relationships among cichlid fishes. J. Mol. Evol. 53, 89–103.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., 2004. Inferring Phylogenies. Sinauer Associates, Inc., Sunderland.

Gatesy, J., Matthee, C., DeSalle, R., Hayashi, C., 2002. Resolution of a supertree/supermatrix paradox. Syst. Biol. 51, 652–664.

Geyer, C., Thompson, E., 1995. Annealing Markov Chain Monte Carlo with applications to ancestral inference. J. Am. Stat. Assoc. 90, 909–920.

Graybeal, A., 1998. Is it better to add taxa or characters to a difficult phylogenetic problem?. Syst. Biol. 47, 9–17.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.

Hodges, S., 1997. Rapid radiation due to a key innovation in columbines (Ranunculaceae: *Aquilegia*). In: Givnish, T.J., Sytsma, K.J. (Eds.), Molecular Evolution and Adaptive Radiation. Cambridge University Press, Cambridge, pp. 391–406.

Hoelzer, G.A., Melnick, D.J., 1994. Patterns of speciation and limits to phylogenetic resolution. Trends Ecol. Evol. 9, 104–107.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Huelsenbeck, J.P., Imennov, N., 2002. Geographic origin of human mitochondrial DNA: accommodating phylogenetic uncertainty and model comparison. Syst. Biol. 51, 155–165.

Jackman, T.R., Larson, A., de Queiroz, K., Losos, J.B., 1999. Phylogenetic relationships and tempo of early diversification in *Anolis* lizards. Syst. Biol. 48, 254–285.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Kontula, T., Kirilchik, S., Väinöla, R., 2003. Endemic diversification of the monophyletic cottoid fish species flock in Lake Baikal explored with mtDNA sequencing. Mol. Phylogenet. Evol. 27, 143–155.

Kullander, S.O., 1986. Cichlid Fishes of the Amazon River Drainage of Peru. Swedish Museum of Natural History, Stockholm.

Kullander, S.O., 1990. *Mazarunia mazarunii* (Teleostei: Cichlidae), a new genus and species from Guyana, South America. Ichthyol. Explor. Freshwaters 1, 3–14.

Kullander, S.O., 1998. A phylogeny and classification of the Neotropical Cichlidae (Teleostei: Perciformes). In: Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M., Lucena, C.A.S. (Eds.), Phylogeny and Classification of Neotropical Fishes. Editora Universitaria, Pontificia Universidad Catolica do Rio Grande do Sul, Porto Alergre, pp. 461–498.

Kullander, S.O., 2003. Family Cichlidae (Cichlids). In: RE, R., Kullander, S.O., Ferraris, C.J. (Eds.), Check List of the Freshwater Fishes of South and Central America. Museu de Ciências e Tecnologia, Pontifí cia Universidade Católica do Rio Grande do Sul, Porto Alegre, pp. 605–656.

Kullander, S.O., Ferreira, E.J.G.., 1988. A new *Satanoperca* species (Teleostei, Cichlidae) from the Amazon River basin in Brazil. Cybium 12, 343–355.

Kullander, S.O., Nijssen, H., 1989. The Cichlids of Surinam. E.J. Brill, Leiden.

Kumar, S., Tamura, K., Jakobsen, I., Nei, M., 2001. MEGA 2: Molecular Evolutionary Genetics Analysis Software. Arizona State University, Tempe, Arizona.

Kumazawa, Y., Yamaguchi, M., Nishida, M., 1999. Mitochondrial molecular clocks and the origin of euteleostean biodiversity: familial radiation of perciforms may have predated the Cretaceous/Terciary boundary. In: Kato, M. (Ed.), The Biology of Biodiversity. Springer, Tokyo, pp. 35–52.

Larson, A., 1994. The comparison of morphological and molecular data in phylogenetic systematics. In: Schierwater, B., Streit, B.,

Wagner, A., DeSalle, R. (Eds.), 1996, Molecular Ecology and Evolution: Approaches and Applications. Birkhä user Verlag, Basel, Switzerland, pp. 371–390.

Leaché, A., Reeder, T., 2002. Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): a comparison of Parsimony, Likelihood and Bayesian approaches. Syst. Biol. 51, 44–68.

López-Fernández, H., 2004. Phylogeny of geophagine cichlids from South America (Perciformes: Labroidei). Ph.D. Dissertation, Texas A&M University, Texas.

López-Fernández, H., Taphorn, D.C., 2004. *Geophagus abalios*, *G. dicrozoster* and *G. winemilleri* (Perciformes: Cichlidae), three new species from Venezuela. Zootaxa 439, 1–27.

Lovejoy, N.R., Collette, B.B., 2001. Phylogenetic relationships of New World needlefishes (Teleostei: Belonidae) and the biogeography of transitions between marine and freshwater habitats. Copeia 2001, 324–338.

Lowe-McConnell, R.H., 1991. Ecology of cichlids in South American and African waters, excluding the African Great Lakes. In: Keenleyside, M.H.A. (Ed.), Cichlid Fishes: Behavior, Ecology and Evolution. Chapman Hall, London, pp. 62–85.

Maddison, W.P., 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5, 365–377.

Maddison, W.P., Maddison, D.R., 2000. MacClade: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Inc., Sunderland.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford Uniersity Press, New York.

Oliver, M., 1984. Systematics of African cichlid fishes: determination of the most primitive taxon, and studies on the haplocromines of lake Malawi. Ph.D. Dissertation, Yale University.

Poe, S., Chubb, A.L.., 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. Evolution 58, 404–415.

Posada, D., Crandall, K., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Robinson, M., Gouy, M., Gautier, C., Mouchiroud, D., 1998. Sensitivity of the relative rate test to taxonomic sampling. Mol. Biol. Evol. 15, 1091–1098.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sanger, F., Nicklen, S., Coulson, A., 1977. DNA sequencing with chain-termination inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Schluter, D., 2000. The Ecology of Adaptive Radiation. Oxford University Press, New York.

Simmons, M., Pickett, K., Miya, M., 2004. How meaningful are Bayesian support values. Mol. Biol. Evol. 21, 188–199.

Sitnikova, T., 1996. Bootstrap method of Interior-Branch test for phylogenetic trees. Mol. Biol. Evol. 13, 605–611.

Sitnikova, T., Rzhetsky, A., Nei, M., 1995. Interior-branch and bootstrap tests of phylogenetic trees. Mol. Biol. Evol. 12, 319–333.

Sorhannus, U., Van Bell, C., 1999. Testing for equality of molecular evolutionary rates: a comparison between a relative-rate test and a likelihood ratio test. Mol. Biol. Evol. 16, 849–855.

Stiassny, M.L.J., 1987. Cichlid familial intrarelationships and the placement of the Neotropical genus *Cichla* (Perciformes, Labroidei). J. Nat. Hist. 21, 1311–1331.

Stiassny, M.L.J., 1991. Phylogenetic intrarelationships of the family Cichlidae: an overview. In: Keenleyside, M.H. (Ed.), Cichlid Fishes: Behaviour, Ecology and Evolution. Chapman Hall, London, pp. 1–35.

Streelman, J., Karl, S., 1997. Reconstructing labroid evolution with single-copy nuclear DNA. Proc. R. Soc. Lond. B 264, 1011–1020.

Suzuki, Y., Glazko, G., Nei, M., 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA 99, 16138–16143.

Swofford, D.L., 2002. PAUP* Phylogenetic Analysis Using Parsimony (* and other Methods). Sinauer Associates, Inc., Sunderland.

Swofford, D.L., Olsen, G., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer Associates Inc., Sunderland, pp. 407–514.

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

Tamura, K., Nei, M., 1993. Estimaton of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

Tang, K.L., 2001. Phylogenetic relationships among damselfishes (Teleostei: Pomacentridae) as determined by mitochondrial DNA data. Copeia 2001, 591–601.

Thompson, J., Higgins, D., Gibson, T., 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.

Walsh, H.E., Kidd, M.G., Moum, T., Friesen, V.L.., 1999. Polytomies and the power of phylogenetic inference. Evolution 53, 932–937.

Weidner, T., 2000. South American Eartheaters. Cichlid Press, El Paso.

Wiens, J., Reeder, T., 1995. Combining data sets with different numbers of taxa for phylogenetic analysis. Syst. Biol. 44, 548–558.

Wimberger, P.H., Reis, R.E., Thornton, K.R., 1998. Mitochondrial phylogenetics, biogeography, and evolution of parental care and mating systems in *Gymnogeophagus* (Perciformes: Cichlidae). In: Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M., Lucena, C.A.S. (Eds.), Phylogeny and Classification of Neotropical Fishes. Editora Universitaria, Pontificia Universidad Catolica do Rio Grande do Sul, Porto Alergre, pp. 509–518.

Winemiller, K.O., Kelso-Winemiller, L.C., Brenkert, A.L., 1995. Ecomorphological diversification and convergence in fluvial cichlid fishes. Environ. Biol. Fishes. 44, 235–261.

Zardoya, R., Vollmer, D.M., Craddock, C., Strelman, J.T., Karl, S., Meyer, A., 1996. Evolutionary conservation of microsatellite flanking regions and their use in the phylogeny of cichlid fishes (Pisces: Perciformes). Proc. R. Soc. Lond. B 263, 1589–1598.

Zwickl, D., Hillis, D.M., 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. 51, 588–598.