Zoological Journal of the Linnean Society, 2023, **197**, 514–531. With 7 figures.

Testing palaeodrainage hypotheses in south-eastern Brazil: phylogeography of the sinistral livebearer fish of the genus *Phalloceros* (Cyprinodontiformes: Poeciliidae)

IGOR C. A. SOUTO-SANTOS[1,*], W. BRYAN JENNINGS[1] and PAULO A. BUCKUP[1]

[1]*Universidade Federal do Rio de Janeiro, Museu Nacional, Departamento de Vertebrados, Quinta da Boa Vista, 20490-040 Rio de Janeiro, RJ, Brazil*

Received 16 October 2021; revised 28 February 2022; accepted for publication 7 March 2022

The 'sinistral *Phalloceros* group' consists of three species, *P. aspilos*, *P. leptokeras* and *P. tupinamba*, that belong to *Phalloceros*, a genus of freshwater fish endemic to South America. They inhabit the Paraíba do Sul Basin and coastal drainages in south-eastern Brazil. This group is diagnosed by large hooks in the medial portion of the gonopodial appendices in males and the sinistral direction of the urogenital papilla in females. We conducted a phylogeographic analysis of mitochondrial (*COI* and *Cytb*) and nuclear (*RAG1*) haplotypes of 36 individuals sampled from 11 localities to test the hypothesis that the biogeographic history of sinistral *Phalloceros* was mediated by connections of coastal basins caused by lowered sea-levels during the Quaternary. We evaluated the taxonomic status of these nominal taxa by integrating molecular species delimitation methods and morphological data. Our results suggest that the three nominal taxa are synonyms, and *P. leptokeras* is designated as the valid species name. The geographic expansion of the *P. leptokeras* lineage began in the coastal region (~2.3 Mya). The phylogeographic relationships among populations partially corroborate the palaeodrainage model, but also suggest that *P. leptokeras* colonized the inland Paraíba do Sul drainage, overcoming the Serra do Mar mountains.

ADDITIONAL KEYWORDS: ABGD – ASAP – BIN – biogeographic history – BPP – coastal river drainages – GMYC – molecular species delimitation – vicariance.

INTRODUCTION

Phalloceros Eigenmann, 1907, is a monophyletic genus (Lucinda & Reis, 2005) of small livebearing fish distributed in the freshwaters of south-eastern South America from the southern portion of the Brazilian State of Bahia to Uruguay and to the La Plata drainage in Argentina and Paraguay, and to the Araguaia drainage in the Tocantins Basin (Lucinda, 2008). *Phalloceros* was originally treated as a monotypic genus, with *Phalloceros caudimaculatus* (Hensel, 1868) regarded as a single, widely distributed species for almost a century. However, in 2005, a morphology-based phylogenetic analysis of Poeciliinae (Poeciliidae *sensu* Bragança *et al.*, 2018) led to the recognition of 21 additional *Phalloceros* species (Lucinda & Reis, 2005), which were formally described three years later (Lucinda, 2008). Like other poeciliid livebearers, species of *Phalloceros* have pronounced sexual dimorphism. Mature males have expanded third, fourth and fifth anal-fin rays forming a copulatory organ called a gonopodium.

Phalloceros is diagnosed by: (1) the presence of an anterior pair of appendices in the tip of the third ray of the gonopodium (Lucinda, 2008; character state 93-1 in Lucinda & Reis, 2005); (2) the partially open preopercular canal between pores 8, 9 and 10 forming a deep groove; (3) the elongate pore 11 (sometimes confluent with the eighth, ninth and tenth grooves); and (4) the canal between pores 12, U–V closed (Lucinda, 2008; character state 8–5 in Lucinda & Reis, 2005). Most of the 22 described species differ in the morphology of the gonopodium. The gonopodium has a pair of appendages, which may be symmetric or asymmetric in size and may bear bony processes called 'hooks' by Lucinda (2008). Like the observed variation in male genitalia, females exhibit variable urogenital papillae. This reproductive structure in females is a fleshy process located posterior to the anus. The tip

*Corresponding author. E-mail: icass.ufrj@gmail.com.

of the papilla is directed caudally (straight) in nine species of *Phalloceros* (e.g. *P. harpagos* Lucinda, 2008), directed to the right side of the body in ten species (e.g. *P. anisophallos* Lucinda, 2008) or to the left in three nominal species (e.g. *P. leptokeras* Lucinda, 2008). We refer to the latter as the group of sinistral species.

These sinistral *Phalloceros* species include three nominal species: *P. aspilos* Lucinda, 2008, *P. leptokeras* and *P. tupinamba* Lucinda, 2008 (Fig. 1). In addition to the orientation of the urogenital papilla, the 'sinistral *Phalloceros* group' is diagnosed by a unique gonopodial morphology: the gonopodial appendices are slender, and the appendicular hooks are large and located in the medial portion of the appendices (Fig. 2; Lucinda, 2008). This group is geographically restricted to river drainages of the Atlantic Rainforest associated with the Serra do Mar mountains in south-eastern Brazil. *Phalloceros aspilos* was described based on specimens from the Parati-Mirim River, a coastal river that flows into Ilha Grande Bay. *Phalloceros leptokeras* was described based on specimens from the Paquequer and Piabanha River drainages, two tributaries of the Paraíba do Sul River. *Phalloceros tupinamba* has a disjunct distribution, which includes coastal drainages in the States of Rio de Janeiro and São Paulo, and thus overlaps the type locality of *P. aspilos*. The holotype of *P. tupinamba* was collected in the Macacu drainage, a coastal river that flows into Guanabara Bay in the State of Rio de Janeiro, while most of the paratypes were collected from the Itamambuca, Indaiá and Grande coastal drainages of the State of São Paulo (Lucinda, 2008). Souto-Santos *et al.* (2019) recorded the presence of sinistral *Phalloceros* in additional drainages near the type locality of *P. aspilos*, and in the middle of the distribution range for *P. tupinamba* (Fig. 3). Our morphological examination of these samples revealed overlap of diagnostic characters of *P. tupinamba* and *P. leptokeras*. Thomaz *et al.* (2019) recorded the presence of a presumably undescribed species of sinistral *Phalloceros* in a stream near Puruba Road in the municipality of Ubatuba, a coastal area in the State of São Paulo. However, these authors did not name the species due to morphological similarity with *P. tupinamba*. In the present work, we test whether these morphological and genetic differences correspond to species or population units.

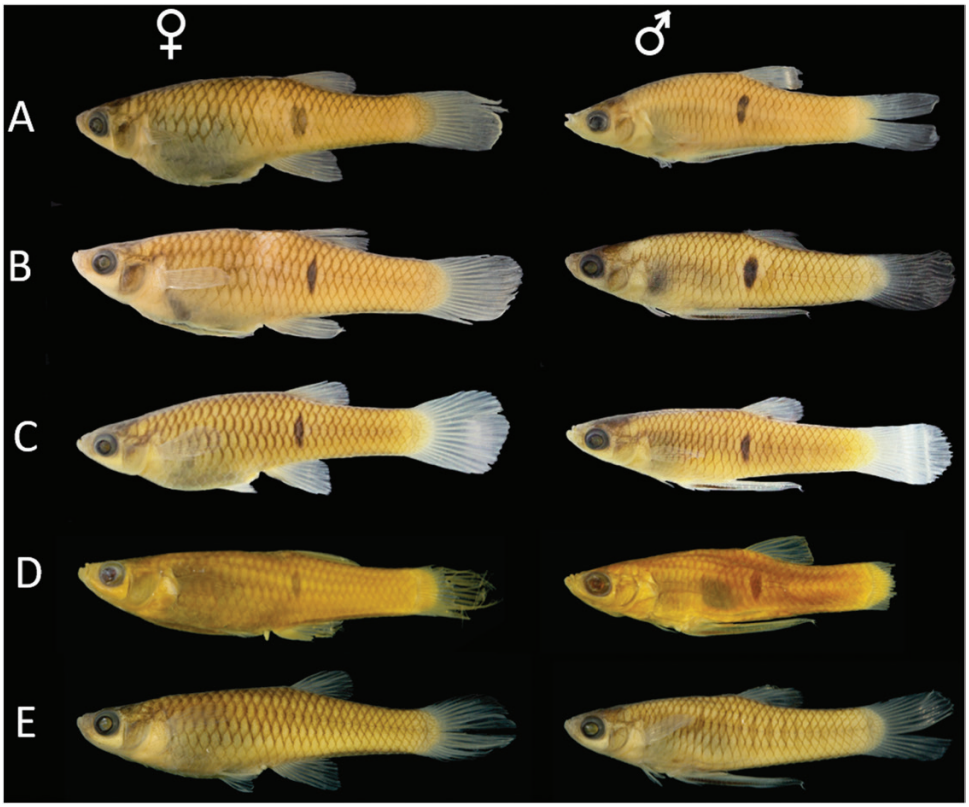

Figure 1. Morphological variation within females and males of the sinistral *Phalloceros* group (= *Phalloceros leptokeras*). A, Paquequer drainage (MNRJ 43596, ♀ 30.4 mm SL, ♂ 22.5 mm SL). B, Macacu drainage (MNRJ 43600, ♀ 28.1 mm SL, ♂ 17.4 mm SL). C, Saco drainage (MNRJ 51337, ♀ 27.8 mm SL, ♂ 18.1 mm SL). D, Japuíba drainage (MNRJ 4226, ♀ 32.0 mm SL, ♂ 21.2 mm SL). E, Parati-Mirim drainage (MNRJ 43505, ♀ 31.3 mm SL, ♂ 20.9 mm SL).

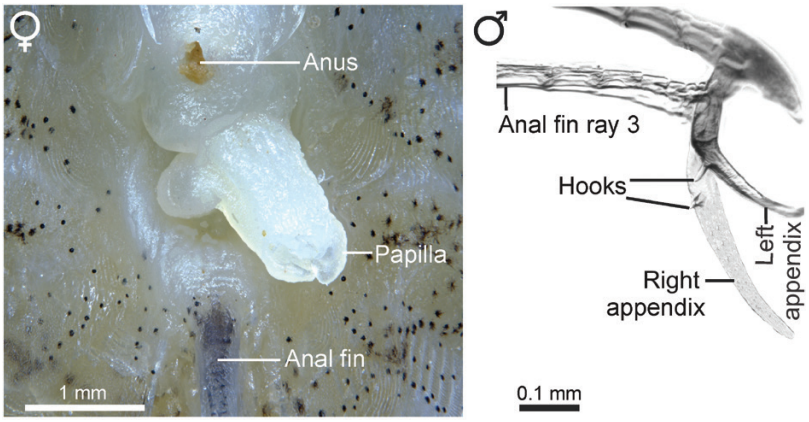

Figure 2. Genital morphology of sinistral *Phalloceros* (= *Phalloceros leptokeras*). Ventral view of female urogenital papilla (MNRJ 52301, 28.3 mm SL, Pocinhos drainage) and left lateroventral view of male gonopodium tip skeleton, skin removed (MNRJ 24398, Guandu drainage).

Two phylogenetic hypotheses have been proposed for *Phalloceros* (Fig. 4). One was based on morphological data (Lucinda, 2008), while the other was based on restriction site-associated DNA sequencing (RADseq; Thomaz *et al.*, 2019). These hypotheses produced conflicting trees, but both support the monophyly of the group of nominal species with papilla directed to the left (Fig. 4).

The phylogenetic relationships among species in the sinistral *Phalloceros* group are still poorly supported. The morphological phylogeny in Lucinda (2008) recovered *P. aspilos* as the sister-species of *P. leptokeras*, but the molecular phylogeny in Thomaz *et al.* (2019) did not include topotypes of *P. leptokeras* and *P. tupinamba*.

Here we present an assessment of phylogeographic relationships among sinistral *Phalloceros* species and use the resulting tree to test biogeographic hypotheses about freshwater fish biogeography in the Serra do Mar. The Serra do Mar consists of scarps that extend from the north-eastern coast of the State of Rio de Janeiro to the coast of the State of Santa Catarina. In its north-eastern region, a tectonic rift resulted in the Paraíba do Sul River valley, and the Serra do Mar forms a mountain range that extends along the coast of the States of Rio de Janeiro and São Paulo (Oliveira, 2003). Two hypotheses have been invoked in the literature to explain the relationships among fish populations in currently isolated coastal drainages of the Serra do Mar: (1) temporary river connections during Quaternary sea-level lowering (Buckup, 2011; Thomaz *et al.*, 2015; Lima *et al.*, 2016; Thomaz & Knowles, 2018); and (2) tectonic and erosive events resulting in headwater stream capture across watersheds (Ribeiro, 2006; Buckup, 2011; Lima *et al.*, 2016, 2017).

Palaeodrainage connections among coastal rivers in south-eastern Brazil have been proposed since the beginning of the 20th century to explain similarities between aquatic organisms across land barriers. Ihering (1927), for example, hypothesized the existence of a palaeoriver that may have connected coastal drainages between the Brazilian States of Rio de Janeiro and Rio Grande do Sul. This hypothetical stream, dubbed the Ameghino River, would have later been submerged by the Atlantic Ocean (Ihering, 1927). Palaeodrainage connections occur when the sea level is lowered due to climate change. During periods of low sea-level, exposed palaeodrainages may facilitate faunal exchanges between adjacent drainages, which are isolated from each other when sea levels are high. Thomaz & Knowles (2018) hypothesized that 15 palaeodrainages associated with the Paraíba do Sul and adjacent coastal drainages were important drivers of freshwater fish diversification (Fig. 3). These palaeodrainages reached their maximum exposure between 19 and 26 thousand years ago during the Last Glacial Maximum (LGM) of the Pleistocene, a time when the sea level dropped 125 m below the current level (Thomaz *et al.*, 2015). This hypothesis predicts that fish lineages from rivers that shared a palaeodrainage are more closely related to each other than to lineages from rivers that were not part of the same palaeodrainage. Thus, phylogenetic trees constructed from independent data can be used to test this and other palaeodrainage hypotheses.

Headwater capture occurs when a stream changes its course and connects to a different drainage because of a geomorphological event (Oliveira, 2003; Buckup, 2011). Headwater capture thus allows the sharing of fauna between the rivers involved in the capture. A prediction of this hypothesis is that fish populations from rivers involved in the capture event are more closely related to each other than to populations found in other rivers. The hypothesis is corroborated when lineages from

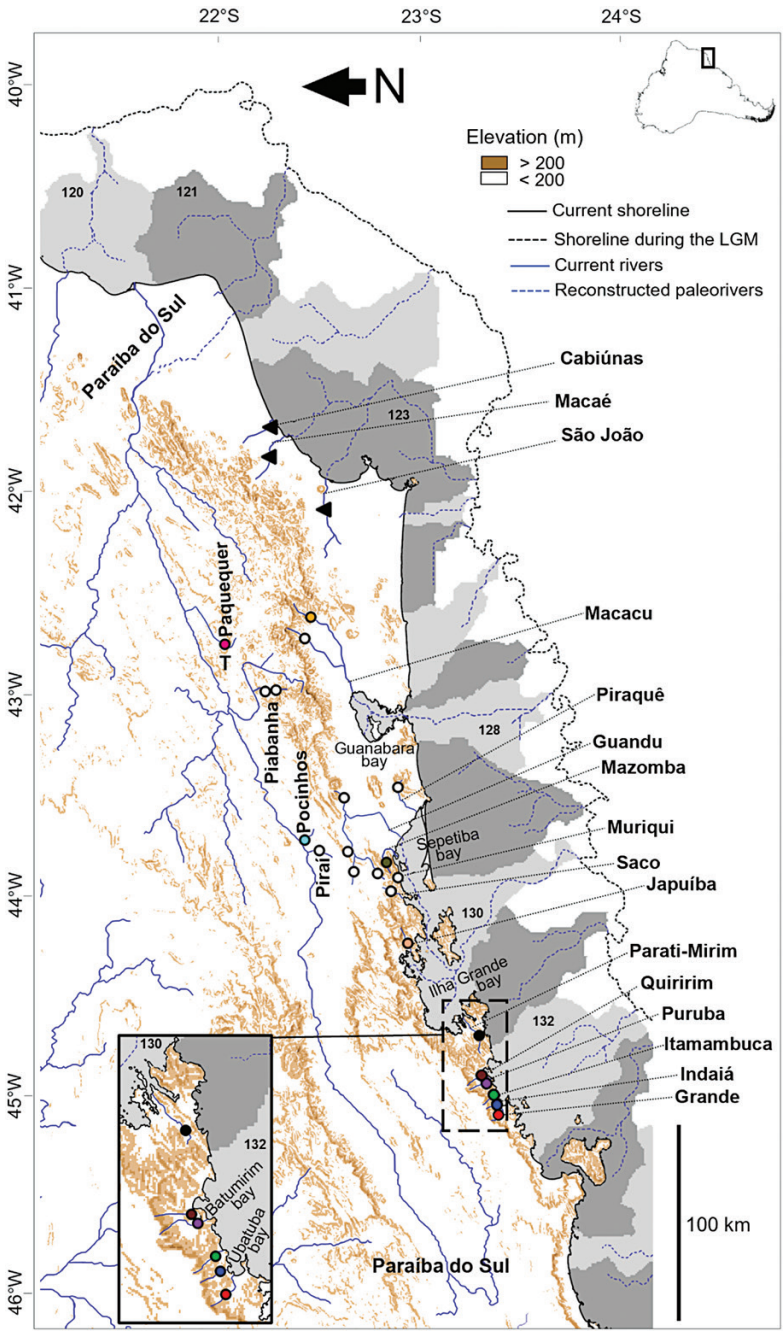

Figure 3. Geographic distribution of the sinistral *Phalloceros* group (= *Phalloceros leptokeras*) (circles) and its sister-group *Phalloceros* sp. 'Macaé' (black triangles). White circles correspond to material not associated with DNA sequences. The colours of solid circles correspond to samples associated with DNA sequences (see Figs 5–7). T = type locality according to Lucinda (2008). Grey colours indicate the reconstruction of palaeodrainages exposed with the sea level lowering during the LGM. Palaeodrainege numbers correspond to those established by Thomaz & Knowles (2018).

rivers involved in capture events occur in both rivers and is rejected when these lineages are absent in one of the rivers. Headwater capture events among drainages from the Serra do Mar in south-eastern Brazil are only known for a few basins (Menezes *et al.*, 2008; Torres *et al.*, 2008; Torres & Ribeiro, 2009; Zamudio *et al.*, 2009; Buckup, 2011). Despite this limitation, some authors have proposed headwater capture events to explain the distribution of species, even in the absence of geological evidence (Lima *et al.*, 2016, 2017).

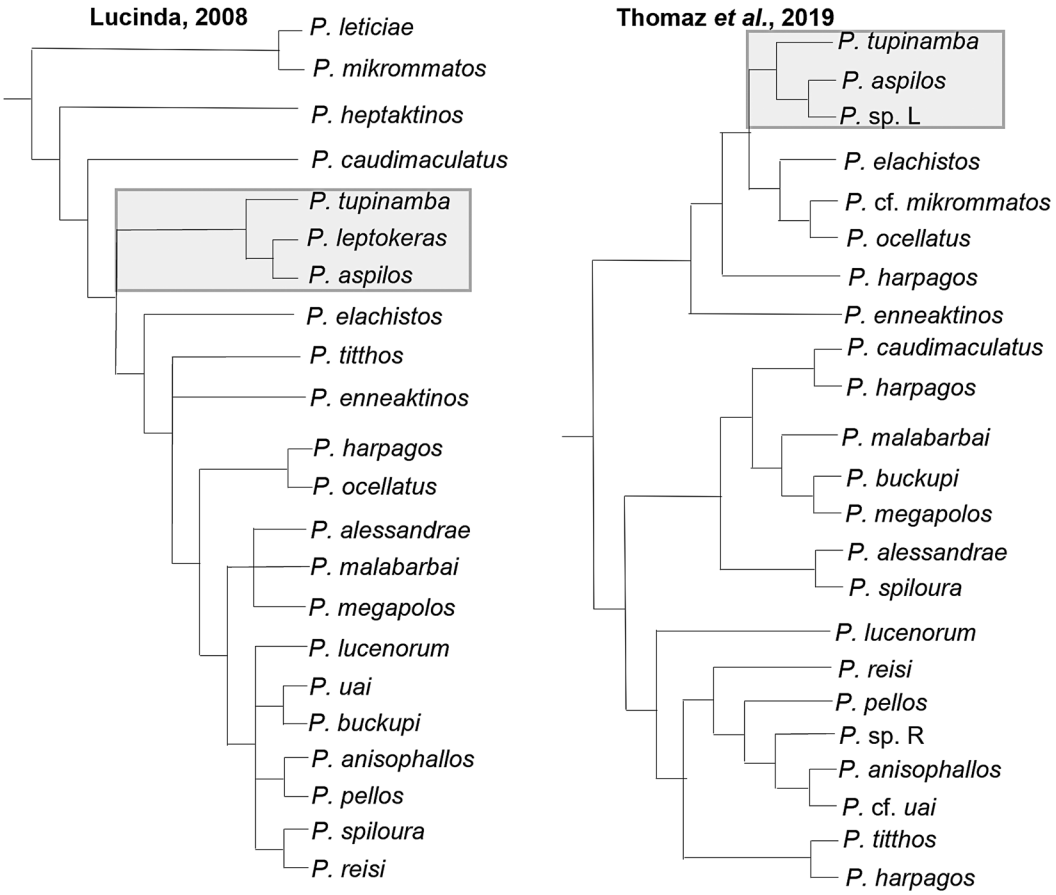

Figure 4. Morphological (Lucinda, 2008) and molecular (Thomaz *et al.*, 2019) phylogenetic hypotheses proposed for *Phalloceros*. Shaded rectangles indicate the lineage of the sinistral *Phalloceros* (= *Phalloceros leptokeras*).

Another means of geographic expansion of fish distributions involves dispersal through coastal lowlands. *Phalloceros* species are commonly found in shallow, temporary, wetland habitats, suggesting that lowland migration among neighbouring drainages is a third alternative hypothesis to explain species distribution patterns between adjacent drainages connected by lowland areas. These ephemeral environments are formed during rainy periods and offer a suitable habitat to small Cyprinodontiformes (Costa, 2009). One prediction of this hypothesis is that lineages of fish from rivers that are temporarily connected by wetlands are more closely related to each other than to lineages from other rivers. Dispersal hypotheses are difficult to test, as there are no robust models to specify how, when and where the dispersal events occurred. In general, dispersal is a random phenomenon that can affect each species or lineage differently from others, making it difficult to establish common predictions for all species (Tzeng *et al.*, 2006).

Phylogeographic approaches, which investigate evolutionary relationships among conspecific populations and closely related species, have been applied to understand the evolutionary boundaries of several freshwater fish lineages from south-eastern Brazil. Examples include lineages of *Phalloceros* (Crepaldi, 2015; Oliveira, 2017) and other taxonomic groups (e.g. Lima *et al.*, 2016, 2021).

Here, we performed a multilocus phylogeographic analysis based on mitochondrial and nuclear genes to understand the diversity of sinistral *Phalloceros* throughout the geographic distribution of this species group. One of the genes we employed is the standard DNA barcode 'Folmer fragment' of the mitochondrial cytochrome *c* oxidase subunit I (*COI*) gene (Hebert *et al.*, 2003a). *COI* barcodes have enabled many studies to solve taxonomic questions and identify Neotropical freshwater fish species boundaries. The effectiveness of this technique has been widely demonstrated in the literature (e.g. Hebert *et al.* 2003a, b), including *Phalloceros* (Amaral *et al.*, 2015) and other fish groups (Pereira *et al.*, 2013; Henriques *et al.*, 2015; Rossini *et al.*, 2016; Shimabukuro-Dias *et al.*, 2017).

Our goals were (1) to sample as many drainages as possible in order to understand the intraspecific and interspecific boundaries among populations of

sinistral *Phalloceros*, including the type localities for *P. aspilos*, *P. leptokeras* and *P. tupinamba*; and (2) to correlate the patterns of genetic variation with past climatic and geomorphological processes in order to explain colonization across currently isolated basins. We explicitly tested hypotheses of river basin interchanges through sea-level fluctuations based on the palaeodrainage model proposed by Thomaz & Knowles (2018) and tested models for headwater capture and wetlands dispersal.

MATERIAL AND METHODS

Sampling

Specimens examined in this study, which included the type series of *Phalloceros aspilos*, *P. leptokeras* and *P. tupinamba*, are in the ichthyological collections of the Museu Nacional, Universidade Federal do Rio de Janeiro (MNRJ) and Pontifícia Universidade Católica do Rio Grande do Sul (MCP) (Supporting Information, File S1). Our study included samples from the known distribution of sinistral *Phalloceros*, including the Paraíba do Sul River and adjacent coastal drainages along the Serra do Mar (Fig. 3). The Paraíba do Sul River Basin is limited by the Mantiqueira, Caparaó and Santo Eduardo Mountain chains in the north, and the Serra do Mar in the south. The latter separates the Paraíba do Sul from various small independent coastal basins that flow directly into the ocean (SEMADS, 2001).

We included 36 individuals of sinistral *Phalloceros* specimens from 11 localities and four outgroup species: *Phalloceros* sp. 'Macaé', *P. ocellatus* Lucinda, 2008, *P. mikrommatos* Lucinda, 2008 and *Poecilia vivipara* Bloch & Schneider, 1801. Samples from each coastal drainage, and two subdrainages of the Paraíba do Sul, were treated as isolated populations. DNA sequence data were obtained from tissue samples collected by the authors, except for one outgroup (*Poecilia vivipara*, GenBank accession number KU684422). One to six individuals were sampled from each locality (Supporting Information, Table S1). Choice of outgroup species was based on Thomaz *et al.* (2019), which showed that *P. mikrommatos* and *P. ocellatus* were closely related to the sinistral *Phalloceros* clade (Fig. 4). Additionally, we included in the analysis an undescribed species (*Phalloceros* sp. 'Macaé', straight urogenital papilla), because it is more closely related to sinistral *Phalloceros* than any other species considered by Thomaz *et al.* (2019) according to an unpublished DNA barcode phylogeny that included 20 nominal species of *Phalloceros* (Buckup *et al.*, in prep.).

We sequenced segments of two mitochondrial genes and one nuclear gene. The standard DNA barcode comprising the Folmer region of cytochrome *c* oxidase subunit I (*COI*) and cytochrome *b* (*Cytb*) allow the discrimination of closely related species and phylogeographic groups of a single species (Avise, 2000; Hebert *et al.*, 2003a; Pereira *et al.*, 2013; Henriques *et al.*, 2015; Lima *et al.*, 2016, 2017; Rossini *et al.*, 2016; Shimabukuro-Dias *et al.*, 2017). The nuclear recombination activating gene subunit 1 (*RAG1*) has resolved species-level relationships and its mutation rate is lower than the mitochondrial genes, allowing resolution of older divergences (Bermingham & Martin, 1998; Arroyave & Stiassny, 2011; Arroyave *et al.* 2013, Hirschmann *et al.*, 2015). These three markers have been successfully employed in phylogenetic studies of poecilid fishes (Reznick *et al.*, 2017).

All sequences and associated data, including geographic coordinates, chromatograms, sequence data and primer details, are available in the Barcode of Life Data Systems (BOLD Systems, http://boldsystems.org/; Ratnasingham & Hebert, 2007). Sample details are listed in the Supporting Information, Table S1, along with GenBank accession numbers. Figure 3 shows the localities of voucher specimens used to obtain molecular data for this study.

Molecular procedures and analyses

We extracted genomic DNA from muscle tissue using the Qiagen Blood and Tissue kit following the manufacturer's instructions, or the salting out method by Miller *et al.* (1988). DNA quality was verified with standard agarose gel electrophoresis, while DNA concentration was measured using a NanoDrop ND-2000 spectrophotometer. Partial sequences of mitochondrial (*COI* and *Cytb*) and nuclear (*RAG1*) genes were amplified via the polymerase chain reaction (PCR). The *COI* primers were previously developed for Neotropical freshwater fish (Jennings *et al.*, 2019); the *Cytb* primers were developed for characiform fishes but have since worked well for many other South American fish species, including Poeciliidae (Buckup *et al.*, in prep.); and the *RAG1* primers were developed in this study specifically for Poeciliidae (Supporting Information, Table S2).

Amplified products were checked using 2% agarose gel electrophoresis. PCR products were purified using Exo-SAP (Handy *et al.*, 2011) or PEG (Lis, 1980; Jennings, 2017). We sequenced each PCR product in both directions on an ABI3730xl (Applied Biosystems) automated sequencer at the Fundação Oswaldo Cruz – FIOCRUZ.

The resulting sequences were aligned to a reference sequence (GenBank accession numbers NC_011379.1 for *COI* and *Cytb*, NC_024336.1 for *RAG1*) using the GENEIOUS v.6 software (http://www.geneious.com) and manually edited to fine-tune base calls and ensure codon alignment. Aligned sequences of the three loci were concatenated for joint analyses, resulting

in concatenated sequences of 2695 bp (no gaps or missing data) for 42 individuals. Evolutionary models of nucleotide substitution for each gene were selected using PARTITION FINDER 2.1.0 software (Lanfear *et al.*, 2016) via the CIPRES online platform (http://www.phylo.org/). The analysis was set to run under a greedy search, unlinked branch lengths and used the corrected Akaike information criterion (AIC).

We conducted the phylogenetic analysis using a Bayesian inference method. A Bayesian Markov chain Monte Carlo algorithm (MCMC) was applied using MRBAYES 3.2.7 (Ronquist *et al.*, 2012) with *Poecilia vivipara* as the outgroup, using four chains in two parallel runs of 10 million generations, sampling every 1000 trees and discarding the first 20% of the total sample trees as the burn-in subset. We considered that the analysis yielded enough independent samples when the log-likelihood (lnL) values reached a plateau, and when the effective sample size (ESS) values of each parameter exceeded 200 in TRACER 1.7.1 (Rambaut *et al.*, 2018). The posterior probability values, which we visualized in FIGTREE v.1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/), were used to evaluate the support of the branches of the resulting tree. Branches with posterior probability value greater than 0.95 were regarded as well-supported clades.

A Bayesian analysis was also conducted in the software BEAST 2.6.3 (Bouckaert *et al.*, 2014) to produce an ultrametric tree for estimating divergence times using the *COI*, *Cytb* and *RAG1* genes together, but with each gene submitted in its own FASTA file. The StarBeast package was used to estimate divergence times, employing a calibrated Yule model, with 42 million years of divergence between *Phalloceros* and *Poecilia* (Reznick *et al.*, 2017) as the prior. We also specified 0.5 as the gene ploidy value for the mitochondrial genes, and 2.0 for the nuclear gene (Drummond & Bouckaert, 2015). Two independent MCMC runs were conducted with different random seeds following the same criteria described above for the MRBAYES analyses. Trees and parameters of the two runs were combined through LOGCOMBINER 2.6.3 (Rambaut & Drummond, 2007) and the parameters were then checked in TRACER following the same criteria as used in the previous analysis. A maximum clade credibility tree was obtained through TREEANNOTATOR 2.6.3 (Bouckaert *et al.*, 2014).

Single nucleotide polymorphisms (SNPs) were checked to detect synapomorphic and autapomorphic mutations using MEGA X software (Kumar *et al.*, 2018). The SNPs of the sinistral *Phalloceros* were then mapped on the tree using the sequence of *Phalloceros* sp. 'Macaé' as the outgroup to polarize the character transformations. The SNPs were numbered sequentially according to their position in the concatenated matrix of genes, in the following order: *COI*, *Cytb* and *RAG1*. A haplotype network for each gene was inferred using the TCS method (Templeton *et al.*, 1992) on PopART (Leigh & Bryant, 2015), and each mutation represented in the haplotype network was also identified by the SNP sequential number.

To investigate the best threshold for discrimination of species of sinistral *Phalloceros* we estimated evolutionary distances among *COI* sequences following Rossini *et al.* (2016) in the usage of pairwise Kimura-2-parameter (K2P) distances in Mega X (Kumar *et al.*, 2018). This analysis compared these distances with distances among well-delimited non-sister species of *Phalloceros*. We also compared the results of five automated methods of lineage partitioning: automatic barcode gap discovery (ABGD, Puillandre *et al.*, 2012), assemble species by automatic partitioning (ASAP; Puillandre *et al.*, 2020), barcode index number (BIN; Ratnasingham & Hebert, 2013), general mixed Yule coalescent model (GMYC; Pons *et al.*, 2006) and Bayesian phylogenetics and phylogeography (BPP; Zhang *et al.*, 2011). These analyses were based on the *COI* dataset, except for BPP, which was based on all three genetic markers. The ABGD analysis was performed on the website https://bioinfo.mnhn.fr/abi/public/abgd/, with parameters used by Rossini *et al.* (2016) (K2P substitution model, relative value gap X = 0.1, Pmin = 0.005 and Pmax = 0.1) (Supporting Information, File S2). The ASAP analysis was conducted on the website https://bioinfo.mnhn.fr/abi/public/asap/asapweb.html, with a K2P substitution model and default parameter values (Supporting Information, File S3). The BIN analysis was performed using the BOLD-systems (http://www.boldsystems.org), which also include sequences submitted by other researchers. For the GMYC, an ultrametric tree was generated in BEAST, as described above, except for the usage of the Yule speciation model and rooting in *Phalloceros* sp. 'Macaé'. The GMYC analysis was implemented on the website http://species.h-its.org/gmyc/ with the 'single threshold' (sGMYC; Supporting Information, File S4) and 'multi threshold' (mGMYC, Supporting Information, File S5) options. Only unique haplotypes were used in GMYC, as recommended by Rossini *et al.* (2016). In the BPP analysis, we used the phylogenetic hypothesis of Figure 5 as a guide tree. As BPP does not allow for polytomies, we were unable to test all populations. Instead, based on the topotypes of *P. aspilos*, *P. leptokeras* and *P. tupinamba*, and the phylogenetic relationships of all sinistral *Phalloceros* (Fig. 5), we used as input the hypothesis that the sinistral *Phalloceros* includes four species, comprising: (1) populations of the Paquequer and Pocinhos Rivers; (2) populations of the Macacu, Mazomba, Quiririm, Puruba and Japuíba Rivers; (3) populations of the Grande, Indaiá and Itamambuca Rivers; and (4) the population found in the Parati-Mirim River.

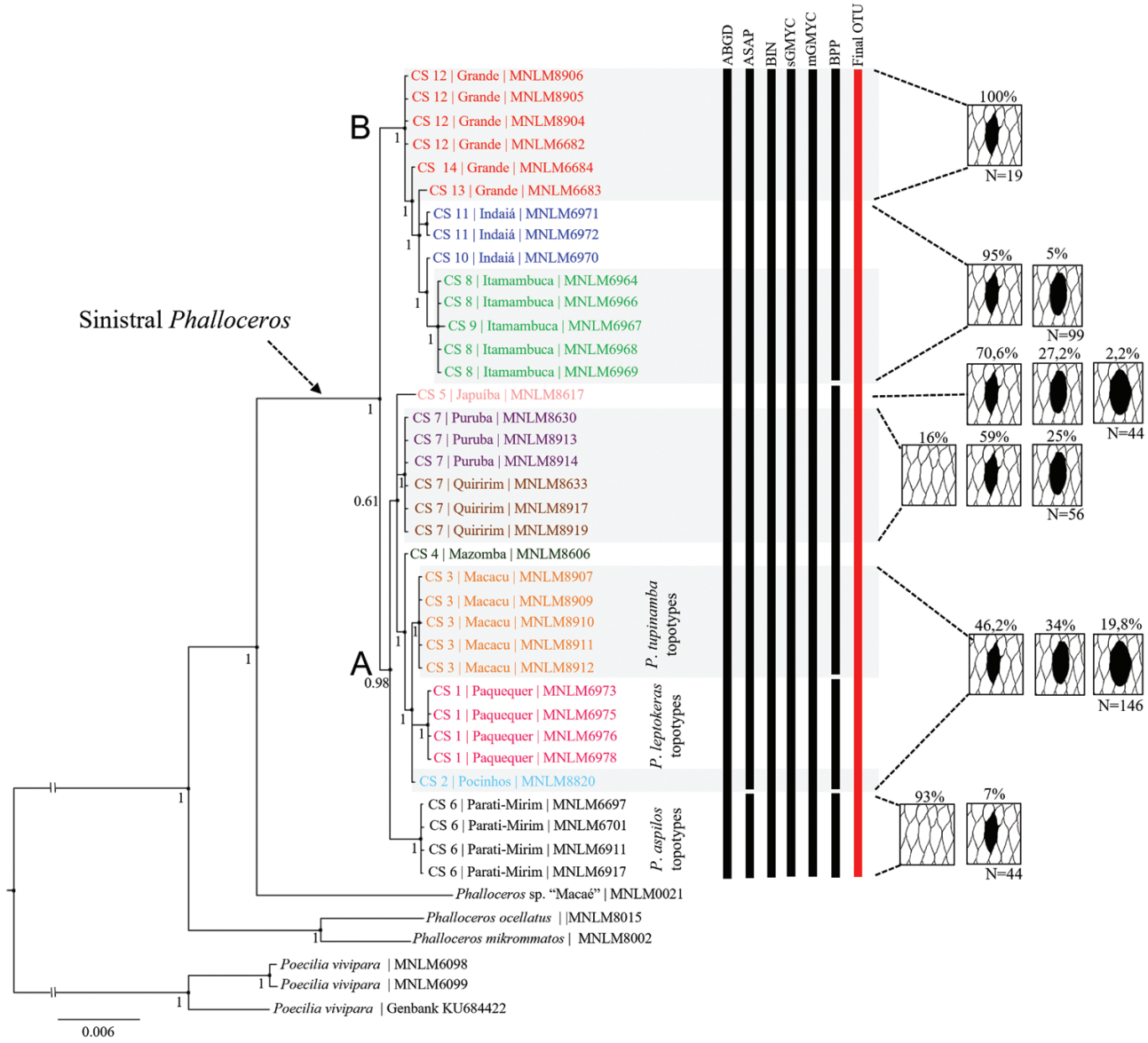

Figure 5. Phylogenetic relationships showing two main clades (A, B) within the sinistral *Phalloceros* group (= *Phalloceros leptokeras*) according to Bayesian inference (BI) based on concatenated sequences (CS) from the mitochondrial *COI* and *Cytb* regions and the nuclear *RAG1* gene. Values at nodes represent posterior probabilities. Samples from Parati-Mirim, Paquequer and Macacu are topotypes of *Phalloceros aspilos*, *P. leptokeras* and *P. tupinamba*, respectively. Vertical black bars show the results of species delimitation analyses (see details in Supporting Information, Files S2–S6). Vertical red bar shows our taxonomic decision. Frequencies in the size of the lateral spot are indicated above the drawings and the number of individuals examined (N) are below (see details in Supporting Information, Table S3).

We specified *heredity = 1* to indicate independent partitions (Yang, 2015) and *seed = –1* to start the analysis with a random seed. The other parameters were default values.

To test the palaeogeographic drainage hypotheses, we used the shapefiles made available by Thomaz & Knowles (2018), plotted with QUANTUM GIS software (QGIS Development Team, 2017) (Fig. 3). The vector representations of current rivers were drawn from Google Earth satellite images. The elevation shapefile was extracted from the digital elevation model GEBCO_2020 at 15 arc-second resolution (http://www.gebco.net/) with the *Slope Special Analyst* tool in ARCGIS PRO software.

RESULTS

Species delineation

A total of 4403 specimens of sinistral *Phalloceros* from 14 isolated drainages were examined, including type material of *P. aspilos*, *P. leptokeras* and *P. tupinamba* (Supporting Information, File S1). Morphological characters supported the hypothesis of group monophyly. In mature males, the tip of the gonopodium is symmetrical with slender appendices and a large hook in the medial portion of each appendix. Mature females have the urogenital papilla turned to the left side of the body.

Our morphological examination failed to recognize unambiguous diagnostic characters supporting the validity of nominal species. Specimens without a black lateral spot did not form a group (Fig. 5), despite Lucinda's (2008) reliance on this character as a diagnostic feature for *P. aspilos*. The black lateral spot in 464 sexually mature males and females was absent or covered by one, one-and-a-half or two scales in the midlateral series (Supporting Information, Table S3). The absence of the lateral spot is modal in the Parati-Mirim material (Supporting Information, Table S3). Our examination of lateral spot pigmentation revealed only three (of 44) specimens with the lateral spot present in the Parati-Mirim drainage (Supporting Information, Table S3). One of these specimens was included in our phylogenetic analysis (MNLM 6701) and was genetically identical to others collected in the same basin (Fig. 5). Nine specimens without pigmentation in the lateral spot were found in the drainages of the Quiririm and Puruba (Supporting Information, Table S3), small coastal drainages located approximately 33 km (in a straight line) west of the Parati-Mirim River. However, the samples from Quiririm and Puruba were not nested with those from Parati-Mirim. Instead, they belong to a distinct lineage according to our phylogenetic analysis (Fig. 5).

The position of the lateral spot on the flank and the shape of the urogenital papilla of females were not useful to differentiate species of sinistral *Phalloceros*. We observed that both holotypes of *Phalloceros tupinamba* and *P. leptokeras* have the black lateral spot on the 15th scale of the midlateral series. Among female paratypes of *P. tupinamba*, 11 possessed flattened urogenital papilla and four had the papilla not flattened.

We generated 42 *COI* sequences, which included 12 distinct *COI* haplotypes of *Phalloceros* and *Poecilia vivipara*. Eight haplotypes occur in the sinistral *Phalloceros* group (Supporting Information, Tables S4, S5). The pairwise K2P distances among sinistral *Phalloceros* (Hap *COI* 1–8) varied from 0.15 to 1.23%. Among the haplotypes of topotypes of the three nominal species, the average pairwise distances were 0.15% (*P. leptokeras* vs. *P. tupinamba*), 0.77% (*P. aspilos* vs. *P. tupinamba*) and 0.92% (*P. aspilos* vs. *P. leptokeras*). The distance from the sinistral *Phalloceros* group (Hap *COI* 1–8) to the closest outgroup, *Phalloceros* sp. 'Macaé' (Hap *COI* 9), varied from 6.4 to 7.32%.

The delimitation analyses of *COI* sequences of the sinistral *Phalloceros* yielded contradictory results (Fig. 5). The ABGD analysis grouped all sinistral *Phalloceros* in a single unit (Supporting Information, File S2). In contrast, the ASAP analysis produced four clustering options (Supporting Information, File S3). The option with the highest ASAP score divided sinistral *Phalloceros* into two operational taxonomic units (OTUs): (1) the population from the Parati-Mirim River and (2) a group containing the remaining sinistral *Phalloceros*. The BIN analyses grouped all sinistral *Phalloceros* in BOLD:AAB5571 as a single unit. Both GMYC delimitation analyses (sGMYC and mGMYC) were congruent with the ABGD and BIN analyses and grouped all sinistral *Phalloceros* into a single OTU (Supporting Information, Files S4, S5). The BPP analyses corroborated the partition of the sinistral *Phalloceros* into four groups with 0.993560 posterior probability (Supporting Information, File S6).

Phylogenetic relationships

The concatenated sequences (CS) comprised 2695 bp (656 for *COI*, 484 for *Cytb* and 1555 for *RAG1*) and included 29 variable sites among sinistral *Phalloceros*. The best evolutionary models found for *COI*, *Cytb* and *RAG1* were HKY+G, TVM+G and K80, respectively. The best log-likelihood score for the Bayesian inference (BI) analysis was –lnl 6179.148.

Among the sinistral *Phalloceros* we recovered 14 CSs (concatenated sequences of *COI*, *Cytb* and *RAG1*), grouped into two main clades (Clades A and B) with well-supported branch values (Fig. 5). Clade A (seven distinct CSs) is widely distributed, grouping a total of eight localities of the Paraíba do Sul River and coastal drainages from the Macacu to the Puruba drainages. The topotypes of *P. aspilos*, *P. leptokeras* and *P. tupinamba* nested in Clade A (Fig. 5). Clade A is supported by three mitochondrial molecular synapomorphies (Fig. 6). Each locality grouped in Clade A has a unique and exclusive CS, except for the Quiririm and Puruba Basins, which share a CS. The Parati-Mirim lineage (CS 6) was recovered as sister-group of the other species of Clade A. The Parati-Mirim samples formed the longest branch of the phylogeny (Fig. 5) due to six mitochondrial autapomorphies (Fig. 6). Despite samples from Batumirim Bay (CS 7) sharing the same palaeodrainage with those from Ubatuba Bay (CS 8–14, Clade B), they unexpectedly grouped in Clade A (Figs 3, 6).

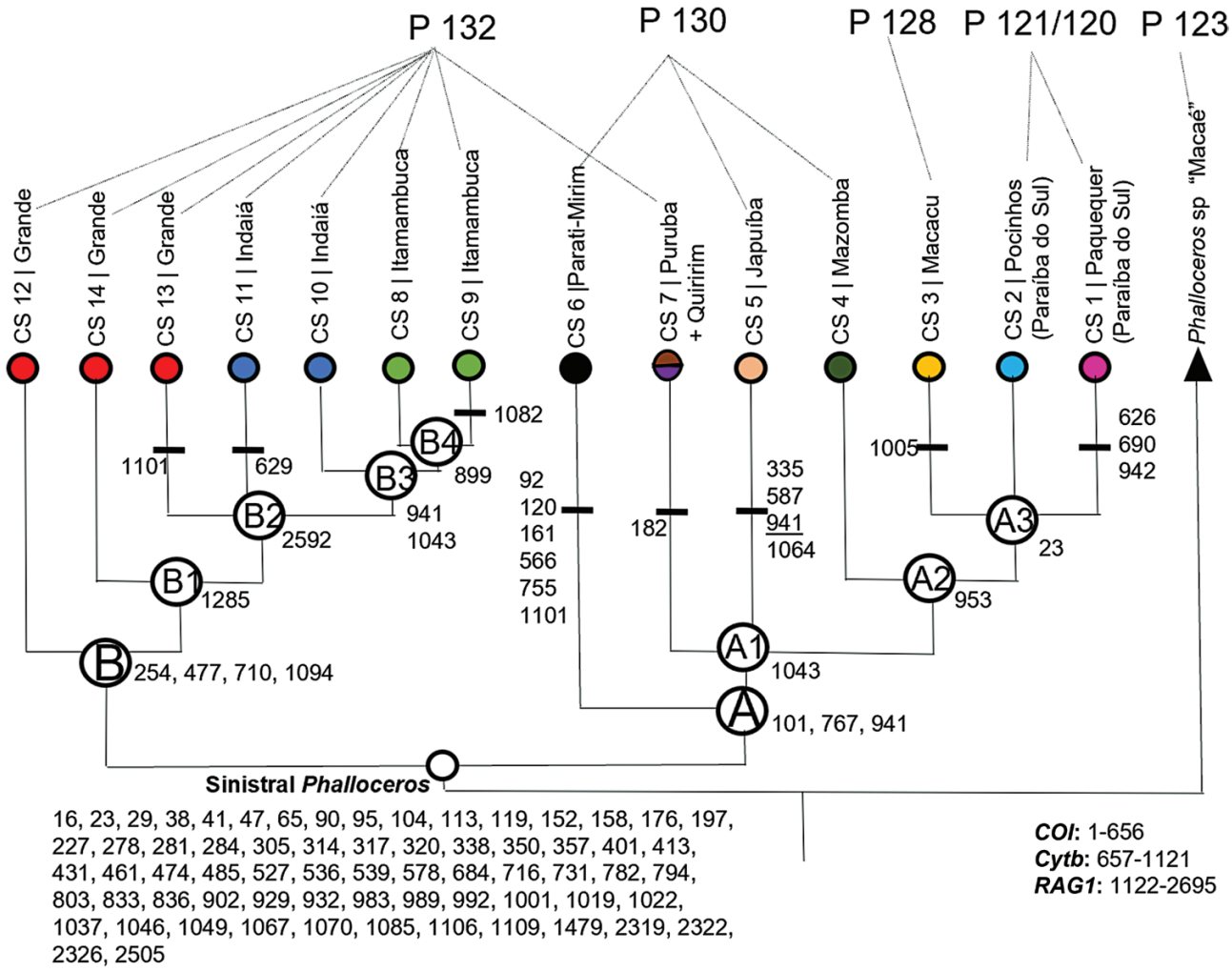

Figure 6. Phylogenetic relationships among the 14 sinistral *Phalloceros* (= *Phalloceros leptokeras*) with nucleotide substitutions mapped. Letter codes within circles indicate clade names as listed in the main text. Substitutions are identified according to the concatenated (CS) alignment of genes (*COI*: 1–656, *Cytb*: 657–1121, *RAG1*: 1122–2695), polarized based on *Phalloceros* sp. 'Macaé' as outgroup. The concatenated sequence alignment has 2695 sites, 29 of which are informative among sinistral *Phalloceros*. The underlined substitution of Genotype 5 from Japuíba represents a reversal to the ancestral character state of Clade A. This topological relationship was extracted from Figure 5. Palaeodrainage (P) numbers correspond to those proposed by Thomaz & Knowles (2018). According to those authors, the Paraíba do Sul River belonged to palaeodrainage 121 during the LGM, but currently its mouth is related to palaeodrainage 120.

We recovered a monophyletic group consisting of individuals sampled in the Paraíba do Sul drainage and the coastal drainages of Guanabara Bay (Fig. 3). Samples from Paraíba do Sul were found to be closely related to samples from the Macacu coastal drainage. The material from the subdrainage of the Paquequer River formed a polytomy with the material from Córrego dos Pocinhos and Macacu. The sister-group of the Paraíba do Sul + Macacu lineage is the lineage of the Mazomba River, a small coastal drainage that flows into Sepetiba Bay (Figs 3, 6).

Clade B (seven distinct CSs) is restricted to the Palaeodrainage 132, occurring in the Itamambuca, Indaiá and Grande River drainages at the south-western end of the sinistral *Phalloceros* group distribution (Figs 3, 6). Clade B is supported by three mitochondrial synapomorphies and one nuclear synapomorphy (Fig. 6). Unlike Clade A, Clade B includes samples from localities that have more than one CS (Figs 4, 5). The samples from Grande River possess three CSs (CS 12–14), forming paraphyletic lineages at the base of Clade B. The two CSs of Indaiá (CS 10–11) are also paraphyletic. The two CSs from Itamambuca (CS 8–9) form a monophyletic group, sister to CS 10 from Indaiá (Fig. 5).

As expected, the mitochondrial DNA sequences (*COI* and *Cytb*) were more variable than the nuclear (*RAG1*)

sequences (Fig. 7). We detected eight haplotypes of *COI*, 11 of *Cytb* and three of *RAG1*. The structures of the *COI* and *Cytb* haplotypes networks are congruent with the phylogenetic analysis. Haplotype 1 of *RAG1* is present in all sampled localities, which suggests that this is the ancestral haplotype, retained in all populations. All three nuclear haplotypes occur in the Grande River, but the mitochondrial haplotypes are unique to this drainage.

The geographic expansion of the sinistral *Phalloceros* group began approximately 2.3 million years ago (Mya) (Supporting Information, Fig. S1). This divergence-time estimate corresponds to the boundary between the Late Pliocene and Early Pleistocene in the Quaternary period (http://quaternary.stratigraphy.org/major-divisions). The diversification of Clade A began about 1.7 Mya, while clade B diverged about 0.5 Mya. The age of the clade that unites the sinistral *Phalloceros* group with its sister-group is estimated to be 11.1 Mya (6–15 Mya) (Supporting Information, Fig. S1).

DISCUSSION

SPECIES DELIMITATION

Previous knowledge concerning phylogenetic relationships within the sinistral *Phalloceros* group is restricted to interpretations from morphological data (Lucinda, 2008) and a partial analysis based on molecular data of 12 specimens identified as *P. aspilos*, *P. tupinamba* and an unidentified species (Thomaz *et al.*, 2019). In the present study, we covered the entire known distribution of the sinistral *Phalloceros* group, including the type locality of *P. leptokeras*, to produce a comprehensive reinterpretation of the phylogenetic diversity of this group. As discussed below, we conclude that *P. aspilos*, *P. tupinamba* and *P. leptokeras* constitute a single species.

Phalloceros aspilos was originally diagnosed by a unique autapomorphic trait of the caudal skeleton, which has been described as 'hypural plate almost bipartite, with very large aperture' (character state 131-3 in: Lucinda & Reis, 2005; Lucinda 2008). However, Souto-Santos *et al.* (2019) examined the caudal skeleton of 32 specimens and concluded that the configuration of the hypural plate varies in samples from the Parati-Mirim drainage. The character state described by Lucinda & Reis (2005) was observed in only two males in the sample of the Parati-Mirim River examined by Souto-Santos *et al.* (2019). The remaining specimens exhibited a condition in which the hypural aperture is longer, almost eliminating the distal bony bridge connecting the two portions of the hypural plate [described as 'hypural plate partially fused with an

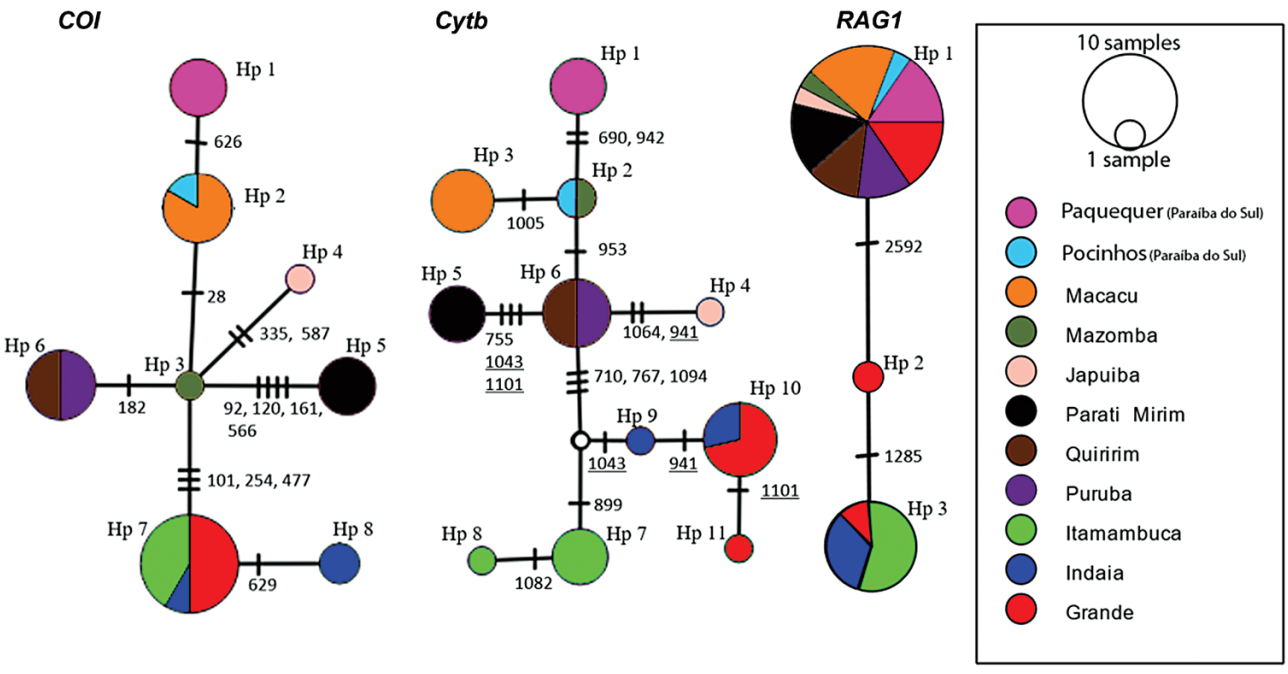

Figure 7. Haplotype (Hp) network of sinistral *Phalloceros* group (= *Phalloceros leptokeras*). Mutations are identified according to the concatenated alignment of genes (*COI*: 1–656, *Cytb*: 657–1121, *RAG1*: 1122–2695), non-exclusive nucleotide substitutions are underlined. Samples from Parati-Mirim, Paquequer and Macacu are topotypes of *Phalloceros aspilos*, *P. leptokeras* and *P. tupinamba*, respectively.

elongate aperture' character state 131-1 in Lucinda & Reis (2005)].

The absence of the lateral spot was the only remaining character diagnosing *P. aspilos* as a valid species (Souto-Santos *et al.*, 2019), but this character was found to be variable in the present study. Specimens of sinistral *Phalloceros* without lateral spots were observed in the Parati-Mirim, Quiririm and Puruba drainages. All specimens from the other basins have a black spot on the flank. The lateral spot size is variable, and the spot can cover one, one-and-a-half or two scales on the midlateral series (Figs 1, 5; Table S3). According to Lucinda (2008), *P. leptokeras* has a densely pigmented rectangular spot covering two or three scales (positioned between scales 14 and 16), while *P. tupinamba* has a narrow, vertically elongate, dark lateral spot covering one scale horizontally (positioned between scales 16 and 19). None of the specimens examined have the lateral spot covering three scales, and our examination of the type series showed the ineffectiveness of the lateral spot position for the differentiation of these species.

Lucinda (2008) also stated that *P. leptokeras* can be distinguished from *P. tupinamba* by the flattened urogenital papilla in females (vs. not flattened in *P. tupinamba*). This character varies even in the type series of *P. tupinamba* (MCP 20585). Among the female paratypes of *P. tupinamba*, 11 possess flattened urogenital papilla and four have the papilla not flattened.

Analyses of genetic delimitation of species show the great similarity among sinistral *Phalloceros*, including topotypes of *Phalloceros aspilos*, *P. leptokeras* and *P. tupinamba*. The results of the ABGD, BIN, sGMYC and mGMYC analyses proposed that all 11 populations (drainages) of sinistral *Phalloceros* belong to the same species. GMYC analysis is one of the most popular species-delimitation tools (Puillandre *et al.*, 2020), largely applied in studies of freshwater fish (e.g. Rossini *et al.*, 2016). ABGD, BIN and ASAP analyses are exploratory delimitation methods that do not require prior relationship hypotheses (Puillandre *et al.*, 2020). The other analyses (ASAP, BPP) divided the group into two and four groups, respectively (Fig. 5). According to the ASAP analysis, the Parati-Mirim River samples belong to a species distinct from the other sinistral *Phalloceros*. However, recognition of the Parati-Mirim lineage as a valid species renders the remaining group of populations paraphyletic. The results of the BPP analysis are useful to confirm the genetic structure of lineages but cannot be used to assert that these lineages represent species. This is because multispecies coalescent methods delimit genetic structure, not species (Sukumaran & Knowles, 2017).

Molecular species-delineation methods must be applied in taxonomy in an integrative way (Puillandre *et al.*, 2020). The Parati-Mirim material revealed four autapomorphies in the Folmer region of the *COI* gene, more than the other sampled localities (Fig. 5). ASAP analysis is a method based on genetic distance and, therefore, produces groups based on phenetic similarity without considering phylogenetic relationships. Although most individuals from Parati-Mirim lack a lateral spot (Supporting Information, Table S3), the ASAP result is not congruent with our hypothesis of phylogenetic relationships (Fig. 5). We follow Nixon & Wheeler's (1990) species concept, also adopted by Lucinda (2008): species are diagnosable groups of organisms representing lineages within a phylogenetic hypothesis. The results of the ASAP analysis are, therefore, not accepted because they violate the principle of monophyly implied by the results of our phylogenetic analyses. Instead, the results of the ABGD, BIN and GMYC analyses are accepted because they agree with the phylogenetic results.

Fish specimens with *COI* genetic distances less than 2% K2P distance are usually conspecific (Ward *et al.*, 2009; Pereira *et al.*, 2013). The highest pairwise divergence value among sinistral *Phalloceros* species was 1.23%, which was found between the samples from Parati-Mirim and Indaiá drainages. In contrast, the distance between the sinistral *Phalloceros* and their sister-group exceeded 6% (Supporting Information, Table S5). Accordingly, the BIN analysis supports the hypothesis that the sinistral *Phalloceros* are conspecific, as all our samples have been clustered within the same barcode index number, as estimated by the OTU designation through refined single linkage (RESL) analysis (Ratnasingham & Hebert, 2013).

As originally described, *P. tupinamba* comprised populations from the Macacu River and the Ubatuba Bay drainages (Lucinda, 2008). However, there is no morphological or molecular evidence supporting the monophyly of these disjunct populations. Samples associated with the type localities of the three nominal species of sinistral *Phalloceros* were nested in Clade A (Fig. 5). Topotypes of *P. tupinamba* (from Macacu) are more related to topotypes of *P. leptokeras* (from Paquequer) than to other sinistral *Phalloceros* from the Ubatuba Bay drainage (Clade B) where *P. tupinamba* paratypes were collected (Itamambuca, Indaiá and Grande).

Considering the combined molecular and morphological evidence, we propose synonymizing the names *P. aspilos*, *P. leptokeras* and *P. tupinamba*. The three nominal species were described at the same time (Lucinda, 2008) and as first reviewers (ICZN Article 24.2, https://www.iczn.org/the-code/the-international-code-of-zoological-nomenclature/the-code-online/) here we give

precedence to *P. leptokeras* as the valid name for sinistral *Phalloceros*. We chose the name *P. leptokeras* because it alludes to the slender gonopodial appendix (Fig. 3) shared by all sinistral *Phalloceros* (Lucinda, 2008).

Thomaz *et al.* (2019) proposed a lineage composed of *Phalloceros elachistos*, *P. mikrommatos* and *P. ocellatus* as the sister-group of *P. leptokeras* (= all sinistral *Phalloceros*). Here, samples of *P. mikrommatos* and *P. ocellatus* were included in the phylogenetic analysis. We also added in the analysis an undescribed species of *Phalloceros* that occurs in the drainages of Macaé, São João and Cabiúnas on the east coast of the State of Rio de Janeiro. *Phalloceros leptokeras* is more closely related to this undescribed species informally named here *Phalloceros* sp. 'Macaé' than to *P. mikrommatos* and *P. ocellatus*. *Phalloceros* sp. 'Macaé' can be distinguished from *P. leptokeras* by the straight urogenital papilla (vs. turned to the left in *P. leptokeras*). *Phalloceros* sp. 'Macaé' belongs to the *P. harpagos* species complex (Lucinda, 2008) and is being described elsewhere.

Phylogeographic patterns

The correlation among the geographic distribution of the haplotypes and their phylogenetic relationships is remarkable. The vicarious isolation of lineages A and B (Fig. 5) corresponds to the water divide separating the drainages of the Puruba and Itamambuca (Fig. 3). All drainages to the north-east of this limit belong to Clade A and the drainages to the south-west belong to Clade B.

A phylogeographic approach was used to evaluate the origin and relationships among *Phalloceros leptokeras* (= all sinistral *Phalloceros*) populations. The lineages of Clade B and the basal lineages of Clade A inhabit the coastal side of the Serra do Mar, as well as *Phalloceros* sp. 'Macaé', the sister-species of *P. leptokeras* (Figs 3, 5). This phylogenetic pattern suggests that the species has a coastal origin and later invaded the drainage of the Paraíba do Sul River on the northern slope of the Serra do Mar. The coastal origin of *Phalloceros leptokeras* and the beginning of the species diversification (2.3 Mya) pre-dates the LGM (0.26–0.19 Mya; Thomaz *et al.*, 2015), corroborating the hypothesis that ancestors of *P. leptokeras* inhabited exposed palaeochannels during periods of low sea-level during the Pleistocene period. The *Phalloceros leptokeras* clade groups samples from ten isolated drainages associated with four palaeodrainages (Figs 3, 6). Clade A is associated with all these palaeodrainages (i.e. Palaeodrainages 121, 128, 130 and 132) and Clade B is only associated with Palaeodrainage 132 (Table 1; Fig. 6).

Palaeodrainage 121 (or 120, see Fig. 3) is associated with the drainage of the Paraíba do Sul River, represented in our phylogeny by concatenated sequences (CS) 1 and 2 (Fig. 5). These CSs are related to CS 3 of the coastal drainage of the Macacu River that belongs to Palaeodrainage 128. There is no evidence of past relationships between the mouth of the Paraíba do Sul and the mouth of the Macacu Rivers, thus leading to the rejection of the palaeodrainages hypothesis as an explanation for the evolutionary relationship among lineages represented by CS 1–3. There are seven palaeodrainages between Palaeodrainage 121 and Palaeodrainage 128 (Fig. 3) and *Phalloceros leptokeras* has never been recorded in basins that flow into those palaeodrainages. There are also no records of sinistral *Phalloceros* in the lower portions of the Paraíba do Sul Basin. In contrast, the drainages of the Piabanha River (tributary of Paraíba do Sul) and the Macacu are adjacent and flow in opposite directions, isolated by the Serra dos Órgãos, an extension of the Serra do Mar. The relationships among the samples from the Paraíba do Sul and the Macacu suggest that *Phalloceros leptokeras* colonized the Paraíba do Sul drainage, overcoming the Serra dos Órgãos geographical barrier. Another possible scenario is lowland dispersal followed by local extinction in Palaeodrainages 122–127 region. The Serra dos Órgãos has a rugged relief of crystalline rocks and its maximum elevation can exceed 2000 m (Melo, 2001). The rocks that form the Serra dos Órgãos date from the Precambrian period and thus it would be more plausible to consider that *Phalloceros leptokeras* overcame this barrier through stretches of lower elevation in the Serra do Mar further west. The disjunct distribution of species isolated by the topography in the Serra do Mar is often associated with headwater capture events, when the hypothesis of dispersal through Pleistocene palaeodrainages is rejected (e.g. Lima *et al.*, 2016). However, there is no geological evidence in the literature corroborating

Table 1. Result of the palaeodrainage hypothesis test. Clades restricted to the same palaeodrainage (*sensu* Thomaz & Knowles, 2018) are marked in palaeodrainage dispersal. Clades with geographic distribution in different palaeodrainages are indicated in the inter palaeodrainage dispersal. Node ages are indicated in Mya

Clade from Figure 6	Node age	Palaeodrainage dispersal	Inter palaeodrainage dispersal
Sinistral *Phalloceros*	2.3		X
A	1.7		X
A1	1.1		X
A2	0.5		X
A3	0.4		X
B	0.5	X	

headwater capture events in the Serra dos Órgãos. Geological studies in this stretch of the Serra do Mar are necessary to detect evidence of headwater capture events that may explain the presence of *P. leptokeras* in the Paraíba do Sul Basin. Additionally, the great genetic similarity between the samples from the Paraíba do Sul and Macacu Rivers might be explained by human-mediated dispersal.

Palaeodrainage 130 is the largest in the study area. According to Thomaz & Knowles (2018), in the LGM this palaeodrainage connected all coastal drainages from Ilha Grande Bay and Sepetiba Bay, including the sampled drainages from Piraquê, Guandu, Mazomba, Muriqui, Saco, Japuíba and Parati-Mirim (Fig. 3). The Mazomba CS (CS 4, Palaeodrainage 130) is more related to the Paraíba do Sul CSs (CS 1–2, Palaeodrainage 121) and Macacu CS (CS 3, Palaeodrainage 128) than to the CSs sampled from the Japuíba (CS 5, Palaeodrainage 130) and Parati-Mirim (CS 6, Palaeodrainage 130) (Fig. 6). Despite sharing Palaeodrainage 130, the Japuíba and Parati-Mirim CSs do not form a monophyletic group. Therefore, the relationship of CS 4-6 (Fig. 6) cannot be explained by the palaeodrainage model.

The phylogenetic positioning of the Mazomba River CS suggests that sinistral *Phalloceros* dispersed along adjacent drainages, from the coastal basins of Sepetiba Bay (Palaeodrainage 130) towards the drainages of Guanabara Bay (Palaeodrainage 128) (Figs 3, 6). The large lowland area (Baixada Fluminense region) that stretches between the drainages of Sepetiba Bay and Guanabara Bay is a possible route of dispersal. These lowlands are occupied by extensive wetlands, which provide suitable habitats for small species of fish such as *Phalloceros*. Additionally, intense summer floods are frequent in the area and further support the hypothesis of lowland dispersal.

On the other hand, the three drainages in Ubatuba Bay that flow into Palaeodrainage 132 (Itamambuca, Indaiá and Grande) are grouped in Clade B (Figs 3, 6). The evolutionary origin of Clade B dates back to 0.5 Mya (Supporting Information, Fig. S1) and corroborates the hypothesis of faunal interchange through palaeodrainages exposed during Pleistocene marine regressions (Thomaz & Knowles, 2018). The lowering of sea level would have temporarily allowed the connection of these three rivers, prior to the isolation of the current drainages that promoted population differentiation.

The basins of Batumirim Bay (Puruba and Quiririm) flow into the east border of Palaeodrainage 132 (Figs 3, 6). However, the CS shared by Puruba and Quiririm (CS 7) is more closely related to the CSs of Clade A than to the CSs of other drainages associated with Palaeodrainage 132 (Fig. 5). The phylogenetic positioning of CS 7 suggests that the Clade A lineage colonized Palaeodrainage 132 after the establishment of Clade B lineage in that region. Serra do Parati, a stretch of Serra do Mar, is the main watershed divide between the drainages of Batumirim Bay and the drainages grouped by Clade A. Lima *et al.* (2016) hypothesized that there was a headwater capture event in the Serra do Parati between the Parati-Mirim and Camburi drainages to explain the colonization of the coastal basins of this southern slope of the Serra do Mar by the catfish *Trichogenes longipinnis* Britski & Ortega, 1983. However, Lima *et al.* (2016) did not specify the geological evidence for the proposed stream capture event nor its location. There are no records of *Phalloceros leptokeras* in the Camburi River, and it is equally plausible to explain the presence of lineage A in the basins of Batumirim Bay by dispersal instead of headwater capture. There is no geological, ecological or anthropic evidence to explain the dispersal of Clade A to Palaeodrainage 132. The drainages of the Quiririm and Puruba do not share a watershed divide with the Parati-Mirim. The Puruba headwaters share a watershed divide with the Perequê-Açu drainage, a basin adjacent to Parati-Mirim. However, the species of *Phalloceros* found in the Perequê-Açu drainage are not related to *P. leptokeras* (Souto-Santos *et al.*, 2019).

The area of occurrence of Clade A is dominated by mountain ranges, valleys, islands and bays (Fig. 3). This geographical complexity may have produced the complex faunal relationships that encompass various palaeodrainages. The area of occurrence of Clade B is more isolated and restricted to a narrow strip of high-gradient slope of the Serra do Mar (Fig. 3), thus confining the resident lineage to Palaeodrainage 132.

CONCLUSIONS

The group of sinistral *Phalloceros* corresponds to a single species. Here we propose *P. aspilos* and *P. tupinamba* as junior synonyms of *P. leptokeras*. The diversification of the *P. leptokeras* lineage started in the coastal region through a process of vicarious speciation. The watershed divide between the Puruba and Itamambuca River drainages isolates Clade A with a distribution to the north-east from Clade B to the south-west. The test of the palaeodrainage model proposed by Thomaz & Knowles (2018) explains only the evolution of Clade B, restricted to Palaeodrainage 132. *Phalloceros leptokeras* from Clade A dispersed along the coastal basins of Ilha Grande, Sepetiba and Guanabara Bays, and colonized the Paraíba do Sul Basin, thereby overcoming the geographical barrier of Serra dos Órgãos. There were at least three dispersal events between different palaeodrainages: (1) Palaeodrainage 130 to 132, (2) Palaeodrainage 130 to 128, and (3) Palaeodrainage 128 to 121.

ACKNOWLEDGEMENTS

This research is part of an M. Sc. thesis study conducted by ICASS at Programa de Pós-Graduação em Ciências Biológicas (Zoologia), Museu Nacional, Universidade Federal do Rio de Janeiro. We received support from Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq grants (proc. 312801/2017-3, 423526/2018-9, 131148/2019-2), Fundação de Amparo à Pesquisa no Estado do Rio de Janeiro – FAPERJ grants (proc. E-15 200.063/2019), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES grant (proc. 88882.156885/2016-01). Collecting activities were performed under permit number 12129-1, issued by the Sistema de Autorização e Informação em Biodiversidade – SISBIO and Instituto Chico Mendes de Conservação da Biodiversidade – ICMBio. We are grateful to Andrea Thomaz and Carlos Figueiredo for comments and suggestions on an earlier version of the manuscript. We are grateful to Aline Moreira and Thiago Parente for supporting DNA sequencing at the Genomic Platform – DNA Sequencing – RPT01A (Rede de Plataformas Tecnológicas, Fundação Oswaldo Cruz – FIOCRUZ). We thank Axel Katz, Cecília Malanski, Cristiano Moreira, Daniela Takiya, Decio Moraes, Emanuel Neuhaus, Eugenia Zandona, Gabriel Araújo, Gustavo Ferraro, Jefferson Ribeiro, Joana Zanol, Karina Ferreira, Manuela Dopazo, Marcelo Britto, Sérgio dos Santos, Thiago Barros, Victor de Brito and Vinicius Padula for help during field, laboratory and bioinformatics activities. Carlos Lucena and Roberto Reis provided access to the type series of *Phalloceros tupinamba* at the Museu de Ciências e Tecnologia, Pontifícia Universidade Católica do Rio Grande do Sul. Roberto Reis provided suggestions for improvement of the manuscript. Sergio Lima and Thomas Litz provided access to old literature about palaeogeographic hypotheses.

CONFLICT OF INTEREST

The authors declare they have no conflict of interest.

DATA AVAILABILITY

All sequences and associated data are available in the Barcode of Life Data Systems (BOLD Systems, http://boldsystems.org/), and GenBank (https://www.ncbi.nlm.nih.gov/genbank/) under process IDs and accession numbers listed in listed in the Supporting Information, Table S1.

REFERENCES

Amaral CRL, **Maciel VA**, **Pereira F**, **Mazzoni R**, **Silva DA**, **Amorim A**, **Carvalho EF. 2015.** Genetic diversity of freshwater fishes from the South American Atlantic Rainforest: the case study of the genus Phalloceros. *Forensic Science International: Genetics Supplement Series* **5:** e608608–e60e610.

Arroyave J, **Stiassny MLJ. 2011.** Phylogenetic relationships and the temporal context for the diversification of African characins of the family Alestidae (Ostariophysi: Characiformes): evidence from DNA sequence data. *Molecular Phylogenetics and Evolution* **60:** 385–397.

Arroyave J, **Denton JSS**, **Stiassny MLJ. 2013.** Are characiform fishes Gondwanan in origin? Insights from a time-scaled molecular phylogeny of the Citharinoidei (Ostariophysi: Characiformes). *PLoS One* **8:** e77269.

Avise JC. 2000. *Phylogeography: the history and formation of species*. Cambridge: Harvard University Press.

Bermingham E, **Martin AP. 1998.** Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology* **7:** 499–517.

Bouckaert R, **Heled J**, **Kuhnert D**, **Vaughan T**, **Wu CH**, **Xie D**, **Suchard MA**, **Rambaut A**, **Drummond AJ. 2014.** BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Bragança PHN, **Amorim PF**, **Costa WJEM. 2018.** Pantanodontidae (Teleostei, Cyprinodontiformes), the sister group to all other cyprinodontoid killifishes as inferred by molecular data. *Zoosystematics and Evolution* **94:** 137–145.

Buckup PA. 2011. The Eastern Brazilian Shield. In: Albert J, Reis RE, eds. *Historical biogeography of Neotropical freshwater fishes*. Berkeley: University of California Press, 203–210.

Costa WJEM. 2009. *Peixes aploqueilóideos da Mata Atlântica brasileira: história, diversidade e conservação*. Rio de Janeiro: Série Livros 34, Museu Nacional, Universidade Federal do Rio de Janeiro.

Crepaldi C. 2015. *Relações filogenéticas e filogeográficas de populações de* Phalloceros reisi *Lucinda, 2008 (Cyprinodontiformes, Poeciliidae) da bacia do alto rio Paraná e bacia costeira*. Unpublished M. Sc. Thesis, Universidade Estadual Paulista.

Drummond AJ, **Rambaut A. 2007.** BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Drummond AJ, **Bouckaert RR. 2015.** *Bayesian evolutionary analysis with BEAST*. Cambridge: Cambridge University Press.

Handy SM, **Deeds JR**, **Ivanova NV**, **Hebert PDN**, **Hanner RH**, **Moore MM**, **Yancy HF. 2011.** A single-laboratory validated method for the generation of dna barcodes for the identification of fish for regulatory compliance. *Journal of AOAC International* **94:** 201–210.

Hebert PDN, **Cywinska A**, **Ball SL**, **deWaard JR. 2003a.** Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* **270:** 313–321.

Hebert PDN, **Ratnasingham S**, **deWaard JR. 2003b.** Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences* **270:** 596–599.

Henriques JM, **Costa-Silva GJ**, **Ashikaga FY**, **Hanner R**, **Foresti F**, **Oliveira C. 2015.** Use of DNA barcode in the identification of fish species from Ribeira de Iguape Basin and coastal rivers from São Paulo State (Brazil). *DNA Barcodes* **3:** 118–128.

Hensel R. 1868. Beiträge zur Kenntniss der Wirbelthiere Südbrasiliens. *Archiv für Naturgeschichte* **34:** 323–375.

Hirschmann A, **Malabarba LR**, **Thomaz AT. 2015.** Riverine habitat specificity constrains dispersion in a Neotropical fish (Characidae) along southern Brazilian drainages. *Zoologica Scripta* **44:** 374–382.

von Ihering H. 1927. *Die Geschichte des Atlantischen Ozeans*. Jena: G. Fischer.

Jennings WB. 2017. *Phylogenomic data acquisition: principles and practice*. Boca Raton: CRC Press/Taylor and Francis.

Jennings WB, **Ruschi PA**, **Ferraro G**, **Quijada CC**, **Silva-Malanski ACG**, **Prosdocimi F**, **Buckup PA. 2019.** Barcoding the Neotropical freshwater fish fauna using a new pair of universal *COI* primers with a discussion of primer dimers and M13 primer tails. *Genome* **62:** 77–83.

Kumar S, **Stecher G**, **Li M**, **Knyaz C**, **Tamura K. 2018.** MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35:** 1547–1549.

Lanfear R, **Frandsen PB**, **Wright AM**, **Senfeld T**, **Calcott B. 2016.** PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **34:** 772–773.

Leigh JW, **Bryant D. 2015.** PopArt: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6:** 1110–1116.

Lima SMQ, **Vasconcellos AV**, **Berbel-Filho WM**, **Lazoski C**, **Russo CAM**, **Sazima I**, **Solé-Cava AM. 2016.** Effects of Pleistocene climatic and geomorphological changes on the population structure of the restricted-range catfish Trichogenes longipinnis (Siluriformes: Trichomycteridae). *Systematics and Biodiversity* **14:** 155–170.

Lima SMQ, **Berbel-Filho WM**, **Araújo TFP**, **Lazzarotto H**, **Tatarenkov A**, **Avise JC. 2017.** Headwater capture evidenced by paleo-rivers reconstruction and population genetic structure of the armored catfish (*Pareiorhaphis garbei*) in the Serra do Mar mountains of southeastern Brazil. *Frontiers in Genetics* **8:** 199.

Lima SMQ, **Berbel-Filho WM**, **Vilasboa A**, **Lazoski C**, **Volpi TA**, **Lazzarotto H**, **Russo CAM**, **Tatarenkov A**, **Avise J**, **Solé-Cava AM. 2021.** Rio de Janeiro and other palaeodrainages evidenced by the genetic structure of an Atlantic Forest catfish. *Journal of Biogeography* **48:** 1475–1488.

Lis JT. 1980. Fractionation of DNA fragments by polyethylene glycol induced precipitation. *Methods in Enzymology* **65:** 347–353.

Lucinda PHF. 2008. Systematics and biogeography of the genus *Phalloceros* Eigenmann, 1907 (Cyprinodontiformes: Poeciliidae: Poeciliinae), with the description of twenty-one new species. *Neotropical Ichthyology* **6:** 113–158.

Lucinda PHF, **Reis RE. 2005.** Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. *Neotropical Ichthyology* **3:** 1–60.

Menezes NA, **Ribeiro AC**, **Weitzman S**, **Torres RA**. **2008.** Biogeography of Glandulocaudinae (Teleostei: Characiformes: Characidae) revisited: phylogenetic patterns, historical geology and genetic connectivity. *Zootaxa* **1726:** 33–48.

Melo FAG. 2001. Revisão taxonômica das espécies do gênero *Astyanax* Baird e Girard, 1854, (Teleostei: Characiformes: Characidae) da região da Serra dos Órgãos. *Arquivos do Museu Nacional* **59:** 1–46.

Miller SA, **Dykes DD**, **Polesky HFRN. 1988.** A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* **16:** 1215.

Nixon KC, **Wheeler QD. 1990.** An amplification of the phylogenetic species concept. *Cladistics* **6:** 211–223.

Oliveira D. 2003. *A captura do Alto Rio Guaratuba: uma proposta metodológica para o estudo da evolução do relevo na Serra do Mar, Boracéia-SP*. Unpublished Ph. D. Thesis, Universidade de São Paulo.

Oliveira IJ. 2017. *Filogeografia de um complexo de espécies de Phalloceros harpagos Lucinda, 2008 (Cyprinodontiformes: Poeciliidae)*. Unpublished M. Sc. Thesis, Universidade Estadual de Maringá.

Pereira LHG, **Hanner R**, **Foresti F**, **Oliveira**. **2013.** Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna. *BMC Genetics* **14:** 20.

Pons J, **Barraclough TG**, **Gomez-Zurita J**, **Cardoso A**, **Duran DP**, **Hazell S**, **Kamoun S**, **Sumlin WD**, **Vogler AP. 2006.** Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55:** 595–609.

Puillandre N, **Lambert A**, **Brouillet**, **Achaz G. 2012.** ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* **21:** 1864–1877.

Puillandre N, **Brouillet S**, **Achaz G. 2020.** ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources* **21:** 609–620.

QGIS Development Team. **2017.** *QGIS Geographic Information System. Open Source Geospatial Foundation Project*. Available at: http://qgis.org. Accessed 23 March 2022.

Rambaut A, **Drummond AJ**, **Xie D**, **Baele G**, **Suchard MA. 2018.** Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67:** 901–904.

Ratnasingham S, **Hebert PDN. 2007.** BOLD: the barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes* **7:** 355–364.

Ratnasingham S, **Hebert PDN. 2013.** A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS One* **8:** e66213.

Reznick DN, **Furness AI**, **Meredith RW**, **Springer MS. 2017.** The origin and biogeographic diversification of fishes in the family Poeciliidae. *PLoS One* **2:** e0172546.

Ribeiro AC. 2006. Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. *Neotropical Ichthyology* **4:** 225–246.

Ronquist F, **Teslenko M**, **Van der Mark P**, **Ayres DL**, **Darling A**, **Hohna S**, **Larget B**, **Liu L**, **Suchard MA**, **Huelsenbeck JP. 2012.** MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Rossini BC, **Oliveira CAM**, **Melo FAG**, **Bertaco VA**, **Astarloa JMD**, **Rosso JJ**, **Foresti F**, **Oliveira C. 2016.** Highlighting *Astyanax* species diversity through DNA barcoding. *PLoS One* **11:** e0167203.

SEMADS. **2001.** *Bacias hidrográficas e rios fluminenses, síntese informativa por macrorregião ambiental*. Rio de Janeiro: SEMADS.

Shimabukuro-Dias CK, **Costa-Silva GJ**, **Ashikaga FY**, **Foresti F**, **Oliveira C. 2017.** Molecular identification of the fish fauna from the pantanal flood plain area in Brazil. *Mitochondrial DNA Part A* **28:** 588–592.

Souto-Santos ICA, **Ferraro GA**, **Jennings WB**, **Vergara GLS**, **Buckup PA. 2019.** Geographic distribution of *Phalloceros* Eigenmann, 1907 (Cyprinodontiformes, Poeciliidae) in the Ilha Grande Bay Hydrographic Region, Rio de Janeiro, Brazil. *Check List* **15:** 181–192.

Sukumaran J, **Knowles LL. 2017.** Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences of the USA* **144:** 1607–1612.

Templeton AR, **Crandall KA**, **Sing CF. 1992.** A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132:** 619–633.

Thomaz AT, **Knowles LL. 2018.** Flowing into the unknown: inferred paleodrainages for studying the ichthyofauna of Brazilian coastal rivers. *Neotropical Ichthyology* **16:** e180019.

Thomaz AT, **Malabarba LR**, **Bonatto SL**, **Knowles LL. 2015.** Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: study of a Neotropical fish of the Brazilian coastal Atlantic Forest. *Journal of Biogeography* **42:** 2389–2401.

Thomaz AT, **Carvalho TP**, **Malabarba LR**, **Knowles LL. 2019.** Geographic distributions, phenotypes, and phylogenetic relationships of *Phalloceros* (Cyprinodontiformes: Poeciliidae): insights about diversification among sympatric species pools. *Molecular Phylogenetics and Evolution* **132:** 265–274.

Torres RA, **Ribeiro J. 2009.** The remarkable species complex *Mimagoniates microlepis* Characiformes: Glandulocaudinae) from the southern Atlantic rain forest (Brazil) as revealed by molecular systematic and population genetic analyses. *Hydrobiologia* **617:** 157–170.

Torres RA, **Motta TS**, **Nardino DN**, **Adam ML**, **Ribeiro J. 2008.** Chromosomes, RAPDs and evolutionary trends of the Neotropical fish *Mimagoniates microlepis* (Teleostei: Characidae: Glandulocaudinae) from coastal and continental regions of the Atlantic forest, southern Brazil. *Acta Zoologica* **89:** 253–259.

Tzeng CS, **Lin YS**, **Wang TY**, **Wang FY. 2006.** The phylogeography and population demographics of selected freshwater fishes in Taiwan. *Zoological Studies* **45:** 285–297.

Ward RD, **Hanner R**, **Hebert PDN. 2009.** The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology* **74:** 329–356.

Yang Z. 2015. The BPP program for species tree estimation and species delimitation. *Current Zoology* **61:** 854–865.

Zamudio KR, **Robertson JM**, **Chan LM**, **Sazima I. 2009.** Population structure in the catfish *Trichogenes longipinnis*: drift offset by asymmetrical migration in a tiny geographic range. *Biological Journal of the Linnean Society* **97:** 259–274.

Zhang C, **Zhang DX**, **Zhu T**, **Yang Z. 2011.** Evaluation of a Bayesian coalescent method of species delimitation. *Systematic Biology* **60:** 747–761.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

File S1. Material examined. Species are listed in alphabetical order and river drainages are listed from north to south.

File S2. ABGD results following parameters recommended by Rossini *et al*. (2016). All sinistral *Phalloceros* were grouped together. Details of the 36 *COI* samples are in Table S4.

File S3. ASAP results following default parameters of Puillandre *et al*. (2020). The partition with the best ASAP score (3.50) divided the sinistral *Phalloceros* into two groups. Details of the 36 *COI* samples are given in Table S4.

File S4. GMYC single threshold result. A single sequence was selected to represent each *COI* haplotype (see details in the methodology and Table S4). Black colour represents distinct OTU's and red represents the same OTU. Sinistral *Phalloceros* are nested in a unique OTU.

File S5. GMYC multi threshold results. A single sequence was selected to represent each *COI* haplotype (see details in the methodology and Table S4). Black colour represents distinct OTU's and red represents the same OTU. Sinistral *Phalloceros* are nested in a unique OTU.

File S6. BPP results. The partition with the highest probability is indicated in yellow. Number 1 in model represents lineages (nodes) of distinct partitions. Number 0 represents lineages that grouped in the same partition.

Figure S1. Estimates of gene divergence time for the sinistral *Phalloceros* group based on Bayesian inference for the mitochondrial *COI* and *Cytb* region, and the nuclear RAG1 region, calibrated with the age of 42 Mys attributed to the divergence between *Phalloceros* and *Poecilia* (Reznick *et al.*, 2017). Scale bar in millions of years before present. The 95% credibility intervals for node ages are shown with horizontal blue bars. Values at

nodes represent posterior probabilities. Samples from Parati-Mirim, Paquequer and Macacu represent topotypes of *Phalloceros aspilos*, *P. leptokeras*, and *P. tupinamba*, respectively.

Table S1. List of specimens used to obtain molecular data. Sample ID refers to the DNA Extract Collection of the Laboratório de Pesquisa em Biodiversidade Molecular, Museu Nacional (MNLM) (http://www.boldsystems.org/). Voucher code refers to the whole-organism catalog numbers deposited in the ichthyological collection of the Museu Nacional (MNRJ), Specimen/tissue code refers to tissue catalog number from the collection of fish tissues of the Museu Nacional (MNTI).

Table S2. Primers and PCR profiles used in the study.

Table S3. Size of the lateral spot in adults of sinistral *Phalloceros* group (*sensu* Lucinda, 2008). Holotypes were examined on their left (L), and right (R) sides; other specimens were examined on the left side. Drainages associated with DNA sequencing samples are indicated with an asterisk. Bold voucher numbers indicate paratypes. Drainages 1 to 4 flow to the Rio Paraíba do Sul, and drainages 5 to 17 flow directly into the Atlantic Ocean. Palaeodrainage numbers correspond to those proposed by Thomaz & Knowles (2018). According to those authors, the Rio Paraíba do Sul belonged to palaeodrainage 121 during the LGM, but currently its mouth is related to palaeodrainage 120.

Table S4. *COI* haplotypes associated with sinistral *Phalloceros* group. Haplotypes nine to 14 correspond to samples used as outgroup in the phylogenetic analyses. Locality codes correspond to the way each drainage was identified in the ABGD, ASAP and GMYC analyses (Supplementary Files S2-S5). Palaeodrainage numbers correspond to those proposed by Thomaz & Knowles (2018). According to those authors, the Rio Paraíba do Sul belonged to palaeodrainage 121 during the LGM, but currently its mouth is related to palaeodrainage 120. Haplotypes of the outgroup are shaded.

Table S5. Pairwise estimates of evolutionary sequence divergence (K2P genetic distance) between samples. Haplotypes correspond to those listed in Table S4. Values given as percent. Haplotypes of the outgroup are shaded.