



Patterns of diversification in the discus fishes (*Symphysodon* spp. Cichlidae) of the Amazon basin

Izeni Pires Farias^a, Tomas Hrbek^{a,b,*}

^aUniversidade Federal do Amazonas (UFAM), Departamento de Biologia, Laboratório de Evolução e Genética Animal (LEGAL), Av. Rodrigo Octávio Jordão Ramos, 3000, 69077-000 Manaus, AM, Brazil

^bDepartment of Biology, University of Puerto Rico, Rio Piedras, San Juan, PR, Puerto Rico

ARTICLE INFO

Article history:

Received 21 September 2007

Revised 17 May 2008

Accepted 20 May 2008

Available online 3 June 2008

Keywords:

Discus fish

Symphysodon spp.

Phylogeography

Population genetics

Amazonian diversification

mtDNA control region

nuDNA RAG1

ABSTRACT

We carried out a phylogeographic and population genetic analysis of fishes of the taxonomically contentious genus *Symphysodon* from the Amazon basin in order to test hypotheses of relationships among taxonomic units, and potential processes driving diversification within this genus. We sampled 334 individuals of the genus *Symphysodon* from 24 localities that span the complete geographic distribution of this genus. The sampling scheme included all known phenotypic groups, species and subspecies. Analyses were based on 474 bp of the mitochondrial control region and 1443 bp of the exon 3 of RAG1 gene. We observed 102 mtDNA haplotypes defined by 89 segregating sites, and 5 nuDNA alleles defined by three segregating sites. Maximum-likelihood, Bayesian-inference and statistical parsimony analyses revealed three well defined monophyletic groups. These clades corresponded to the 'green' and 'blue' groups of *Symphysodon aequifasciatus*, and to a previously morphologically unrecognized clade from the Xingu River drainage. These three clades were nested within a paraphyletic assemblage consisting of the 'brown' group of *S. aequifasciatus* and of both described subspecies of *S. discus*, the 'Heckel' and the 'abacaxi' discus. Nuclear allele sharing was observed among groups, but there were significant differences in frequencies. We inferred several processes including past fragmentation among groups, and restricted gene flow with isolation by distance within the paraphyletic 'brown + Heckel + abacaxi' groups, and suggest that differences among the 'blue', 'Heckel' and 'brown' groups are potentially maintained by differences in water chemistry preferences. We further inferred colonization of the western Amazon basin by an ancestor of the 'green' clade. The 'green' group was the only group with a pattern of haplotype distribution consistent of a demographic expansion, and the divergence of this clade from other groups of discus was consistent with recent geologic evidence on the breach of the Purus Arch which separates western Amazon from eastern Amazon. We further hypothesized that the differentiation of the 'Xingu' clade could be due to vicariance events resulting from Pleistocene sea level, and thus Amazon River level fluctuations. We discuss the bearings of our results on the current taxonomy of this group, and on the biological reality of the different forms, subspecies and species of *Symphysodon* concluding that we are probably observing a process of diversification, and therefore taxonomy will remain contentious.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

The Amazon basin is undoubtedly one of the most species rich regions of this planet. Many major groups of plants and animals have exceptionally high alpha as well as beta diversities within this region, and a number of mechanistic hypotheses have been proposed to have generated this diversity. As early as 1852, the British naturalist Alfred Russel Wallace put forward an evolutionary hypothesis to explain the history of diversification of Amazonian vertebrates (Wallace, 1852), and since then, several alternative hypotheses have been proposed (e.g. Ayres and Clutton-Brock,

1992; Bush, 1994; Colinvaux, 1993; Endler, 1977; Fjeldsá, 1994; Haffer, 1969, 1997; Lundberg et al., 1998), although nearly all these hypotheses focus on processes of diversification on land.

Most relevant events for the diversification of the South American fish fauna appear to be those tied to the Neogene orogenic history of South American, and Plio-Pleistocene climatic events. Preceding the establishment of modern drainage patterns in the late Miocene, the western Amazon was covered by the freshwater Lake Pebas from 15 to 10 mya (Hoorn, 1996; Lundberg et al., 1998; Marshall and Lundberg, 1996) which drained north into the Caribbean Sea though what is now the Maracaibo basin of Venezuela. Lake Pebas was also intermittently connected to the Paraná basin and thus connected the south-central Atlantic to what is today's Rio de la Plata region. Western and central Amazonian rivers drained into Lake Pebas (Hoorn et al., 1995; Lundberg et al.,

* Corresponding author. Address: Department of Biology, University of Puerto Rico, Rio Piedras, San Juan, PR, Puerto Rico. Fax: +1 787 764 3875.

E-mail address: hrbek@uprrp.edu (T. Hrbek).

1998), while eastern Guyana and Brazilian Shield rivers drained into the Atlantic Ocean through what is now the lower Amazon River channel (Costa et al., 2001; Hoorn et al., 1995). The western Amazon occupying the Lake Pebas was separated from the central and eastern Amazon by the Purus Arch. The uplift of the Venezuelan Andes and the Vuapes arch between 8 and 5 mya (Gregory-Wodzicki, 2000; Hoorn, 1993; Hoorn et al., 1995) blocked off the outflow of Lake Pebas into the Caribbean Sea. A subsequent rise of the central Andes (Gregory-Wodzicki, 2000) resulted in the breach of the Purus Arch and the connection of the western Amazon with the central and eastern Amazon. The initial independence of the western Amazon from the central and eastern Amazon basin could have promoted allopatric diversification, while the reconnection of the basins would have allowed for reciprocal basin colonization by endemic faunas.

Pliocene and Pleistocene climatic oscillations also appeared to have had a significant impact on the Amazonian fish fauna. An early Pliocene marine incursion into the eastern Amazon basin associated with global sea level changes occurred 5 million years ago and lasted for approximately 800,000 years (Haq et al., 1987; Nores, 1999). This incursion effectively eliminated all freshwater habitats in the low-lying areas of the lower and central Amazon basin. More recent Pleistocene global sea level changes (Haq et al., 1987; Liu and Herbert, 2004; Pillans et al., 1998) could potentially have resulted in marine incursions, or would have affected the extent of the Amazonian floodplain especially in the eastern Amazon basin. The Amazon River channel in the eastern Amazon basin is relatively narrow, and compared to the western Amazon basin has a relatively steep gradient (Costa et al., 2001). Low sea levels would have resulted in an increase in current velocity, decrease in the width of the Amazon River and the elimination of large portions of the narrow floodplain of the central and eastern Amazon. Conversely, higher sea levels insufficiently high to result in marine incursions would have resulted in the backing up of the Amazon River with the concomitant result of decreased flow velocity, an increase in the width of the Amazon River and the expansion of the floodplain. Higher sea levels would thus have facilitated floodplain connectivity between regions, while low sea levels would have promoted fragmentation and isolation.

It is clear that tectonic and climatic events in the middle to late Cenozoic of South America have drastically modified the Amazonian landscape. The western Amazon basin was a lacustrine habitat, while the freshwater habitat of eastern and central Amazon basin was at least once, but possibly repeatedly destroyed by marine incursions. These marine incursions into the eastern and central Amazon basin also cause the isolation of Guyana and Brazilian Shield tributaries, while more moderate sea level rises not resulting in incursions would have facilitated the colonization of the central and eastern Amazon basin.

Species of the genus *Symphysodon* are distributed throughout the low-lying areas of the Amazon basin. These fishes are found in western Amazon from the border of Peru and Brazil to the lower Araguaia–Tocantins River system which empties into the delta of the Amazon River. *Symphysodon* is found in sediment-free calm backwater lakes and river arms in all three major Amazonian water types, the white, the black and the clear-water types (Sioli, 1984). In the western Amazon, *Symphysodon* is found along the length of most tributaries. Western Amazon is situated in an area previously occupied by the freshwater Lake Pebas (Lundberg et al., 1998). In the central and eastern Amazon, *Symphysodon* appears to be largely restricted to lower portions of the Amazon's tributaries which during the early Pliocene would have been covered by a marine incursion for nearly a million years (Haq et al., 1987; Nores, 1999).

In addition to their Amazon basin-wide distribution, *Symphysodon* also exhibits a large amount of morphological variation which is partitioned into species and subspecies. *Symphysodon discus* (the

Heckel discus) is found in the Negro River basin, a major northern black-water tributary of the Amazon River; populations from the Trombetas River, also a northern tributary of the Amazon, are also considered *S. discus* (Kullander, 1996). The subspecies *Symphysodon discus willischwartzi* (the abacaxi discus) has been described from the Abacaxis River which is a minor tributary of the Madeira River, itself a major southern muddy tributary of the Amazon. Other portions of the Amazon basin are occupied by *Symphysodon aequifasciatus*. The subspecies *S. aequifasciatus haraldi*, the 'green' discus, is found in the western Amazon basin (type locality Tabatinga), and *S. aequifasciatus axelrodi*, the 'brown' discus, is found in the eastern Amazon River basin (type locality Belém). The nominal subspecies *S. aequifasciatus aequifasciatus*, the 'blue' discus, is found in the central portion of the Amazon basin (note: in the popular literature the 'green' western Amazon fishes are considered *S. a. aequifasciatus*, while the 'blue' central Amazon fishes are considered *S. a. haraldi*). *Symphysodon* species and especially subspecies are predominantly diagnosed by color and color patterns, and geographic distribution. The main characters that distinguishes *S. discus* from *S. aequifasciatus* is the presence of the distinctive broad black vertical bar in the central portion of the body of *S. discus*, faint horizontal silvery striations and the higher mean number of scales (E series), dorsal rays and spines in *S. discus* (Kullander, 1996). *Symphysodon d. discus* and *S. d. willischwartzi* are allopatrically distributed north and south of the Amazon River, respectively, and also differ in color and color pattern, *S. d. discus* being purplish grey and *S. d. willischwartzi* being brownish yellow. The three subspecies of *S. aequifasciatus* differ from each other in color and color pattern, and are allopatrically distributed; *S. a. haraldi* (the green discus) has nine distinct vertical bars, overall greenish hue over light brown background, and dominant individuals have red spots on the ventral portion of their body; *S. a. aequifasciatus* (the blue discus) has distinctive iridescent silvery-blue striations on the head and dorsal portion of the body, overall bluish sheen and weak to non-existent vertical bars; *S. a. axelrodi* (the brown discus) is a brownish to reddish-brown fish with occasional very faint dorsal striations and lack of vertical bars except for a weak to no-present vertical bar across the eye.

Although the above listed criteria can be used to assign individuals to species and subspecies, the taxonomy of the genus *Symphysodon* is contentious. The most commonly accepted classification of Schultz (1960) and Burgess (1981) who divide the genus into two species with five subspecies (see previous paragraph). Unfortunately Pellegrin (1904), the original describer of *S. aequifasciatus*, never designated a holotype, and two of the three syntypes are from the geographical area of *S. a. haraldi* while the third syntype is from the geographical area of *S. aequifasciatus axelrodi*, which invalidates the subspecific descriptions of Schultz (1960), see Ready et al. (2006). Other authors such as Allgayer (2001) consider *Symphysodon* a monotypic genus with three subspecies (*S. d. discus*, *S. d. willischwartzi* and *S. d. aequifasciatus*), while Kullander (1996, 2003) considers the genus to comprise two species (*S. discus* and *S. aequifasciatus*). A recent study of Ready et al. (2006) concluded that *Symphysodon* comprises three species, as did the study of Bleher et al. (2007). Ready et al. (2006) suggested that red-spotted individuals largely restricted to the western Amazon represent a distinct species *Symphysodon tarzoo*, while Bleher et al. (2007) considered the western Amazon fishes *S. aequifasciatus*, and renamed the 'blue' and 'brown' fishes of central and eastern Amazon basin *S. haraldi*. The name *S. tarzoo* is a new usage of *Symphysodon discus tarzoo* described by Lyons (1959). Lyons's (1959) description is, however, flawed in that it fails to designate a holotype and is based on specimens held at a Peruvian exporting facility without certain geographic providence. Therefore Ready et al. (2006) designate a neotype (INPA 25960) from the Jutá River, restrict the type series of *S. aequifasciatus* to Santarém (MNHN 1902–130) and synonymize *S. a. haraldi* and *S. a. axelrodi* with *S. aequifasciatus*. Although recog-

nizing the existence of a third species in the western Amazon basin, the taxonomic treatment of *S. aequifasciatus* and *S. haraldi* by Bleher et al. (2007) suffers from similar taxonomic difficulties as does the subspecific description of Schultz (1960). Ready et al. (2006) also question the taxonomic validity *S. discus willischwartzii* from the Abacaxis River, considering these fish to be hybrids between *S. discus* and *S. aequifasciatus*. However, it is difficult to imagine how could *S. d. willischwartzii* which is found within the distributional range of *S. aequifasciatus*, and the nearest *S. discus* are several hundred kilometers away in the Negro River have arisen from a hybridization event between *S. aequifasciatus* and *S. discus*.

In spite of these recent taxonomic developments, we use the above described phenotypic characteristics to assign individuals to the five phenotypic groups corresponding to the species and subspecies of Schultz (1960) and Burgess (1981). The taxonomic scheme of Schultz (1960) and Burgess (1981) represents a finer partitioning of observed phenotypic variation in the genus *Symphysodon*, and it is not incongruent with the taxonomic treatments of *Symphysodon* by Kullander (1996, 2003), Ready et al. (2006) and Bleher et al. (2007). This phenotypic diversity and Amazon basin-wide geographic distribution provide an ideal study system for the testing of hypotheses of diversification of fishes of the Amazon basin.

2. Materials and methods

2.1. Species sampling

We sampled 334 individuals from 24 localities which span the geographic range of the genus *Symphysodon*. The sampling also encompassed all the observed phenotypic variation. Average sample size per locality was 14 individuals (Fig. 1). All individuals have been vouchered either at the ichthyological collection at Instituto Nacional de Pesquisas da Amazonia (INPA), or in the tissue collection at UFAM-LEGAL. For the purposes of the present work we defined the following five groups: 'green' group (*S. a. haraldi*) from the western Amazon basin, 'blue' group (*S. a. aequifasciatus*) from central Amazon basin, 'brown' group (*S. a. axelrodi*) from eastern Amazon basin, 'Heckel' group (*S. discus*) from the Negro and Trombetas River basins and 'abacaxi' group (*S. d. willischwartzii*) from the Abacaxis River basin. The closely related heroine species *Uaru amphiacanthoides*, *Mesonauta insignis* and *Pterophyllum scalare* (Farias et al., 2000), were used as outgroups in phylogenetic analyses.

2.2. Laboratory protocols

Specimens were collected, and tissue sample from the right caudal peduncle or the right pectoral fin was preserved in 95% ethanol. Specimens were then quick-frozen or preserved in 10% formalin, and deposited at museum and tissue collections of INPA and UFAM. Total genomic DNA was extracted using a standard phenol/chloroform method and precipitated with 70% ethanol (Sambrook et al., 1989), or alternately using Qiagen™ DNeasy spin-column tubes (Valencia, CA).

Polymerase chain reaction (PCR) amplification was performed on total genomic DNA. Negative controls were run for all reactions. General amplification strategy for mitochondrial DNA follows Sivasundar et al. (2001) while nuclear gene amplification strategy follows Hrbek et al. (2007). Mitochondrial DNA sequences used in this study encoded the control region, while nuclear DNA sequences consist of the nearly complete third exon on the recombination activating gene 1 (RAG1).

Mitochondrial DNA PCR reactions were carried out in a final volume of 25 μ l and contained 11.8 μ l of ddH₂O, 1.7 μ l of 10 mM MgCl₂, 2.0 μ l of 10 mM dNTPs, 2.5 μ l of 10 \times buffer (100 mM Tris-HCl, 500 mM KCl, 15 mM MgCl₂), 2.5 μ l of each primer (2 μ M), 0.3 μ l of Taq DNA Polymerase (5 U/ μ l) and 1 μ l of DNA (concentration varied between 50 and 100 ng). PCR conditions were as follows: pre-heating at 72 $^{\circ}$ C for 60 s, denaturation at 92 $^{\circ}$ C for 15 s, primer annealing at 55 $^{\circ}$ C for 35 s and primer extension at 72 $^{\circ}$ C for 90 s; the last three steps were repeated 35 times and followed by a final extension at 72 $^{\circ}$ C for 5 min. Primers 5'-AAACTGGGATTAGATACCCCACTA-3' (F-TTF) and 5'-GGCGGATACTTGCATGT-3' (12Sr.5) were used to amplify the complete mitochondrial control region. To amplify the third exon of the nuclear coded recombination activating gene 1 (RAG1), we used a semi-nested PCR strategy. PCR was performed in 25 μ l reaction volumes containing 11.8 μ l of ddH₂O, 1.7 μ l of 10 mM MgCl₂, 2.5 μ l of 10 \times buffer (200 mM Tris-HCl, 20 mM MgSO₄, 100 mM KCl, 100 mM (NH₄)₂SO₄, 1% Triton[®] X-100, 1 mg/ml nuclease-free BSA), 2.5 μ l of each primer (2 μ M), 2.0 μ l dNTP mix (10 mM), 3 U KlenTaqLA DNA Polymerase and 1 μ l of DNA template (concentration varied between 50 and 100 ng). A semi-nested PCR consists of a first round of PCR followed by a second round of PCR on the unpurified PCR product with a new set of primers; one primer is same as the original while the second primer is placed internally to the one

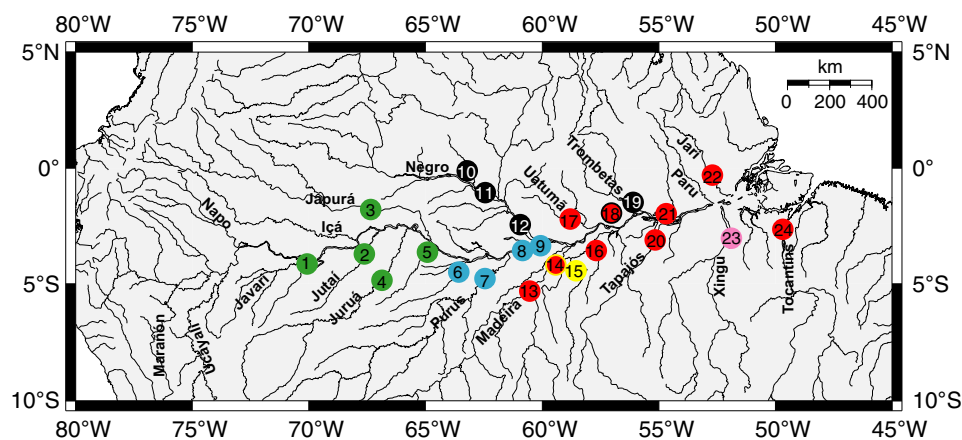


Fig. 1. Distribution and collecting areas of varieties of the genus *Symphysodon*. Numbers represent: (1) Tabatinga (Caldeirão River), (2) lower Jutai River, (3) Lake Santa Maria (Japurá River), (4) Bauana (Juruá River), (5) Igarapé Bowona (Tefé River), (6) Lake Coari (Coari River), (7) Lake Castanho (Purus River), (8) Lake Manacapuru (Manacapuru River), (9) Iranduba (Solimões River), (10) lower Demini River, (11) Igarapé Bui-Bui (Negro River), (12) Novo Airão (Negro River), (13) Novo Aripuanã (Madeira River), (14) Nova Olinda do Norte (Madeira River), (15) central Abacaxis River, (16) Maues (Maués River), (17) Itapiranga (Uatumã River), (18) Nhamundá (Nhamundá River), (19) Porto Trombetas (Trombetas River), (20) Belterra (Tapajós River), (21) Lake Grande (Amazonas River), (22) Igarapé Arapiranga (Jari River), (23) Vitória do Xingu (Xingu River), (24) Cametá (Tocantins River). Colors correspond to the varieties and groups analyzed in this study: green = 'green' group; blue = 'blue' group; black = 'Heckel' group; yellow = 'abacaxi' group; red = 'brown' group and purple = 'Xingu' group. Localities 14 and 18 contained the 'abacaxi' and 'brown' group phenotypes, and the 'Heckel' and 'brown' group phenotypes, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

used in the first PCR. The temperature profile for the first and nested PCRs consisted of (1) pre-heating at 68 °C for 60 s, (2) denaturation at 93 °C for 10 s, (3) annealing at 55 °C for 35 s, (4) extension at 68 °C for 150 s and (5) a final extension at 68 °C for 10 min. Steps 2–4 were repeated 25 times. The Rag1 gene transcript was amplified with an external primer pair 5'-CCWGCTGTITGYYTG GCCATIMG-3' (Rag1.HB.F.L1) and 5'-GTGTAGAGCCARTGRTGYTT-3' (Rag1.Mart.R6), and then in the semi-nested reaction the primer Rag1.Mart.R6 was replaced with 5'-TGYTTTCAGCACGTCCTCCARY TC-3' (Rag1r.6).

PCR products were evaluated on a 1% agarose gel, and then purified using the GFX PCR DNA Kit (GE Healthcare). The mitochondrial DNA product was sequenced using the primer F-TTF (5'-AAAC TGGGATTAGATACCCACTA-3') while the RAG1 product was sequenced using the internal primers Rag1f.4 (5'-GAACGCT AYGATATATGGAG-3'), Rag1r.7 (5'-GCAGACRCCCTTTGACTCTGT C-3') and Rag1r.2 (5'-GAGAARCGRACAGCCTTYTC-3'). Cycle sequencing PCR followed manufacturer's recommended protocol for DYEnamic ET Dye Terminator mix (GE Healthcare); primer annealing temperature was at 50 °C and we used ~30 ng of purified PCR product. Sequences were determined on an Amersham MegaBACE automatic DNA sequencer. Sequence products were edited and aligned using BioEdit Program (Hall, 1999). We sequenced 474 bp of the mitochondrial control region and 1443 bp of the exon 3 of RAG1 gene.

2.3. Data analysis

All electrophoregrams were inspected visually for quality, orthologous gene regions were aligned using the program Clustal W (Thompson et al., 1996) implemented in the program BioEdit (Hall, 1999), and alignment was confirmed visually. Any insertions and deletions, as well as SNPs were rechecked against original electrophoregram data. Mitochondrial gene regions were tested for an anti-G bias characteristic of the mitochondrial DNA genes, but not of the nuclear genome, to support our conclusion that we have collected genuine mitochondrial DNA data (Zhang and Hewitt, 1996). Alignment and integrity of the third exon of the RAG1 gene was confirmed by translating DNA data into putative aminoacid sequences in BioEdit (Hall, 1999). Since each electrophoregram represented two alleles due to the diploid nature of the *Symphysodon* nuclear genome, we partitioned the nuclear data into two alleles per individual. Allelic states could be inferred unambiguously. All sequence data have been deposited in GenBank under Nos. EU497298–EU497304 (RAG1 gene) and EU497305–EU497407 (mtDNA control region).

2.4. Phylogenetic reconstruction

Phylogenetic relationships were estimated on two separate datasets. The first dataset contained the mtDNA control region haplotypes only, and the second dataset contained combined mtDNA and nuDNA genes. Mitochondrial and nuclear data were tested for phylogenetic incongruence via the Partition Homogeneity test (Farris et al., 1994). For this test and for phylogenetic analyses based on combined mitochondrial and nuclear data, each individual was coded to contain one nuclear allele with polymorphic sites coded by a standard IUPAC code.

Maximum-likelihood topology for molecular data was estimated in Treefinder (Jobb et al., 2004), and its robustness was assessed with 2000 non-parametric bootstrap replicates. We used the Hasegawa–Kishino–Yano (HKY85) model of molecular evolution (Hasegawa et al., 1985) with a portion of the sites treated as invariable, and in the remaining sites the probability of change from one nucleotide state to another was assumed to follow a gamma distribution. The HKY85 model was suggested as the

most appropriate model of molecular evolution for this dataset by the software MODELTEST 3.7 (Posada and Crandall, 1998) using the Akaike information criterion to select among alternate models.

Bayesian-inference analyses were performed in the software MrBayes 3.01 (Ronquist and Huelsenbeck, 2003). Data were partitioned into four categories representing the control region mtDNA (one partition) and the first, second and third positions of the nuDNA RAG1 gene (three partitions). We ran 5,000,000 generations using default long and short chain and heating parameters, sampling trees and branch-length every 100 generations. Convergence of all estimated parameters was inferred using the Gelman and Rubin (1992) criterion. Log-likelihoods stabilized within the first 5% of the run, and therefore we discarded these initial 100,000 trees as burn in steps in the computation of a 50% majority rule consensus tree. Following Huelsenbeck et al. (2001), the frequency of a clade among the sampled trees was interpreted as the posterior probability of that clade. Although the interpretation of Bayesian posterior probabilities has been questioned (e.g. Suzuki et al., 2002), the posterior probabilities are true probabilities under the assumed model of substitution (Rannala and Yang, 1996); thus we considered clades to be significantly supported when Bayesian posterior probabilities were >95%.

2.5. Tests of alternate phylogenetic hypotheses

Alternate hypotheses of phylogenetic relationships derived from taxonomic hypotheses were tested under the likelihood framework as implemented in Treefinder (Jobb et al., 2004) by searching for the most likely phylogenetic hypothesis under the alternate phylogenetic hypothesis constraint. Tests were implemented for both mtDNA only and for combined mtDNA and nuDNA data. These alternative hypotheses were compared with most likely unconstrained topologies, and significance of topological differences were assessed using the parametric Shimodaira–Hasegawa test (Shimodaira and Hasegawa, 1999).

2.6. Nested clade analysis

We used the program TCS 1.21 (Clement et al., 2000) to estimate mtDNA haplotype and nuDNA allelic networks following the cladogram estimation rules laid out in Templeton et al. (1992) and elaborated in Templeton (1998, 2004b). This method gives an estimate of the maximum number of differences among the haplotypes as a result of single substitutions with the statistical confidence of 95% (Posada and Crandall, 2001). Ambiguous connections resulting from homoplastic mutations were resolved using information from the maximum-likelihood topology inferred under the HKY85+I+ Γ model of molecular evolution in Treefinder (Jobb et al., 2004). The program GeoDis 2.5 (Posada et al., 2000) was used to test significant changes in haplotype and nested clade geographic distribution relative to other haplotypes and nested clades within their higher level nesting clades (Templeton, 1995). Nested clade analysis (NCA) takes into account gene genealogies, haplotype frequencies and geographical data to inferentially discriminate between historical events (such as fragmentation or range expansion events) and ongoing processes (such as gene flow). To implement NCA, the haplotype network was manually nested into increasingly inclusive clades following the rules described by Templeton and collaborators (Templeton et al., 1987; Templeton and Sing, 1993). NCA analysis estimates the clade distance, which measures the geographical spread of a clade, and the nested clade distance, which measures how a clade is geographically distributed relative to other clades in the same higher level nesting category. Significance of both distances was calcu-

lated using the method of random permutations (Templeton and Sing, 1993) implemented in GeoDis 2.5 (Posada et al., 2000). Geographic distances among regions were calculated using Google-Earth™. The interpretation of significant contrasts results follows the November 11, 2005 Inference Key for the Nested Haplotype Tree Analysis of Geographical Distance available at <http://darwin.uvigo.es>. Although recently criticized (e.g. Knowles and Maddison, 2002; Panchal and Beaumont, 2007; Petit, 2008) we share the opinion with other authors (e.g. Garrick et al., 2008; Templeton, 2004a) that NCA can provide valuable inferences about historical and ongoing processes that shape patterns of population genetic diversity, however, whenever possible, such inferences should be validated using other statistical tests.

2.7. Population diversity analyses

Additional characterizations of the population genetic parameters were carried out in the software package Arlequin 3.1.1 (Excoffier et al., 2005). Genetic diversity within regions was measured as the number of nuclear DNA alleles and mitochondrial DNA haplotypes, haplotype diversity (H), nucleotide diversity (π) calculated by Nei's method (Nei, 1987) and number of polymorphic sites. We also performed Tajima's D test (Tajima, 1989) and the more sensitive Fu's F_s test (Fu, 1997) to examine whether samples from different localities are at equilibrium with respect to mtDNA, considering that a significant deviation from mtDNA genetic equilibrium is presumably a result of recent population expansion or bottleneck in situations where no selective advantage among haplotypes exists (Rand, 1996). Tajima's test compares the number of segregating sites with nucleotide diversity (defined as the average number of nucleotide differences per site between any two sequences), while Fu's F_s test considers the probability of an estimated θ given a mutational pattern. Additional information regarding past and present population sizes and time to coalescence was derived from mismatch analysis (Rogers and Harpending, 1992). The Mantel test (Mantel, 1967) as implemented in the program Arlequin 3.1.1 (Excoffier et al., 2005) was used to assess the significance of association between genetic and geographical distances. Statistical significance was tested using 10,000 random permutations.

3. Results

The dataset of 334 individuals contained 102 mtDNA haplotypes defined by 89 segregating sites, and 5 nuDNA alleles defined by four segregating sites (Table S1). The pattern of nucleotide substitution in the mtDNA and combined mtDNA and nuDNA datasets best fits the HKY85 (Hasegawa et al., 1985) model of molecular evolution, with transition/transversion substitution = 13.54:1 based on analysis in the program ModelTest (Posada and Crandall, 1998). Uncorrected pair-wise genetic distance matrix shows that the sequence divergence among all groups varies from 0.50% to 4.11%, while within group divergence varies from 0.04% to 0.42% (Table 1).

Table 1
Mean pair-wise p -divergences between major groups of *Symphysodon*

	green (%)	blue (%)	Heckel (%)	abacaxi (%)	brown (%)	Xingu (%)
green	0.40	—	—	—	—	—
blue	3.71	0.36	—	—	—	—
Heckel	2.79	3.15	0.42	—	—	—
abacaxi	2.46	3.05	0.50	0.04	—	—
brown	2.48	3.01	0.60	0.17	0.29	—
Xingu	3.41	4.11	2.06	1.90	1.98	0.15

Diagonal values are within group divergences.

3.1. Phylogenetic and phylogeographic analyses

Statistical parsimony analysis of mtDNA data indicated that *Symphysodon* is composed of four interconnected networks (Fig. 2). These four networks corresponded to the 'green' group from upper Amazon basin, to the 'blue' group from central Amazon basin, to a group from the Xingu River, and to a group from the lower Amazon basin including the Negro River and all areas downstream of its confluence with Solimões River. The last group contained the 'brown', the 'Heckel' and the 'abacaxi' phenotypes. An outgroup rooted maximum-likelihood phylogeny (Fig. 3) indicated that monophyletic 'green' and 'blue' clades are sister, and are nested within the 'brown + Heckel + abacaxi' group, and the root was placed on the branch separating haplotypes from the Xingu River plus two haplotypes found in the Maues River from all other haplotypes (Fig. 3). Bayesian-inference of phylogenetic relationships placed the root on the branch separating the Xingu River haplotypes from all other haplotypes. Outgroup rooting of haplotype networks is difficult (Castelloe and Templeton, 1994; Posada and Crandall, 2001), and therefore we also used a coalescent based inference (Castelloe and Templeton, 1994) which placed the root on the haplotype 39 (Fig. 3). This haplotype shared by the 'brown', the 'Heckel' and the 'abacaxi' phenotypes, was centrally positioned in the haplotype network and was the most frequent haplotype (Fig. 2). Although rooting was not consistent among inferences, all three methodologies inferred the sister taxon relationship of the 'green' and 'blue' clades, and their placement within the 'brown + Heckel + abacaxi' group. When *S. discus* ('Heckel' and 'abacaxi' group) and *S. aequifasciatus* ('green', 'blue' and 'brown') were forced into monophyly, the resulting phylogenetic hypotheses were significantly less likely (Shimodaira–Hasegawa test; $P < 0.05$).

Nested clade analysis suggested that these four groups were the result of allopatric fragmentation. At the highest level of nesting within the 'brown + Heckel + abacaxi' group, an additional allopatric fragmentation event was inferred, and this event pertained to the difference between the genetically divergent Xingu River locality and all other localities of the 'brown + Heckel' group. Restricted gene flow with isolation by distance was also inferred for the contrast between fishes of the Negro, Trombetas and Nhamunda River and all other localities of the lower Amazon. No significant inferences were made for the 'blue' group. At the highest level of nesting, long distance colonization with some short distance dispersal was inferred for the 'green' group.

3.2. Population analyses

When the mtDNA haplotypes of *Symphysodon* were partitioned in six groups corresponding to the 'green', 'blue', Xingu, 'Heckel', 'abacaxi' and 'brown' phenotypes, 87.54% of the total genetic variance accounted for among group variance ($\Phi_{ST} = 0.8754$, $P < 0.0001$). For individual groups we observed significant difference among localities, albeit less than 25% of the total genetic variance accounted for among locality variance in the 'brown' group,

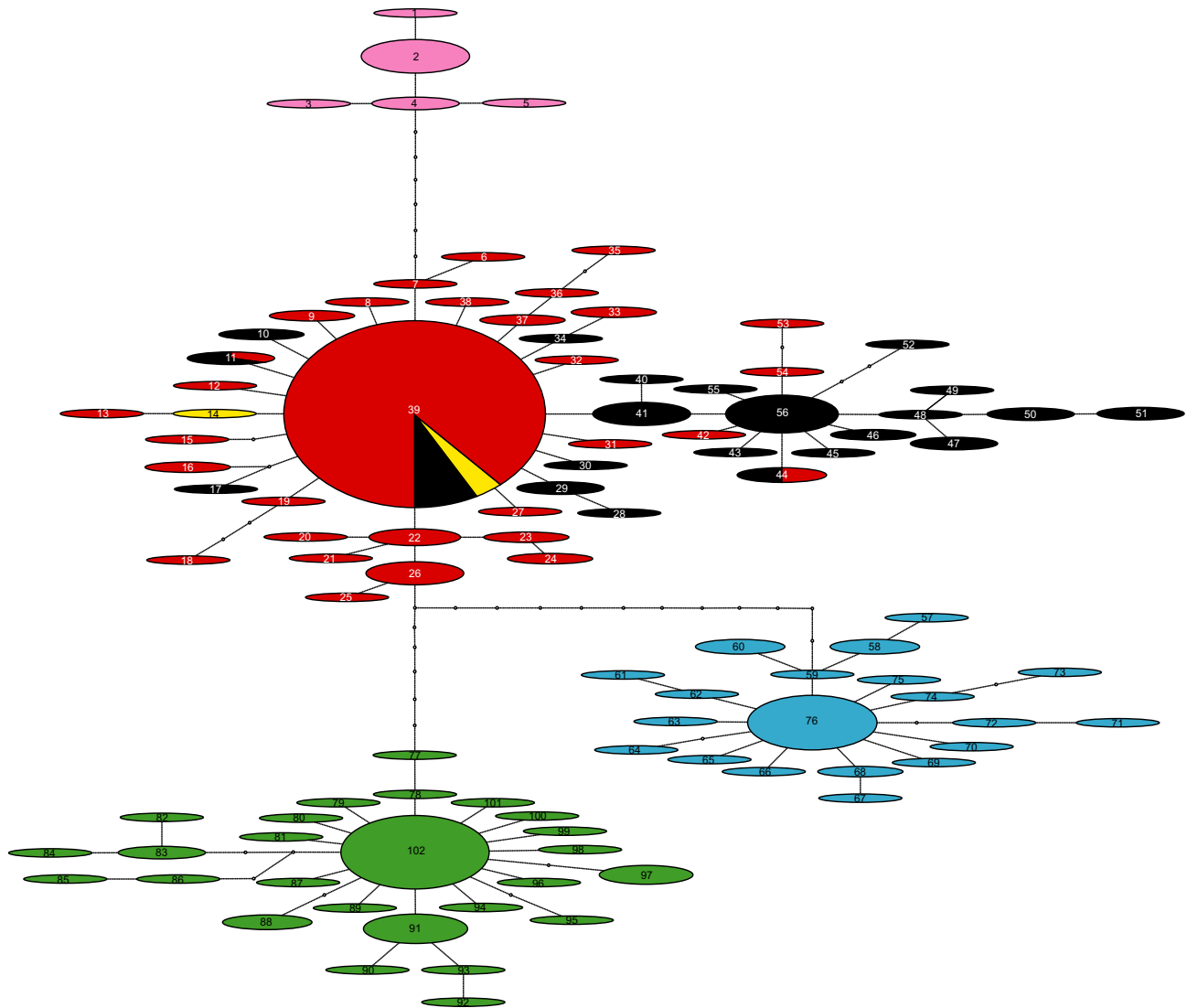


Fig. 2. Maximum parsimony networks of mtDNA haplotype relationships. The three networks cannot be connected parsimoniously. The tree networks are not drawn to the same scale. Colors correspond to the varieties and groups analyzed in this study: green = 'green' group; blue = 'blue' group; black = 'Heckel' group; yellow = 'abacaxi' group; red = 'brown' group and purple = 'Xingu' group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and less than 15% of the total genetic variance in the 'green' and 'blue' groups (Table 2). In the 'Heckel' group approximately 55% of the observed variance was due to among sampling locality differences, largely accounted for by the difference between the Trombetas locality and all other localities. Within no group was there a significant correlation of genetic and geographic distances. Although RAG1 alleles were largely shared among different groups (Fig. 4), there were also significant differences in allelic frequencies among groups ($\Phi_{ST} = 0.1364$, $P = 0.0013$).

The null hypothesis of selective neutrality in the mtDNA was rejected for majority of groups using both Tajima's D (Tajima, 1989) and Fu's F_s (Fu, 1997) tests. Significant negative values in these two tests result from an excess alleles separated by small mutational distances or an excess of rare alleles, respectively, and in cases of selective neutrality can be interpreted as a signature of a demographic expansion (e.g. Hartl and Clark, 2006). Hence, a mismatch distribution analysis (Rogers and Harpending, 1992) was performed on all groups. For the 'green' group the observed distribution did not deviate significantly from the null hypothesis of population expansion (SSD = 0.00019, $P = 0.985$). The Harpending's Raggedness index was significantly low ($r = 0.01785$, $P = 0.965$),

indicating a smooth distribution which also consistent with the hypothesis of population expansion (Harpending, 1994). The null hypothesis of population expansion was rejected for all other groups, and their haplotype distributions were ragged.

4. Discussion

Evolutionary history of the genus *Symphysodon* is clearly complex resulting in a complicated taxonomic history. Although there are differences in color and color patten among the different varieties of *S. aequifasciatus*, the traditional grouping into three 'green', 'blue' and 'brown' phenotypes also encompasses large amount of intragroup phenotypic variation. These phenotypic groups also represent geographic units with the 'green' group occurring upstream of the city of Coari, the 'blue' group occurring from Coari to the junction of the Negro River and the Amazon River, and the 'brown' group occurring downstream of this junction. The 'abacaxi' discus, *S. d. willischwartzi*, is found in the Abacaxis River, a southern tributary of the Madeira River, a southern tributary of the Amazon River. The nominal 'Heckel' discus, *S. discus*, is thought to occur in the Negro River, although Kullander (1996) also lists it as occurring in the

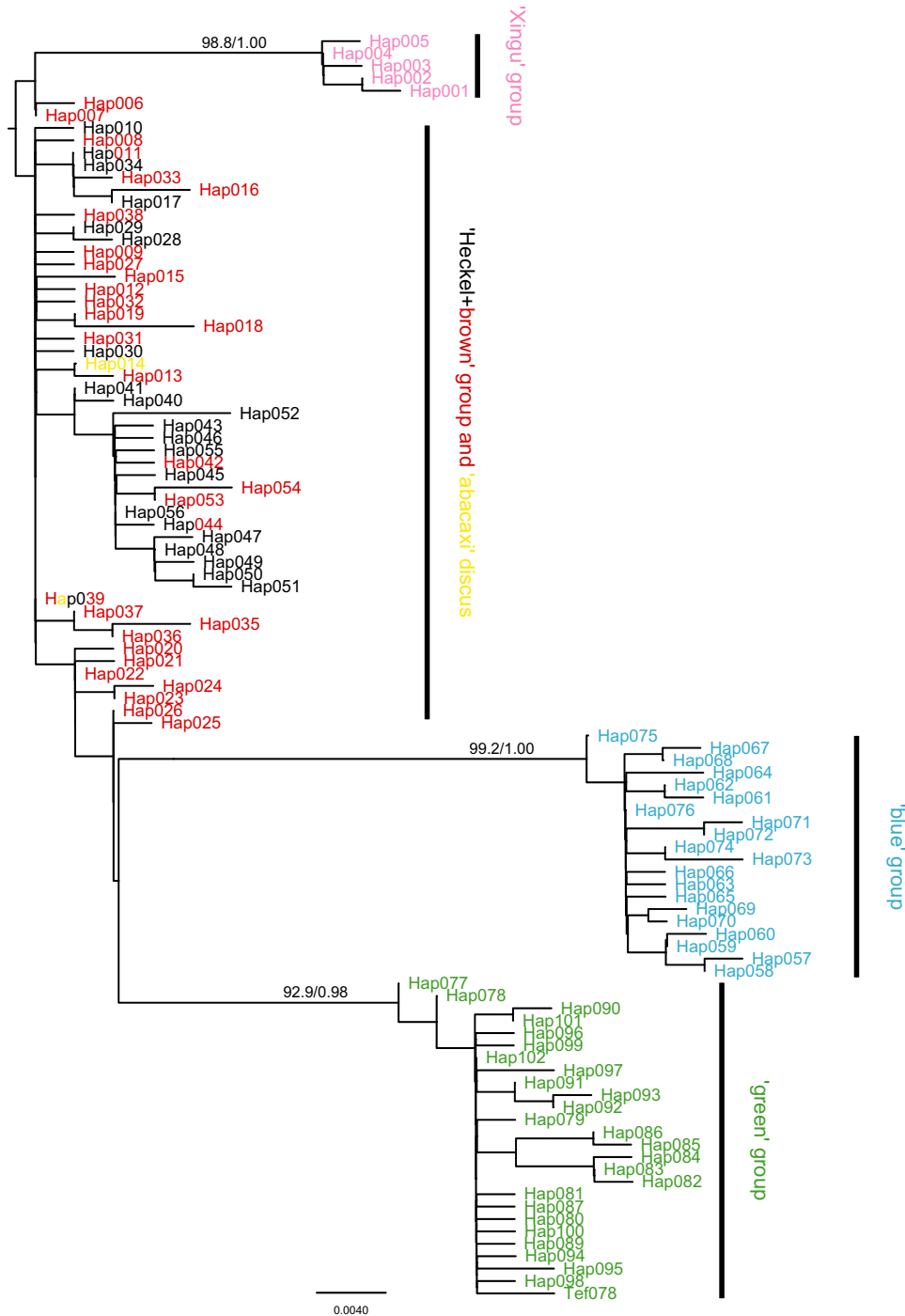


Fig. 3. A maximum-likelihood phylogenetic hypothesis. Only unique haplotypes were analyzed. Position of the root is derived from outgroup rooting with *Uaru amphicanthoides*, *Mesonauta insignis* and *Pterophyllum scalare*. Statistically most likely (Crandall and Templeton, 1993) ancestral haplotype is the haplotype Cam198. Colors correspond to the varieties and groups analyzed in this study: green = 'green' group; blue = 'blue' group; black = 'Heckel' group; yellow = 'abacaxi' group; red = 'brown' group and purple = 'Xingu' group. Multicolored haplotype designations correspond to those haplotypes that were found in more than one phenotypic group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Trombetas River where we also collected it. We also observed fish with a characteristic 'Heckel' discus vertical bar in the Nhamundá River, which also descends the Guiana Shield. From a mitochondrial DNA perspective, the 'green' and 'blue' varieties of *S. aequifasciatus* are monophyletic, but are nested within a paraphyletic assemblage containing the 'brown' variety of *S. aequifasciatus* and both subspecies of *S. discus*. Furthermore, we observe an additional divergent lineage from the Xingu River basin that based on outgroup rooting, appears to be the sister clade to all other *Symphys-*

odon groups (Fig. 3). Nuclear DNA provides relatively little information, with most of the phylogenetic signal coming from mitochondrial data. Nuclear haplotypes are not exclusive to any particular group or species, however, significant differences in haplotype distribution among phenotypic groups exists.

The 'green' and 'blue' groups have a geographic distribution roughly upstream and downstream, respectively, of the Purus Arch. The Purus Arch is a major geological feature situated just upstream of Coari, and it is correlated with the faunistic division of

Table 2
Molecular characteristics of major groups of *Symphysodon*

	green	blue	Heckel	abacaxi	brown	Xingu	All
Individuals	77	45	77	11	105	19	334
Haplotypes	26	18	21	2	33	5	102
Alleles	3	2	1	1	5	1	5
Theta k	13.3882 (8.1725–21.6009)	10.6239 (5.7890–19.1953)	7.6576 (4.3462–13.1503)	0.6911 (0.1486–3.1790)	16.1551 (10.6425–24.1896)	1.8636 (0.6694–4.8624)	49.6859 (38.6706–63.5221)
Theta H	3.2472 ± 0.9258	2.6947 ± 1.0553	5.3607 ± 1.2805	0.4965 ± 0.4920	1.6160 ± 0.3785	0.8306 ± 0.4280	9.2647 ± 1.4667
Theta S	5.6974 ± 1.7721	3.8877 ± 1.4148	3.3407 ± 1.1797	0.4800 ± 0.4800	6.7053 ± 1.8967	1.1446 ± 0.6638	13.4652 ± 3.0070
Theta pi	1.9016 ± 1.2148	1.5970 ± 1.0723	2.0878 ± 1.3095	0.4000 ± 0.5086	1.5212 ± 1.0190	0.7251 ± 0.6638	9.4108 ± 4.7912
H	0.8035 ± 0.0415	0.7747 ± 0.0636	0.8661 ± 0.0252	0.4000 ± 0.2373	0.6798 ± 0.0488	0.5263 ± 0.1266	0.9141 ± 0.0113
π	0.0040 ± 0.0026	0.0034 ± 0.0023	0.0044 ± 0.0027	0.0009 ± 0.0011	0.0032 ± 0.0021	0.0015 ± 0.0013	0.0199 ± 0.0101
Tajima's D	−2.0770	−1.8722	−1.1071	−0.8165	−2.3670	−1.0788	−1.0038
P (D)	0.003	0.015	0.114	0.279	<0.001	0.166	0.149
Fu's Fs	−22.9397	−14.9067	−9.1658	0.0902	−28.2355	−2.0230	−24.0549
P (Fs)	<0.001	<0.001	<0.001	0.302	<0.001	<0.025	0.004
SSD	0.0002	0.0056	0.0002	0.0072	0.0035	0.0020	0.0175
P (SSD)	0.985	0.484	0.945	0.782	0.779	0.766	0.653
Ragg.	0.0179	0.0389	0.0298	0.2000	0.0306	0.0719	0.0105
P (Ragg.)	0.965	0.758	0.735	0.933	0.913	0.849	0.759
Va	12.61%	14.89%	56.80%	—	22.72%	—	87.85%
Vb	87.39%	85.11%	43.20%	—	77.46%	—	12.15%
Φ _{ST}	0.1261	0.1489	0.5680	—	0.2254	—	0.8785
P (Φ _{ST})	<0.0001	0.0019	<0.0001	—	<0.0001	—	<0.0001
Corr. r	−0.5136	0.0925	0.9060	—	0.1890	—	—
P (r)	0.9024	0.3746	0.0525	—	0.7816	—	—

Haplotypes = mtDNA haplotypes; alleles = nuDNA alleles; theta k = theta from the infinite-allele equilibrium (Ewens, 1972); theta H = theta from homogeneity (Chakraborty and Weiss, 1991); theta S = theta from number of segregating sites (Watterson, 1975); theta pi = theta from pair-wise differences (Tajima, 1983); H = haplotype diversity (Nei, 1987); π = nucleotide diversity (Nei, 1987); Tajima's D (Tajima, 1989), Fu's Fs (Fu, 1997); SSD = sum of square deviations statistic for fit between observed and expected mismatch expansion pattern (Schneider and Excoffier, 1999); Ragg. = raggedness index of a mismatch distribution (Harpending, 1994); Va = AMOVA among group variance component; Vb = AMOVA within group variance component; Φ_{ST} = a sequence data analogue to F_{ST} (Cockerham and Weir, 1993); Corr. r = correlation coefficient (Zar, 1999).

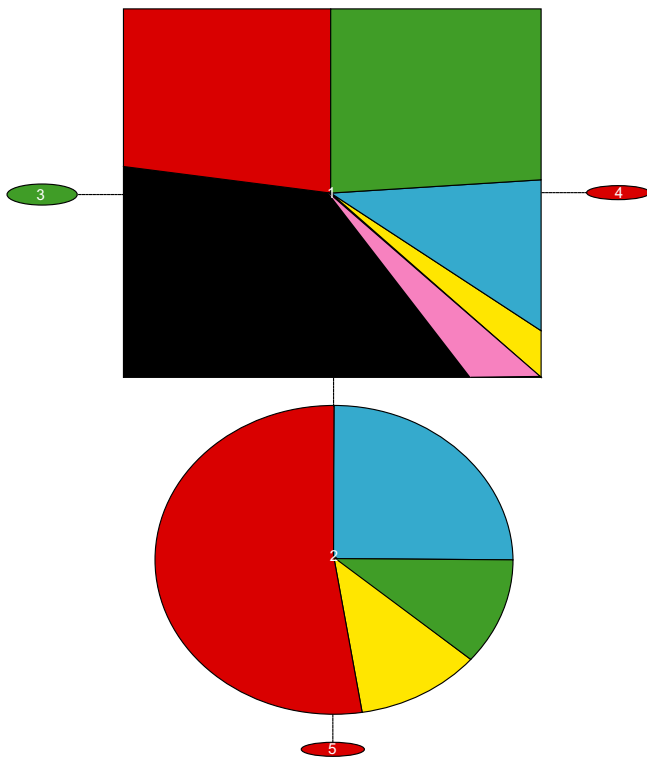


Fig. 4. Maximum parsimony network of nuDNA allele relationships. Colors correspond to the varieties and groups analyzed in this study: green = 'green' group; blue = 'blue' group; black = 'Heckel' group; yellow = 'abacaxi' group; red = 'brown' group and purple = 'Xingu' group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

western and eastern Amazon basins (e.g. Hubert and Renno, 2006; Wesselingh et al., 2002; Wesselingh and Salo, 2006). The Purus Arch formed during the last phases of the South American Andean uplift (Cordani et al., 2000; Gregory-Wodzicki, 2000), and was probably breached when the Amazon River reversed its direction and commenced flowing eastward (Lundberg et al., 1998), its current course. The geological overview of South America by Lundberg et al. (1998) suggests that the breach of the arch is approximately 8–6 mya, however, more recent evidence (Cordani et al., 2000; Vega, 2007) suggest the breach occurred much more recently. Based on tectonic evidence, Cordani et al. (2000) suggest a 3 mya breach, while Vega (2007) based on exhaustive study of sedimentary deposits between the cities of Tefé and Coari suggests the end of the Neogene period, i.e. the end of the Pliocene which terminates at 1.8 mya.

Ready et al. (2006) suggest that the Purus Arch could have acted as a barrier to gene flow in a previously continuous population, initiating allopatric speciation between *S. tarzoo*, the red-spotted 'green' group individuals, and all remaining *Symphysodon* groups. The breach of the arch would have then brought these two groups into secondary contact, resulting in the observed sympatry of *S. tarzoo* and *S. aequifasciatus*. The relatively low genetic divergence between different groups of *Symphysodon* (Table 1), and specifically the 3.71% mtDNA control region divergence between the 'green' clade and 'blue' clade of *Symphysodon* seems too low to have preceded the formation of the Purus Arch—this supposition is based on a neutral substitution rate of 2.0×10^{-8} substitutions per site per year for the control region which represents an average derived from a wide range of teleost species (Donaldson and Wilson, 1999; Sato et al., 2003). Although we have no means by which to rigorously date divergence time between the 'green' clade and 'blue' clades, there appears to be a broad concordance between

the observed genetic divergence and the apparent late Pliocene breach of the Purus Arch. Under this scenario, the colonization of the western Amazon basin would have occurred after the breach of the Purus Arch in the late Neogene. The hypothesized colonization of the newly available habitat in the western Amazon basin is based on the phylogenetically nested position of the 'green' clade, and also a signature of a demographic expansion in the 'green' clade as would be expected after a colonization event. A similar signal of demographic expansion is also observed in organism that are thought to have colonized western Amazon after the breach of the Purus Arch (e.g. Aleixo, 2006).

Processes that potentially resulted in the differentiation and maintenance of the other groups are less clear. There was a signature of significant restriction of gene flow between the lower Amazon 'brown' discus and the 'Heckel' discus of the Negro River. Large differences in the chemical properties of these two waters types exist (Sioli, 1984), and these could potentially have resulted in restricted gene flow due to local adaptation. Similarly, the signal of past fragmentation observed between the 'blue' clade and the 'brown' plus 'Heckel' group could also have resulted from differences in water chemistry. Although a tributary of the Amazon River, the Negro River is also the fifth largest river in the world (Goulding et al., 2003). The confluence results in large chemical differences of the Amazon upstream and downstream of the confluence. These and other differences are colloquially recognized through different names given to the Amazon River upstream (Solimões River) and downstream (Amazon River) of the confluence with the Negro River. The confluence corresponds to the three-way division of the 'Heckel', 'blue' and 'brown' discus occurring in the Negro, Solimões and Amazon Rivers, respectively. Further modification to the chemistry of the Amazon River is due to the discharge of the Madeira River which contributes approximately 50% of the total sediment load of the Amazon River short distance downstream of the junction of the Amazon with the Negro. The phenotypic and genotypic difference among the 'Heckel', 'blue' and 'brown' groups could be maintained by differences in water chemistry, although it is unclear if these difference would be causal factors in initiating the differentiation. If they were, they would give credence to ecological speciation hypothesis proposed by Endler (1977), in addition to selection acting to reinforce difference among groups.

The mechanisms responsible for the differentiation of the Xingu River clade are less clear, but may be a result of more general processes responsible for the high levels of ichthyofaunal endemism in the Xingu River (Zuanon, 1999). The lower Amazon basin would have been influenced by Plio–Pleistocene sea level fluctuations (Billuops, 2004; Liu and Herbert, 2004; Pillans et al., 1998) which would have repeatedly isolated and connected the tributary rivers of the lower Amazon, as well as radically changed their stream gradients. This could have resulted in isolation of the Xingu basin for a sufficient amount of time to allow allopatric differentiation. However, these sea level fluctuations clearly have not had the same effect on all lower Amazon localities. It should be noted that fishes from Cameté on the Araguaia–Tocantins Rivers which represent the eastern-most distributional limit of *Symphysodon*, and occur in a hydrologically isolated drainage, show no differentiation with respect to the remaining lower Amazon *Symphysodon*. Similarly, fishes from the Jarí River, the eastern-most northern tributary of the Amazon where *Symphysodon* is known to occur naturally also appear undifferentiated from other lower Amazon *Symphysodon*.

One of the more intriguing hypotheses for the diversification of Amazonian biota is the museum hypothesis. This hypothesis states that species originate allopatrically in refugia during marine highstands, and later accumulate by dispersal lowlands which accumulate species and act as 'museums' (Fjeldså, 1994; Nores, 1999). Hubert and Renno (2006) suggest this pattern for South American

characid fishes, noting high levels of endemism in the western Amazon basin and on the Guyana and Brazilian shield, and high species diversity in the central and eastern Amazon basins, an observation compatible with the 'museum' hypothesis. Applied to one species, a population genetic perspective of the 'museum' hypothesis would predict lower genetic diversity within, and differentiation between areas of refugia, and higher genetic diversity within the colonized areas, as well as potentially lack of differentiation of the colonized and refugial areas. This pattern is also expected in a scenario of peripatric speciation (Mayr, 1970), except that under the 'museum' hypothesis a pattern of population amalgamation should be observed in the colonized area which is rejected (via Chakraborty's test of population amalgamation). Our data rejects the 'museum' hypothesis, and seems to provide little support for peripatric speciation. Haplotypes found in refugial areas are phylogenetically nested within and therefore are derived from within the eastern Amazon region. *Symphysodon* of the western Amazon refugium show a signature of a demographic expansion, while expansion is not observed in the hypothetically colonized areas of the central and eastern Amazon basin. And finally, the 'Heckel' group and the 'Xingu' clade which occur in Guyana and Brazilian Shield drainages, respectively, occur in areas that were inundated by the Pliocene marine incursion (Haq et al., 1987; Nores, 1999).

4.1. Speciation and species status

Although there are phenotypic differences between the five subspecies of *Symphysodon* of Schultz (1960) and Burgess (1981), and the tree species recognized by Ready et al. (2006) and Bleher et al. (2007), there is evidence of mtDNA haplotype sharing between different phenotypic groups and species. According to Kullander (1996), *S. d. discus* (the 'Heckel' discus) is restricted to the Negro River basin and the Trombetas River, both of which are black-water rivers of the Guyana Shield. However, we have also found the 'Heckel' discus phenotypes in the Guyana Shield Rivers Nhamundá and Uatumã, and fishes with *S. discus* color pattern, a distinct black vertical bar in the center of its body, have been reported from the Jarí River (J. Junior, pers. com.), an eastern-most Guyana Shield tributary of the Amazon River (Fig. 1). Approximately 50% of the 'Heckel' discus from Trombetas and all 'Heckel' discus from the Nhamundá and Uatumã River drainages have haplotypes more typical of the 'brown' discus *S. a. aequifasciatus*. The 'brown' discus occurs sympatrically with the Nhamundá and Uatumã River 'Heckel' discus. Some of these sympatrically occurring 'brown' discus have haplotypes typical of other 'brown' discus haplotypes, while other 'brown' discus have haplotypes typical of 'Heckel' discus. The phenotypically distinct subspecies *S. d. willischwartzi* (the 'abacaxi' discus) described from the Abacaxis River, an affluent of the southern Amazon River tributary, the Madeira River (Fig. 1) is molecularly indistinguishable from the 'brown' discus.

The possibility of haplotype sharing as a result of genetic exchange is reinforced by the observation that different species and phenotypic groups of *Symphysodon* hybridize in the aquarium. Furthermore, isozyme studies demonstrate sharing of isozyme patterns among phenotypic groups and species, and further suggest there is as much genetic differentiation between different localities of *S. aequifasciatus* as there is between *S. aequifasciatus* and *S. discus* (Kokoscha and Greven, 1996). Sharing of mitochondrial cytochrome *b* and nuclear Rhodopsin is also reported for *S. discus* and *S. aequifasciatus* by Ready et al. (2006).

Several studies (Bleher et al., 2007; Ready et al., 2006) including our study identified the existence of a unique lineage present in the western Amazon basin which corresponds to *S. tarzoo*, *S. aequifasciatus* or *S. a. haraldi* according to different authors and the 'green'

clade in our study. Our study also identified the presence of a two additional unique lineages corresponding to the central Amazonian 'blue' clade and the 'Xingu' clade. The 'green' and 'blue' clades are sister, and are nested within the 'Heckel + abacaxi + brown' group. Depending on the location of the root, the 'Xingu' clade might be nested within this group as well, or alternately, the 'Xingu' clade or the 'Xingu' clade and two haplotypes from the Maues River might be sister to all other *Symphysodon* (Figs. 2 and 3).

These phylogenetic relationships clearly present a taxonomic problem. McKittrick's and Zink's (1988) reformulation of the phylogenetic species concept (PSC) states that a species: "is the smallest diagnosable cluster of individual organisms forming a monophyletic group within which there is a parental pattern of ancestry and descent". In other words, the PSC requires that species are defined by at least one apomorphic character (diagnostic character) with the implication that a species cannot be phylogenetically nested within another species (Baum and Donoghue, 1995). The 'green' group is defined by two molecular apomorphies, the 'blue' group by five molecular apomorphies and the 'Xingu' group by three molecular apomorphies. No molecular apomorphies define the 'brown' group or the 'Heckel' discus *S. discus* or its two subspecies *S. d. discus* and *S. d. willischwartzii* individually or in any combination. Additionally, while the 'green' and the 'blue' groups are monophyletic, they are nested within the remaining *Symphysodon* haplotypes as most likely is the 'Xingu' groups. The PSC does not reject the hypothesis that all species and phenotypic varieties of *Symphysodon* represent a single species, albeit a highly variable one with a complex evolutionary history. Neither does Mayr's (1942) biological species concept (BSC) which states that: "a species is a group of potentially or actually interbreeding populations that is reproductively isolated from other such groups" since all forms of *Symphysodon* interbreed under captive or semi-natural conditions, and apparent hybrid forms have been reported from nature. Other species concepts such as Templeton's, 1981 cohesion species concept (CSC) also do not reject the hypothesis that *Symphysodon* is just one species. Different morphs are genetically exchangeable, i.e. they hybridize when given the chance, and they are also ecologically exchangeable. Whether occurring in black-water, white-water or clear-water types, *Symphysodon* is found in lentic environments with minimal suspended sediment. The three water types are chemically distinct, however, large-scale variation on small spatial and temporal scale exist with each water type, and all three water types tend to be found within the geographic distribution of all groups. Both genetic and ecological exchangeability is further evidenced from numerous unintentional as well as intentional releases and colonization events such the unintentional release and establishment of *Symphysodon* originating from many different localities of the Brazilian Amazon in the Nanay River near Iquitos, Peru.

Adherence to the diagnosability criteria of the species concepts leads us to the conclusion that only one evolutionary entity (one species) is present in the genus *Symphysodon* based on the above species concepts. However, it is also clear that patterns of strong geographic structuring exist. This is clearly evident in the NCA inferences of past fragmentation or restricted gene flow with isolation by distance between the major groups of *Symphysodon*. Postzygotic isolating mechanisms may exist at least between some pairs of phenotypes. A recent study by Gross (2006) has shown that *S. discus* from Novo Airão differ in chromosomal complement and meiotic organization from *S. aequifasciatus* from Manacapuru Lake. These two sampling points represent the 'Heckel' and 'blue' phenotypes. However, a karyotypic analysis of *Symphysodon* from Jarí River and Tabatinga, the 'brown' and 'green' phenotypes, respectively, revealed no differences to each other or to the Manacapuru Lake fishes (J.C. Pieczarka, pers. com.). Thus although molecular data do not support the hypothesis of the existence of more than

one species, the data also reject the hypothesis of one homogeneous evolutionary entity. Based on our results, on field observations and the reading of the literature, we conclude that *Symphysodon* is in the process of diversification. Species are natural entities that embody the evolutionary process. Species go extinct, just as ancestral species give rise to descendent species. There should be no implicit expectation that we should not observe the speciation process. During the process of speciation, species concepts will fail for the simple reason that they are meant to identify species. During the speciation process we no longer are observing the ancestral species, nor are we observing the descendant species. Neither one species exists, nor multiple descendent species exist. Therefore we should accept and become comfortable with the fact that some organismal groups are inherently unclassifiable, and move on to the interesting questions of which processes drive the patterns of biological diversification in these groups.

Acknowledgments

This research was supported by grants from CNPq, PNOFG, FE-PIM, IFS and DCG. Permits for field collection and DNA accession was given by IBAMA/CGEN No. 045/DIFAP/2005 and 148/DIFAP/2006. K. Yamamoto, J. Alonso, J. Junior and J.R. Carvalho Jr. collected many additional valuable specimens. TH acknowledges FAPEAM and the J. William Fulbright Foundation for financial support.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ymp.2008.05.033.

References

- Aleixo, A., 2006. Historical diversification of floodplain forest specialist species in the Amazon: a case study with two species of the avian genus *Xiphorhynchus* (Aves: Dendrocolaptidae). *Biol. J. Linn. Soc.* 89, 383–395.
- Allgayer, R., 2001. Les Discus. Les Guides de l'Aquarium, Paris, France.
- Ayres, J.M., Clutton-Brock, T.H., 1992. River boundaries and species range size in Amazonian primates. *Am. Nat.* 140, 531–537.
- Baum, D.A., Donoghue, M.J., 1995. Choosing among alternative "phylogenetic" species concepts. *Syst. Bot.* 20, 560–573.
- Billops, K., 2004. Low-down on a rhythmic high. *Nature* 427, 686–687.
- Bleher, H., Stölting, K.N., Salzburger, W., Meyer, A., 2007. Revision of the genus *Symphysodon* Heckel, 1840 (Teleostei: Perciformes: Cichlidae) based on molecular and morphological characters. *Aqua* 12, 133–174.
- Burgess, W.E., 1981. Studies on the family Cichlidae: 10. New information on the species of the genus *Symphysodon* with the description of a new subspecies of *S. discus* Heckel. *Trop. Fish Hobby.* 29, 32–42.
- Bush, M.B., 1994. Amazonian speciation: a necessarily complex model. *J. Biogeogr.* 21, 5–17.
- Castelloe, J., Templeton, A.R., 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. *Mol. Phylogenet. Evol.* 3, 102–113.
- Chakraborty, R., Weiss, K.M., 1991. Genetic variation of the mitochondrial DNA genome in American Indians is at mutation-drift equilibrium. *Am. J. Hum. Genet.* 86, 497–506.
- Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9, 1657–1659.
- Cockerham, C.C., Weir, B.S., 1993. Estimation of gene flow from *F*-statistics. *Evolution* 47, 855–863.
- Colinvaux, P.A., 1993. Pleistocene biogeography and diversity in tropical forests of South America. In: Goldblatt, P. (Ed.), *Biological Relationships Between Africa and South America*. Yale University Press, New Haven, CT, pp. 473–499.
- Cordani, U.G., Sato, K., Teixeira, W., Tassinari, C.C.G., Basei, M.A.S., 2000. Crustal evolution of the South American platform. In: Cordani, U.G., Milani, E.J., Thomaz-Filho, A., Campos, D.d.A. (Eds.), *Tectonic Evolution of South America*. Academia Brasileira de Ciências, and Departamento Nacional da Produção Mineral (DNPM), Rio de Janeiro, RJ, Brazil, pp. 19–40.
- Costa, J.B.S., Léa Bemerguy, R., Hasui, Y., da Silva Borges, M., 2001. Tectonics and paleogeography along the Amazon river. *J. South Am. Earth Sci.* 14, 335–347.
- Crandall, K., Templeton, A.R., 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* 134, 959–969.
- Donaldson, K.A., Wilson Jr., R.R., 1999. Amphipanamic germinates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. *Mol. Phylogenet. Evol.* 13, 208–213.

- Endler, J.A., 1977. Geographic variation, speciation, and clines. Princeton University Press, Princeton, NJ.
- Ewens, W.J., 1972. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* 3, 87–112.
- Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform.* 1, 47–50.
- Farias, I.P., Orti, G., Meyer, A., 2000. Total evidence: molecules, morphology, and the phylogenetics of cichlid fishes. *J. Exp. Zool. (Mol. Dev. Evol.)* 288, 76–92.
- Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. *Cladistics* 10, 315–319.
- Fjeldså, J., 1994. Geographical patterns for relict and young species of birds in Africa and South America and implications for conservation priorities. *Biodivers. Conserv.* 3, 207–226.
- Fu, Y.-X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.
- Garrick, R.C., Dyer, R.J., Beheregaray, L.B., Sunnucks, P., 2008. Babies and bathwater: a comment on the premature obituary for nested clade phylogeographical analysis. *Mol. Ecol.* 17, 1401–1403.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7, 457–472.
- Goulding, M., Barthem, R.B., Ferreira, E.J.G., 2003. The Smithsonian Atlas of the Amazon. Smithsonian Institution Press, Washington, DC.
- Gregory-Wodzicki, K.M., 2000. Uplift history of the central and northern Andes: a review. *Geol. Soc. Am. Bull.* 112, 1091–1105.
- Gross, M.C., 2006. Comportamento cromossômico meiotico emitorico de acarisca (*Symphysodon aequifasciatus* e *Symphysodon discus*, Cichlidae, Perciformes) da Amazonia Central. *Genética, Conservação e Biologia Evolutiva*. Instituto Nacional de Pesquisas da Amazônia (INPA) and Universidade Federal do Amazonas (UFAM), Manaus, AM, Brazil, p. 82.
- Haffer, J., 1969. Speciation in Amazonian forest birds. *Science* 165, 131–166.
- Haffer, J., 1997. Alternative models of vertebrate speciation in Amazonia: an overview. *Biodivers. Conserv.* 6, 451–476.
- Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Haq, B.U., Hardenbol, J., Vail, P.R., 1987. Chronology of fluctuating sea levels since the Triassic. *Science* 235, 1156–1167.
- Harpending, H.C., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66, 591–600.
- Hartl, D.L., Clark, A.G., 2006. Principles of Population Genetics. Sinauer Associates, Sunderland, MA.
- Hasegawa, M., Kishino, H., Yano, T.A., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174.
- Hoorn, C., 1993. Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia: results of a palynostratigraphic study. *Palaeogeogr. Palaeoclim. Palaeoecol.* 105, 267–309.
- Hoorn, C., 1996. Miocene deposits in the Amazonian foreland basin. *Science* 273, 122–123.
- Hoorn, C., Guerrero, J., Sarmiento, G.A., Lorente, M.A., 1995. Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology* 23, 237–240.
- Hrbek, T., Seckinger, J., Meyer, A., 2007. A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. *Mol. Phylogenet. Evol.* 43, 986–998.
- Hubert, N., Renno, J.-F., 2006. Historical biogeography of South American freshwater fishes. *J. Biogeogr.* 33, 1414–1436.
- Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294, 2310–2314.
- Jobb, G., Haeseler, A.v., Strimmer, K., 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* 4, 1–9.
- Knowlton, L.L., Maddison, W.P., 2002. Statistical phylogeography. *Mol. Ecol.* 11, 2623–2635.
- Kokoscha, M., Greven, H., 1996. Gibt es nur eine Art? Isoenzym-Elektrophorese bei Diskusfischen. *D. Aqua. Terr. Z. Sonderheft Diskus* 1, 20–21.
- Kullander, S.O., 1996. Eine weitere Übersicht der Diskusfische, Gattung *Symphysodon* Heckel. *D. Aqua. Terr. Z. Sonderheft Diskus* 1, 10–16.
- Kullander, S.O., 2003. Family Cichlidae. In: Reis, R.E., Kullander, S.O., Ferraris, C.J. (Eds.), Check List of the Freshwater Fishes of South and Central America. EDIPUCRS, Porto Alegre, Brazil, pp. 605–654.
- Liu, Z., Herbert, T.D., 2004. High-latitude influence on the eastern equatorial Pacific climate in the early Pleistocene epoch. *Nature* 427, 720–723.
- Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C.S.L., Wesselingh, F.P., 1998. The stage for Neotropical fish diversification: a history of tropical South American rivers. In: Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M.S., Lucena, C.A.S. (Eds.), Phylogeny and Classification of Neotropical Fishes. EDIPUCRS, Porto Alegre, Brazil, pp. 13–48.
- Lyons, E., 1959. *Symphysodon discus* Tarzoo. New blue discus electrify aquarium world. *Trop. Mag.* 4, 6–8, 10.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220.
- Marshall, L.G., Lundberg, J.G., 1996. Technical comment: Miocene deposits in the Amazonian foreland basin. *Science* 273, 123–124.
- Mayr, E., 1942. Systematics and the Origin of Species. Columbia University Press, New York, NY.
- Mayr, E., 1970. Populations, Species, and Evolution. Belknap Press, Cambridge, MA.
- McKittrick, M.C., Zink, R.M., 1988. Species concepts in ornithology. *Condor* 90, 1–14.
- Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, NY.
- Nores, M., 1999. An alternative hypothesis for the origin of Amazonian bird diversity. *J. Biogeogr.* 26, 475–485.
- Panchal, M., Beaumont, M.A., 2007. The automation and evaluation of nested clade phylogeographic analysis. *Evolution* 61, 1466–1480.
- Pellegrin, J., 1904. Contribution à l'étude anatomique, biologique et taxinomique des Poissons de la famille des Cichlides. *Mém. Soc. Zool. France* 16, 41–399.
- Petit, R.J., 2008. The coup de grâce for the nested clade phylogeographic analysis? *Mol. Ecol.* 17, 516–518.
- Pillans, B., Chappell, J., Naish, T.R., 1998. A review of the Milankovitch climatic beat: template for Plio-Pleistocene sea-level changes and sequence stratigraphy. *Sediment. Geol.* 122, 5–21.
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Posada, D., Crandall, K.A., 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol. Evol.* 16, 37–45.
- Posada, D., Crandall, K.A., Templeton, A.R., 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol. Ecol.* 9, 487–488.
- Rand, D.M., 1996. Neutrality tests of molecular markers and the connections between DNA polymorphism, demography, and conservation biology. *Conserv. Biol.* 10, 665–671.
- Rannala, B., Yang, Z., 1996. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J. Mol. Evol.* 43, 304–311.
- Ready, J.S., Ferreira, E.J.G., Kullander, S.O., 2006. Discus fishes: mitochondrial DNA evidence for a phylogeographic barrier in the Amazonian genus *Symphysodon* (Teleostei: Cichlidae). *J. Fish Biol.* 69, 200–211.
- Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9, 552–569.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Springs Harbor Laboratory Press, Cold Springs Harbor, NY.
- Sato, A., Takezaki, N., Tichy, H., Figueroa, F., Mayer, W.E., Klein, J., 2003. Origin and speciation of haplochromine fishes in East African crater lakes investigated by the analysis of their mtDNA, Mhc genes, and SINES. *Mol. Biol. Evol.* 20, 1448–1462.
- Schneider, S., Excoffier, L., 1999. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152, 1079–1089.
- Schultz, L.P., 1960. A review of the pompadour or discus fishes, genus *Symphysodon* of South America. *Trop. Fish Hobby.* 8, 5–17.
- Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.
- Sioli, H., 1984. The Amazon and its main affluents: hydrography, morphology of the river courses and river types. In: Sioli, H. (Ed.), The Amazon. Limnology and Landscape Ecology of a Mighty Tropical River and its Basin. Springer-Verlag, New York, NY, pp. 127–165.
- Sivasundar, A., Bermingham, E., Orti, G., 2001. Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. *Mol. Ecol.* 10, 407–417.
- Suzuki, Y., Glazko, G.V., Nei, M., 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenies. *Proc. Natl. Acad. Sci. USA* 99, 16138–16143.
- Tajima, F., 1983. Evolutionary relationships of DNA sequences in finite populations. *Genetics* 105, 437–460.
- Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595.
- Templeton, A.R., 1981. Mechanisms of speciation—a population genetic approach. *Annu. Rev. Ecol. Syst.* 12, 23–48.
- Templeton, A.R., 1995. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping or DNA sequencing. V. Analysis of case/control sampling designs: Alzheimer's disease and the apoprotein E locus. *Genetics* 140, 403–409.
- Templeton, A.R., 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Mol. Ecol.* 7, 381–397.
- Templeton, A.R., 2004a. Statistical phylogeography: methods of evaluating and minimizing inference errors. *Mol. Ecol.* 13, 789–809.
- Templeton, A.R., 2004b. Using haplotype trees for phylogeographic and species inference in fish populations. *Environ. Biol. Fishes* 69, 7–20.
- Templeton, A.R., Boerwinkle, E., Sing, C.F., 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics* 117, 343–351.
- Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data: III. Cladogram estimation. *Genetics* 132, 619–633.
- Templeton, A.R., Sing, C.F., 1993. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping: IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* 134, 659–669.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1996. 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.
- Vega, A.M.L., 2007. Reconstituição paleoambiental dos depósitos Miocenos na região centro-oriental da Bacia do Solimões, Amazonas. *Geociencias. Universidade Federal do Amazonas (UFAM), Manaus, AM, Brazil*, p. 92.
- Wallace, A.R., 1852. On the monkeys of the Amazon. *Proc. Zool. Soc. Lond.* 20, 107–110.
- Watterson, G.A., 1975. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* 7, 256–276.
- Wesselingh, F.P., Räsänen, M.E., Irion, G., Vonhof, H.B., Kaandorp, R.J.G., Renema, W., Romero Pittman, L., Gingras, M., 2002. Lake Pebas: a palaeoecological reconstruction of a Miocene, long-lived lake complex in western Amazonia. *Cainozoic Res.* 1, 35–81.
- Wesselingh, F.P., Salo, J.A., 2006. A Miocene perspective on the evolution of the Amazonian biota. *Script. Geol.* 133, 439–458.
- Zar, J.H., 1999. *Biostatistical Analysis*, fourth ed. Prentice Hall, Upper Saddle River, NJ.
- Zhang, D.-X., Hewitt, G.M., 1996. Nuclear integrations: challenges for mitochondrial DNA markers. *Trends Ecol. Evol.* 11, 247–251.
- Zuanon, J.A.S., 1999. *História natural da ictiofauna de corredeiras do rio Xingú, na região de Altamira, Pará*. Biology. University of Campinas, Campinas, Brazil (p. 190).