

# The complete mitochondrial genome of *Corydoras nattereri* (Callichthyidae: Corydoradinae)

Daniel A. Moreira<sup>1</sup>, Paulo A. Backup<sup>2</sup>, Marcelo R. Britto<sup>2</sup>, Maithê G. P. Magalhães<sup>1</sup>,  
Paula C. C. de Andrade<sup>1</sup>, Carolina Furtado<sup>3</sup> and Thiago E. Parente<sup>1</sup>

The complete mitogenome of *Corydoras nattereri*, a species of mailed catfishes from southeastern Brazil, was reconstructed using next-generation sequencing techniques. The mitogenome was assembled using mitochondrial transcripts from the liver transcriptomes of three individuals, and produced a circular DNA sequence of 16,557 nucleotides encoding 22 tRNA genes, two rRNA genes, 13 protein-coding genes and two noncoding control regions (D-loop, OrigL). Phylogeographic analysis of closely related sequences of Cytochrome Oxidase C subunit I (COI) demonstrates high diversity among morphologically similar populations of *C. nattereri*. *Corydoras nattereri* is nested within a complex of populations currently assigned to *C. paleatus* and *C. ehrhardti*. Analysis of mitogenome structure demonstrated that an insertion of 21 nucleotides between the ATPase subunit-6 and COIII genes may represent a phylogenetically informative character associated with the evolution of the Corydoradinae.

O mitogenoma completo de *Corydoras nattereri*, uma espécie de bagres encouraçados do sudeste do Brasil, foi reconstruído através de técnicas de sequenciamento de DNA de próxima geração. O mitogenoma foi produzido a partir de produtos de transcrição mitocondrial dos transcriptomas hepáticos de três indivíduos, resultando numa sequência de DNA circular de 16.557 nucleotídeos abrangendo 22 genes de tRNA, dois genes de rRNA, 13 genes codificadores de proteínas e duas regiões de controle não codificadoras (D-loop, OrigL). A análise filogenética de sequências proximamente relacionadas da subunidade I do gene Citocrome Oxidase C (COI) demonstrou a existência de elevada diversidade entre populações morfologicamente similares de *C. nattereri*. *Corydoras nattereri* está inserida num complexo de populações atualmente identificadas como *C. paleatus* e *C. ehrhardti*. A análise da estrutura do mitogenoma demonstra que a inserção de uma sequência de 21 nucleotídeos entre os genes da subunidade 6 da ATPase e do COIII representa um caráter filogeneticamente informativo associado à evolução de Corydoradinae.

**Keywords:** Barcode, DNA, Mitogenome, Molecular diversity, mtRNA.

## Introduction

Corydoradinae are a species-rich group of armored freshwater catfishes that inhabit streams, rivers and floodplains throughout South America (Alexandrou *et al.*, 2011). Together with the Callichthyinae, they comprise the Callichthyidae, a family of catfishes diagnosed by the presence of two series of bony plates on the sides of the body and one pair of barbels at lips junction (Reis, 1998). The Corydoradinae comprises 227 nominal taxa and 188 valid species (Eschmeyer & Fong, 2015), assigned to the *Aspidoras* Ihering, 1907, *Corydoras* Lacépède, 1803, and *Scleromystax* Gunther, 1864 (Britto, 2003). *Corydoras* is the most species-rich genus of catfishes with over 160 described and nearly as many undescribed species (Alexandrou *et al.*,

2011; Eschmeyer & Fong, 2015). According to Reis (2003) about two new species of *Corydoras* are described each year. While the Callichthyinae is relatively well-known based on morphological (Reis, 1997, 1998, 2003) and molecular studies (Mariguela *et al.*, 2013), the Corydoradinae remains poorly known, despite their great interest to the aquarium hobby. Phylogenetic studies that included species of *Corydoras* have been performed, primarily by Britto (2003) based on morphological characters, and more recently by Alexandrou *et al.* (2011), based on molecular data. In the later study, however, a large proportion of the 52 taxa recognized could not be associated to a valid name, with many species being referenced to informal “C-Numbers” (Fuller & Evers, 2005) available from the aquarium industry or to their geographic origin.

<sup>1</sup>Laboratório de Toxicologia Ambiental, Fundação Oswaldo Cruz (FIOCRUZ), Avenida Brasil, 4036 sala 101, 2104-031 Rio de Janeiro, RJ, Brazil. (DAM) daniel.moreira@ioc.fiocruz.br, (MGPM) magalhaes.maithe@gmail.com, (PCCA) paula.andrade.bio@gmail.com, (TEP) parente@ensp.fiocruz.br

<sup>2</sup>Departamento de Vertebrados, Museu Nacional, Universidade Federal do Rio de Janeiro (UFRJ), Quinta da Boa Vista, 20940-040 Rio de Janeiro, RJ, Brazil. (PAB) backup@acd.ufrj.br (corresponding author), (MRB) mrbritto2002@yahoo.com.br

<sup>3</sup>Divisão de Genética, Instituto Nacional do Câncer (INCA), André Cavalcanti, 37, 4º andar, 20231-050 Rio de Janeiro, RJ, Brazil. (CF) cfurtado@inca.gov.br

*Corydoras nattereri* Steindachner, 1876, is a widespread species of *Corydoras* in southeastern Brazil, ranging from rio Mucuri, in Bahia, to the Paranaguá Bay, in Paraná (Britto, 2007; Shimabukuro-Dias *et al.*, 2004a). *Corydoras nattereri* and *Scleromystax prionotos* (Nijssen & Isbrücker, 1980) form a pair of color mimics where their distribution overlaps (Alexandrou *et al.*, 2011). The geographic range of *C. nattereri* represents a distributional range of about 1,350 km, encompassing numerous isolated coastal river drainages. With such widespread distribution it is not surprising that significant variability has been found among various populations of *C. nattereri*. Oliveira *et al.* (1990) identified three different cytotypes among these populations, that differ in number of chromosomes (2n numbers of 40, 43 and 44), suggesting that more than one species is represented by the taxon. That cytogenetic variation among populations was later correlated with variation in DNA sequence data (Simabukuro-Dias *et al.*, 2004b).

Currently partial sequences of the Cytochrome Oxidase C subunit I (COI) of *Corydoras nattereri*, a widely used DNA barcode marker, are available publicly from only two localities (Pereira *et al.*, 2011, 2013). Sequences of additional mitochondrial genes have been made available by Alexandrou *et al.* (2011), but only specimens with imprecise locality data have been listed in that study. Within the Corydoradinae, complete mitochondrial sequence data is only available for *C. rabauti* from a specimen without associated locality data (Saitoh *et al.*, 2003). Herein, we present the complete mitochondrial genome for *C. nattereri* based on three specimens with precise geographic provenance, thus providing a significant increment in DNA data available for phylogenetic studies of the Corydoradinae.

## Material and Methods

Specimens of *Corydoras nattereri* were collected in the rio Suruí (22.600556 S, -43.091667 W) at the Santo Aleixo district, Magé, Rio de Janeiro, Brazil. The fish were deposited at the Museu Nacional, Rio de Janeiro, UFRJ (MNRJ 41520) and tissues from three individuals (MNTI 8664-8666) were preserved in ethanol and RNALater. Total RNA was extracted from the liver tissue following conventional phenol-chloroform extraction. RNA quality was assessed using the RNA Nano kit for Bioanalyzer (Agilent). Three individual cDNA libraries were constructed using the TruSeq RNA Sample kit v.2 (Illumina). Libraries were accessed for quality (Bioanalyzer, DNA1000 kit, Agilent) and quantity (Kapa Biosystems). Two separated runs (single-end and paired-end) were performed in an Illumina HiSeq 2500 using the TrueSeq SBS kit v.3 (Illumina). Raw Illumina data were demultiplexed using the BCL2FASTQ software (Illumina). Reads were trimmed for Illumina adaptors by Trimmomatic (Bolger *et al.*, 2014) and its quality was evaluated using FastQC (Babraham Bioinformatics). Only reads with Phred score over 30 were used for the transcriptome assembly. Cleaned reads from the three individual fish were used for the *de novo* assembly

of transcriptomes using the default parameters of Trinity (v. 2.0.2) (Haas *et al.*, 2013). Mitochondrial genomes were assembled using the mitochondrial transcripts from the liver transcriptome, following the approach described by Moreira *et al.* (2015a, 2015b). Briefly, mitochondrial transcripts were retrieved running a BLASTN search against the mitogenome of the closest related species with a complete mitogenome available, *Corydoras rabauti* (GI: 29501080) (Saitoh *et al.*, 2003). Mitochondrial transcripts were edited according to the information of strand orientation given by the BLASTN result, and aligned with SeaView using the built-in CLUSTAL alignment algorithm and the mitogenome of *C. rabauti* (Gouy *et al.*, 2010). The sequence of each CONTIG was manually checked for inconsistencies and gaps. Small gaps at mitogenomes, which code for transfer RNAs, were completed with Sanger sequencing data from PCR with specific designed primers. As the identity of the three assembled mitogenomes was higher than 99.8%, the three sequences were concatenated in one consensus sequence. The consensus mitogenome was annotated using the web-based services MitoFish and MITOS (Iwasaki *et al.*, 2013; Bernt *et al.*, 2013), and the origin of the L-strand replication was identified based on Wong *et al.* (1983). In order to determine sequencing depth of each base in the mitogenome, Bowtie v. 1.0.0 was used to align the reads on the assembled mitogenome. The aligned reads were viewed using Integrated Genome Viewer (IGV), Tablet (Langmead *et al.*, 2009; Milne *et al.*, 2010; Robinson *et al.*, 2011; Thorvaldsdóttir *et al.*, 2012), and Geneious version 6 (<http://www.geneious.com>, Kearse *et al.*, 2012).

A Maximum Likelihood (ML) analysis of relationships was performed to position the mitogenomes within a taxonomic and phylogenetic context. All publicly available sequences of COI that exhibited nucleotide identity greater than 90% in a BLAST search of the GenBank Nucleotide Database and Bold Systems were included in the analysis, as well as additional sequences of *Corydoras nattereri*, *Callichthys callichthys*, *Scleromystax barbatus*, and *Aspidoras lakoi* (the latter three used as outgroups) produced in the Museu Nacional (MNLM) laboratory (Table 1) using Sanger sequencing methods, and the corresponding COI sequence extracted from the *C. rabauti* mitochondrial genome (GenBank Accession AB054128). To ensure uniformity of coverage, only nucleotides from positions 58 to 699, and only sequences with full coverage of that segment were included in the analysis, producing a matrix of nucleotide sequences from 35 fish. The ML analysis was performed using Mega 6.06 (Tamura *et al.*, 2013) under the Hasegawa-Kishino-Yano model (Hasegawa *et al.*, 1985), selected by the corrected Akaike information criterion using jModeltest v.2.1.7 (Darriba *et al.*, 2012). Initial tree(s) for the heuristic search were obtained by Neighbor-Join and BioNJ algorithms applied to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The discrete Gamma distribution was used to model evolutionary differences among sites, with 5 categories (+G, parameter = 0.1563). Branch support was estimated with the Bootstrap method, using 350 replications. The tree was rooted using *Callichthys callichthys* as outgroup.

**Table 1.** List of samples used in this study, ordered according to GenBank Accession codes. Voucher museum catalog number is provided only for specimens that are being made available for the first time.

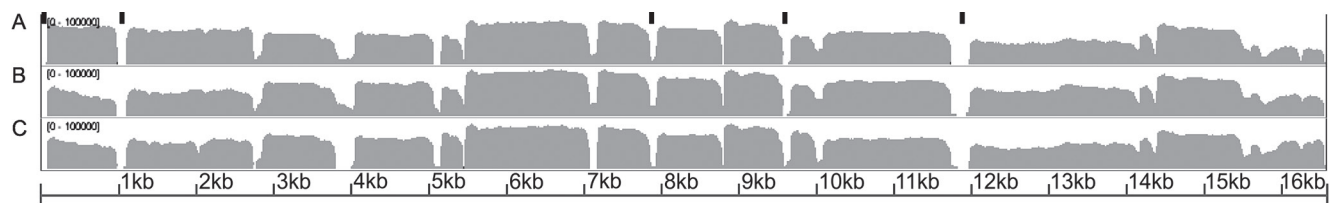
GenBank Accession	BOLD ProcessID	Sample	Voucher	Species	Publication/Source	Latitude	Longitude	Locality
GU701815	FUPR517-09	LBP-32757		<i>Corydoras ehrhardti</i>	Pereira <i>et al.</i> , 2013	-23,9383	-50,7290	Upper Parana basin, Brazil
GU701816	FUPR516-09	LBP-32756		<i>Corydoras ehrhardti</i>	Pereira <i>et al.</i> , 2013	-23,9383	-50,7290	Upper Parana basin, Brazil
GU701817	FUPR818-09	LBP-36128		<i>Corydoras ehrhardti</i>	Pereira <i>et al.</i> , 2013	-23,9383	-50,7290	Upper Parana basin, Brazil
GU701818	FUPR817-09	LBP-36126		<i>Corydoras ehrhardti</i>	Pereira <i>et al.</i> , 2013	-23,9383	-50,7290	Upper Parana basin, Brazil
GU701819	FUPR816-09	LBP-36124		<i>Corydoras ehrhardti</i>	Pereira <i>et al.</i> , 2013	-23,9383	-50,7290	Upper Parana basin, Brazil
GU701809	FUPR819-09	LBP-36114		<i>Corydoras paleatus</i>	Pereira <i>et al.</i> , 2013	-	-	Upper Parana basin, Brazil
GU701810	FUPR822-09	LBP-36119		<i>Corydoras paleatus</i>	Pereira <i>et al.</i> , 2013	-	-	Upper Parana basin, Brazil
GU701812	FUPR824-09	LBP-36122		<i>Corydoras paleatus</i>	Pereira <i>et al.</i> , 2013	-	-	Upper Parana basin, Brazil
GU701813	FUPR823-09	LBP-36121		<i>Corydoras paleatus</i>	Pereira <i>et al.</i> , 2013	-	-	Upper Parana basin, Brazil
GU701814	FUPR515-09	LBP-32755		<i>Corydoras ehrhardti</i>	Pereira <i>et al.</i> , 2013	-23,9383	-50,7290	Upper Parana basin, Brazil
GU701871	FUPR820-09	LBP-36116		<i>Corydoras paleatus</i>	Pereira <i>et al.</i> , 2013	-	-	Upper Parana basin, Brazil
GU702213	FPSR082-09	LBP-29094		<i>Corydoras nattereri</i>	Pereira <i>et al.</i> , 2011	-23,3690	-46,0240	[Rio Paraiíba do Sul upstream from Jacare], SP, Brazil
JN988818	FUPR285-09	LBP-32330		<i>Corydoras nattereri</i>	Pereira <i>et al.</i> , 2013	-23,5112	-45,8591	[Rio Paraitinguinha, Tietê drainage], upper Parana Basin, SP, Brazil
JN988819	FUPR286-09	LBP-32331		<i>Corydoras nattereri</i>	Pereira <i>et al.</i> , 2013	-23,5112	-45,8591	[Rio Paraitinguinha, Tietê drainage], upper Parana Basin, SP, Brazil
JN988821	FUPR288-09	LBP-32333		<i>Corydoras nattereri</i>	Pereira <i>et al.</i> , 2013	-23,5112	-45,8591	[Rio Paraitinguinha, Tietê drainage], upper Parana Basin, SP, Brazil
JN988822	FUPR289-09	LBP-32334		<i>Corydoras nattereri</i>	Pereira <i>et al.</i> , 2013	-23,5112	-45,8591	[Rio Paraitinguinha, Tietê drainage], upper Parana Basin, SP, Brazil
JX1111730	FARGB202-11	UNMDP-T 0370		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-36,2680	-59,9800	Arroio Tapalque, Buenos Aires, Argentina
JX1111731	FARGB238-11	UNMDP-T 0406		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-38,3350	-61,6014	El Divisorio Stream, Buenos Aires, Argentina
JX1111732	FARGB358-11	UNMDP-T 0526		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-34,9700	-61,0433	Adeg Puente Santa Cruz (Ascencion), Buenos Aires, Argentina
JX1111733	FARGB239-11	UNMDP-T 0407		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-38,3350	-61,6014	El Divisorio Stream, Buenos Aires, Argentina
JX1111734	FARGB201-11	UNMDP-T 0369		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-36,2680	-59,9800	Arroyo Tapalque, Buenos Aires, Argentina
JX1111735	FARGB200-11	UNMDP-T 0368		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-36,2680	-59,9800	Arroyo Tapalque, Buenos Aires, Argentina
JX1111736	FARGB184-11	UNMDP-T 0352		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-35,6056	-59,9914	Chascomus lagoon, Buenos Aires, Argentina
JX1111737	FARGB183-11	UNMDP-T 0351		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-35,6056	-59,9914	Chascomus lagoon, Buenos Aires, Argentina
JX1111738	FARGB172-11	UNMDP-T 0340		<i>Corydoras paleatus</i>	Rosso <i>et al.</i> , 2012	-35,5411	-58,0894	Vitel lagoon, Buenos Aires, Argentina

Mitogenome of *Corydoras nattereri*

(Table 1: Cont.)

GenBank Accession	BOLD ProcessID	Sample	Voucher	Species	Publication/Source	Latitude	Longitude	Locality
KT874579	MNRJ531-15	MNTT9093/ MNL.M5113	MNRJ41901	<i>Callichthys callichthys</i>	This study	-19,9753	-40,5478	rio Valssungana Velha, Reis Magos drainage, near bridge of Santa Teresa - Santa Leopoldina road, Santa Teresa, ES, Brazil
KT874580	MNRJ533-15	MNTT554/ MNL1354	MNRJ31639	<i>Aspidoras lakoi</i>	This study	-20,7500	-46,2122	Left bank tributary of ribeirão do Turvo, between Fazenda da Serra de Cima and Fazenda Baixadão, Capitólio, MG, Brazil
KT874581	MNRJ538-15	MNTT8665/ MNL.M5103	MNRJ41520	<i>Corydoras nattereri</i>	This study	-22,6006	-43,0917	Tributary of rio Suruí and rio Suruí downstream of Piabeta - Santo Aleixo road, Magé, RJ, Brazil
KT874582	MNRJ532-15	MNTT7422/ MNL.M4992	MNRJ40948	<i>Scleromystax barbatus</i>	This study	-26,1972	-49,9219	Rio Lindo (right margin tributary of rio Cubatão), SC-301, bairro Dona Francisca, Joinville, SC, Brazil
KT874583	MNRJ534-15	MNTT3387/ MNL.M935	MNRJ37470	<i>Corydoras nattereri</i>	This study	-22,4675	-42,2975	Córrego Aldeia Velha, under bridge on road Aldeia Velha, near RPPN Fazenda do Bom Retiro, Casimiro de Abreu, RJ, Brazil
KT874584	MNRJ539-15	MNTT8666/ MNL.M5104	MNRJ41520	<i>Corydoras nattereri</i>	This study	-22,6006	-43,0917	Tributary of rio Suruí and rio Suruí downstream of Piabeta - Santo Aleixo road, Magé, RJ, Brazil
KT874585	MNRJ537-15	MNTT8664/ MNL.M5102	MNRJ41520	<i>Corydoras nattereri</i>	This study	-22,6006	-43,0917	Tributary of rio Suruí and rio Suruí downstream of Piabeta - Santo Aleixo road, Magé, RJ, Brazil
KT874586	MNRJ536-15	MNTT7464/ MNL.M5000	MNRJ41030	<i>Corydoras ehnhardti</i>	This study	-26,2803	-49,3244	rio Vermelho, Itapocu drainage, downstream of dam, upstream of aqueduct, São Bento do Sul, SC, Brazil
KT874587	MNRJ535-15	MNTT3832	MNRJ37693	<i>Corydoras nattereri</i>	This study	-18,3106	-40,2411	rio do Sul under bridge of ES-130 between Vinhático and pinheiros, on municipal border between Montanha and Pinheiros, ES, Brazil





**Fig. 2.** Sequencing depth over the complete mitogenomes of the three individuals of *Corydoras nattereri*: KT239008 (A), KT239009 (B), and KT239010 (C). Read counts (y-axis) are shown in logarithmic scale and sharp decreases correspond to the punctuation model of mitochondrial transcription (positions correspond to those shown in Fig. 1 and Table 3). Black vertical bars indicate position of gaps that were filled with Sanger sequencing (Table 2). Reads were mapped to the mitogenomes using Bowtie and visualized at the Integrative Genome Viewer (IGV, Bernt *et al.*, 2013; Thorvaldsdóttir *et al.*, 2013).

**Table 3.** Positioning of genes in the mitochondrial genome of *Corydoras nattereri*. Negative gap values indicate overlap.

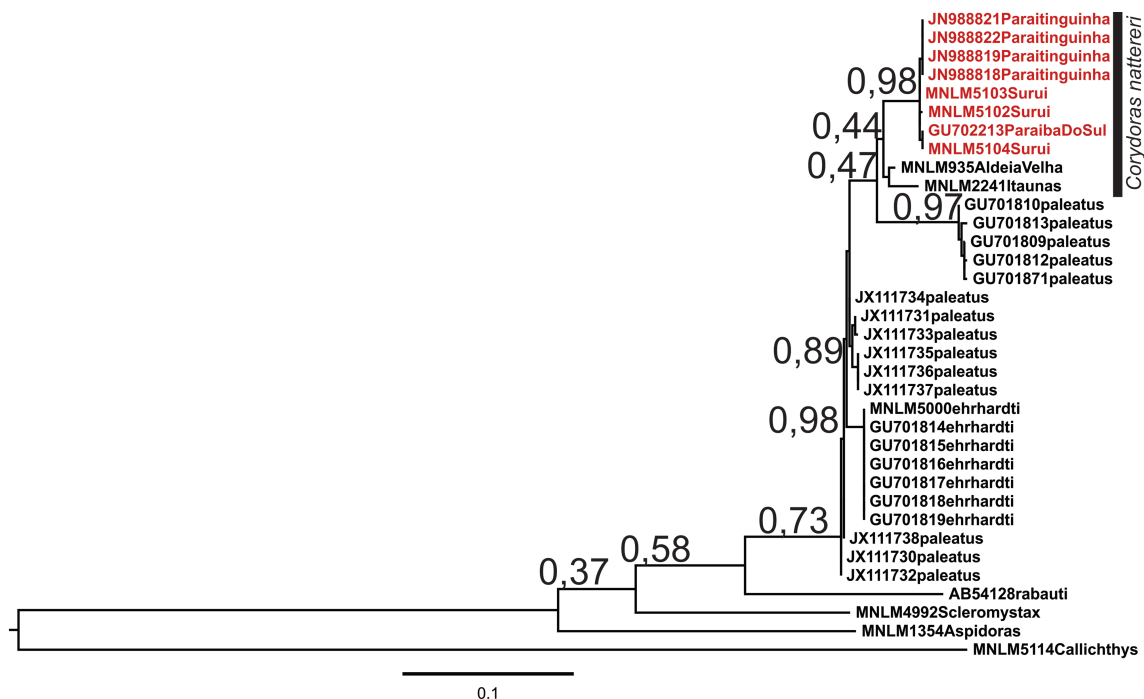
Gene	Start	End	Length	Gap	Direction	Start Codon	Stop Codon
tRNA-Phe	1	68	68	0			
12S rRNA	69	1013	945	0			
tRNA-Val	1014	1085	72	0			
16S rRNA	1086	2752	1667	0			
tRNA-Leu	2753	2827	75	0			
ND1	2828	3799	972	8		ATG	TAG
tRNA-Ile	3808	3879	72	-2			
tRNA-Gln	3878	3948	71	-1	complement		
tRNA-Met	3948	4017	70	0			
ND2	4018	5062	1045	0		ATG	T--
tRNA-Trp	5063	5133	71	1			
tRNA-Ala	5135	5203	69	1	complement		
tRNA-Asn	5205	5277	73	0	complement		
origin-L	5278	5312	35	-3	complement		
tRNA-Cys	5310	5377	68	-1	complement		
tRNA-Tyr	5377	5446	70	1	complement		
COI	5448	7007	1560	-13		GTG	AGG
tRNA-Ser	6995	7065	71	4	complement		
tRNA-Asp	7070	7138	69	6			
COII	7145	7835	691	0		ATG	T--
tRNA-Lys	7836	7909	74	1			
ATPase 8	7911	8078	168	-10		ATG	TAA
ATPase 6	8069	8752	684	21		ATG	TAA
COIII	8774	9557	784	0		ATG	T--
tRNA-Gly	9558	9629	72	0			
ND3	9630	9978	349	0		ATG	T--
tRNA-Arg	9979	10048	70	0			
ND4L	10049	10345	297	-7		ATG	TAA
ND4	10339	11719	1381	0		ATG	T--
tRNA-His	11720	11789	70	0			
tRNA-Ser	11790	11856	67	1			
tRNA-Leu	11858	11930	73	0			
ND5	11931	13757	1827	-4		ATG	TAA
ND6	13754	14269	516	0	complement	ATG	TAA
tRNA-Glu	14270	14337	68	3	complement		
Cyt b	14341	15478	1138	0		ATG	T--
tRNA-Thr	15479	15550	72	-2			
tRNA-Pro	15549	15618	70	0	complement		
D-loop	15619	16557	939				

**Genome organization.** The complete mitochondrial genome sequence of *Corydoras nattereri* contains the typical vertebrate features: 22 tRNA genes, 2 rRNA genes, 13 protein-coding genes and two noncoding control regions (D-loop, OrigL) (Table 3, Fig. 2). The majority of genes are encoded on the heavy strand, whereas ND6 and eight tRNAs are found on the light strand. All protein-coding genes used ATG start codons, except for COI that used GTG. Seven protein-coding genes are terminated with the complete stop codon, of which five ended with TAA (ATP8, ATP6, ND4L, ND5 and ND6), one with TAG (ND1) and one with AGG (COI). The remaining protein-coding genes are ended by incomplete stop codons, T (COII, COIII, ND2, ND3, ND4 and Cytb), which are completed by post-transcriptional polyadenylation. The 12S and the 16S rRNA genes are separated by tRNA-Val gene and their lengths are 945 and 1,667 bp, respectively. The 22 tRNA genes had sizes ranging from 67 to 75 nucleotides and the control region, located between tRNA<sup>Pro</sup> and tRNA<sup>Phe</sup> genes, is 939 bp long. A 21-nucleotide insertion sequence between the ATPase subunit-6 and COIII genes was found in *C. nattereri*. The nucleotide composition for the heavy strand was 32.3% A, 25.7%T, 15.1% G, and 27.0% C.

**Phylogenetic context.** The phylogeographic analysis of publicly available COI together with additional COI sequences generated by our research group confirmed our

identification of the samples of the rio Suruí as *Corydoras nattereri* (Fig. 3). Our three samples form a monophyletic group with specimens of *C. nattereri* from the Paraíba do Sul and the Paraitinguinha rivers (upper Tietê drainage, upper Paraná basin) (Pereira *et al.*, 2011, 2013). Contrasting with the high similarity of the samples from these three basins, our samples from the rio Aldeia Velha (rio São João coastal basin), and rio Itaúnas are considerably different. The latter are morphologically similar to *C. nattereri*, but their sequences are 2.5%-3.6% divergent in relation to the rio Suruí samples. Such high divergence suggests that the populations of *Corydoras* from the São João and Itaúnas river basins represent cryptic species.

Our phylogenetic analysis also demonstrates that the *C. nattereri* clade is nested within a large clade of samples identified in the literature (Pereira *et al.*, 2011, 2013; Rosso *et al.*, 2012) as *C. paleatus* and *C. ehrhardti*. Samples of *C. ehrhardti* (including new sequences produced here) form a monophyletic clade also included among this larger clade, but samples of *Corydoras paleatus* form a complex of non-monophyletic populations. Within this large complex, samples of *C. paleatus* GU701809, GU701810, GU701812, GU701813, and GU701871, from the upper rio Paraná basin, form the monophyletic subunit most closely related to *C. nattereri*, but the bootstrap value for this sister group relationship is low, indicating that further study of *C. paleatus* species complex is still necessary.



**Fig. 3.** Maximum likelihood tree (log likelihood = -2410.0522) of *Corydoras* samples with at least 90% similarity to the mitochondrial cytochrome oxidase I sequences of *C. nattereri* from the rio Suruí. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap robustness is indicated next to selected branches. Samples of *C. nattereri* have the locality name appended to the sample ID (those for which mitogenomes were produced are from de rio “Suruí”), outgroups have the genus name and other samples of *Corydoras* have the species epithet name appended to the ID code (Table 1).

## Discussion

Despite the great diversity of *Corydoras*, this is the first mitogenome with voucher specimens sampled in their native habitat and deposited in a permanent collection. A mitogenome of *C. rabauti* has been reported by Saitoh *et al.* (2003), but that study was based on a specimen without locality data obtained from the aquarium trade, and the study does not mention a registration number of a voucher specimen in any biological collection.

Our analysis of COI sequences revealed high levels of genetic divergence and taxonomic complexity among samples closely related to *C. nattereri*. Specimens from the São João and Itaúnas river basins may represent cryptic species, that are currently undistinguishable from topotype specimens of *C. nattereri*. Their level of divergency (2.5% - 3.6%) far exceeds the maximum intraspecific divergence (1.6%) reported among six species of *Corydoras* from the upper Paraná basin (Pereira *et al.*, 2013). The analysis also demonstrated the complexity of relationships among populations of *C. nattereri*, *C. paleatus*, and *C. ehrhardti*. Within this context, our newly produced mitogenome *C. nattereri* is likely to provide a solid base to identify additional mitochondrial markers to be used in future studies designed to clarify the relationships among these and other callichthyid taxa.

Comparison of the mitogenome of *Corydoras nattereri* with that of *C. rabauti* (Saitoh *et al.*, 2003) reveals features that are likely to be phylogenetically informative and useful in future studies. A 21-nucleotide insertion sequence between the ATPase subunit-6 and COIII genes was found in *C. nattereri*. This insertion corresponds to a 17-nucleotide insertion previously detected in *C. rabauti* (Saitoh *et al.*, 2003). Most vertebrate mitochondrial genomes have a head-to-tail junction between the ATPase subunit-6 and COIII genes, and this insertion was considered phylogenetically uninformative in the study of Saitoh *et al.* (2003). Our discovery of an insertion in the homologous position of *C. nattereri* is interpreted as an apomorphic trait shared by the two species. Further investigation about the distribution and length of this insertion among Corydoradinae is likely to yield significant insights about the phylogeny of the group.

## Acknowledgements

Authors acknowledge the use of the computer cluster Titan from the Woods Hole Oceanographic Institution (WHOI), where TP holds a guest investigator position. Carla Christie Dijban Quijada, Emanuel Neuhaus, Decio Ferreira de Moraes Junior (UFRJ/MN), and José R. Gomes participated in the field work that yielded the specimens from the rio Surui. William Bryan Jennings (UFRJ/MN) and Gustavo Ferraro (FAPERJ fellowship) provided essential support for laboratory activities at the Laboratório de Pesquisa em Biodiversidade Molecular (UFRJ/MN). This work was supported by Partnerships for Enhanced

Engagement in Research (PEER) grant from U.S. Agency for International Development (PGA-2000003446) to TP, and grants from CNPq (307610/2013-6, 564940/2010-0, 476822/2012-2) and FAPERJ (E-26/111.404/2012, E-26/200.697/2014) to PAB.

## References

- Alexandrou, M. A., C. Oliveira, M. Maillard, R. A. R. McGill, J. Newton, S. Creer & M. I. Taylor. 2011. Competition and phylogeny determine community structure in Müllerian mimics. *Nature*, 469: 84-88.
- Bernt, M., A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritsch, J. Pütz, M. Middendorf & P. F. Stadler. 2013. MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69: 313-319.
- Bolger, A. M., M. Lohse & B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30: 2114-2120.
- Britto, M. R. 2003. Phylogeny of the subfamily Corydoradinae Hoedeman, 1952 (Siluriformes: Callichthyidae), with a definition of its genera. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 153: 119-154.
- Britto, M. R. 2007. Família Callichthyidae. Pp.75-81. In: Backup, P. A., N. A. Menezes & M. S. Ghazzi (Eds.). *Catálogo das espécies de peixes de água doce do Brasil*. Rio de Janeiro, Museu Nacional. (Série Livros, 23).
- Darriba, D., G. L. Taboada, R. Doallo & D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9: 772.
- Eschmeyer, W. N. & J. D. Fong. 2015. *Catalog of fishes: genera, species, references*. Electronic Version. San Francisco, CA, California Academy of Sciences. Available from: <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp> (6 August 2015).
- Fuller, I. A. M. & H.-G. Evers. 2005. Identifying Corydoradinae catfish. *Aspidoras - Brochis - Corydoras - Scleromystax & C-numbers*. 1st ed. Worcestershire, Ian Fuller Enterprises; Rodgau, A.C.S. GmbH (Aqualog), 384p.
- Gouy, M., S. Guindon & O. Gascuel. 2010. SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, 27: 221-224.
- Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman & A. Regev. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8: 1494-1512.
- Hasegawa, M., H. Kishino & T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22: 160-174.
- Iwasaki, W., T. Fukunaga, R. Isagozawa, K. Yamada, Y. Maeda, T. P. Satoh, T. Sado, K. Mabuchi, H. Takeshima, M. Miya & M. Nishida. 2013. Mitofish and mitoannotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. *Molecular Biology and Evolution*, 30: 2531-2540.



- Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes & A. Drummond. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28: 1647-1649.
- Langmead, B., C. Trapnell, M. Pop & S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10: R25.
- Mariguela, T. C., M. A. Alexandrou, F. Foresti & C. Oliveira. 2013. Historical biogeography and cryptic diversity in the Callichthyinae (Siluriformes, Callichthyidae). *Journal of Zoological Systematics and Evolutionary Research*, 51: 308-315.
- Milne, I., M. Bayer, L. Cardle, P. Shaw, G. Stephen, F. Wright & D. Marshall. 2010. Tablet-next generation sequence assembly visualization. *Bioinformatics*, 26: 401-402.
- Moreira, D. A., C. Furtado & T. E. Parente. 2015a. The use of transcriptomic next-generation sequencing data to assemble mitochondrial genomes of *Ancistrus* spp. (Loricariidae). *Gene*, 573: 171-175.
- Moreira, D. A., M. G. P. Magalhães, P. C. C. Andrade, C. Furtado, A. L. Val & T. E. Parente. 2015b. An RNA-based approach to sequence the mitogenome of *Hypoptopoma incognitum* (Siluriformes: Loricariidae). *Mitochondrial DNA*: 1-3.
- Nijssen, H. & I. J. H. Isbrücker. 1980. On the identity of *Corydoras nattereri* Steindachner, 1877 with the description of a new species, *Corydoras prionotos* (Pisces, Siluriformes, Callichthyidae). *Beaufortia*, 30: 1-9.
- Oliveira, C., L. F. Almeida Toledo & S. A. Toledo Filho. 1990. Comparative cytogenetic analysis in three cytotypes of *Corydoras nattereri* (Pisces, Siluriformes, Callichthyidae). *Cytologia*, 55: 21-26.
- Pereira, L. H. G., G. M. G. Maia, R. Hanner, F. Foresti & C. Oliveira. 2011. DNA barcodes discriminate freshwater fishes from the Paraíba do Sul River Basin, São Paulo, Brazil. *Mitochondrial DNA*, 22(S1): 71-79.
- Pereira, L. H. G., R. Hanner, F. Foresti & C. Oliveira. 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genetics*, 14: 20 (1-14).
- Reis, R. E. 1997. Revision of the Neotropical catfish genus *Hoplosternum* (Ostariophysi: Siluriformes: Callichthyidae), with the description of two new genera and three new species. *Ichthyological Exploration of Freshwaters*, 7: 299-326.
- Reis, R. E. 1998. Anatomy and phylogenetic analysis of the Neotropical callichthyid catfishes (Ostariophysi, Siluriformes). *Zoological Journal of the Linnean Society*, 124: 105-168.
- Reis, R. E. 2003. Family Callichthyidae (Armored catfishes). Pp. 291-309. In: Reis, R. E., S. O. Kullander & C. J. Ferraris, Jr. (Orgs.). Check list of the freshwater fishes of South and Central America. Porto Alegre, Edipucrs.
- Robinson, J. T., H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz & J. P. Mesirov. 2011. Integrative genomics viewer. *Nature Biotechnology*, 29: 24-26.
- Rosso, J. J., E. Mabrugaña, M. González Castro & J. M. Díaz de Astarloa. 2012. DNA barcoding Neotropical fishes: recent advances from the Pampa Plain, Argentina. *Molecular Ecology Resources*, 12: 999-1011.
- Saitoh, K., M. Miya, J. G. Inoue, N. B. Ishiguro & M. Nishida, M. 2003. Mitochondrial genomics of ostariophysan fishes: perspectives on phylogeny and biogeography. *Journal of Molecular Evolution*, 56: 464-472.
- Shimabukuro-Dias, C. K., C. Oliveira & F. Foresti. 2004a. Karyotype variability in eleven species of the catfish genus *Corydoras* (Siluriformes: Callichthyidae). *Ichthyological Exploration of Freshwaters*, 15: 135-146.
- Shimabukuro-Dias, C. K., C. Oliveira, R. E. Reis & F. Foresti. 2004b. Molecular phylogeny of the armored catfish family Callichthyidae (Ostariophysi, Siluriformes). *Molecular Phylogenetics and Evolution*, 32: 152-163.
- Tamura, K., G. Stecher, D. Peterson, A. Filipski & S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30: 2725-2729.
- Thorvaldsdóttir, H., J. T. Robinson & J. P. Mesirov. 2012. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14: 178-192.
- Wong, J. F. H., D. P. Ma, R. K. Wilson & B. A. Roe. 1983. DNA sequence of the *Xenopus laevis* mitochondrial heavy and light strand replication origins and flanking tRNA genes. *Nucleic Acids Research*, 11: 4977-4995.

Submitted October 20, 2015

Accepted January 12, 2016 by Guillermo Orti