PLOS ONE

RESEARCH ARTICLE

DNA Barcoding in Pencilfishes (Lebiasinidae: *Nannostomus*) Reveals Cryptic Diversity across the Brazilian Amazon

Denise Corrêa Benzaquem[1], Claudio Oliveira[2], Jaqueline da Silva Batista[3], Jansen Zuanon[4], Jorge Ivan Rebelo Porto[1]*

1 Laboratório de Genética Animal, Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo, 2.936, Petrópolis, CEP 69067-375, Manaus, AM, Brazil, **2** Instituto de Biociências, Departamento de Morfologia, UNESP, CEP 18618-970 Botucatu, SP, Brazil, **3** Laboratório de Fisiologia Comportamental e Evolução, Laboratório Temático de Biologia Molecular, Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo, 2.936, Petrópolis, CEP 69067-375, Manaus, AM, Brazil, **4** Laboratório de Sistemática e Ecologia de Peixes, Instituto Nacional de Pesquisas da Amazônia, Av. André Araújo, 2.936, Petrópolis, CEP 69067-375, Manaus, AM, Brazil

* jirporto@inpa.gov.br

Abstract

Nannostomus is comprised of 20 species. Popularly known as pencilfishes the vast majority of these species lives in the flooded forests of the Amazon basin and are popular in the ornamental trade. Among the lebiasinids, it is the only genus to have undergone more than one taxonomic revision. Even so, it still possesses poorly defined species. Here, we report the results of an application of DNA barcoding to the identification of pencilfishes and highlight the deeply divergent clades within four nominal species. We surveyed the sequence variation in the mtDNA cytochrome c oxidase subunit I gene among 110 individuals representing 14 nominal species that were collected from several rivers along the Amazon basin. The mean Kimura-2-parameter distances within species and genus were 2% and 19,0%, respectively. The deep lineage divergences detected in *N. digrammus, N. trifasciatus, N. unifasciatus* and *N. eques* suggest the existence of hidden diversity in *Nannostomus* species. For *N. digrammus* and *N. trifasciatus*, in particular, the estimated divergences in some lineages were so high that doubt about their conspecific status is raised.

Introduction

Neotropical ichthyofauna is extremely large and diverse based on the latest survey of the diversity of freshwater fishes of Central and South America [1], it includes 71 families encompassing 4,475 known valid species. Moreover, the authors estimate there are 6,000 species in the neotropics, close to half the number of freshwater fish worldwide.

The Amazon, home of more than two thousand freshwater fish species, is well known as a diversity hotspot. The complex evolutionary history of Amazonian organisms, including fishes, is gradually beginning to be better understood. Postulated reasons for the origin and evolution

of Neotropical fish richness have included tectonic movements and Andes orogenesis leading to the rise and fall of biogeographic barriers, rates of speciation, extinction, and taxa dispersal, and the great antiquity of the Amazonian ecosystem as a whole that dates to the early Cenozoic and Cretaceous periods [2–3].

The identification of Amazonian fish species is a challenging using only morphology as a tool, because fish taxonomists face difficulties daily in the process of identifying species. In this context, there are several debates and statements that molecular data should be encouraged as a supplement to species description and diagnosis, but should not replace morphological data [4–9]. Indeed, species discriminations or descriptions involving the cytochrome c oxidase subunit I (*COI*) appear to be a trend in fish taxonomy [10–16].

The Barcode of Life, referred to as DNA barcoding, has recently emerged as a way to identify species by minimal DNA sequences. While this approach offers all of the advantages of conventional PCR and DNA sequencing, it also connects bioinformatics with the taxonomic identifications in museum and biological repositories, implement a modern and integrated system to ensure that newly reported cryptic species will be described following their discovery [17–18]. This global identification system, based primarily on nucleotide sequences, has been applied to several organisms, and many published works have demonstrated the efficacy of this methodology for species identification [19–27].

By using DNA barcoding, the identification of fish at the species level ranges from 98–100% [18, 28–29] and the discriminatory power of identification has been 93% in freshwater fish and 98% in marine fish species [30].

This worldwide effort often reveals that species richness is still underestimated. As an indicator of cryptic speciation, at least two delimitations have been employed: the "10X rule" and the "cutoff value." In the first, barcoded specimens from the same nominal species diverge by 10 times or more from the mean intraspecific variability of the group. The "barcode gap", that is, the divergence among organisms belonging to the same species being much smaller than the divergence among different organisms, has been clearly demonstrated in different several. In the second, individuals with a genetic distance of 2% or more are considered much more likely to be congeneric than conspecific [30].

The fishes of the family Lebiasinidae constitute a Neotropical group with approximately 76 valid species, and they are widely distributed throughout South and Central America [31]. Phylogenetic relationships among the seven genera of the group were examined in 32 representatives of Lebiasinidae (A. L. Netto-Ferreira, unpublished- data). Belonging to this family, *Nannostomus* is comprised of colorful ornamental species distributed throughout different countries in the Amazon Basin such as Colombia, Venezuela, and the Guyanas in the north; Brazil and Bolivia in the south; Peru in the west; and Brazil in the east [32–34]. Most species are slender and, pencil-shaped, ranging in size from 1.5 cm to approximately 7 cm in length, and highly are prized in the aquarium industry [35]. The species typically inhabit the shorelines of small creeks, streams, pools, lakes, and flooded forests, rarely being found in the open primary channels of Amazonian rivers. In these locations, they are typically found in shallow water, living among aquatic plants, leaf litter, and floating meadows [36–37].

Despite its economical importance, little is known about biology or history of life of *Nannostomus* species. However, In Amazonian streams, they graze on algae on submerged trees and dead trunks, and feed on small invertebrates. It also provides food for animals at higher trophic levels developing an important role in the maintenance of the functioning and health of the whole aquatic system [38].

Popularly known as the pencilfish, *Nannostomus* has 20 taxonomically valid species [31]. It is only genus to have undergone two taxonomic revisions, but other poorly defined species complexes probably remain in need of revision, including *N. beckfordi* Günther, *N. eques*

Steindachner, *N. marginatus* Eigenmann, and *N. trifasciatus* Steindachner (A. L. Netto-Ferreira, unpublished data). Many of the species have conspicuous color variations depending on their geographic location. Indeed, some of these color variants have been described as separate species in the past [39–42].

Recently, molecular analyses were conducted in two *Nannostomus* species (*N. eques* and *N. unifasciatus*), both of which are widely distributed throughout the Amazon basin [43–44]. Strikingly, the divergence time analyses of the lineages of both species allowed the authors to estimate an approximate divergence time in the middle Pliocene for lineages in *N. eques* (around 2.9 Mya) and in the late Miocene for *N. unifasciatus* (around 8.4 Mya). The presence of distinct phylogroups in both species have been observed in representatives from the Negro River Basin, showing that due to hidden diversity, those species should not be treated as having a single stock for management purposes. Such data may be useful in regulating the exploitation of fish species by the aquarium trade [43–44].

Given the biological and economic importance of these species, and as part of a study of the molecular phylogeny of *Nannostomus*, we report herein an application of DNA barcoding to the identification of 12 species of pencilfishes and highlight the deeply divergent clades detected within several nominal species, including the presence of additional cryptic species.

Materials and Methods

Ethics statement

This survey was carried out in strict accordance with the recommendations by the National Council for Control of Animal Experimentation and the Federal Board of Veterinary Medicine. The protocol was approved by the Committee on the Ethics of Animal Use of the Instituto Nacional de Pesquisas da Amazônia (041/2012).

All specimens for this study were collected in accordance with Brazilian laws under a permanent scientific collection license issued in the name of Dr. Jorge I.R. Porto and approved by the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) through the System Authorization and Information on Biodiversity (SISBIO #11489-1).

Sampling data

We sampled 14 species of *Nannostomus* (*N. beckfordi*, *N. britiskii*, *N. digrammus*, *N. eques*, *N. espei*, *N. harrisoni*, *N. limatus*, *N. margnatus*, *N. marylinae*, *N. morthenhaleri*, *N. nitidus*, *N. rubrocaudatus*, *N. trifasciatus and N. unifasciatus*). After, each specimen was identified, it is anesthetized by immersion in Eugenol in water and preserved in 96% ethanol for the molecular studies. Upon identification, morphological vouchers were deposited in the Zoological Collection at the National Institute for Amazonian Research, INPA, Manaus, Brazil.

Sampling sites were chosen at distant points in the Amazon basin in order to cover at a minimum, the distribution range of the species [33] (Fig. 1). Two lebiasinids were used as outgroups (*Lebiasina colombia* and *Copella nigrofasciata*) because they are members of each of the two subfamilies in Lebiasinidae.

All specimens were identified with the help of taxonomists and identification keys [34,37] end all procedures complied with the recommendations of local ethics committees. Voucher specimens were deposited in the collection of the Instituto Nacional de Pesquisas da Amazônia, Manaus, AM, Brazil.
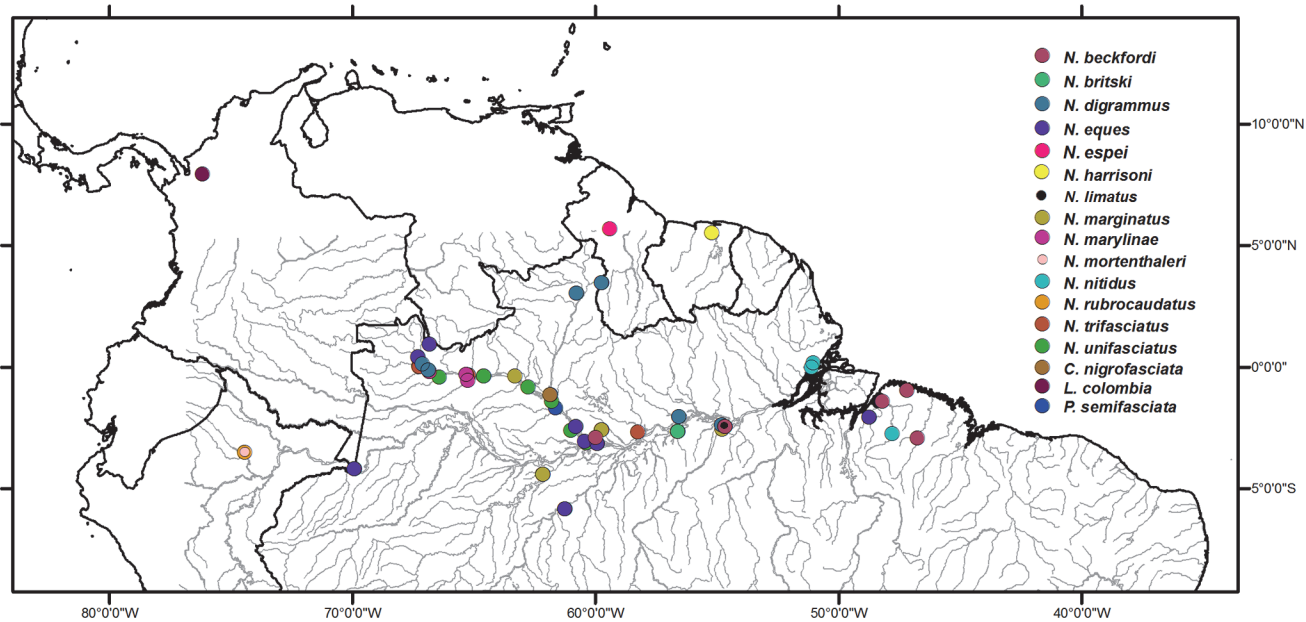
Fig 1. Map of the study area showing the sampling sites the pencilfishes collected in this study.

doi:10.1371/journal.pone.0112217.g001

Extraction, amplification, and DNA sequencing

Genomic DNA was isolated from muscle tissue based on standard techniques and protocols [45]. The partial mitochondrial *COI* gene (648pb) was amplified by PCR using a combination of primers FishF1, FishR1, FishF2 FishR2) [20] or cock-tail C_VF1LF_t1- C_VR1LR_t1 under conditions previously described [46].

Polymerase chain reactions were performed in a total volume of 15μL ($\sim$10–50 ng DNA template, 1X buffer (750 mM Tris-HCl, pH 8.8, 200 mM $(NH_4)_2SO_4$), 1U Taq polymerase (Thermo Scientific, Waltham, USA), 0.2 mM dNTPs, 0.2 μM of each primer, 2 mM $MgCl_2$, and ultrapure water. PCR cycling was performed with the initial denaturation for 2 min at 95ºC followed by 35 cycles of 30s at 95ºC, 30s at 52ºC-54ºC, 1 min at 72ºC and with a final extension for 10 min at 72ºC. PCR products were resolved on 1% agarose gels and purified using polyethylene glycol 8000 (USB, Cleveland, USA). The bi-directional sequencing was performed utilizing an ABI BigDye TM Terminator v.3.1 Cycle Sequencing Ready Reaction Kit and an ABI 3130xl DNA Analyzer (Applied Biosystems, Foster City, USA).

Data sequences, collection sites, primers details and trace files were submitted to the Barcode of Life database (BOLD; http//www.boldsystems.org) in under project "Barcoding of Lebiasinids."

Data analysis

Consensus sequences for the COI gene were generated using the BioEdit program [47], and after editing the sequences, the final matrix was 574bp.

All sequences were analyzed using MEGA 5 to check the occurrence of deletions, insertions, and stop codons. Search tools with local alignment were used to identify the sequence in GenBank and the BOLD. Sequences were aligned using Clustal W [48], and the program DnaSP version 4.0 [49] was used to determine the nucleotide composition, number of polymorphic sites, and haplotypes diversity.

The genetic distance among and within observed clusters was calculated using the Kimura-2-parameter (K2P) model. A Bayesian phylogenetic analysis was conducted using MrBayes 3.2 [50]. For this analysis, Markov chain Monte-Carlo sampling was conducted every 120,000 generations until the standard deviation of split frequencies was <0.01. A burn-in period equivalent to 25% of the total generations was necessary to recapitulate the parameter values and trees. The parameter values were evaluated based on 95% credibility levels to ensure a sufficient number of generations had been run for the analysis. A neighbor-joining (NJ) tree of K2P distances was created to provide a graphic representation of the relationships among specimens and clusters with MEGA 6.0 [51]. Bootstrap resampling [52] was applied to assess the support for individual nodes using 1000 pseudo-replicates. The program Haploviewer[53] was used to construct a tree-based haplotype network. Independent networks were regarded as unconfirmed candidate species.

Results

COI sequence data were obtained for 110 specimens of *Nannostomus* representing 14 species-level taxa. A mean of 8 individuals (range 1–17) represented each species, with only *N. harrissoni* represented by a single specimen.

The amplified product was approximately 650 bp but only 573 nucleotides were considered for this analysis, of which 234 were variable and parsimony informative. These variations defined 68 haplotypes ranging from 1–5 individuals per haplotype. At no time was a detected haplotype shared between different species. There were no deletions, insertions, or stop-codons. As expected for fish *COI*, the nucleotide composition showed a CT bias (means C = 24.0%, T = 32.7%, A = 25.9%, G = 17.3%) within *Nannostomus*.

The results indicated that species could be discriminated by the DNA barcode approach since the samples of distinct species were represented by a unique haplotype, a single tight cluster of haplotypes, or distinct clusters of haplotypes in neighbor-joining tree (Fig. 2).

The mean genetic distances (Kimura—2-parameter distances) among different *Nannostomus* species ranged from 2.2% to 22.5% (Table 1). The congeneric distance values found in *Nannostomus* were high, with an overall mean of 19. 0% ± 1. 3%, with values >22% between *N. britski* and its congeners and 2.2% between *N. limatus* and *N. nitidus* (Table 1).

The conspecific distance showed less than 2% divergence in 8 of the 14 species. Exceptionally deep sequence divergences were evident between individuals identified as the same morphospecies of *N. digrammus* (9.80% ± 1.0%), *N. trifasciatus* (8.1% ± 0.7%), *N. unifasciatus* (7.1% ± 0.7%), and *N. eques* (4.% ± 0.4%). We constructed a haplotype network based on Templeton's method (95% statistical parsimony) [53], and all four species displayed unconnected networks, which was consistent with the clusters identified through the NJ tree. Each of these lineages or phylogroups is presented in more detail below.

Nannostomus eques

Of the 16 specimens identified as *N. eques*, 13 unique haplotypes were identified. These haplotypes were assigned to five lineages (E1—E5). E1 (n = 6) was sampled in Rio Tapajós, Lower Rio Amazonas and Lower Rio Negro, E2 (n = 2) in Upper Rio Amazonas and Upper Rio Negro: E3 (n = 5) in Lower Rio Amazonas and Upper Rio Negro, E4 (n = 3) in Rio Guamá, and E5 (n = 1) in Rio Madeira (Fig. 3A). Low genetic variation was observed within the lineages (E1 = 0.55%, E3 = 0.26%, E4 = 0.52%), with the exception of E2 (2.29%) (Table 2). The mean genetic distance between the lineages was 4.3%. The smallest genetic distance encompassing different lineages was found between E2 and E3 (3.4% ±0.7%), while the largest was between E1 and E5 (7. 7% ± 1.1%).

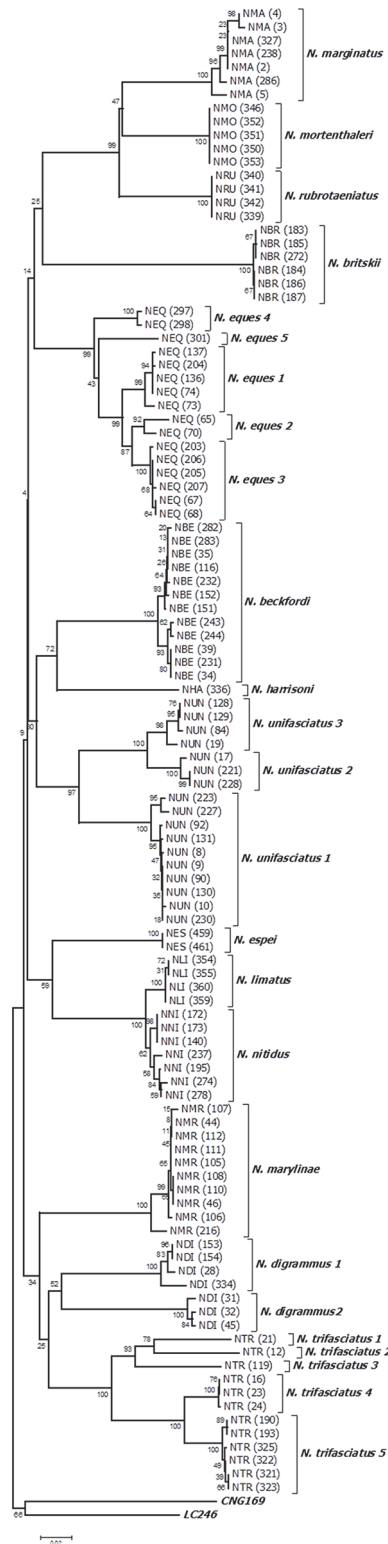**Fig 2. Neighbor-joining tree of 110 mitochondrial cytochrome oxidase subunit I gene sequences from 14 *Nannostomus* species and two outgroups using Kimura 2-parameter.** Deep conspecific divergences shown by 4 species (*N. digrammus*, *N. trifasciatus*, *N. unifasciatus* and *N. eques*) are identified as distinct lineages. Bootstrap values based on 1000 replicates are indicated at the branches.

doi:10.1371/journal.pone.0112217.g002

Table 1. Kimura 2-parameter genetic divergence values in 14 *Nannostomus* species.

	K2P divergence genetic (%)			
	Minimum	Mean	Maximum	SE
Conspecific	0	2,8	9.80	2,0
Congeneric	2,2	19,0	22,5	2,6

doi:10.1371/journal.pone.0112217.t001

Nannostomus unifasciatus

Of the 17 specimens identified as *N. unifasciatus*, 13 unique haplotypes were identified. These haplotypes were assigned to three lineages (U1—U3); U1 (n = 10) was sampled in Rio Purus, Rio Amazonas, and Lower and Middle Rio Negro. U2 (n = 3) in Lower and Middle Rio Negro and U3 (n = 4) in Middle and Upper Rio Negro (Fig. 3B). Moderate genetic variation was observed within the U1, U2 and U3 lineages (0.71%, 0.57% and 0.93%, respectively). The mean genetic distance between the lineages was 7.1%. The smallest genetic distance comprising different lineages was found between U2 and U3 (4.6% ± 0.8%), while the greatest was between U1 and U3 (12.5% ± 1.4%) (Table 2).

Nannostomus digrammus

Of the 7 specimens identified as *N. digrammus*, 6 unique haplotypes were identified. These haplotypes were assigned to lineages (D1 and D2). D1 (n = 3) was sampled in the Tapajós River, and Amazonas and Takutu River tributaries, while D2 (n = 3) was endemic to the Negro River (Fig. 3C). Moderate genetic variation was seen within the D1 and D2 lineages (1.50% and 0.80%, respectively) (Table 2). The mean genetic distance between the lineages was 16.2% ± 1.9.

Nannostomus trifasciatus

Of the 12 specimens identified as *N. trifasciatus*, 9 unique haplotypes were identified. These haplotypes were assigned to 5 lineages (T1—T5). T1 (n = 1) was sampled in Rio Preto da Eva, T2 (n = 1) in Upper Rio Negro, T3 (n = 1) in Upper Rio Negro, T4 (n = 3) in Middle Rio Negro, and T5 (n = 6) in tributaries of Rio Takutu, Uatumã, and Amazonas near Macapá City (Fig. 3D). Low genetic variation was observed within T4 and T5 (0.11% and 0.33%, respectively) (Table 2). The mean genetic distance between the lineages was 8.1%. The smallest genetic distance encompassing different lineages was found between T4 and T5 (5.0% ± 1.0%), while the largest was between T2 and T4 (16.2% ± 1.9%).

Discussion

Molecular methodologies developed rapidly in recent years, and currently DNA barcoding is considered the most useful tool for species identification. Indeed, a great advantage offered by DNA barcoding is the possibility of identifying cryptic species, as can be seen in several publications since its launch in 2003 [54].

According to the Fish Barcode of Life project database (www.fishbol.org), only one *Nannostomus* species had been previously barcoded. In our data set, 14 *Nannostomus* species were DNA barcoded, and they were easily identified by this approach, given that all recognized species formed monophyletic clusters (Fig. 2). However, two species (*N. nitidus* and *N. limatus*) revealed shallow interspecific sequence divergence (2.2%) when compared to other
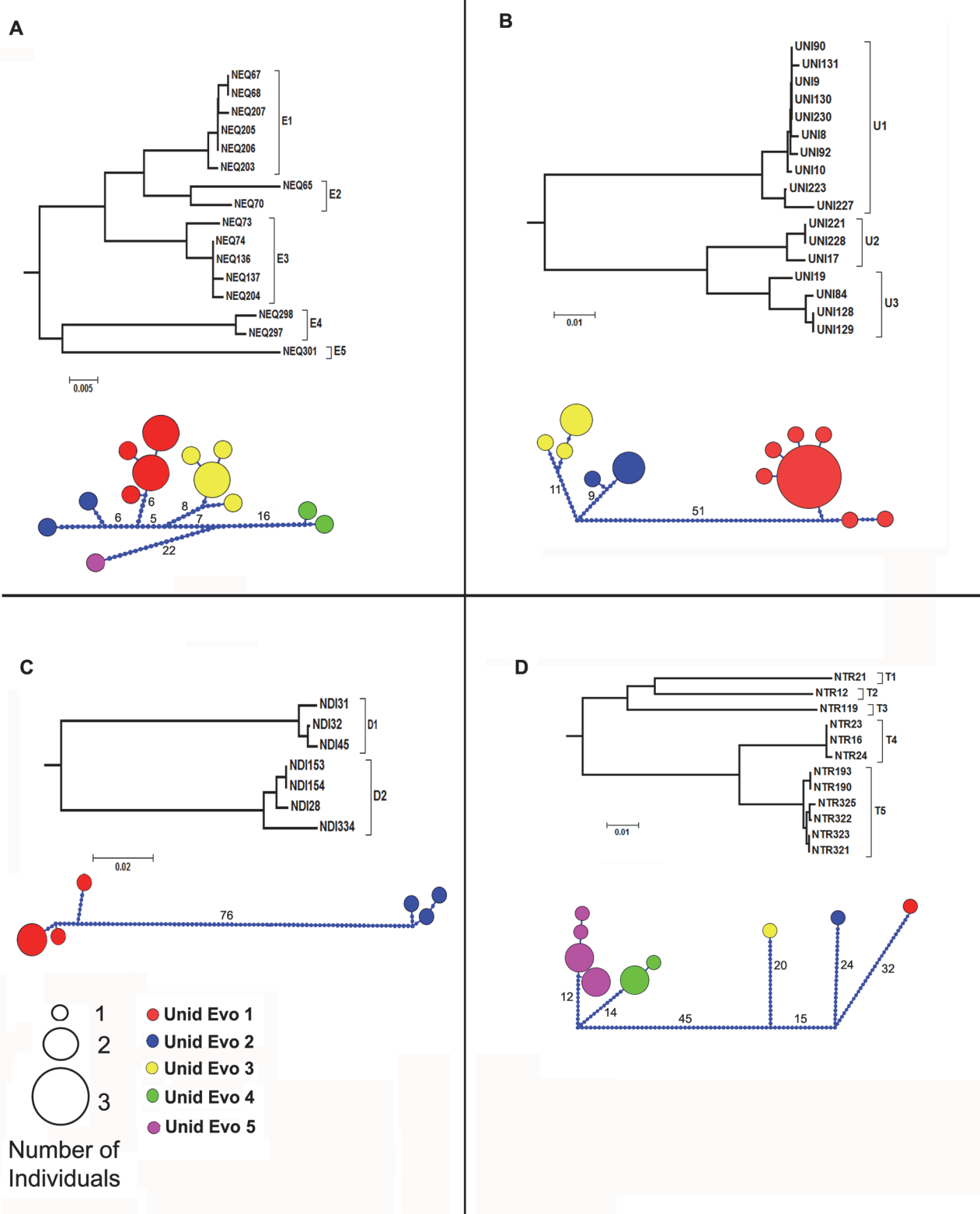
Fig 3. Neighbor-Joining tree and haplotype network of *Nannostomus* species showing distinct lineages: (A) *N. eques*, (B) *N. unifasciatus*, (C) *N. digrammus* and (D) *N. trifasciatus.* In the network each circle represents one haplotype, and theize of the circles is proportional to the haplotype frequency. Color codes represent the lineages found in *Nannostomus*. Numbers on the lines represent the mutational steps between haplotypes.

Table 2. Estimates of pairwise genetic distance between *Nannostomus* species under the Kimura 2- parameter (K2P) model.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. *N. beckfordi*	**0,911**	0,026	0,018	0,021	0,017	0,018	0,019	0,017	0,019	0,017	0,021	0,019	0,020	0,019	0,021	0,023	0,022	0,020	0,021	0,017	0,019	0,019	0,020
2. *N. britskii*	0,266	**0,103**	0,025	0,025	0,022	0,022	0,023	0,022	0,023	0,025	0,023	0,024	0,025	0,024	0,026	0,027	0,026	0,025	0,025	0,022	0,023	0,023	0,025
3. *N. digrammus1*	0,178	0,263	**1,501**	0,018	0,019	0,019	0,019	0,019	0,019	0,022	0,019	0,019	0,018	0,018	0,021	0,021	0,022	0,020	0,021	0,018	0,020	0,019	0,020
4. *N. digrammus2*	0,211	0,262	0,162	**0,813**	0,022	0,020	0,020	0,021	0,020	0,021	0,022	0,023	0,020	0,022	0,022	0,021	0,021	0,020	0,021	0,018	0,020	0,020	0,023
5. *N. eques1*	0,170	0,216	0,184	0,216	**0,557**	0,009	0,008	0,010	0,011	0,018	0,022	0,021	0,019	0,018	0,021	0,022	0,021	0,022	0,022	0,018	0,019	0,019	0,020
6. *N. eques2*	0,177	0,227	0,188	0,196	0,048	**2,293**	0,007	0,011	0,011	0,017	0,021	0,021	0,018	0,016	0,020	0,022	0,021	0,021	0,021	0,017	0,019	0,019	0,019
7. *N. eques3*	0,186	0,226	0,191	0,191	0,041	0,034	**0,265**	0,011	0,011	0,018	0,022	0,021	0,019	0,017	0,021	0,022	0,021	0,022	0,023	0,018	0,019	0,018	0,020
8. *N. eques4*	0,159	0,214	0,186	0,207	0,066	0,076	0,072	**0,521**	0,011	0,018	0,021	0,021	0,019	0,018	0,022	0,023	0,021	0,020	0,021	0,018	0,018	0,018	0,018
9. *N. eques5*	0,171	0,232	0,174	0,182	0,077	0,076	0,074	0,072	**0,000**	0,019	0,021	0,021	0,019	0,018	0,020	0,022	0,021	0,021	0,021	0,017	0,020	0,019	0,020
10. *N. harrisoni*	0,153	0,266	0,225	0,217	0,177	0,179	0,182	0,182	0,190	**0,000**	0,022	0,022	0,021	0,020	0,021	0,021	0,020	0,021	0,021	0,017	0,021	0,021	0,020
11. *N. marginatus*	0,214	0,241	0,189	0,215	0,217	0,214	0,214	0,206	0,202	0,224	**1,252**	0,016	0,021	0,022	0,022	0,023	0,024	0,020	0,020	0,022	0,022	0,022	0,015
12. *N. mortenthaleri*	0,186	0,258	0,188	0,228	0,196	0,201	0,201	0,197	0,199	0,218	0,126	**0,000**	0,022	0,021	0,025	0,023	0,024	0,023	0,023	0,023	0,021	0,021	0,016
13. *N. marylinae*	0,188	0,260	0,170	0,178	0,182	0,174	0,182	0,188	0,178	0,207	0,224	0,215	**0,642**	0,019	0,021	0,021	0,022	0,021	0,023	0,021	0,020	0,020	0,022
14. *N. nitidus*	0,183	0,245	0,178	0,219	0,173	0,147	0,158	0,172	0,184	0,195	0,229	0,208	0,190	**0,881**	0,024	0,024	0,023	0,022	0,023	0,020	0,021	0,020	0,022
15. *N. trifasciatus1*	0,221	0,283	0,210	0,216	0,221	0,218	0,218	0,229	0,214	0,211	0,239	0,269	0,216	0,245	**0,000**	0,015	0,016	0,017	0,018	0,022	0,023	0,023	0,024
16. *N. trifasciatus2*	0,240	0,293*	0,215	0,216	0,226	0,229	0,232	0,233	0,228	0,213	0,248	0,239	0,205	0,253	0,106	**0,000**	0,014	0,019	0,017	0,022	0,022	0,021	0,025
17. *N. trifasciatus3*	0,227	0,283	0,210	0,199	0,202	0,208	0,208	0,207	0,214	0,207	0,257	0,253	0,223	0,229	0,128	0,115	**0,000**	0,018	0,018	0,022	0,022	0,023	0,024
18. *N. trifasciatus4*	0,195	0,289	0,199	0,208	0,219	0,214	0,231	0,199	0,214	0,211	0,209	0,242	0,216	0,240	0,152	0,162	0,149	**0,115**	0,010	0,023	0,024	0,023	0,023
19. *N. trifasciatus5*	0,209	0,286	0,204	0,212	0,231	0,228	0,240	0,213	0,238	0,211	0,217	0,240	0,234	0,238	0,153	0,140	0,148	0,050	**0,335**	0,021	0,023	0,022	0,022
20. *N. unifasciatus1*	0,162	0,237	0,192	0,183	0,178	0,168	0,173	0,160	0,166	0,162	0,226	0,226	0,210	0,187	0,224	0,224	0,230	0,233	0,226	**0,715**	0,015	0,015	0,021
21. *N. unifasciatus2*	0,188	0,245	0,213	0,205	0,187	0,184	0,188	0,185	0,195	0,211	0,237	0,221	0,212	0,199	0,234	0,229	0,234	0,252	0,245	0,124	**0,579**	0,008	0,023
22. *N. unifasciatus3*	0,183	0,239	0,199	0,205	0,184	0,180	0,176	0,178	0,189	0,203	0,222	0,211	0,203	0,192	0,233	0,226	0,247	0,244	0,236	0,125	0,046	**0,931**	0,021
23. *N. rubrotaeniatus*	0,199	0,255	0,214	0,230	0,196	0,196	0,191	0,165	0,190	0,192	0,130	0,123**	0,225	0,225	0,249	0,265	0,245	0,239	0,235	0,206	0,232	0,214	**0,000**

Pairwise congeneric divergence is denoted by the number of base substitutions per site between species (below diagonal) with their corresponding standard error (above diagonal). Complete deletion of all codon positions (1st, 2nd, 3rd and noncoding) was employed in this analysis. Mean conspecific divergences (%) are listed on the diagonal in bold italics.

*Large congeneric K2P distance;

**Small congeneric K2P distance.

Nannostomus species ([Table 1](#)). Despite this, no evidence of shared sequences among both species were observed, suggesting either recent speciation or the need of synonymization.

In contrary, four species (*N. digrammus*, *N. trifasciatus*, *N. unifasciatus*, and *N. eques*) showed deep intraspecific sequence divergence, suggesting the existence of overlooked species within *Nannostomus*, although some lineages were represented by a single individual ([Table 2](#)).

The *COI* delineations of the T1, T2, T3 and E5 *Nannostomus* lineages were achieved with only one individual. We are aware that hypothetical intraspecific lineages can be hindered by inadequate sample size in large geographic areas such as the Amazon Basin. In flathead fishes, for example, the limited number of specimens in a particular lineage and the sparse geographical spread of the samples for some of the proposed lineages restricted the ability to evaluate the extent of genetic diversity across several groups [55]. Thus, we suggest future surveys of these *Nannostomus* lineages to confirm if they do, in fact, represent evolutionary units.

All the lineages found in each of the above mentioned species were well supported by bootstrap values (>80) in the NJ tree, independent of the mutational steps necessary to connect the haplotype networks below the standard statistical parsimony (Figs. [2](#) and [3](#)). The two lineages observed in *N. digrammus* (D1 and D2) diverged by 12.6%. The five lineages of *N. trifasciatus* had a mean divergence of 8.1%, while that of the three lineages of *N. unifasciatus* was of 7.1%. Finally, the five lineages of *N. eques* diverged by 4.3% (mean). The divergence between the lineages of *Nannostomus* was greater than obtained for other marine and freshwater fish species [20, 56–61].

It was evident in this study that the cutoff value of 2% does not apply to *Nannostomus* species, given the mean congeneric (~19%) and conspecific (~3%) distances were high. Indeed, surveys on North American and Neotropical freshwater ichthyofauna have shown that the mean congeneric and conspecific genetic distances are usually > 6.8% and <0.73%, respectively [62–63], with the exception of the Pampa Plain freshwater fishes at, the southernmost distribution range of many Neotropical species, where the mean congeneric genetic distance is 1.67% [64].

Among lebiasinids, only *Nannostomus* has twice been taxonomically revised, in addition to the phylogenetic hypotheses [32–34,36]. Morphology-based taxonomy has shown that *Nannostomus* cannot be considered to have a phenotypic conservatism. Apparently, there are poorly defined *Nannostomus* species complexes based on morphological grounds, including *N. beckfordi*, *N. eques*, *N marginatus*, and *N. trifasciatus* (A. L. Netto-Ferreira, unpublished data). Recent mitochondrial and nuclear DNA data have revealed hypothetical evolutionarily significant units in species of *Nannostomus*. For example, in *N. unifasciatus*, the DNA sequence data of the intron in the S7 ribosomal protein gene revealed two distinct lineages in the Rio Negro basin [43]. In *N. eques*, the mitochondrial DNA control region also revealed the existence of two lineages in the Rio Negro basin [44]. Together, the morphological and genetic studies indicate that species richness in the genus is probably underestimated.

The Amazon aquatic ecosystem is rich in diversity and quick access to information on Amazonian biodiversity is essential. Forest degradation and, fishery over-exploitation enhance the risks of species extinction, and quickly updated information regarding the fishes caught in wild fisheries (such as the *Nannostomus* species) is necessary to implement appropriate practices for to their conservation and, management and to prevent exploitation.

Pencilfishes are commercial ornamental fish and constitute a source of revenue for the riverine people of the Amazon. The discovery of cryptic species becomes very important when they are targets for commercial use. Considering the economic importance of the *Nannostomus* species, its DNA barcoding contributes to conservation policy in two important ways: by enhancing Amazonian biodiversity assessments to prioritize conservation areas (e.g., upper Rio

Negro), and by providing information about evolutionary histories and phylogenetic diversity (e.g., unveiling hidden diversity).

DNA barcoding of ornamental marine fishes has generated data and provided confidence in species identification, opening new avenues for managing business practices [65]. Of the approximately ∼500 Brazilian ornamental fishes allowed to be sold in the ornamental fish trade, 7 are pencilfishes: *N. beckfordi*, *N. digrammus*, *N. eques*, *N. espei*, *N. marginatus*, *N. trifasciatus*, and *N. unifasciatus*. Our study contributes new Amazonian fish barcodes, providing for a more comprehensive species identification of the ornamental pencilfishes, in addition to revealing the hidden diversity in the analyzed *Nannostomus* species. It is certain that future species delimitation of *Nannostomus* should be in accordance with the spirit of integrative taxonomy as well and based on congruence across analyses that utilize multiple data sources.

Acknowledgments

Author Contributions

Conceived and designed the experiments: DCB JIRP. Performed the experiments: DCB. Analyzed the data: DCB CO JSB JZ JIRP. Contributed reagents/materials/analysis tools: DCB CO JSB JZ JIRP. Wrote the paper: DCB JIRP CO.

References

1. Reis RE, Kullander SO, Ferraris CJ Jr (2003) Check list of the freshwater fishes of South and Central America. Porto Alegre, Edipucrs 729pp. PMID: 25057689

2. Lundberg JG, Marshall LG, Guerrero J, Horton B, Malabarba MCSL, et al. (1998) The stage for Neotropical fish diversification: a history of tropical South American rivers. In: Malabarba LR, Reis RE, Vari RP, Lucena ZMS, Lucena CAS. Phylogeny and Classification of Neotropical Fishes. Porto Alegre, Edipucrs 13–48 pp.

3. Albert JS, Petry, Reis RE (2011) Major Biogeographic and Phylogenetic Patterns.In. Albert JS, Reis RE editors. Historical Biogeography of Neotropical Freshwater Fishes Historical Biogeography of Neotropical Freshwater Fishes. Berkeley CA: University of California Press, 388p

4. Moritz C, Cicero C (2004) DNA Barcoding: promises and pitfalls. Public Library of Sci Biol 2: e354

5. Desalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. Philos T Roy Soc B 360: 1905–1916. PMID: 16214748

6. Will KW, Mishler BD, Wheeler QD (2005) The Perils of DNA Barcoding and the Need for Integrative Taxonomy. Syst Biol 5: 844–851. PMID: 16243769

7. Hickerson MJ, Meyer C, Moritz C (2006) DNA-Barcoding will often fail to Discover New Animal Species over Broad Parameter Space. Syst Biol 55: 729–739. PMID: 17060195

8. Padial JM, Miralles M, De la Riva I, Vences M (2010) The integrative future of taxonomy. Front Zool 2010: 7–16.

9. Johnson NF (2011) A collaborative, integrated and electronic future for taxonomy. Invertebr Syst 25: 471–475.

10. Benine RC, Mariguela TC, Oliveira C (2009) New species of *Moenkhausia* Eigenmann, 1903 (Characiformes: Characidae) with comments on the *Moenkhausia* oligolepisspecies complex. Neotrop Ichthyol 7: 161–168.

11. Victor CB, Randall JE (2010) *Gramma dejongi*, a New Basslet (Perciformes: Grammatidae) from Cuba, a Sympatric Sibling Species of *G. loreto*. Zool Stud 6: 865–871.

12. Melo BF, Benine RC, Mariguela TC, Oliveira C (2011) A new species of *Tetragonopterus* Cuvier, 1816 (Characiformes, Characidae: Tetragonopterinae) from the Rio Jari, Amapá, northern Brazil. Neotrop Ichthyol 9: 49–56.

13. Roxo FF, Oliveira C, Zawdzki CH (2012) Three new species of *Neoplecostomus* (Teleostei: Siluriformes: Loricariidae) from the Upper Rio Paraná basin of southeastern Brazil Zootaxa. 3233: 1–21.

14. Allen GR, White WT, Erdmann MV (2013) Two new species of snappers (Pisces: Lutjanidae: Lutjanus) from the Indo-West Pacific. Journal of the Ocean Science Foundation. 6: 33–51.

15. Borsa P, Durand JD, Shen KN, Arlyza IS, Solihin DD, et al. (2013) *Himantura tutul* sp. nov. (Myliobatoidei: Dasyatidae), a new ocellated whipray from the tropical Indo-West Pacific, described from its cytochrome-oxidase I gene sequence. Comptes Rendus Biologies 336: 82–92. doi: 10.1016/j.crvi.2013.01.004 PMID: 23608177

16. Silva GSC, Melo BF, Oliveira C, Benine RC (2013) Morphological and molecular evidence for two new species of *Tetragonopterus* (Characiformes: Characidae) from central Brazil. J Fish Biol 5: 1613–1631. doi: 10.1111/jfb.12094 PMID: 23639157

17. Hebert PDN, Cywinska A, Ball SL, de Waard JR (2003) Biological identification through DNA barcodes. Proc Roy Soc London, Series B: Biol Sci 270: 313–321.

18. Hajibabaei M, Smith MA, Janzen DH, Rodriguez JJ, Whitfield JB, et al. (2006) A minimalist barcode can identify a specimen whose DNA is degraded. Mol Ecol Notes 6: 959–964.

19. Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci 14: 812–14.

20. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philos T Roy Soc B 360: 1847–1857. PMID: 16214743

21. Barret RDH, Hebert PDN (2005) Identifying spiders through DNA barcodes. Can. J Zool 83: 481–491.

22. Kelly RP, Sarkar IN, Eernisse DJ, DeSalle R (2007) DNA barcoding using chitons (genus *Mopalia*). Mol Ecol Notes 7: 177–183

23. Borisenko AV, Lin BK, Ivanova NV, Hanner RH, Hebert PDN (2008) DNA barcoding in surveys of small mammal communities: a field study in Suriname. Mol Ecol Res 8: 471–479. doi: 10.1111/j.1471-8286.2007.01998.x PMID: 21585824

24. Gonzalez MA, Baraloto C, Engel J, Mori SA, Pétronelli P, et al. (2009) Identification of Amazonian trees with DNA barcodes. Plos One 4:e7483 doi: 10.1371/journal.pone.0007483 PMID: 19834612

25. Valdez-Moreno M, Ivanova NV, Elías-Gutierrez M, Contreras-Balderas S, Hebert PDN (2009) Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes. Jourl of Fish Biol 74: 377–402. doi: 10.1111/j.1095-8649.2008.02077.x PMID: 20735566

26. Clare EL, Lim BK, Fenton M, Hebert PDN (2011) Neotropical Bats: Estimating Species Diversity with DNA Barcodes. PLoS One e22648.

27. Renaud AK, Savage J, Adamowicz SJ (2012) DNA barcoding of Northern Nearctic Muscidae (Diptera) reveals high correspondence between morphological and molecular species limits. BMC Ecology 12: 24. doi: 10.1186/1472-6785-12-24 PMID: 23173946

28. Ferri E, Barbuto M, Bain O, Galimberti A, Uni S, et al. (2009) Integrated taxonomy: Traditional approach and DNA barcoding for the identification of filarioid worms and related parasites (Nematoda). Front Zool 6: 1. doi: 10.1186/1742-9994-6-1 PMID: 19128479

29. Galimberti A, Romano DF, Genchi M, Paoloni D, Vercillo F, et al. (2012) Integrative taxonomy at work: DNA barcoding of taeniids harboured by wild and domestic cats. Mol Ecol Res, 13: 403–13.

30. Ward RD, Hanner R, Hebert PDN (2009) The campaign to DNA barcode all fishes, FISH-BOL. J Fish Biol 74: 329–356. doi: 10.1111/j.1095-8649.2008.02080.x PMID: 20735564

31. Froese R, Pauly D, eds. (2014) FishBase: World Wide Web electronic publication. www.fishbase.org, version (accessed 25 October 2014).

32. Weitzman SH (1966) Review of South American characid fishes of the subtribe *Nannostomina*. Proceedings of the United States National Museum, 119: 1–56.

33. Weitzman SH; Cobb JS (1975) A revision of the South American fishes of the genus *Nannostomus* Günther (Family Lebiasinidae). Smithsonian Contribution Zoology, 186: 1–36.

34. Weitzman SH; Weitzman M (2003) Family Lebiasinadae. *In*: Reis R. E.; Kullander S. O.; Ferraris C. J. Jr (Eds). Check list of the freshwater fishes of south and central America. Porto Alegre, Edipucrs101–103pp.

35. Gery J (1877) Characoids of the Worl. Ed TFH Publications. pp 672.

36. Weitzman H, Weitzman M (1982) Biogeography and Evolutionary Diversification in Neotropical Fresh-waters Fishes, with Comments on the Refuge Theory. *In*: Prance G.T. (ed). Biological Diversification in the Tropics. Columbia University Press, New York 404–422 pp.

37. Weitzman SH (1978) Three new species of fishes of the genus *Nannostomus* from the Brazilian states of Pará and Amazonas (Teleostei: Lebiasinidae). Smithsonian Contribution Zoology 263: 1–14.

38. Carvalho LN, Zuanon J, Sazima I (2007) Natural history of amazon fishes. In: Claro PSO and Victor Rico-Gray edd. Tropical biology and conservation management. 1st ed. Oxford: Eolss Publishers; 1–32. doi: 10.1007/s12033-007-0062-9 PMID: 23934597

39. Paepke HJ, Arendt K (2001) *Nannostomus marginatus mortenthaleri* new subspecie from Peru (Tele-ostei: Lebiasinidae). Verh Ges Ichthyol Band 2: 143–154

40. Zarske A (2009) *Nannostomus rubrocaudatus* sp. n.—ein neuer Ziersalmler aus Peru (Teleostei: Char-aciformes: Lebiasinidae). Vertebrate Zoology 59: 11–23.

41. Zarske A (2011) *Nannostomus grandis* spec. nov.—ein neuer Ziersalmler aus Brasilien mit Bemerkungen zu *N. beckfordi* Günther, 1872, *N. anomalus* Steindachner, 1876 und *N. aripirangensis* Meinken, 1931 (Teleostei: Characiformes: Lebiasinidae). Vertebrate Zoology 3: 283–298.

42. Zarske A (2013) *Nannostomus nigrotaeniatus* spec. nov.—ein neuer Ziersalmler aus Venezuela (Tele-ostei: Characiformes: Lebiasinidae). Vertebrate Zoology 2:125–137.

43. Sistrom MJ, Chao NL, Beheregaray LB (2009) Population history of the Amazonian one-lined pencilfish based on intron DNA data. J Zool 4: 287–298.

44. Terencio ML, Schneider CH, Porto JIR (2012) Molecular signature of the D-loop in the brown pencilfish *Nannostomus eques* (Characiformes, Lebiasinidae) reveals at least two evolutionary units in the Rio Negro basin, Brazil. J Fish Biol 81: 110–124. doi: 10.1111/j.1095-8649.2012.03320.x PMID: 22747807

45. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laborato-ry Press. Cold Spring Harbor, New York PMID: 25506954

46. Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN (2007) Universal primer cocktails for fish DNA bar-coding. Mol Ecol Notes 7: 544–548.

47. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/96/NT. Nucleic Acids Symposium Series 41: 95–98.

48. Thompson JD, Higging DG, Gibson TJ (1994) ClustalW: Improving the sensitivity of progressive multi-ple sequence alignment through sequence weighting, position specific gap penalties and matrix choice. Nucl Acids Res 22: 4673–4680. PMID: 7984417

49. Rozas J, Sánchez-Del, Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497. PMID: 14668244

50. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574. PMID: 12912839

51. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genet-ics Analysis version 6.0. Mol Biol Evol 30: 2725–2729. doi: 10.1093/molbev/mst197 PMID: 24132122

52. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–79.

53. Salzburger W, Ewing G, Haeseler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Mol*. *Ecol*. **20**, 1952–1963.

54. Joly S, Davies TJ, Archambault A, Bruneau A, Derry A, et al.(2013) Ecology in the age of DNA barcod-ing: the resource, the promise and the challenges ahead. Mol Ecol Res 14(2): 221–232.

55. Puckridge M, Andreakis N, Appleyard SA, Ward R (2013) Cryptic diversity in flathead fishes (Scorpaeniformes: Platycephalidae) across the Indo-West Pacific uncovered by DNA barcoding. Mol Ecol Res 13: 32–42.

56. Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, et al. (2008) Identifying Canadian freshwater fish-es through DNA barcodes. PLoS One 3: e2490. doi: 10.1371/journal.pone.0002490 PMID: 22423312

57. Lara A, Ponce de Leon JL, Rodriguez R, Casane D, Cote G, et al. (2010) DNA barcoding of Cuban freshwater fishes: Evidence for cryptic species and taxonomic conflicts. Mol Ecol Res 10: 421–430. doi: 10.1111/j.1755-0998.2009.02785.x PMID: 21565041

58. Mabraganã E, Astarloa JMD, Hanner R, Zhang J, Castro MG (2011) DNA Barcoding Identifies Argen-tine Fishes from Marine and Brackish Waters. Plos One 12: e28655. doi: 10.1371/journal.pone.0028655 PMID: 22174860

59. Nwani CD, Becker S, Braid H, Ude EF, Okogwu OI, et al. (2011) DNA barcoding discriminates freshwa-ter fishes from southeastern Nigeria and provides river system level phylogeographic resolution within some species. Mitochondrial DNA (Supp. 1:): 43–51.

60. Wong LL, Peatman E, Lu J, Kucuktas H, He S, et al. (2011) DNA barcoding of catfish: Species authentication and phylogenetic assessment. PLoS One, 6: e17812. doi: 10.1371/journal.pone.0017812 PMID: 21423623

61. Pereira LHG, Hanner R, Foresti F, Oliveira C (2013) Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics 14(20): doi: 10.1186/1471-2156-14-20 PMID: 23497346

62. April J, Mayden RL, Hanner RH, Bernatchez L (2011) Genetic calibration of species diversity among North America's freshwater fishes. PNAS 108:10602–10607. doi: 10.1073/pnas.1016437108 PMID: 21670289

63. Carvalho DC, Oliveira DA, Pompeu PS, Leal CG, Oliveira C, et al. (2011) Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. Mitochondrial DNA 22: 80–86. doi: 10.3109/19401736.2011.624602 PMID: 21985406

64. Rosso JJ, Mabragaña E, Castro MG, Díaz de Astarloa M (2012) DNA barcoding Neotropical fishes: recent advances from the Pampa Plain, Argentina. Mol Ecol Res 12: 999–1011. doi: 10.1111/1755-0998.12010 PMID: 22984883

65. Steinke D, Zemlak TS, Hebert PDN (2009) Barcoding Nemo: DNA-Based Identifications for the Ornamental Fish Trade. Plos One 7: e6300. doi: 10.1371/journal.pone.0006300 PMID: 19621079