Zoological Studies **65:**27 (2026)

Impacts of Outgroup Selection in bPTP Species Delimitation Analyses: a Case Study in Cichlid Fishes of the Genus *Australoheros* Říčan & Kullander, 2006 (Cichliformes: Cichlidae) from Coastal Basins of Southern South America

Rebeca Rosa[1,2], Elisabeth Henschel[1,2,3,*], Pedro H.N. de Bragança[3,4], Grazielle F.E. Gomes[5], and Felipe P. Ottoni[3,6]

[1]*Federal University of Viçosa, Department of Animal Biology, Laboratory of Molecular Systematics and Reproductive Biology (BEAGLE), 36570-900, Viçosa, Minas Gerais, Brazil. *Correspondence: E-mail: elisabeth.henschel@ufv.br (Henschel)*
E-mail: rebeca.rosa@ufv.br (Rosa)

[2]*Graduate program in Animal Biology, Department of Animal Biology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil*

[3]*NRF-South African Institute for Aquatic Biodiversity (NRF-SAIAB) P. Bag 1015, Makhanda 6140, South Africa. E-mail: pedrobra88@gmail.com (Bragança); fpottoni@gmail.com (Ottoni)*

[4]*Department of Ichthyology, American Museum of Natural History, 10024 New York, USA.*

[5]*Federal University of Pará (UFPA), Department of Animal Science, Laboratory of Applied Genetics (LAGA), 68600-000, Bragança, Pará, Brazil. E-mail: grazielle@ufpa.br (Gomes)*

[6]*Federal University of Maranhão (UFMA), Laboratory of Systematics and Ecology of Organisms (LASEOA), 65500-000, Chapadinha, Maranhão, Brazil*

(Received 6 December 2025 / Accepted 4 April 2026 / Published - May 2026)
Communicated by Jen-Pan Huang

ORCiD:
Rosa R: https://orcid.org/0009-0003-6288-9650
Henschel E: https://orcid.org/0000-0001-5507-7396
Bragança PHN: https://orcid.org/orcid.org/0000-0002-8357-7010
Evangelista-Gomes G: https://orcid.org/orcid.org/0000-0001-8898-0311
Ottoni FP: https://orcid.org/orcid.org/0000-0002-9390-0918

The goal of this study is a methodological evaluation of the Bayesian Poisson Tree Processes (bPTP) method for species delimitation in the genus *Australoheros*. We hypothesize that the increase in phylogenetic distance of the outgroup used decreases the number of delimited Operational Taxonomic Units (OTUs). We test the impact of different outgroups, positioned at varying phylogenetic distances from the ingroup, in a dataset extracted from the most recent taxonomic study for *Australoheros*. Lucena and collaborators validated 17 species, but used a dataset containing identical haplotypes, thus violating an important premise of the bPTP method. We reanalyzed the dataset of this work, but followed the assumptions of the bPTP method. We

tested outgroup-distance effects on species delimitation by using seven distinct outgroups from distinct genera, tribes and subfamilies of the Cichlidae, totaling eight COI and eight CYTB matrices. COI analyses delimited between 15 and 9 OTUs; CYTB analyses between 24 and 14, showing consistent OTU reduction as outgroups became more phylogenetically distant. Our results highlight that CYTB delimited a greater number of OTUs, however, this dataset contains a higher number of haplotypes than COI. We consider that the use of phylogenetically distant outgroups may reduce the adequate recovery of OTUs.

Keywords: *Australoheros* species group, Bayesian Inference, Bioinformatics, Coalescence, Mitochondrial Genes

Citation: Rosa R, Henschel E, Bragança PHN, Evangelista-Gomes G, Ottoni FP. 2026. Impacts of outgroup selection in bPTP species delimitation analyses: a case study in cichlid fishes of the genus *Australoheros* Říčan & Kullander, 2006 (Cichliformes: Cichlidae) from coastal basins of southern South America. Zool Stud **65:**27.

BACKGROUND

The species concept is one of the oldest and fundamental topics in the field of Evolutionary Biology (Dobzhansky 1935). The nature of biological species has fascinated scientists from many disciplines for centuries, resulting in an overwhelming amount of literature and debate (de Queiroz 1998 2005; Mayden 1997 2002). However, species delimitation, which is intrinsically operational, has been confused with a conceptual problem involving the concept of species, leading to incompatibility about the limit between and the number of species, hampering a consensus for a common species delimitation. Within this debate, de Queiroz (2007) suggested a reconciliation through the 'Unified Species Concept' (hereafter USC). The USC reduced the conceptual debate about species and turned it into a methodological (operational) discussion, allowing the application of distinct and more effective techniques to delimit species (de Queiroz 2007) a practice which is being called integrative taxonomy. There is currently a general consensus among taxonomists that integrative taxonomy is the best way to delimit species (Taylor 1983; Giagrande 2003; Dayrat 2006). This approach advocates the use of complementary perspectives, incorporating morphological, molecular, phylogenetic, and ontogenetic evidence, among others (Dayrat 2006; Will et al. 2005; Schlick-Steiner et al. 2010; Yeates 2011), as it understands that the complexity of the speciation process requires that species limits should be defined and studied from diverse and complementary perspectives (Dayrat 2006).

Accurate species identification and description are essential for biodiversity conservation (Gomes et al. 2013), especially when we focus on the diversity of the Neotropical region, which

harbors the highest freshwater fish diversity on earth, with an estimated 9,000 species (Birindelli and Sidlauskas 2018), that remains relatively undocumented, with only ~6,200 species formally described (Reis et al. 2016; Albert et al. 2020). Recent research that incorporates DNA-based approaches reveals a high number of undescribed species in the tropics, evidencing an underestimation of species diversity in the region (Gomes et al. 2013). The emergence and popularization of tools, such as DNA-barcoding (Hebert et al. 2003), General Mixed Yule Coalescent (GMYC) (Pons et al. 2006), Poisson Tree Processes (PTP) (Zhang et al. 2013), Automatic Barcode Gap Discovery (ABGD) (Puillandre et al. 2012) and Assemble Species by Automatic Partitioning (ASAP) (Ward 2009), based on molecular methods (Carvalho et al. 2008; Wheller et al. 2013; Piálek et al. 2012; Říčan et al. 2023), has created more opportunities for taxonomists to delimit operational taxonomic units (OTUs), and, in the context of adopting an integrative framework (Říčan et al. 2023) that have been transforming the current Neotropical fish taxonomic panorama, especially for cichlids (Zhang et al. 2013; Říčan et al. 2019 2021; Lozano et al. 2022; Argolo et al. 2020; Ximenes et al. 2021; Alonso et al. 2019; de Oliveira et al. 2025; Willis et al. 2012; Bernardi and Bucciariarelli 1999; Bentancur et al. 2013).

The cichlid genus *Australoheros* Říčan & Kullander, 2006 comprises 17 valid species (Fricke et al. 2025; Lucena et al. 2022) of neotropical fish, commonly referred to as 'Cará' or 'Acará', being typically found in rivers and lakes along the South American Atlantic coast (Lucena et al. 2022). Initially, species of *Australoheros* were classified under the genus *Cichlasoma* Swainson, 1839 until its establishment as a new genus, based on cytochrome *b* (CYTB) sequences and morphology (Říčan and Kullander 2006). Since then, several studies have been published, including new species descriptions based on morphological data only and within an integrative framework, coupling morphology and DNA-based analyses, mainly CYTB-based analyses (for a complete list, see Table S1). Ottoni et al. (2019) and Silva et al. (2022), based on molecular analyses, recovered the *Australoheros* Ottoni & Costa, 2008 species group as a well-supported clade divided into three subgroups: the Northern Mata Atlântica clade, Southern Mata Atlântica clade, and Upper/Middle Paraíba do Sul river basin and adjacent drainages clade (hereafter UMPSAD-clade).

However, the most recent contribution to the taxonomy of *Australoheros*, which proposed an overview of the diversity and phylogeny of the genus through morphological and molecular data on species from coastal rivers of southern South America (Lucena et al. 2022), questioned the aforementioned results, by synonymizing all the species within the Northern Mata Atlântica clade under *A. ipatinguensis* Ottoni & Costa, 2008. This implies the synonymy of *Australoheros*, which would theoretically have taxonomic priority for having been described first according to the International Code of Zoological Nomenclature – ICZN (1999), as well as synonymizing all the

species within the UMPSAD-clade under *A. oblongus* (Castelnau, 1855). The composition of the Southern Mata Atlântica clade was not questioned by Lucena et al. (2022). A complete list of these species, including their geographic distribution is available in table S1.

Lucena et al. (2022) adopted the bPTP method, which is a single-locus species delimitation coalescent method, based on the assumption that interspecific variability exceeds intraspecific variability. Thus, it is possible to delimit species using the number of nucleotide substitutions to model the speciation rate, without requiring ultrametric phylogeny calibration (Zhang et al. 2013). However, the species delimitation results for *Australoheros* presented by Lucena et al. (2022) rely on results potentially compromised by methodological limitations, since a basic assumption of the method was violated: they did not remove identical haplotypes from their DNA matrix and used phylogenetically distant outgroups (Zhang et al. 2013). The input file for bPTP analyses must include only a single individual per haplotype, in order to avoid generating an unrealistic number of species (Zhang et al. 2013; Blair and Bryston 2017; de Souza et al. 2021). Thus, in order to reanalyze the dataset of Lucena et al. (2022), we reduced it to unique haplotypes, creating the database for our study. Moreover, on the bPTP online server (species.h-its:), there is an option that allows researchers to "Remove outgroups that are distant related can improve the delimitation results". Thus, the inclusion of outgroups that are phylogenetically distant from the ingroup may negatively affect the results, leading to the delimitation of fewer lineages. Further, that work (Lucena et al. 2022) concludes that the *COI* gene is more suitable for delimiting species of *Australoheros*, despite the exclusion, without explicit methodological justification, of the results obtained by their CYTB analysis. It is important to emphasize that all taxonomic studies (including species delimitation studies) including a molecular component on the genus *Australoheros*, up to the publication by Lucena et al. (2022), had always relied on CYTB analyses (Lucena et al. 2022; Říčan and Kullander 2006 2008; Ottoni et al. 2019; Silva et al. 2022; Blair and Bryston 2017; de Souza et al. 2021; Říčan and Říčanová 2017; Říčan et al. 2011), including: the genus formal description, and the testing of its monophyly and phylogenetic position (Říčan and Kullander 2006), as well as the proposal of species groups within *Australoheros* (de Souza et al. 2021; Říčan and Říčanová 2017; Říčan and Kullander 2008).

Coalescence methods, such as the bPTP, follow the Coalescence Theory (Kingman 1982a 1982b 1982c). This theory employs genealogical trees to model the evolutionary history of a population in a process referred to as "allele tracking", which determines the point in time in which the analyzed locus coalesce into a most recent common ancestor (MRCA) (Mather et al. 2019; Marjoram and Wall 2006). Among coalescent approaches, bPTP stands out for allowing species delimitation without ultrametric calibration, making it suitable for comparative testing (Zhang et al. 2013). However, to ensure the reliability of these analyses, it is essential that the method's premises

are strictly followed. The lack of adequate rooting (that is, the incorrect choice of the tree's rooting point), as well as poor outgroup selection significantly influences the outcomes of species delimitation (Zhang et al. 2013). Moreover, using evolutionarily distant outgroups may increase phylogenetic uncertainty, leading to alterations in the number of delimited Operational Taxonomy Units (OTUs) (DeSalle and Tessler 2023).

Thus, we hypothesize that, the closer the outgroup is phylogenetically to the ingroup, the better is the analysis resolution (*i.e.*, more OTUs will be delimited), regardless of the marker. The main objective of this study is to test the efficiency of the bPTP method using different outgroups with different phylogenetic positions relative to the ingroup, for both the CYTB and *COI* markers. To test this, we reassess results of the most recent work on the taxonomy of *Australoheros* (Lucena et al. 2022), using their same in-group dataset, haplotype identifications (Table 1), and the same parameters (*e.g.*, best-fit evolutionary model), however, without violating the basic theoretical assumptions of the bPTP method (*i.e.*, grouping identical haplotypes into a single terminal, including only one individual per haplotype). For this, we first conduct a corrected bPTP species delimitation analysis on the datasets of both CYTB and *COI* matrices provided by them (Lucena et al. 2022). Then, we gradually add phylogenetically distant species of cichlids as outgroups to evaluate changes in the number of delimited OTUs and predict a progressive decline in OTU number with more distant outgroups.

Finally, we discuss and compare the differences and similarities of our results, focusing on the impacts of distinct outgroups and genes. Therefore, this work is not intended to make taxonomic decisions, but to reinforce the need for careful outgroup selection and respecting assumptions of species delimitation methods for accurately conducting taxonomic works.

MATERIALS AND METHODS

Taxon sampling

All sequences used in this work were retrieved from GenBank (www.ncbi.nlm.nih.gov/genbank/). As ingroups, we used 107 terminals (Table 1), and, as outgroups, seven species belonging to different Cichlidae tribes and subfamilies (*i.e.*, phylogenetically distant taxa). The chosen outgroups exhibit progressively greater phylogenetic distance in each successive analysis relative to *Australoheros* (Table 2). For outgroup taxa outside of *Australoheros*, we selected species for which both *COI* and CYTB sequences were available on public databases. Two mitochondrial gene fragments were analyzed: cytochrome *b* (CYTB) and

cytochrome *c* oxidase (*COI*), resulting in a total of 254 sequences examined. Institutional acronyms follow Sabaj (2020).

Table 1. Terminals of *Australoheros* utilized in this study for *COI* and CYTB gene matrices, with GenBank accession numbers, collection identifiers, and collection site. The "-" symbol denotes the lack of dataset or information for that gene

Species	Collection	Cat. no	Tissue no.	COI	CYTB
A. acaroides (Hensel, 1870)	MCP	48666	CIC169	MZ984107.1	-
A. acaroides	NRM	48666A	CIC170	MZ969532.1	-
A. acaroides	LBP	14581	CIC186	MZ969540.1	
A. acaroides	NRM	61491	12623	-	OK020133.1
A. acaroides	NRM	53174_D1	D1	MZ969602.1	-
A. acaroides	MCP	21272	CIC31	MZ969576.1	-
A. acaroides	MCP	50424	CIC321	MZ969583.1	OK020116.1
A. acaroides	MCP	49695	CIC311	-	OK020114.1
A. acaroides	MCP	48993	CIC313	MZ969580.1	-
A. acaroides	MCP	48688	CIC34	-	OK020113.1
A. acaroides	MCP	50425_F5	CIC322	-	OK020117.1
A. acaroides	MCP	51163_A	CIC327	MZ969589.1	-
A. acaroides	MCP	70180	12623	-	OK020099.1
A. acaroides	MCP	70181	12786	-	OK020095.1
A. acaroides	MCP	61491	8180	MZ969519.1	OK020094.1
A. acaroides	UFRJ	10499.3	CIC296	-	MK414421.1
A. acaroides	NRM	61427	8187	MZ969520.1	OK020093.1
A. angiru Říčan, Piálek, Almirón & Casciotta, 2011	MCP	51293_ (36)	CIC330	MZ969593.1	-
A. facetus Jenyns, 1842	MCP	55720	4514	-	OK020098.1
A. facetus	-	-	Arg.	-	AY998667.1
A. facetus	NRM	52591	3950	MZ969507.1	OK020097.1
A. facetus	-	-	-	-	AY998665.1
A. facetus	NRM	49538	1163	-	AY998666.1
A. facetus	NRM	52565	3927	MZ969506.1	-
A. facetus	UNMDP-T	359		JX111689.1	-
A. forquilha	-	-	A23	-	HQ197708.1
A. forquilha	-	-	A22	-	HQ197707.1
A. ipatinguensis	UNV	1261	CIC85	MZ969598.1	-
A. ipatinguensis	UNV	1306	CIC86	MZ969599.1	-
A. ipatinguensis	UNV	1312	CIC87	MZ969600.1	-
A. ipatinguensis	UFRGS	19024	CIC130	MZ969526.1	OK020131.1
A. ipatinguensis	UFRGS	19072	CIC125	MZ969522.1	OK020122.1
A. ipatinguensis	UFRGS	19062	CIC126	MZ969523.1	OK020121.1
A. ipatinguensis	UFRGS	19037	CIC127	MZ969524.1	-
A. ipatinguensis	UFRGS	18879	CIC128	MZ969525.1	OK020120.1
A. ipatinguensis	UFRGS	18937	CIC131	MZ969527.1	-
A. ipatinguensis	CZNC	1027	CIC161	MZ969530.1	-
A. ipatinguensis	MNRJ	47279	CIC282	MZ969549.1	-
A. ipatinguensis	MNRJ	47279	CIC283	MZ969550.1	OK020119.1
A. ipatinguensis	CZNC	1099	CIC162	-	OK020118.1
A. ipatinguensis	UFRJ	7567.2	-	-	MK414414.1
A. ipatinguensis	UFRJ	UFRJ 9853.2	-	-	MK414398.1

A. ipatinguensis	UFRJ	UFRJ 9853.2	-	-	MK414396.1
A. ipatinguensis	UFRJ	UFRJ 9853.1	-	-	MK414395.1
A. ipatinguensis	UFRJ	UFRJ 8393.4	-	-	MK414394.1
A. ipatinguensis	UFRJ	8506.1	-	-	MK414390.1
A. ipatinguensis	UFRJ	9511.4	-	-	MK414377.1
A. kaaygua Casciotta, Almirón & Gómez, 2006	-	-	H1	-	HQ197686.1
A. minuano Říčan & Kullander, 2008	MCP	55152	4079	MZ969513.1	OK020105.1
A. minuano	MCP	54196	3911	MZ969505.1	OK020102.1
A. minuano	MCP	61526	8106	-	OK020107.1
A. minuano	NRM	54147	3973	MZ969509.1	-
A. minuano	NRM	55068	4015	MZ969511.1	-
A. oblonga	MNRJ	47247	CIC285	MZ969551.1	OK020124.1
A. oblongus	MNRJ	47365	CIC301	MZ969567.1	-
A. oblongus	MNRJ	47336	CIC295	MZ969561.1	-
A. oblongus	MNRJ	47365	CIC302	-	OK020127.1
A. oblongus	MNRJ	47336	CIC296	-	OK020132.1
A. oblongus	UFRJ	UFRJ 10634.2	-	-	MK414399.1
A. oblongus	UFRJ	UFRJ 8291.3	-	-	MK414386.1
A. oblongus	UFRJ	UFRJ 8565.2	-	-	MK414379.1
A. oblongus	UFRJ	7823.3	-	-	MK414408.1
A. oblongus	UFRJ	9840.1	-	-	MK414367.1
A. ribeirae Ottoni et al., 2008	LBP	2133	CIC189	MZ969543.1	OK020128.1
A. ribeirae	UFRJ	10495.4	-	-	MK414425.1
A. ribeirae	UFRJ	7825.1	-	-	MK414405.1
A. ribeirae	UFRJ	7828.2	-	-	MK414404.1
A. ribeirae	UFRJ	7828.1	-	-	MK414403.1
A. ricani Lucena, Kullander, Norén & Calegari, 2023	MCP	50427_F9	CIC323	MZ969585.1	OK020130.1
A. ricani	MCP	50427_F10	CIC326	-	OK020129.1
A. sanguineus Ottoni, 2013	MCP	40635_E6	E6	MZ969603.1	-
A. sanguineus	MCP	40635_E7	E7	-	OK020109.1
A. sanguineus	UFRJ	10502.4	-	-	MK414418.1
A. scitulus Říčan & Kullander, 2003)	NRM	41626	116	-	AY998662.1
A. scitulus	-	-	A20	-	HQ197705.1
A. scitulus	-	-	6	-	HQ197701.1
A. scitulus	NRM	55338	4858	-	OK020111.1
A. scitulus	-	-	A21	-	HQ197706.1
A. scitulus	-	-	H17	-	HQ197702.1
A. scitulus	NRM	33048	Arg	-	AY998663
A. scitulus	NRM	52247	3321	MZ969503.1	-
A. sp. "Arapey"	NRM	52219	3322	MZ969504.1	-
A. tembe Casciotta, Gomez & Toresanni 1995	-	-	H3	-	HQ197688.1
A. tembe	-	-	H2	-	HQ197687.1
A. tembe	-	-	-	-	AY998660.1
A. tembe	-	-	-	DSFRE200-08	-
A. tembe	-	-	-	DSFRE330-08	-
A. ykeregua Říčan, Piálek, Almirón & Casciotta, 2011	-	-	H15	-	HQ197700.1

A. ykeregua	-	-	H14	-	HQ197699.1
A. ykeregua	-	-	H13	-	HQ197698.1
A. ykeregua	-	-	H12	-	HQ197697.1
A. ykeregua	-	-	H11	-	HQ197696.1
A. ykeregua	-	-	H10	-	HQ197695.1
A. ykeregua	-	-	H9	-	HQ197694.1
A. ykeregua	-	-	H5	-	HQ197690.1

Table 2. Outgroups utilized in this study for *COI* and CYTB gene matrices, along with their corresponding accession numbers in GenBank

N°	Outgroups	*COI*	CYTB
1	*Australoheros scitulus* (*COI*) and *A. kaaygua (*CYTB): Tribe Heroni	MZ969503.1	HQ197686.1
2	*Caquetaia spectabilis* (Steindachner, 1875)*:* Tribe Heroni	OR732852.1	AF370671.1
3	*Cichlasoma portalegrense* (Hensel, 1870): Tribe Cichlasomatini	MZ969528.1	U88854.1
4	*Geophagus diamantinensis* Mattos, Costa & Santos, 2015: Tribe Geophaghini	MH538087.1	KT373995.1
5	*Astronotus crassipinnis* (Heckel, 1840): Tribe Astronotini	GU701862.1	AF370650.1
6	*Cichla piquiti* Kullander & Ferreira, 2006: Tribe Cichlini	KT382897.1	GU295667.1
7	*Oreochromis niloticus* (Linnaeus, 1758) subfamily Pseudocrenilabrinae	KU565863.1	OL989429.1

Sequence edition and alignment

All methodological criteria in this study, such as sequence selection, trimming, and alignment procedures, are identical to Lucena et al. (2022). The sequences were edited using MEGA v.11 (Tamura et al. 2021) and aligned with the MUSCLE algorithm (Edgar 2004). According to Lucena et al. (2022), the sequence ends were manually trimmed, resulting in final alignments of 654 base pairs (bp) for the COI matrix and 1099 bp for the CYTB matrix. The alignment of the *COI* gene included 104 variable sites, of which 84 were phylogenetically informative. The CYTB alignment had 236 variable sites, including 211 informative sites. Subsequently, the aligned fragments were translated to check for the presence of premature stop codons or indels. The initial matrix, which included the same terminals used by Lucena et al. (2022), was reduced to unique haplotypes using DAMBE v. 7.3.32 (Xia 2004). After clustering identical sequences into the same haplotype, the number of haplotypes fell by 66.1% in *COI* and 49.3% in CYTB, creating the first matrices of this study, which contain 69 terminals for CYTB and 40 terminals for the *COI* matrix (Table 3). For the following analyses, the previous outgroup is retained for each new matrix generated, and the next one is added. Thus, starting from the first matrix, the total number of taxa will be the previous total plus one; for example, matrix 2 will have 70 taxa for CYTB and 41 for *COI*. In matrix 8, only the last outgroup was added and used as the root for the analysis.

Table 3. Comparison of the number of terminals used in the previous study and in the present study. Percentages represent the relative amount of haplotypes used in our analyses after reduction to single haplotypes

Gene	Lucena et al.(2022) dataset	Present study dataset
COI	118	40 (33,9%)
CYTB	136	69 (50,7 %)

Phylogenetic reconstruction

As proposed by Lucena et al. (2022), the best-fit evolutionary model selected for both *COI* and CYTB gene matrices was GTR + I. Bayesian Inference (BI) analyses were conducted using the CIPRES interface (www.phylo:) with the MrBayes v3.2.7 tool (Ronquist and Huelsenbeck 2003) to generate the topology for all structured matrices. Each tree was rooted using the most external outgroup corresponding to the eight matrices constructed for each gene. We conducted two Markov Chain Monte Carlo runs of 10 million generations each, with sampling occurring every 1,000 generations. The first 25% of the samples were discarded as burn-in, and the remaining samples were used to construct a consensus tree. All parameters of these phylogenetic analyses were verified using Tracer v1.7.2 (Rambaut et al. 2021), and reached convergence, with effective sample size (ESS) values exceeding 200. The tree topologies were visualized and assessed using FigTree v1.4.4 (Rambaut 2018).

Species delimitation analyses

The resulting BI trees were used as the inputs for species delimitation analyses. The selected method for species delimitation was the Bayesian Poisson Tree Processes (bPTP) analysis, conducted through the bPTP web server interface (species.h-its:), using default parameters: 500,000 MCMC generations, a thinning interval of 100, two independent MCMC chains and a burn-in of 10%. Comparisons between the OTUs of the species delimitation trees from Lucena et al. (2022) with those provided in this study can be found in table S2 and S3.

Calculation of Genetic Distances

Pairwise genetic distances were computed using MEGA v.11 (Tamura et al. 2021) using the p-distance based on aligned *COI* and CYTB sequences for two distinct analyses. First, we calculated the average genetic distance between the most phylogenetically distant outgroup (*i.e.*, the

root) and the ingroup taxa in each analysis. These values were then used to evaluate the influence of outgroup divergence on the number of OTUs recovered. Then, p-distances were also calculated between all pairs of sequences for each CYTB and *COI* matrix.

RESULTS

A total of 16 species delimitation analyses were conducted, with eight analyses performed for each gene based on the various outgroups tested, consistent decrease in OTUs with increasing outgroup distance. Tables 4 and 5 summarize the species delimited by the *COI* and CYTB gene matrices. We analyzed the average p-distance between the outgroup where the analysis was rooted and the ingroup. These data are available in table 4. Figure 1 illustrates the correlation between the number of OTUs recovered in each analysis and genetic distances from the root for both the *COI* and CYTB markers.

Table 4. Lineages of *Australoheros* delimited based on *COI* gene matrices. The term "OTU" stands for Operational Taxonomic Unit. We indicate the number of identified OTUs for each species in the analysis between parentheses. The "+" symbol denotes haplotypes of distinct species that were classified as belonging to the same unique lineage by the species delimitation analysis. The term "(o.g.)" following the names of certain species indicates that they were used as the outgroup in the analyses, while the term "(root)" designates the species at which the Bayesian Inference analysis was rooted

Analysis	Number of OTUs (including outgroups)	Number of OTUs of *Australoheros*	Delimited OTUs of *Australoheros*	Outgroups (o.g.) and root for each analysis	Average *p*-distance between root and ingroups
1	15	15	*A. acaroides* (one OTU); *A. angiru* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* (four OTUs); *A. minuano* (one OTU); *A. oblongus* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (one OTU); *A. sp* "Arapey" (one OTU).	*A. scitulus* (o.g.; root)	0.058
2	11	10	*A. acaroides* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* (one OTU); *A. minuano* + *A. sp* "Arapey" (one OTU); *A. oblongus* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* + *A. angiru* (one OTU); *A. tembe* (one OTU).	*A. scitulus* (o.g) and *Caquetaia spectabilis* (root)	0.148

3	11	9	*A. acaroides* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* + *A. oblongus* (one OTU); *A. minuano* + *A. sp* "Arapey"(one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* + *A. angiru* (one OTU); *A. tembe* (one OTU).	*A. scitulus* (o,g.), *Caquetaia spectabilis* (o.g.) and *Cichlasoma portalegrense* (root)	0.205
4	12	9	*A. acaroides* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* + *A. oblongus* (one OTU); *A. minuano* + *A. sp* "Arapey"(one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* + *A. angiru* (one OTU); *A. tembe* (one OTU).	*A. scitulus* (o.g.), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.) and *Geophagus diamantinensis* (root)	0.236
5	13	9	*A. acaroides* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* + *A. oblongus* (one OTU); *A. minuano* + *A. sp* "Arapey"(one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* + *A. angiru* (one OTU); *A. tembe* (one OTU).	*A. scitulus* (o.g.), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.), *Geophagus diamantinensis* (o.g) and *Astronotus crassipinis* (root)	0.226
6	14	9	*A. acaroides* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* + *A. oblongus* (one OTU); *A. minuano* + *A. sp* "Arapey"(one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* + *A.*	*A. scitulus* (o.g.), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.), *Geophagus diamantinensis* (o.g), *Astronotus crassipinis* (o.g) and *Cichla piquiti* (root)	0.211

			angiru (one OTU); *A. tembe* (one OTU).		
7	15	9	*A. acaroides* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* + *A. oblongus* (one OTU); *A. minuano* + *A. sp* "Arapey"(one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* + *A. angiru* (one OTU); *A. tembe* (one OTU).	*A. scitulus* (o.g.), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.), *Geophagus diamantinensis* (o.g), *Astronotus crassipinis* (o.g), *Cichla piquiti* (o.g) and *Oreochromis niloticus* (root)	0.214
8	11	10	*A. acaroides* (one OTU); *A. facetus* + *A. minuano* (one OTU); *A. ipatinguensis* (one OTU); *A. minuano* + *A. sp* "Arapey" (one OTU); *A. oblongus* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* + *A. angiru* (one OTU); *A. tembe* (one OTU).	*A. scitulus* (o.g) and *Oreochromis niloticus* (root)	0.214

Table 5. Lineages of *Australoheros* delimited based on CYTB gene matrices. The term "OTU" stands for Operational Taxonomic Unit. We indicate the number of identified OTUs for each species in the analysis between parentheses. The "+" symbol denotes haplotypes of distinct species that were classified as belonging to the same unique lineage by the species delimitation analysis. The term "(o.g.)" following the names of certain species indicates that they were used as the outgroup in the analyses, while the term "(root)" designates the species at which the Bayesian Inference analysis was rooted

Analysis	Number of OTUs (including outgroups)	Number of OTUs of *Australoheros*	Delimited OTUs of *Australoheros*	Outgroups (o.g.) and root for each analysis	Average p-distance between root and ingroups
1	24	24	*A. acaroides* (five OTUs); *A. facetus*	*A. kaaygua* (o.g.; root)	0.0411

			(one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (two OTUs), *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* (three OTUs); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (three OTUs).		
2	21	20	*A. acaroides* (five OTUs), *A. facetus* (one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (one OTU); *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (three OTUs).	*A. kaaygua* (o.g) and *Caquetaia spectabilis* (root)	0.128
3	18	16	*A. acaroides* (one OTU); *A. facetus* (one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (one OTU); *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (three OTUs).	*A. kaaygua* (o.g), *Caquetaia spectabilis* (o.g.) and *Cichlasoma portalegrense* (root)	0.157

4	17	14	*A. acaroides* (one OTU); *A. facetus* (one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (one OTU); *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (one OTU).	*A. kaaygua* (o.g), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.) and *Geophagus diamantinensis* (root)	0.154
5	18	14	*A. acaroides* (one OTU); *A. facetus* (one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (one OTU); *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (one OTU).	*A. kaaygua* (o.g), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.), *Geophagus diamantinensis* (o.g) and *Astronotus crassipinis* (root)	0.166
6	19	14	*A. acaroides* (one OTU); *A. facetus* (one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (one OTU); *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (one OTU).	*A. kaaygua* (o.g), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.), *Geophagus diamantinensis* (o.g), *Astronotus crassipinis* (o.g) and *Cichla piquiti* (root)	0.154

7	20	14	*A. acaroides* (one OTU); *A. facetus* (one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (one OTU); *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (one OTU).	*A. kaaygua* (o.g), *Caquetaia spectabilis* (o.g.), *Cichlasoma portalegrense* (o.g.), *Geophagus diamantinensis* (o.g), *Astronotus crassipinis* (o.g), *Cichla piquiti* (o.g) and *Oreochromis niloticus* (root)	0.203
8	17	16	*A. acaroides* (one OTU); *A. facetus* (one OTU); *A. forquilha* (one OTU); *A. ipatinguensis* (one OTU); *A. kaaygua* (one OTU); *A. minuano* (one OTU); *A. oblongus* + *A. ribeirae* (one OTU); *A. ricani* (one OTU); *A. ribeirae* (one OTU); *A. sanguineus* (one OTU); *A. scitulus* (one OTU); *A. tembe* (two OTUs); *A. ykeregua* (three OTUs).	*A. kaaygua* (o.g) and *Oreochromis niloticus* (root)	0.203

Species delimitation analyses based on *COI*

Overall, the number of OTUs delimited in *Australoheros* using the *COI* gene varied from a maximum of 15 (with *A. scitulus* as the outgroup) to a minimum of 9 (with *C. portalegrense*; *G. diamantinensis*, *A. crassipinis*, *C. piquiti* and *O. niloticus* as the outgroups), showing a decrease in OTU count as phylogenetic distance to the outgroup increased (See Table 4 and Fig. 1A). The first COI matrix considered *A. scitulus* as the outgroup and root, delineating 15 OTUs within the genus *Australoheros*. Among these, four distinct OTUs corresponded to *A. ipatinguensis*, while additional 10 OTUs were represented by unique haplotypes, excluding the outgroup (Fig. 2A). The previous outgroups were retained from the second analysis onwards and new outgroups were introduced (Fig. 2B). For this analysis, *Caquetaia spectabilis* was incorporated as the outgroup and tree root; 10 OTUs of *Australoheros* were delimited. In the third analysis, *Cichlasoma portalegrense* was added as the outgroup and root, resulting in the delimitation of 11 OTUs, three of which corresponded to the outgroups utilized in this matrix, while nine were from *Australoheros* (Fig. 2C). The same OTUs from *Australoheros* delimited in the third analysis, were also delimited in the fourth through seventh analyses, resulting in a total of 12, 13, 14 and 15 OTUs, respectively, following the incorporation of their outgroups: *Geophagus diamantinensis*, *Astronotus crassipinnis*, *Cichla piquiti* and *Oreochromis niloticus* (See Table 4).

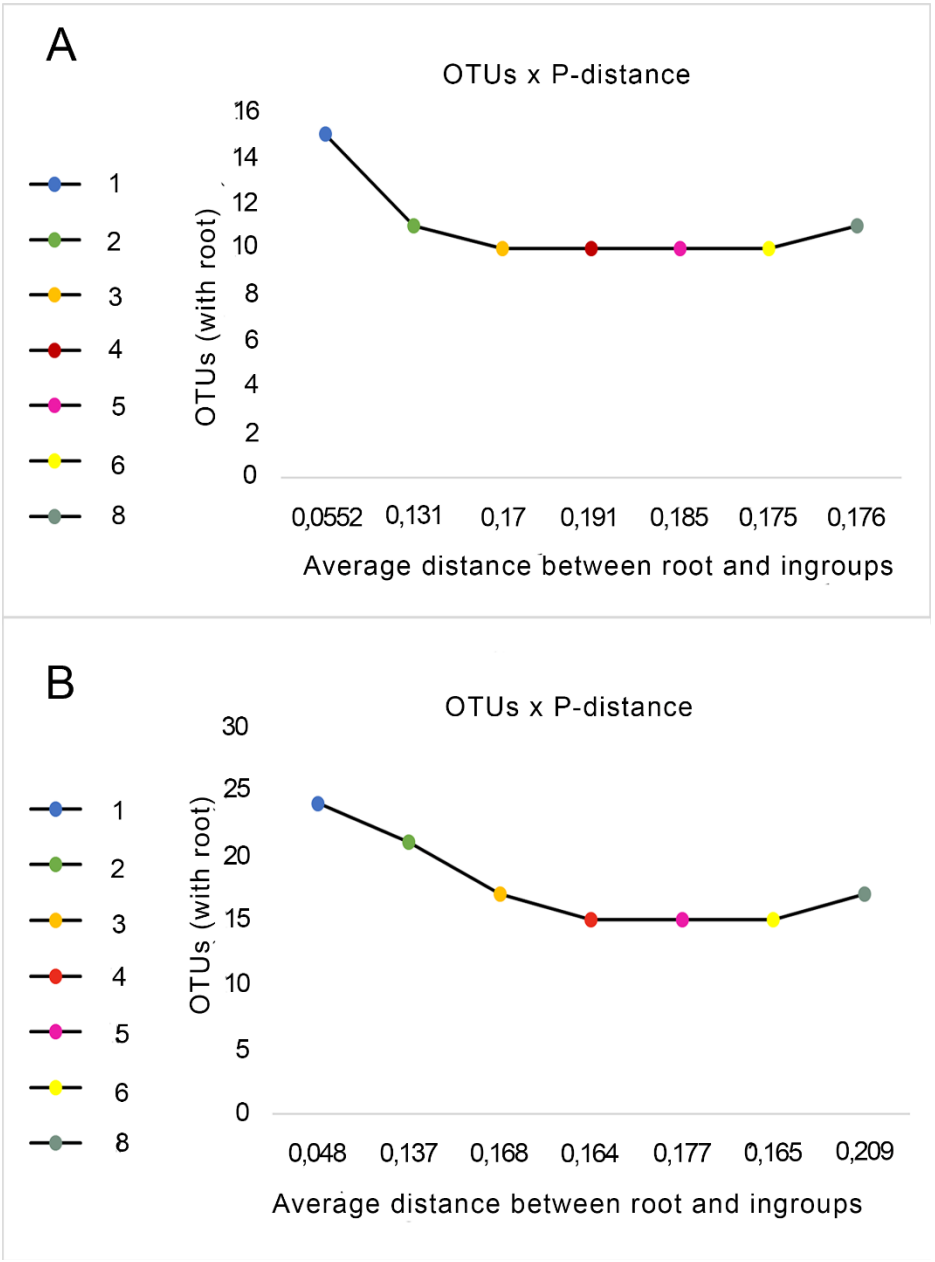

Fig. 1. Relationship between the number of delimited OTUs and the average genetic distance between the ingroup and the root of the analysis. A: *COI* marker; B: CYTB marker. The x-axis indicates the average phylogenetic distance between the ingroup and the root of the analysis, and the y-axis shows the corresponding number of OTUs recovered. The colored dots on the graph indicate each of the different analyses performed in the study. Analysis 7 is not represented in the figure because we chose to analyze only the most distant outrgroup (*i.e.*, the root of each analysis); therefore, analyses 7 and 8 coincide. The dash indicates the gene represented in the graph (*COI* or CYTB).

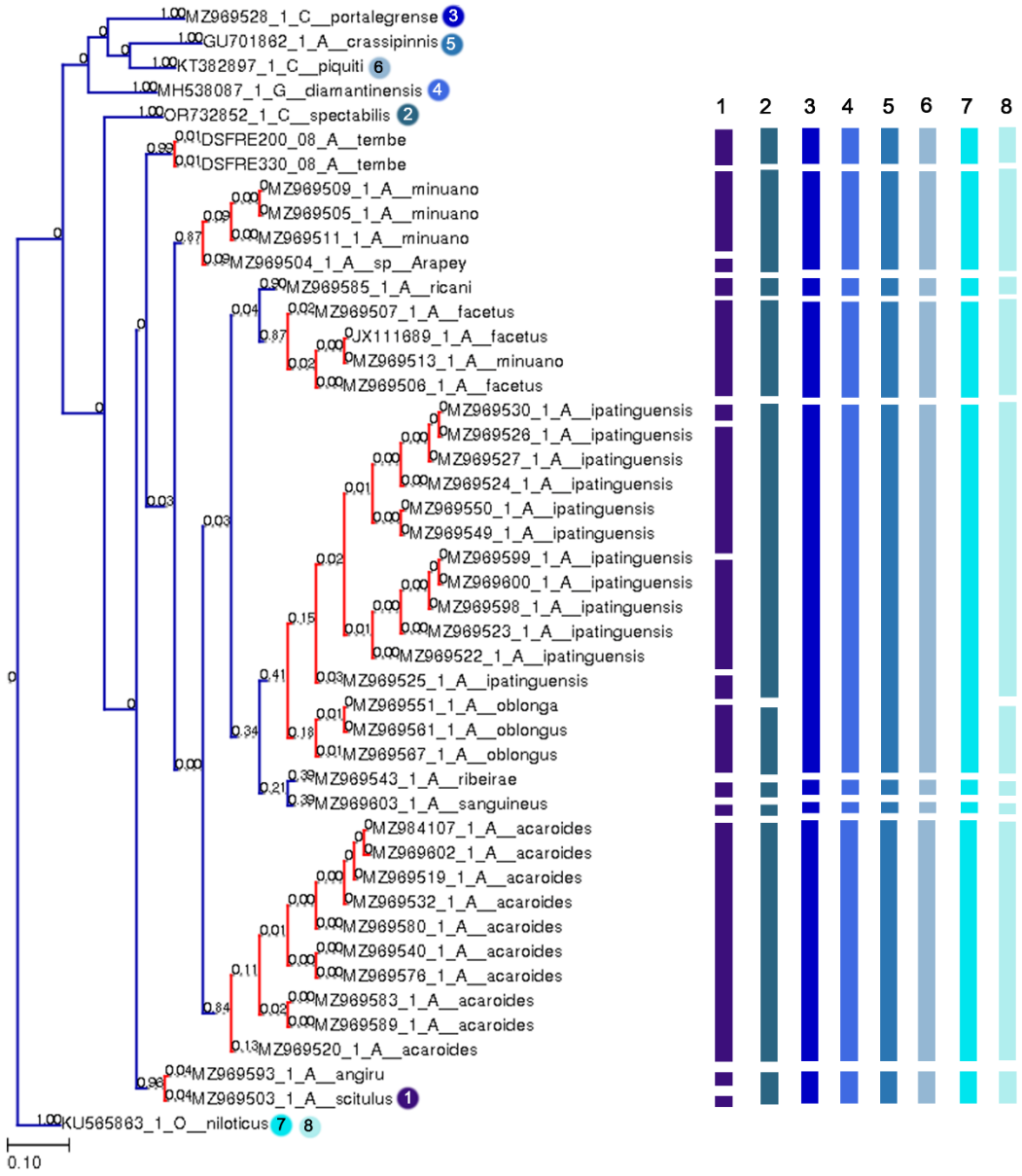

Fig. 2. Species delimitation analyses for the gene *COI*. The term "r" designates the species in which the analysis was rooted. Node values represent Bayesian Support (BS). Branch lengths are not proportional to genetic distances; no scale bar is shown. A: Tree 1 with 15 delimited OTUs of *Australoheros*; *A. scitulus* (r). B: Tree 2 (similar to analysis 8 output) with 10 OTUs of *Australoheros*; *Caquetaia spectabilis* (r). C: Tree 3 (also represents outputs from analyses 4, 5, 6 and 7) with 9 OTUs of *Australoheros*; *Cichlasoma portalegrense, Geophagus diamantinensis*, *Astronotus crassipinnis*, *Cichla piquiti* and *Oreochromis niloticus*, (r, respectively). For an extensive detailment of outputs of all analyses (1-8), see fig. S1.

In the eighth species delimitation analysis, the outgroups from matrices 2–7 were excluded, with the exception of *Oreochromis niloticus*. The results indicated the presence of 10 OTUs of *Australoheros* (Fig. 2B). The outputs of the eighth analyses are available as figure S1.

Species delimitation analyses based on CYTB

The CYTB analyses also exhibit a decrease in OTU count as the phylogenetic distance to the outgroup increased. OTUs of *Australoheros* ranged from 14 to 24, with the highest number (24) observed when *A. kaaygua* (an *Australoheros* species) was used and the lowest (14) when *G. diamantinensis*, *C. piquiti* and *O. niloticus* was selected (more distantly related outgroups) (Table 5; Fig. 1B). The first delimitation analysis of the CYTB matrices identified 24 OTUs of *Australoheros*, using *A. kaaygua* as the outgroup and root instead of *A. scitulus*, which was chosen as outgroup on the *COI* matrices. This choice was taken when collapsing identical haplotypes into a single terminal, which revealed the presence of seven distinct haplotypes of *A. scitulus*, rather than a single haplotype (Fig. 3A). The second analysis incorporated *Caquetaia spectabilis* as the outgroup and root, culminating in a total of 21 OTUs and resulting in the delimitation of 20 OTUs of *Australoheros* (Fig. 3B). The introduction of a third outgroup, *Cichlasoma portalegrense*, led to the delimitation of 18 OTUs, of which 16 corresponded to *Australoheros* (Fig. 3C). In the fourth analysis, the addition of *Geophagus diamantinensis* as the outgroup and root indicated 17 OTUs, of which 14 belong to *Australoheros* and three of which corresponded to the outgroups utilized in matrix (Fig. 3D). Analyses fifth to seventh delimited 14 OTUs from *Australoheros*, with the following outgroups used as roots, respectively: *Astronotus crassipinnis*, *Cichla piquiti* and *Oreochromis niloticus*. In the eighth analysis, only *Oreochromis niloticus* was retained as the outgroup (Fig. 3C) delimiting 16 OTUs of *Australoheros*. The outputs of the eight analyses are available as figure S2.

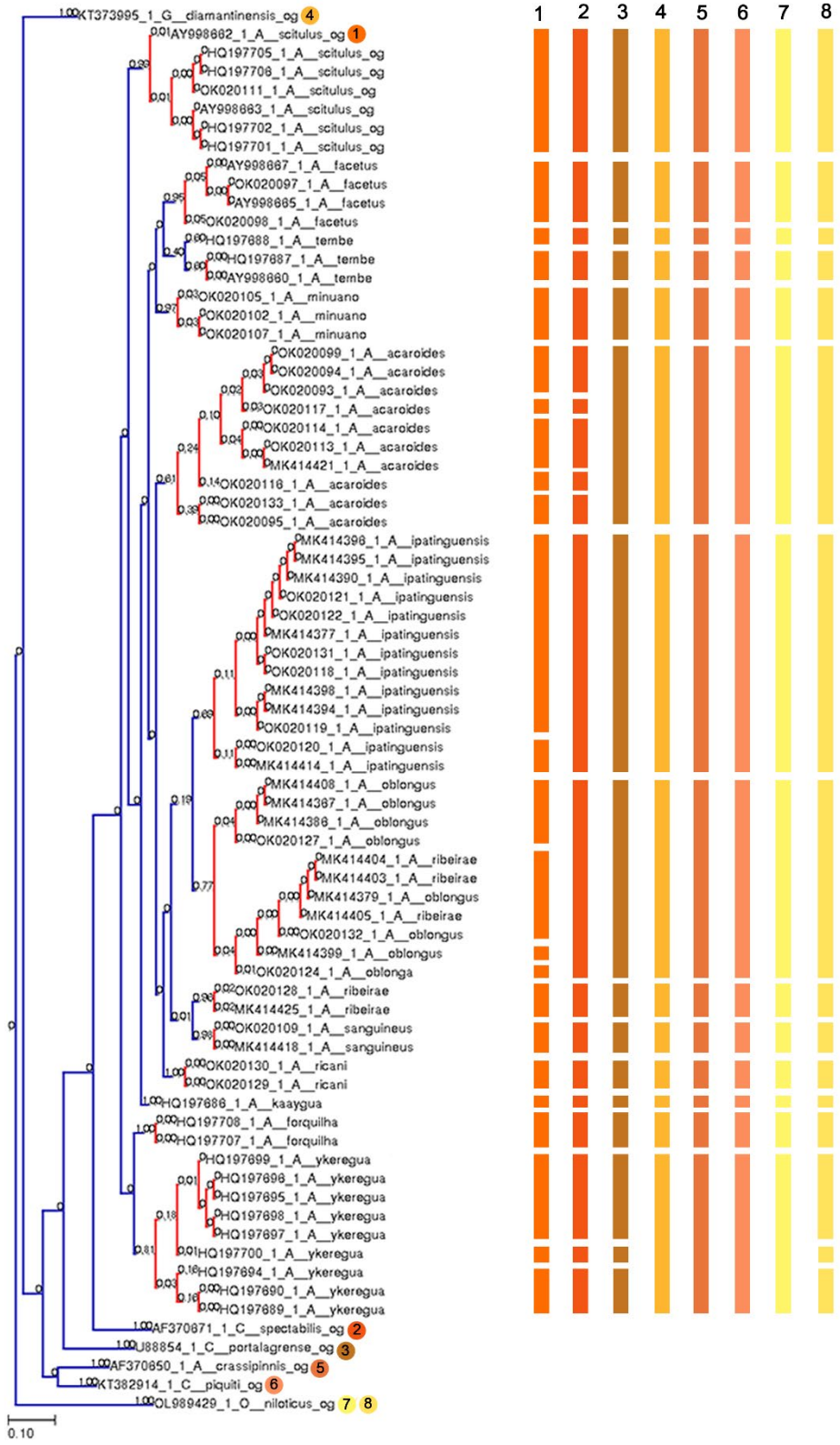

Fig. 3. Species delimitation analyses for CYTB. The term "r" designates which species the analysis was rooted. Node values represent Bayesian Support (BS). Branch lengths are not proportional to genetic distances; no scale bar is shown. A: Tree 1 with 24 delimited OTUs of *Australoheros*; *A. kaaygua* (r). B: Tree 2 with 20 delimited OTUs of *Australoheros, Caquetaia spectabilis* (r). C: Tree 3 (also represent tree 8) with 16 delimited

OTUs of *Australoheros*, *Cichlasoma portalegrense* and *Oreochromis niloticus* (r, respectively). D: Tree 4 (also represents trees 5, 6 and 7) with 14 delimited OTUs of *Australoheros*; *Geophagus diamantinensis*, *Astronotus crassipinis*, *Cichla piquiti* and *Oreochromis niloticus* (r, respectively). For an extensive detailment of outputs of all analyses (1-8), see fig. S2. For haplotypes details see table 1.

DISCUSSION

Our results corroborate our hypothesis that the use of phylogenetically distant species as outgroups decreases the number of delimited OTUs, indicating a possible impact on the sensitivity of the method. The analysis of the average genetic distance between the most distant outgroup in each analysis and the number of delimited OTUs reveals a negative trend, indicating that more distantly related outgroups tend to reduce the number of recovered OTUs. However, analysis 8 presented a distinct trend, suggesting that the exclusive use of a distant outgroup (without intermediate lineages) may affect the algorithm differently than a gradual inclusion of "midterms outgroups". These findings allowed us to test one of the premises of the bPTP species delimitation method, which suggests that the inclusion of closely related outgroup can enhance delimitation results (Zhang et al. 2013). When uploading your dataset to the bPTP online server (species.h-its:) there is a remark emphasizing: "Remove outgroups that are distant related can improve the delimitation results". Here, we discuss the underlying differences between our results (*i.e.*, the distinct lineages obtained in each analysis according to outgroup selection and gene) and between our analyses and that from Lucena et al. (2022).

Species delimitation analysis based on *COI* considering phylogenetically distinct outgroup selection

Eight analyses were conducted to investigate the impact of choosing phylogenetically distinct outgroups and the effectiveness of the *COI* gene in delimiting *Australoheros* species (Fig. S1). Table 4 summarizes the species delimited for the COI gene matrices in all analyses performed. In Analysis 1, using *A. scitulus* as the outgroup and root of the tree, 15 OTUs of *Australoheros* were delimited. These included *A. scitulus*, four distinct OTUs of *A. ipatinguensis* and 10 unique OTUs (see Fig. 2A and Table 4). In Analysis 2, with *Caquetaia spectabilis* as the outgroup and root, 10 OTUs of *Australoheros* were delimited. In this analysis, the previous four OTUs of *A. ipatinguensis* coalesced into a single OTU. *Australoheros scitulus* and *A. angiru* were grouped into the same

OTU, as were *A. minuano* and *A.* sp. "Arapey". The other OTUs remained unchanged from the previous analysis (see Fig. 2B and Table 4). The use of *Cichlasoma portalegrense* as outgroup and root of analysis 3 delimited nine OTUs of *Australoheros* (see Fig. 2C). Notably, *A. ipatinguensis* and *A. oblongus* were identified as belonging to the same OTU for the first time. The other OTUs were consistent with the previous analysis. Following analysis 3, up to analysis 7, the delimitation trees remained stable, with the same nine OTUs of *Australoheros*, consistently grouping according to their respective outgroups: *Geophagus diamantinensis*, *Astronotus crassipinis*, *Cichla piquiti* and *Oreochromis niloticus* [see Fig. 2D; see Fig. S1 (Fig. S1-1 to S1-7)]. In the final *COI* analysis, the use of *Oreochromis niloticus* as the only outgroup and root, separated *A. ipatinguensis* and *A. oblongus* into independent OTUs, consistent with the results obtained in Analysis 2 and delimited 10 OTUs within *Australoheros*.

These analyses demonstrate how the sensitivity of species delimitation is affected by outgroup selection, especially in taxa that have recently diverged or are closely related, such as *A. ipatinguensis* and *A. oblongus*. The minimum genetic distance (*p*-distance) between these two clades is approximately 1.69%, significantly lower than the 2–3% threshold commonly observed between distinct species of the genus *Australoheros*, according to Lucena et al. (2022), suggesting that these species are genetically very similar. These factors contribute to the recovery of *A. ipatinguensis* and *A. oblongus* as a single clade in analyses with phylogenetically distant outgroups, due to their low relative divergence. However, despite the low genetic distance between these clades, *A. oblongus* and *A. ipatinguensis* are considered valid species, based on morphological and phylogenetic analyses (Fig. 11 in Lucena et al. 2022). Pairwise *p*-distance data for *COI* marker is available in table S4.

Species delimitation analysis based on CYTB considering phylogenetically distinct outgroup selection

The experimental patterns applied to the *COI* gene were also tested for CYTB, resulting in eight analyses to delimit *Australoheros* species (see Fig. S2). Table 5 summarizes the species delimited for the CYTB gene matrices in all analyses performed. Analysis 1 utilized *A. kaaygua* as outgroup and root of the tree, delimiting 24 OTUs within the genus *Australoheros* (See Fig. 3A and Table 5). Notably, two haplotypes of *A. oblongus* and three of *A. ribeirae* coalesced into a single OTU, although (there were another two distinct OTUs corresponding to *A. ribeirae* and *A. oblongus*). This could possibly be due to a misidentification of the specimens that generated the haplotypes. In the second analysis, with *Caquetaia spectabilis* as the root, the delimitation resulted

in 20 OTUs (see Fig. 3B). *Australoheros ipatinguensis* were combined into a single OTU, while the three OTUs of *A. oblongus* were grouped with some haplotypes of *A. ribeirae*, while other haplotypes of *A. ribeirae* were recovered as valid independent OTUs (there were two distinct OTUs including haplotypes of *A. ribeirae*) (see Fig. 3B). This may also be due to a misidentification of the specimens that generated the haplotypes.

The five *A. acaroides* OTUs identified in Analyses 1 and 2 as distinct were consolidated into the same OTU in Analysis 3, with the addition of *Cichlasoma portalegrense* as the root (see Fig. 3C). The remaining OTUs remained unchanged, resulting in a total of 16 OTUs delimited for *Australoheros*. *Australoheros ykeregua* was recovered as a single OTU upon the inclusion of *Geophagus diamantinensis* as outgroup and root for Analysis 4 (see Fig. 3D). The delimitation trees remained stable from the fourth to the seventh analysis (see Fig. S2; Fig. S2-4 to S2-7). The seventh and eighth analyses were rooted in *Oreochromis niloticus*, but only the seventh analysis utilized the previously added outgroups. In the eighth analysis, *A. ykeregua* were recovered as representing three valid OTUs, upon the removal of previously added outgroups and the use of *Oreochromis niloticus* as the only outgroup and root for this analysis (see Fig. 3C).

The presence of some haplotypes identified as *A. ribeirae* within a clade composed primarily of *A. oblongus* haplotypes was recurrent in all analyses (see Fig. 3), prompting us to investigate the possible causes of this incongruence. We found that the haplotype associated with code OK020132 is registered in GenBank as *Australoheros* sp.; however, in table 1 provided by Lucena et al. (2022), this same code is identified as *A. oblongus*. Similarly, the haplotype MK414379, identified in GenBank as *Australoheros* cf. *montanus*, shows two different identifications in table S1 of Lucena et al. (2022). It is associated with *A. oblongus* when referring to the tissue sample used for DNA extraction (tissue species) and with *A. robustus* when referring to the voucher specimen or collection code from which the tissue originated (source ID). This type of identification conflict compromises phylogenetic interpretation, directly inferring the delimitation of OTUs in molecular analyses. Pairwise *p*-distance data for CYTB marker is available in table S4.

Comparing Species Delimitation analyses results between *COI* and CYTB

Here, the *COI* and CYTB datasets will be compared; however, it is important to note that these matrices differ in composition. The CYTB dataset includes 69 haplotypes, while the *COI* dataset contains only 40 haplotypes. There are fewer sequencing studies available for the *COI* within the same set of species compared to CYTB; consequently, there are fewer *Australoheros* sequences for the *COI* gene in GenBank. Thus, our results indicate that the matrices based on CYTB delimited a greater number of OTUs for *Australoheros* compared to those based on the *COI*

gene. However, this result is not particularly informative, as the CYTB dataset contains a higher number of haplotypes than *COI*. Therefore, the quantitative analysis of the delimited species based on the *COI* and CYTB genes is not feasible due to the discrepancy between dataset. There is evidence that CYTB, characterized by a rapid rate of evolution, is suitable for species delimitation approaches because it tends to accumulate more mutations over shorter time intervals (Reid and Carstens 2012; Avise 2000). Pairwise p-distance comparisons revealed smaller mean divergence values for the *COI* gene (mean = 0.039 ± 0.02) and larger values for CYTB gene (mean = 0.054 ± 0.02), supporting this evidence. However, it is important to note that the datasets for these two genes differ in size and composition (see Table S4 and S5). Furthermore, it is important to clarify that single-locus data alone are insufficient to capture inconsistencies between gene trees and the evolutionary history of species. Therefore, future studies should incorporate multilocus or genomic approaches or species tree analyses to address this limitation.

Although it is difficult to compare the sensitivity of the two genes using the bPTP method, the results obtained for *COI* and CYTB showed notable similarities. The results obtained for the *COI* gene indicate changes in the number of delimited OTUs from the first to the last test (see Table 4). Following the second *COI* analysis, outgroups outside the Heroini tribe were included. Analyses 3, 4, 5, 6 and 7 consistently identified the same nine OTUs: *A. acaroides*, *A. facetus* + *A. minuano*, *A. ipatinguensis* + *A. oblongus*, *A. minuano* + *A. sp*. "Arapey", *A. ricani*, *A. ribeirae*, *A. sanguineus*, *A. scitulus* + *A. angiru* and *A. tembe* (see Fig. S1). The addition of *Oreochromis niloticus* from the subfamily Pseudocrenilabrinae as the root in analysis 7 did not alter the number of OTUs. However, in analysis 8, when using only *Oreochromis niloticus* as root and single representative outside *Australoheros*, *A. ipatinguensis* and *A. oblongus* were once again identified as distinct OTUs, similar to the results obtained in analyses 2 (see Fig. S1). For *COI*, the stabilization of the species delimitation trees is achieved between the three and seventh analysis (see Fig. S1). Regarding the performance of CYTB, the delimited OTUs underwent changes from the first to the last analysis (see Table 5). In the fourth analysis, with the addition of *Geophagus diamantinensis* from the tribe Geophagini as an outgroup, 14 OTUs were recovered up to the seventh analysis (see Table 5). However, in the eighth analysis, which involved the removal of the previously added outgroups, *A. ykeregua* was identified as three distinct OTUs in the species delimitation results, similar to the results obtained in analyses 3 (See Fig. 3B). Therefore, for CYTB, the stabilization of the species delimitation trees is achieved between the fourth and seventh analysis (see Fig. S2).

For both *COI* and CYTB, we observed a decrease in the number of delimited OTUs from the second analysis on, influenced by the selection of outgroups (see Fig. 1). According to the principles of the bPTP method (Zhang et al. 2013), the use of outgroups that are taxonomically more distant from the genus affects the number of delimited species; consequently, the more

phylogenetically distant outgroup to which the tree is rooted, fewer the number of delimited species. The analyses tend to stabilize due to the method's loss of sensitivity. This means that short internal branches in the input tree may collapse and the bPTP software may fail to delimit them, instead recovering a single OTU that, actually, includes multiple distinct species when this occurs, the recovered OTUs will not change, even with the addition of new phylogenetically distant outgroups (see Table 4 and Table 5), likely because these OTUs that show no alteration are those that diverged earlier, at least more than 2.5 million years ago (m.y.a) according to Ottoni et al. (2019) and consequently have higher genetic divergences.

It has been observed that certain OTUs are recovered more frequently than others. Some OTUs have remained unchanged until the latest analyses. *Australoheros facetus*, *A. minuano*, *A. ricani*, *A. sanguineus* and *A. tembe* were consistently recovered in all analyses for both *COI* and CYTB genes. This recovery pattern is not coincidental; it is related to the divergence times of the species within the genus. Specifically, there is a correlation between the age of the OTUs, the consequently accumulation of nucleotidic changes, and the results obtained using the bPTP method. Ottoni et al. (2019) utilized molecular data to construct a dated phylogeny of the *A.* species group, which enhances the understanding of species divergence and coalescence analysis methods (Fig. 4 in Ottoni et al. 2019).

The authors indicate that the genus *Australoheros* emerged approximately 46 million years ago (m.y.a.) during the middle Eocene, and that species diversification started around 17.5 m.y.a, during middle Miocene (Ottoni et al. 2019). The lineages identified in the dating analyses proposed by Ottoni et al. (2019) that remain unchanged here were *A. facetus*, *A. minuano*, *A. ribeirae*, *A. sanguineus* and *A. tembe* for *COI* and CYTB, as well as *A. forquilha*, *A. kaaygua*, *A. ribeirae*, *A. sanguineus* and *A. tembe* for CYTB. Ottoni et al. (2019) present these lineages as having diverged up to 2.5 m.y.a., with the exception of *A. minuano*, which diverged only 0.7 million years ago, a result that suggests that *A. minuano* may be a species with a faster rate of evolution (Ottoni et al. 2019)

The divergence time of the lineages whose haplotypes clustered into single OTUs exhibited the greatest variation with the inclusion of different outgroups—*A. barbosae*, *A. macacuensis* and *A. robustus* (corresponding to *A. oblongus*); *A.autrani*, *A. ipatinguensis*, *A. muriae* and *A. macaensis* (corresponding to *A. ipatinguensis*)—date back to 2.6 million years ago or less (Ottoni et al. 2019). According to the divergence results presented by Ottoni et al. (2019), these same lineages are classified as more recent, indicating a shorter divergence time. Coalescent model-based delimitation methods, such as bPTP, are more effective in recovering lineages with slightly older divergence times (Ottoni et al. 2019), as demonstrated in the dated phylogeny of Ottoni et al. (2019). For species that have recently diverged or for clades undergoing rapid speciation, these

methods are unlikely to yield identifiable results under this model (Ottoni et al. 2019).

Comparison with the work by Lucena et al. (2022)

We compare our results with those of Lucena et al. (2022), with our analyses 1 conducted for both *COI* and CYTB because these analyses use external groups most closely related to our delimited group, thus providing more reliable results as dicussed above. However, it is important to emphasize that Lucena et al. (2022) did not reduce their matrices to include only one individual per haplotype, which may generate an unrealistic number of species, as previously discussed.
To understand the relationship between the OTUs used by Lucena et al. (2022) and those used in this present work, refer to tables S2 and S3. The *COI* species delimitation analysis conducted by Lucena et al. (2022) identified 12 OTUs within the genus *Australoheros*. The aforementioned study used *Cichlasoma portalegrense*, from the tribe Cichlasomatini, as outgroup; however, this taxon is not represented in their tree (fig. 12 in Lucena et al. 2022). The following OTUs were recovered in the *COI* analysis by the Lucena et al. (2022): *Australoheros tembe*; one OTU corresponding to *A. scitulus* and *A. angiru*; two OTUs of *A. canterai*; one corresponding to *A. minuano* and the other one to *A. sp*. "Arapey" in our analysis; *A. acaroides*; *A. ribeirae*; *A. sanguineus*; two OTUs of *A. ipatinguensis*; *A. oblongus*; *A. facetus* and *A. ricani* (see Table S2 to comparisons between the OTUs). In our first *COI* analysis, 15 OTUs were delimited, in contrast to the 12 OTUs delimited by Lucena et al. (2022). *Australoheros* sp. "Arapey" and *A. ricani* were found to be consistent with the findings of Lucena et al. (2022). The differences included the delimitation of four distinct OTUs of *A. ipatinguensis*, compared to the two OTUs indicated by Lucena et al. (2022). The main differences in analysis 2, relative to the results of Lucena et al. (2022), included a unique OTU for *A. ipatinguensis* in our analysis, as well as *A. minuano* and *A.* sp. "Arapey" as classified as a unique OTU, and the inclusion of *A. minuano* in the *A. facetus* clade (see Fig. 2B).

The CYTB analysis conducted by Lucena et al. (2022) delimited 22 OTUs within the *Australoheros*, utilizing *Amphilophus citrinellus* (Gunther, 1864) and *Neetroplus nematopus* (Gunther, 1867) as outgroups, both of which belong to the tribe Heroini. The analysis revealed the following OTUs: *A. scitulus*; two OTUs for *A. canterai* ceibal, corresponding to *A. minuano*; three OTUs for *A. acaroides*; *A. sanguineus*; *A. ribeirae*; an OTU comprising *A. perdi*, *A. ipatinguensis*, *A.* cf. *capixaba*, *A. macaenses* and *A. muriae*, all corresponding to *A. ipatinguensis*; an OTU of *A. cf*. and *A. saquarema*, also corresponding to *A. ipatinguensis*; two OTUs for *A. macacuensis*, corresponding to *A. oblongus*; an OTU consisting of *A. robustus* cachoeirinha, *A. mattosi*, *A.* sp. ouro branco, *A.* cf. *montanus* (corresponding to *A. oblongus*) and *A. robustus* (corresponding to *A. ribeirae*); an OTU including CIC Juiz de Fora, *A. tavaresi*, *A.* cf. *parabae*, *A. barbosae*, *A.* cf.

barbosae, and *A. parabae*, all corresponding to *A. oblongus*; *A. ricani*; *A. facetus*; two OTUs for *A. tembe*; *A. kaaygua*, three OTU for *A. ykeregua* and *A. forquilha* (see Table S3 to comparisons between the OTUs).

Twenty-four OTUs were recovered in our initial CYTB analysis, compared to the 22 obtained by Lucena et al. (2022). *Australoheros ykeregua*, *A. ipatinguensis*, *A. ribeirae*, *A. tembe*, *A. forquilha*, *A. scitulus*, *A. oblongus* + *A. ribeirae*, *A. sanguineus*, *A. ricani*, *A. kaaygua* and *A. oblongus* were consistently identified in both analyses. The main differences are related to the delimitation of five OTUs within *A. acaroides* in our analysis, as opposed to three in Lucena et al. (2022) and one OTU of *A. minuano* instead of two in Lucena et al. (2022). In the second analysis, 20 OTUs were delimitated, in contrast to the 22 OTUs recovered by Lucena et al. (2022). In our analysis, *Australoheros ipatinguensis* was classified as a single OTU and *A. oblongus* was not treated as two distinct OTUs, with its haplotypes grouped with those of *A. ribeirae*.

We do not intend to describe or question the validity of species at this time, although Lucena et al. (2022) considered *A.*, *A. saquarema*, *A. capixaba*, *A. macaensis*, *A. perdi*, and *A. muriae* as junior synonyms of *A. ipatinguensis*, even though several of these species are supported by morphological and molecular diagnoses (Lucena et al. 2022; Ottoni 2008 2010 2011); rather, we propose a review of the species delimitation analyses that have been conducted. In general, the initial species delimitation analyses for both COI and CYTB were not significantly discrepant compared to the results obtained by Lucena et al. (2022). Notably, the outgroup listed in the material and methods by Lucena et al. (2022) for the species delimitation of *COI* was not utilized, and this alteration was not justified. Lucena et al. (2022) indicates that the results obtained from the analysis of the *COI* gene are more satisfactory for the delimitation of *Australoheros* than the results obtained with the use of the CYTB gene, without presenting proper justification. As previously mentioned, however, it is important to emphasize this once again, all molecular and taxonomic studies (including species delimitation studies) on the genus *Australoheros*, as well as the taxonomy of the genus, up to the publication by Lucena et al. (2022) had always been based on CYTB analyses or results (Říčan,and Kullander 2006; Ottoni et al. 2019; da Silva et al. 2022; Blair and Bryston 2017; de Souza et al. 2021; Říčan and Říčanová 2017) including the genus *Australoheros* formal description, and its test of monophyly and phylogenetic position (Říčan,and Kullander 2006), as well as the proposal of species groups within *Australoheros* (Ottoni et al. 2019; da Silva et al. 2022; Říčan and Říčanová 2017).

This study by Lucena et al. (2022) describes two new species: *Australoheros mboapari* Lucena, Kullander, Norén & Calegari 2023, which was identified only through morphological comparisons with the parapatric *A. acaroides*, without a DNA-based analysis, as it is known only from formalin-fixed specimens with no available DNA sequences; and *A. ricani* Lucena, Kullander,

Norén & Calegari 2023, which was proposed based on its position in the *COI* and CYTB phylogenetic trees and the uncorrected p-distance from other species of *Australoheros* (Lucena et al. 2022, p. 103), lacking clear diagnostic morphological characters states. Therefore, it is not accurate to assert that the congruence of molecular and morphological analyses solely determined the selection of *COI* as the most suitable gene for delimiting species within this genus.

The authors cited that CYTB analysis failed to find a credible gene tree that could account for the higher number of delimited lineages. This phenomenon is believed to stem from the limitations of the data available in GenBank (possibly including misidentifications and data curated by no experts in the studied group), coupled with the fact that CYTB sequences are shorter than those of *COI*, potentially influencing the results obtained (Lucena et al. 2022, p. 46). Despite the mention of the size of the CYTB fragments, the CYTB sequences utilized in this study were 1099 bp, while the *COI* sequences were 654 bp. Lucena and collaborators dismissed the CYTB analysis, claiming they were unable to construct a reliable genetic tree due to the limitations cited before. The acceptance of the COI analysis was justified based on the congruence of the BI trees from both *COI* and CYTB, as well as the uncorrected p-distances. However, the rejection of the hypothesis proposed by the CYTB analysis was not clearly articulated, making it difficult to understand the rationale behind rejecting the CYTB gene results. Additionally, the failure to remove identical haplotypes may have impacted the performance of the bPTP.

Lucena et al. (2022; p. 5) mentioned that *COI* is an overlooked gene in cichlid systematics to date, but mention that it has proven to be efficient and that new studies (DeSalle and Goldstein 2019; Kullander et al. 2019) are emerging for its application in species delimitation and phylogenetic analysis. However, as discussed in this work, CYTB sequences have proven to be more sensitive for bPTP analysis because it is a less conserved gene, as well as it has demonstrated to be appropriate for delimiting OTUs that have diversified recently (less than about 3.5 m.y.a). In addition, other studies carried out with the same parameters, such as Ottoni et al. (2019) and Silva et al. (2022), also used CYTB to delimit *Australoheros* species and theorize this use for the same reasons, which reinforces the hypothesis that CYTB is a mitochondrial gene with faster evolutionary rates.

CONCLUSIONS

Several studies have approached the taxonomy of *Australoheros* in recent years, presenting controversial results. The Coalescence Theory has transformed this landscape by providing a theoretical foundation for the development of new algorithms capable of delimiting species using

rooted unilocus phylogenetic trees. Among these models, the bPTP method has proven to be particularly sensitive for more recent lineages, with divergence times of around 3.5 m.y.a in the case of *Australoheros* (Říčan et al. 2011). It is evident that the selection of outgroups for the roots of the trees significantly impacts the delimitation of OTUs. According to the bPTP server (Zhang et al. 2013), outgroups should be chosen as closely related as possible to the delimited/focal species group. Furthermore, the database used to generate the delimitation trees must have identical haplotypes reduced to only one to ensure the method's effectiveness and prevent the generation of false results (Fujisawa and Barraclough 2013; Monaghan et al. 2009). Lucena et al. (2022), when not collapsing identical haplotypes into a single terminal, performed analyses that did not strictly follow the method's assumptions and subsequently excluded the CYTB results obtained during species delimitation. In addition to bPTP (Zhang et al. 2013), other analytical methods, such as the tree-based method proposed by Wiens and Penkrot (2002), GMYC (Pons et al. 2006), ABGD (Puillandre et al. 2012), and ASAP (Ward 2009) can also be employed to enrich research and enhance understanding of the taxonomy of the genus *Australoheros*. Finally, it is important to highlight that in both bPTP analyses (CYTB and *COI*) conducted by Lucena et al. (2022) [Figs 10, 11 and 12], *Australoheros* was recovered as a valid species, corroborating the results previously found by Ottoni et al. (2019) and Silva et al. (2022). Nevertheless, Lucena et al. (2022) considered *A. autrani* as a junior synonym of *A. ipatinguensis*, despite *A.* having been described first, and thus possessing taxonomic priority. Therefore, based on the previous results found by Ottoni et al. (2019) and Silva et al. (2022), as well as the bPTP analyses (both based on CYTB and *COI*) conducted by Lucena et al. (2022; figs. 10, 11 and 12), we suggest that the taxonomic status of *A. autrani* should be reassessed in a future taxonomic work.

Thus, we consider that the use of phylogenetic distant outgroups may reduce the adequate recovery of OTUs. We also conclude that further studies are needed to determine whether the CYTB gene is indeed more informative than the *COI* for the genus *Australoheros*, especially considering the current discrepancy in the amount of sequences available for CYTB and that adherence to the premises of the bPTP model (especially the deletion of identical haplotypes) is fundamental to ensure the method does not lose sensitivity. Finally, we emphasize that single-locus analyses have limitations and can fail to rebuild evolutionary relationships more robustly. Integrative studies using multilocus or genomic data may be useful to understand the species delimitation in *Australoheros*.

Acknowledgments: Thanks are given to members of Laboratory of Molecular Systematics and Reproductive Biology- BEAGLE, from Federal University of Viçosa for support on development of this work. This study was supported by the Conselho Nacional de Desenvolvimento Científico e

Tecnológico - CNPq (CNPq grants number 307974/2021–9 and 306490/2024-2 to FPO) and Deutsche Cichliden-Gesellschaft - DCG.

Authors' contributions: Rebeca Rosa: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing—original draft preparation, Writing—review and editing, Visualization, Project administration. Elisabeth Henschel: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing—review and editing, Visualization, Supervision Project administration. Pedro Henrique Negreiros de Bragança: Data curation and Writing—review and editing. Grazielle Fernanda Evangelista Gomes: Writing—review and editing. Felipe Polivanov Ottoni: Conceptualization, Validation, Formal analysis, Investigation, Resources, Data curation, Writing—review and editing, Visualization, Supervision, Project administration. All authors have read and agreed to the published version of the manuscript.

Competing interests: Competing interests with all authors and reviewers of the paper 'Conjectures and refutations: Species diversity and phylogeny of *Australoheros* from coastal rivers of southern South America (Teleostei: Cichlidae)' from Lucena et. al (2022). Authors: Carlos A. Santos de Lucena; Sven Kullander; Michael Norén and Bárbara Calegari. Editor: Roberto E. Reis, Pontifical Catholic University of Rio Grande do Sul, Brazil

Availability of data and materials: All data used in this study are openly available at GenBank. table 1 of our paper provides a comprehensive list of all used terminals.

Consent for publication: Not applicable**.**

Ethics approval consent to participate: Not applicable.

REFERENCES

Albert JS, Tagliacollo VA, Dagosta F. 2020. Diversification of Neotropical Freshwater Fishes. Annu Rev Ecol Evol Syst **51:**27–53. doi:10.1146/annurev-ecolsys-011620-031032.

Alonso F, Terán GE, Aguilera G, Říčan O, Casciotta J, Serra WS et al. 2019. Description of a new species of the Neotropical cichlid genus *Gymnogeophagus* Miranda Ribeiro, 1918 (Teleostei: Cichliformes) from the Middle Paraná basin, Misiones, Argentina. PLoS ONE **14:**e0210166.

doi:10.1371/journal.pone.0210166.

Argolo LA, López-Fernández H, Batalha-Filho H, Affonso PRAM. 2020. Unraveling the systematics and evolution of the *Geophagus brasiliensis* (Cichliformes: Cichlidae) species complex. Mol. Phylogenet Evol **150:**106855. doi:10.1016/j.ympev.2020.106855.

Avise JC. 2000. Phylogeography: The History and Formation of Species. Oxford University Press, Cambridge.

Bentancur-R R, Broughton RE, Wiley EO, Carpenter KE, López JA et al. 2013. The tree of life and a new classification of bony fishes. PLoS Cur Tree of Life **5:**Edition 1. doi:10.1016/j.ympev.2012.02.018.

Bernardi G, Bucciarelli G. 1999. Molecular phylogeny and speciation of the surfperches (Embiotocidae, Perciformes). Molec Phylog Evol **13:**77–81. doi:10.1006/mpev.1999.0629.

Birindelli JLO, Sidlauskas BL. 2018. Preface: How far has Neotropical Ichthyology progressed in twenty years? Neotrop Ichthyol **16:**e180128. doi:10.1590/1982-0224-20180128.

Blair, C, Bryson RW Jr. 2017. Cryptic diversity and discordance in single-locus species delimitation methods within horned lizards (Phrynosomatidae: *Phrynosoma*). Mol Ecol Resour **17:**1168–1182. doi:10.1111/1755-0998.12658.

Carvalho DC, Seerig A, de Melo DC, deSousa AB; Pimenta DA et al. 2008. Identificação molecular de peixes: o caso do Surubim (*Pseudoplatystoma* spp.). Rev. Bras. Reprod. Anim., **32**, 215–219.

CZN. 1999. International Code of Zoological Nomenclature, 4th ed, The International Trust for Zoological Nomenclature: London, UK.

da Silva MA, Ottoni, FP, Mattos JLO et al. 2022. Species diversity and phylogeography of the *Australoheros* group (Teleostei, Cichlidae) in eastern Brazil. Ichthyol Res **70:**378–385. doi:10.1007/s10228-022-00888-9.

Dayrat B. 2005. Towards integrative taxonomy. Biol J Linn Soc **85:**407–415. doi:10.1002/aqc.584.

de Oliveira RC, Britzke R, de Oliveira C et al. 2025. Species delimitation in Neotropical fishes of *Mesonauta* Günther 1862 (Cichliformes: Cichlidae). Org Divers Evol **25:**343–355. doi:10.1007/s13127-025-00675-0.

de Queiroz K. 1998. The General Lineage Concept of Species, Species Criteria, and the Process of Speciation. In Endless Forms: Species and Speciation; Howard, D.J., Berlocher, S.H., Eds, Oxford University Press: New York, pp. 57–75.

de Queiroz K. 2005. Different species problems and their resolution. BioEssays **27:**1263–1269. doi:10.1002/bies.20325.

de Queiroz K. 2007. Species concepts and species delimitation. Syst Biol **56:**879–886. doi:10.1080/10635150701701083.

de Sousa JLP, de Araújo Bitencourt J, Sampaio I, Schneider H, de Mello Affonso PRA. 2021. "More than meets the eye": Phylogeographic inferences and remarkable cryptic diversity in endemic catfish *Parotocinclus* (Loricariidae: Hypoptopomatinae) from neglected and impacted basins in South America. Conserv Genet **22:**411–425. doi:10.1007/s10592-021-01336-3.

DeSalle R, Narechania A, Tessler M. 2023. Multiple outgroups can cause random rooting in phylogenomics. Mol Phylogenet Evol **184:**107806. doi:10.1016/j.ympev.2023.107806.

DeSalle R, Goldstein P. 2019. Review and Interpretation of Trends in DNA Barcoding. Front Ecol Evol **7:**302. doi:10.3389/fevo.2019.00302.

Dobzhansky Th. 1935. A Critique of the Species Concept in Biology. Philosophy of Science. **2:**344–355. doi:10.1086/286379.

Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792–1797. doi:10.1093/nar/gkh340.

Fricke R, Eschmeyer WN, Van der Laan R. (eds). 2025. Eschmeyer's Catalog of Fishes: Genera, Species, References. Electronic version accessed on 30 June.

Fujisawa T, Barraclough TG. 2013. Delimiting Species Using Single-Locus Data and the Generalized Mixed Yule Coalescent Approach: A Revised Method and Evaluation on Simulated Data Sets. Syst Biol **62:**707–724. doi:10.1093/sysbio/syt033.

Giangrande A. 2003. Biodiversity, conservation, and the 'Taxonomic impediment'. Aquatic Conserv: Mar Freshw Ecosyst **13:**451–459. doi:10.1002/aqc.584.

Gomes LC, Pessali TC, Sales NG, Pompeu PS, Carvalho DC. 2015. Integrative taxonomy detects cryptic and overlooked fish species in a neotropical river basin. Genetica **143:**581–588. doi:10.1007/s10709-015-9856-z.

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. Proc R Soc B **270:**313–321. doi:10.1098/rspb.2002.2218.

Kingman JFC. 1982a. The coalescent. Stochastic Processes and Their Applications **13:**235–248.

Kingman JFC. 1982b. Exchangeability and the evolution of large populations. *In*: Koch G, Spizzichino F (Eds) Exchangeability in Probability and Statistics. North-Holland: Amsterdam, pp. 97–112.

Kingman JFC. 1982c. On the genealogy of large populations. Journal of Applied Probability **19A:**27–43.

Kullander S, Norén M, Rahman MM, Mollah AR. 2019. Chameleonfishes in Bangladesh: Hipshot Taxonomy, Sibling Species, Elusive Species, and Limits of Species Delimitation (Teleostei: Badidae). Zootaxa **4586:**301–337. doi:10.11646/zootaxa.4586.2.7.

Lozano AP, Lasso-Alcalá OM, Bittencourt PS, Taphorn DC, Perez N et al. 2022. A new species of *Astronotus* (Teleostei, Cichlidae) from the Orinoco River and Gulf of Paria basins, northern

South America. ZooKeys **1113:**111–152. doi:10.3897/zookeys.1113.81240.

Marjoram P, Wall JD. 2006. Fast "coalescent" simulation. BMC Genetics **7:**16. doi:10.1186/1471-2156-7-16.

Mather N, Traves SM, Ho SYW. 2019. A practical introduction to sequentially Markovian coalescent methods for estimating demographic history from genomic data. Ecol Evol **10:**579-589. doi:10.1002/ece3.5888.

Mayden RL. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In Species: The units of diversity; Claridge, M.F., Dawah, H.A., Wilson, M.R., Eds, Chapman & Hall: London, pp. 381–423.

Mayden RL. 2002. On biological species, species concepts and individuation in the natural world. Fish and Fisheries **3:**171–196. doi:10.1046/j.1467-2979.2002.00086.x.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M et al. 2009. Accelerated Species Inventory on Madagascar Using Coalescent-Based Models of Species Delineation. Syst Biol **58:**298–311. doi:10.1093/sysbio/syp027.

Ottoni FP. 2010. *Australoheros capixaba*, a new species of Australoheros from southeastern Brazil (Labroidei: Cichlidae: Cichlasomatinae). Vert Zool **60:**19–25. doi:10.3897/vz.60.e30989.

Ottoni FP. 2011. Revisão taxonômica do grupo de espécies *Australoheros* (Teleostei: Cichlidae: Heroini). Dissertação (Mestrado), Programa de Pós-graduação em Zoologia, Museu Nacional, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 168 pp.

Ottoni FP. 2008. Taxonomic revision of the genus *Australoheros* Říčan & Kullander, 2006 (Teleostei: Cichlidae) with descriptions of nine new species from south-eastern Brazil. Vert Zool **58:**207–232. doi:10.3897/vz.58.e30934.

Ottoni FP, Mattos JLO, Katz AM, Bragança PHN. 2019. Phylogeny and species delimitation based on molecular approaches on the species of the *Australoheros* group (Teleostei, Cichlidae), with biogeographic comments. Zoosyst Evol **95:**49–64. doi:10.3897/zse.95.31658.

Piálek L, Říčan O, Casciotta J, Almirón A, Zrzavý J. 2012. Multilocus phylogeny of *Crenicichla* (Teleostei: Cichlidae), with biogeography of the *C. lacustris* group: species flocks as a model for sympatric speciation in rivers. Mol Phylogenet Evol **62:**46–61. doi:10.1016/j.ympev.2011.09.006.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP et al. 2006. Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. Syst Biol **55:**595–609. doi:10.1080/10635150600852011.

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol Ecol **21:**1864–1877. doi:10.1111/j.1365-294X.2011.05239.x.

Rambaut A. 2018. FigTree, Version 1.4.4; Institute of Evolutionary Biology, University of Edinburgh: Edinburgh. Available at: http://tree.bio.ed.ac.uk/software/figtree.

Rambaut A, Suchard M, Drummond A. 2021. Tracer, Version 1.7.2. Available at: http://beast.community/tracer.

Reid NM, Carstens BC. 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol Biol **12:**196. doi:10.1186/1471-2148-12-196.

Reis RE, Albert JS, Di Dario F, Mincarone MM, Petry P et al. 2016. Fish biodiversity and conservation in South America. J Fish Biol **89:**12–47. doi:10.1111/jfb.13016.

Říčan O, Dragová K, Almirón A, Casciotta J, Gottwald J et al. 2021. MtDNA species-level phylogeny and delimitation support significantly underestimated diversity and endemism in the largest Neotropical cichlid genus (Cichlidae: *Crenicichla*). PeerJ **9:**e12283. doi:10.7717/peerj.12283.

Říčan O, Kullander SO. 2006. Character- and tree-based delimitation of species in the *Cichlasoma facetum* group. J Zool Syst Evol Res **44:**136–152.

Říčan O, Kullander SO. 2008. The *Australoheros* (Teleostei: Cichlidae) species of the Uruguay and Paraná River drainages. Zootaxa **1724:**1–51. doi:10.5281/zenodo.181173.

Říčan O, Piálek L, Almirón A, Casciotta J. 2011. Two new species of *Australoheros* (Teleostei: Cichlidae), with notes on diversity of the genus and biogeography of the Río de la Plata basin. Zootaxa **2982:**1–26. doi:10.11646/zootaxa.2982.1.1.

Říčan O, Říčanová Š. 2017. Live coloration, habitat, biogeography and phylogenetic position of *Australoheros guarani* Říčan & Kullander, 2008 (Teleostei: Cichlidae). Ichthyol Contrib Peces Criollos **57:**1–6.

Říčan O, Říčanová Š, Dragová K et al. 2019. Species diversity in *Gymnogeophagus* (Teleostei: Cichlidae) and comparative biogeography of cichlids in the Middle Paraná basin, an emerging hotspot of fish endemism. Hydrobiologia **832:**331–354. doi:10.1007/s10750-018-3691-z.

Říčan O, Říčanová Š, Rodriguez Haro LR, Rodriguez Haro CE. 2023. Unrecognized species diversity and endemism in the cichlid genus *Bujurquina* (Teleostei: Cichlidae) together with a molecular phylogeny document large-scale transformation of the western Amazonian river network and reveal complex paleogeography of the Ecuadorian Amazon. Hydrobiologia **850:** 2199–2229. doi:10.1007/s10750-022-05019-z.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19:**1572–1574. doi:10.1093/bioinformatics/btg180.

Sabaj MH. 2020. Codes for Natural History Collections in Ichthyology and Herpetology. Copeia **108:**593–669. doi:10.1643/ASIHCODONS2020.

Santos de Lucena CA, Kullander S, Norén M, Calegari B. 2022. Conjectures and refutations: Species diversity and phylogeny of *Australoheros* from coastal rivers of southern South America (Teleostei: Cichlidae). PLoS ONE **17:**e0261027. doi:10.1371/journal.pone.0261027.

Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E et al. 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. Annu Rev Entomol **55:**421–438. doi:10.1146/annurev-ento-112408-085432.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis version 11. Mol Biol Evol **38:**3022–3027. doi:10.1093/molbev/msab120.

Taylor RW. 1983. Descriptive taxonomy: past, present and future. In Australian systematic entomology: a bicentenary perspective; Highley, E., Taylor, R.W., Eds, CSIRO: Melbourne, pp. 93–134.

Ward RD. 2009. DNA barcode divergence among species and genera of fishes. Mol Ecol Res **9:**1077–1085. doi:10.1111/j.1755-0998.2009.02541.x.

Wheeler QD, Houghton DJ, Wilson JA, Simons AM, Carvalho MR. 2013. The New Taxonomy: Integrative Systematics and the Biodiversity Crisis. Front Ecol Evol. (please confirm)

Wiens JJ, Penkrot TA. 2002. Delimiting Species Using DNA and Morphological Variation and Discordant Species Limits in Spiny Lizards (*Sceloporus*). Syst Biol **51:**69–91. doi:10.1080/106351502753475880.

Will KW, Mishler BD, Wheeler QD. 2005. The Perils of DNA Barcoding and the Need for Integrative Taxonomy. Syst Biol **54:**844–851. doi:10.1080/10635150500354878.

Willis SC, López-Fernández H, Montaña CG, Farias IP, Ortí G. 2012. Species-level phylogeny of 'Satan's perches' based on discordant gene trees (Teleostei: Cichlidae: *Satanoperca* Günther 1862). Mol Phylogenet Evol **63:**798–808. doi:10.1016/j.ympev.2012.02.018.

Xia X. 2018. DAMBE7: New and improved tools for data analysis in molecular biology and evolution. Mol Biol Evol **35:**1550–1552. doi:10.1093/molbev/msy073.

Ximenes AM, Bittencourt PS, Machado VN, Hrbek T, Farias IP. 2021. Mapping the hidden diversity of the *Geophagus* sensu stricto species group (Cichlidae: Geophagini) from the Amazon basin. PeerJ **9:**e12443. doi:10.7717/peerj.12443.

Yeates DK, Seago A, Nelson L, Cameron SL, Joseph L et al. 2011. Integrative taxonomy, or iterative taxonomy? Syst Entomol **36:**209–217. doi:10.1111/j.1365-3113.2010.00558.x.

Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics **29:**2869–2876. doi:10.1093/bioinformatics/btt499.

Supplementary materials

Table S1. Australoheros species classified according to approach type based on taxonomic literature along with their geographic distribution.

Table S2. Terminals of Australoheros utilized for COI gene matrices with reduce of identical haplotypes and our comparison with dataset of Lucena et al., 2022

Table S3. Terminals of Australoheros utilized for CYTB gene matrices with reduce of identical haplotypes and our comparison with dataset of Lucena et al., 2022

Table S4. Values of p-distance for taxons of COI marker generated through MEGA 11 using the Kimura model (1980)

Table S5. Values of p-distance for taxons of C marker generated through MEGA 11 using the Kimura model (1980)

Figs. S1. Outputs of the eight analyses of COI gene. (download)

Figs. S2. Outputs of the eight analyses of CYTB gene. (download)