UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO

INSTITUTO DE BIOCIÊNCIAS DE BOTUCATU

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS (ZOOLOGIA)

TESE DE DOUTORADO

CHARACINAE (ACTINOPTERYGII: CHARACIFORMES: CHARACIDAE): IDENTIFICAÇÃO MOLECULAR E ESTUDO DAS RELAÇÕES FILOGENÉTICAS ENTRE ESPÉCIES

CAMILA DA SILVA DE SOUZA

BOTUCATU-SP

2022

UNIVERSIDADE ESTADUAL PAULISTA

INSTITUTO DE BIOCIÊNCIAS DE BOTUCATU

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS (ZOOLOGIA)

CHARACINAE (ACTINOPTERYGII: CHARACIFORMES: CHARACIDAE): IDENTIFICAÇÃO MOLECULAR E ESTUDO DAS RELAÇÕES FILOGENÉTICAS ENTRE ESPÉCIES

Discente: Camila da Silva de Souza

Orientador: Dr. Claudio de Oliveira

Co-orientador: Dr. Bruno Francelino de Melo

Tese apresentada ao Programa de Pós-Graduação em Ciências Biológicas (Zoologia) do Instituto de Biociências de Botucatu, Universidade Estadual Paulista "Júlio de Mesquita Filho", como parte dos requisitos para a obtenção do título de Doutorado.

BOTUCATU-SP

2022

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Souza, Camila da Silva de.
Characinae (actinopterygii: characiformes: characidae) : identificação molecular e estudo das relações filogenéticas entre espécies / Camila da Silva de Souza. - Botucatu, 2022

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu
Orientador: Claudio de Oliveira
Coorientador: Bruno Francelino de Melo
Capes: 20400004

1. Characinae. 2. Biodiversidade. 3. Filogenia. 4. Zoologia - Classificação. 5. Código de barras de DNA taxonômico.

Palavras-chave: Biodiversidade; DNA Barcode; Filogenia; Sistemática; Taxonomia.

AVISO

Esta tese é parte dos requerimentos necessários à obtenção do título de Doutor em Ciências Biológicas, Zoologia, e como tal, não deve ser vista como uma publicação no senso do Código Internacional de Nomenclatura Zoológica (apesar de disponível publicamente sem restrições). Desta forma, quaisquer informações inéditas, opiniões e hipóteses, bem como nomes novos, mudanças taxonômicas não estão disponíveis na literatura zoológica. Pessoas interessadas devem estar cientes de que referências públicas ao conteúdo deste estudo, na sua presente forma, somente devem ser feitas com a aprovação prévia do autor.

NOTICE

This thesis is a partial requeriment for the PhD degree in Biological Sciences with emphasis in Zoology and, as such, should not be considered as a publication in the sense of the International Code of Zoological Nomenclature (although it is available without restrictions). Therefore, any new information, opinions, and hypotheses, as well as new names, taxonomic changes ae not available in the zoological literuature. Interested people are advised that any public reference to this study, in its current form, should only be done after previous acceptance of the author.

Esta pesquisa foi financiada com recursos da
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e
Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) - Processo
Número: 2017/06551-0 e 2019/04602-1

"It seems to me that the natural world is the greatest source of excitement; the greatest source of visual beauty; the greatest source of intellectual interest. It is the greatest source of so much in life that makes life worth living." (*David Attenborough*)

À vida

AGRADECIMENTOS

Deixo meus sinceros agradecimentos às instituições e pessoas que de forma direta ou indireta, contribuíram para a concretização deste trabalho.

Ao auxílio financeiro concedido pelo processo nº 2017/06551-0 e 2019/04602-1, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Ao Dr. Claudio de Oliveira, pela orientação acadêmica em todas as fases do projeto, e às oportunidades oferecidas desde a iniciação científica, essenciais para a minha formação.

Ao Dr. Bruno F. Melo pela orientação, amizade e imensurável motivação!

Ao Dr. Ricardo Benine pela oportunidade em realizar o estágio de docência sob sua orientação e pelas discussões e ensinamentos.

Ao Dr. George M.T. Mattox e Dr. Mônica Toledo-Piza pelas grandiosas colaborações e conhecimento compartilhado.

Ao Dr. Brian L. Sidlauskas, por ter me acolhido e orientado na Oregon State University - USA durante o doutorado-sanduíche e pelos valiosos exemplos e ensinamentos sobre a vida profissional e pessoal.

Ao Dr. Adamm Summers, pela oportunidade de obter dados para o meu projeto do doutorado-sanduíche no Friday Harbor Laboratories - University of Washington - USA.

Ao Dr. Thaddaeus Buser, Álvaro Cortés, Hakan Aydogan e Jonathan Huie pela ajuda e colaboração durante o doutorado-sanduíche.

À Dr. Luz Ochoa e Dr. Nadayca Mateussi pela amizade e por compartilharem seus conhecimentos para a execução desse trabalho e terem sido minhas maiores referências próximas da representatividade feminina na ictiologia.

À Dr. Rafaela Ota e ao Dr. Victor Tagliacollo, Dr. Fernando Carvalho e Dr. Fernando Jerep por todos os ensinamentos e sugestões na última etapa deste trabalho.

A todos integrantes e ex-integrantes do Laboratório de Biologia e Genética de Peixes, em especial a Angélica Dias, Beatriz Dorini, Bruno Mora, Cristiane Shimabukuro-Dias, Cristhian Conde, Diogo Freitas, Daniela Oliveira, Duílio Zerbinato, Érica Serrano, Fabilene Gomes, Fabio Roxo,

Gabriel Costa e Silva, Guilherme J. da Costa Silva, Junior Flausino, Lais Reia, Maria Ligia, Natalia Mendez, Sâmia Mouallem, Silvana Melo e Renato Devidé, por todos os momentos de descontração, incentivo e colaboração.

As grandes amigas Aline Veríssimo, Carolina Magalhães, Laura Borgatto e Milene Pereira, por todo companheirismo e motivação ao longo da vida acadêmica e pessoal.

À Eric. V. Ywamoto por ter chegado na minha vida no começo desse projeto e ser um dos motivos de eu seguir em frente e finalizá-lo. Meus sinceros agradecimentos por toda ajuda, companheirismo e por sempre acreditar em mim, muitas vezes mais do que eu mesma.

A toda minha família que são fonte de motivação e inspiração e em especial a Giulia, Mirella, Miguel, Bruna, Tiago, David, Felipe, Cinthia, Rosa e William.

E, finalmente, deixo meu eterno agradecimento aos meus pais, Valdomiro e Silvana, pelo amor, suporte e por sempre incentivarem minhas aventuras pelo mundo em busca de experiências e sabedoria.

RESUMO

Characinae é uma das subfamílias mais diversas de Characidae com 85 espécies válidas amplamente distribuídas pelas América do Sul e Central. A subfamília é representada por espécies de pequeno a médio porte que adotam diferentes estratégias alimentares, tais como carnivoria, onivoria e lepidofagia. As relações filogenéticas dentro de Characinae e desta com outras subfamílias de Characidae têm sido objeto de alguns estudos morfológicos, porém, não há, até o momento, hipóteses de relacionamento incluindo uma significativa representatividade de espécies de Characinae usando caracteres morfológicos ou moleculares. Além disso, os estudos publicados ainda apresentam incongruências nas relações intergenéricas e interespecíficas. Neste contexto, o presente trabalho teve como objetivos testar as hipóteses de relações filogenéticas da subfamília Characinae através de análises filogenômicas utilizando os elementos ultraconservados (UCEs) e construir um banco de dados genéticos de DNA *barcode* para identificar geneticamente as espécies de Characinae. Os resultados das análises filogenéticas corroboram a hipótese morfológica em relação à monofilia de Characinae e revelam novas hipóteses de relações intergenéricas e interespecíficas. Com base na filogenia obtida, analisamos origem e diversificação, assim como a evolução do tamanho e formato do corpo. Os resultados das identificações moleculares reconheceram uma diversidade antes subestimada que deverá contribuir para ampliação do conhecimento sobre a diversidade das espécies de Characinae.

PALAVRAS-CHAVE: DNA Barcode, Biodiversidade, Filogenia, Sistemática, Taxonomia, Peixes.

ABSTRACT

Characinae is one of the most speciose subfamilies of Characidae and widely distributed throughout South and Central America. The subfamily contains small to medium-sized species that adopt distinct feeding strategies as carnivory, omnivory and lepidophagy. Phylogenetic relationships within the Characinae and with other subfamilies have been subject of a few morphological studies, although none hypothesis included a representative number of Characinae species using either morphological or molecular characters. Furthermore, the proposed hypotheses still present inconsistencies at intergeneric and interspecific levels. In this context, the present study aimed to test the hypotheses of phylogenetic relationships of the subfamily Characinae through phylogenomic analyzes using ultraconserved elements (UCEs) and to generate a barcode DNA genetic database for molecular identification of Characinae. The results of the phylogenetic analyses corroborate the morphological hypothesis about the monophyly of Characinae and reveal new hypotheses of intergeneric and interspecific relationships. Based on the phylogeny obtained, we analyzed the origin, diversification, and the evolution of body size and body shape. The results of molecular identification recognized a previously underestimated diversity that should contribute to the expansion of knowledge about the diversity of Characinae.

Keywords: DNA Barcode, Biodiversity, Phylogenomics, Systematics, Taxonomy, Fish.

SUMÁRIO

INTRODUÇÃO GERAL

Subfamília Characinae

Characinae Eigenmann, 1910 representa a terceira maior subfamília de Characidae (Teleostei: Characiformes) (Fricke *et al.*, 2022) e *sensu* Mattox & Toledo-Piza, 2012 compreende 85 espécies (Fricke *et al.*, 2022) agrupadas em sete gêneros: *Acanthocharax* Eigenmann, 1912, *Acestrocephalus* Eigenmann 1910, *Charax* Scopoli 1777, *Cynopotamus* Valenciennes 1850, *Phenacogaster* Eigenmann 1907, *Galeocharax* Fowler 1910 e *Roeboides* Günther, 1864 (Tabela 1). As espécies de Characinae são conhecidas popularmente como peixes-cachorra, peixes-cigarra, dentudos e tetra-vidros (*glass tetras*), dentre outros nomes (Géry 1977; Mattox & Toledo-Piza 2012) e estão amplamente distribuídas pela região Neotropical desde o sul do México até afluentes da bacia dos rios Paraguai e Uruguai na Argentina, com ocorrência principalmente em rios de fluxo lento associados com a vegetação marginal (Lucena & Menezes 2003; Mattox *et al.*, 2017). A maioria das espécies são facilmente reconhecidas devido ao formato alto do corpo, especialmente na região anterior onde uma gibosidade é característica, embora esta esteja ausente em *Phenacogaster* e *Acestrocephalus* (Lucena & Menezes 2003). São espécies de pequeno a médio porte que apresentam diferentes hábitos alimentares, como a carnivoria adquirida pela maioria dos gêneros, a onivoria presente nas espécies do gênero *Phenacogaster* e a lepidofagia nas espécies de *Roeboides*, que apresentam modificações especializadas como dentes mamiliformes externos para esse tipo de habito alimentar (Géry 1977; Sazima & Machado 1982; Sazima 1984; Lucena & Menezes 2003). As espécies de Characinae não exercem grande importância para a economia pesqueira, mas apresentam importância para a pesca de subsistência de populações ribeirinhas e é de interesse da aquariofilia, sendo reconhecidas como espécies com potencial ornamental (Venere & Garutti 2001; Hercos *et al.*, 2009).

Na sistemática Characinae tem uma notória importância por conter *Charax*, o gênero-tipo da família e da ordem. Apesar dessa notoriedade a história taxômica de Characinae foi baseada por muito tempo a agrupamentos que incluíram representantes da família Characidae (Eigenmann 1910, 1912; Myers 1960; Géry 1966, 1977; Weitzman & Vari 1987; Lucena 1998, Lucena & Menezes 2003). A última definição de agrupamento foi proposto por Lucena & Menezes, 2003 que incluiu 12 gêneros *Acanthocharax*; *Acestrocephalus*; *Charax*; *Cynopotamus*; *Galeocharax*; *Gnathocharax* Fowler 1913; *Heterocharax* Eigenmann 1912; *Hoplocharax* Géry 1966; *Lonchogenys* Myers 1927; *Phenacogaster*; *Priocharax* Weitzman & Vari 1987; *Roeboides*) baseados no formato do corpo, presença de mais de 20 dentes cônicos na maxila, pseudotímpano na frente da primeira costela pleural e nadadeira peitoral larval em espécimes de até 41,0 mm de comprimento.

Mirande (2009, 2010) baseado em dados filogenéticos morfológicos e da literatura recuperou a monofilia de Characinae incluindo oito dos 12 gêneros (*sensu* Lucena & Menezes, 2003): *Acanthocharax, Acestrocephalus, Charax, Cynopotamus, Galeocharax, Phenacogaster, Priocharax* e *Roeboides*. Adicionalmente, propôs três gêneros, *Bryconexodon* Géry, 1980, *Exodon* Müller & Troschel 1844 e *Roeboexodon* Géry 1959 como pertencente a Characinae formando um clado-irmão de todos os gêneros remanescentes da subfamília. No entanto, mais recentemente com um conjunto mais amplo de táxons e caracteres, reinterpretou esses resultados e atribuiu *Bryconexodon, Exodon* e *Roeboexodon* para sua própria subfamília, Exodontinae (Mirande 2019: 296).

Mattox & Toledo-Piza (2012) realizaram o primeiro estudo filogenético de Characinae baseado em caracteres morfológicos. Seus resultados indicaram Characinae irmã de *Tetragonopterus* Cuvier, 1816 e internamente restringida ao clado (*Phenacogaster* ((*Charax+Roeboides*) (*Acanthocharax* (*Cynopotamus* (*Acestrocephalus+Galeocharax*))))). *Priocharax* é recuperado numa politomia composta por *Lonchogenys*, *Heterocharax* e *Hoplocharax* dentro de Heterocharacinae. Os autores ainda propõem/redefinem as tribos Phenacogasterini (*Phenacogaster*), Characini (*Charax* e

Roeboides) e Cynopotamini (*Acanthocharax, Acestrocephalus, Cynopotamus* e *Galeocharax*). Os demais gêneros propostos por Lucena & Menezes, 2003 foram incluídos em Heterocharacinae e constituíram a tribo Heterocharacini (*Gnatocharax* (*Lonchogenys* (*Heterocharax+Hoplocharax*))) irmã de Roestini (*Roestes* Günther 1864 + *Gilbertolus* Eigenmann 1907) e distante dos Characinae (*sensu* Lucena & Menezes 2003).

Outros estudos de Characiformes também incluíram representantes de Characinae usando dados moleculares (Javonillo *et al.*, 2010; Oliveira *et al.,* 2011; Tagliacollo *et al.,* 2012; Melo *et al.*, 2016; Betancur- R *et al.,* 2019) e análise de evidência total (Mirande 2019). Estes estudos concordam sobre a relação entre Characinae e Tetragonopterinae (Javonillo *et al.,* 2010; Oliveira *et al.,* 2011; Tagliacollo *et al.,* 2012; Melo *et al.,* 2016; Mirande, 2019; Betancur- R *et al.,* 2019); no entanto, os estudos moleculares não corroboram a monofilia de *Cynopotamus* (Oliveira *et al.,* 2011; Tagliacollo *et al.,* 2012; Melo *et al.,* 2016) e de *Roeboides* (Betancur-R *et al.,* 2019). Adicionalmente, *Acanthocharax* é hipotetizado como o grupo irmão de (*Charax+Roeboides*) ou (*Cynopotamus* (*Acestrochephalus+Galeocharax*)) em estudos morfológicos (Lucena 1998; Mattox & Toledo-Piza 2012), mas ainda não foram incluídos em estudos moleculares.

Relações interespecíficas em Characinae

Hipóteses de relações interespecíficas em Characinae ainda são escassas e restritas às análises morfológicas, revisões taxonômicas ou descrições de espécies. *Phenacogaster*, um dos gêneros mais diversos, ainda necessita de descrições de muitas espécies e da definição de subgrupos (Lucena & Malabarba 2010). Das 23 espécies válidas, quatro (*P. beni* Eigenmann 1911, *P. microstictus* Eigenmann 1909, *P. pectinata* (Cope 1870) e *P. suborbitalis* Ahl 1936) apresentam ampla distribuição e compõem o complexo *P. pectinata* (Lucena & Malabarba 2010). Além delas, várias espécies não descritas compõem este grupo (Lucena & Malabarba 2010) e a dificuldade em

diagnosticar essas espécies tem sido explicada pela ampla distribuição geográfica e ausência de diferenças morfológicas consistentes (Géry 1972; Lucena 2003; Lucena & Malabarba 2010).

Lucena (1998) propõe quatro subunidades de *Roeboides*: grupo *dispar* (*R. dispar* Lucena 2001), grupo *microlepis* (*R. araguaito* Lucena, 2003, *R. margareteae* Lucena 2003, *R. microlepis* (Reinhardt 1851) e *R. myersii* Gill 1870), grupo *affinis* [*R. affinis* (Günther, 1868), *R. biserialis* (Garman, 1890), *R. descalvadensis* Fowler 1932, *R. numerosus* Lucena 2000, *R. oligistos* Lucena 2000, *R. paranensis* Pignalberi 1975 (= *R. descalvadensis*), *R. prognathus* (Boulenger 1895) (=*R. affinis*), *R. thurni* Eigenmann 1912 (=*R. affinis*), *R. xenodon* (Reinhardt 1851) e *R. sazimai* Lucena 2007)]) e grupo *guatemalensis* (*R. bouchellei* Fowler 1923, *R. carti* Lucena 2000, *R. dayi* (Steindachner 1878), *R. dientonito* Schultz 1944, *R. guatemalensis* (Günther 1864), *R. ilsea* Bussing 1986, *R. occidentalis* Meek & Hildebrand 1916 e *R. loftini* Lucena 2011). O grupo *microlepis* é hipotetizado como mais relacionado às espécies transandinas do grupo *guatemalensis* (Lucena 2000). Alternativamente, Mattox & Toledo-Piza (2012), mostraram a relação (*R. dientonito* ((*R. affinis* + *R. descalvadensis*) (*R. myersii* (*R. occidentalis* + *R. xenodon*))))).

Recentemente publicada, a revisão de *Charax* inclui a redescrição de todas as espécies além da descrição de *C. delimai* Menezes & Lucena 2014 e a sinonimização de *C. unimaculatus* Lucena 1989 em *C. michaeli* Lucena 1989 (Menezes & Lucena 2014). A única hipótese de relacionamento é apresentada baseada em caracteres morfológicos e inclui quatro das 17 espécies (Mattox & Toledo-Piza 2012). Esta filogenia recupera a relação ((*C. condei* + *C. stenopterus*) (*C. gibbosus* + *C. pauciradiatus*)).

Cynopotamus contém 12 espécies (Fricke *et al*., 2022) e apenas a relação (*C. xinguano* Menezes 2007 (*C. gouldingi* Menezes 1987 (*C. kincaidi* (Schultz 1950) (*C. juruenae* Menezes 1987 + *C. tocantinenses* Menezes 1987))) foi hipotetizada através de caracteres morfológicos (Mattox & Toledo-Piza 2012). Dados moleculares não recuperaram a monofilia de *Cynopotamus*, sendo *C.*

kincaidi irmão de *Roeboides guatemalensis* e *C. venezulae* (Schultz 1944) irmão de *Acestrocephalus+Galeocharax* (Oliveira *et al.,* 2011).

Acestrocephalus teve cinco espécies descritas recentemente num total de oito (Menezes 2006) e a única hipótese existente consiste na relação (*A. acutus* Menezes 2006 (*A. sardina* (Fowler 1913) (*A. pallidus* Menezes 2006 + *A. stigmatus* Menezes 2006))) (Mattox & Toledo-Piza 2012).

Galeocharax foi recentemente revisado (Giovannetti *et al.,* 2017) e apenas três espécies foram consideradas como válidas (*G. gulo* (Cope 1870), *G. goeldii* (Fowler 1913), *G. humeralis* (Valenciennes 1834), e com *G. knerii* (Steindachner 1879) considerada sinônima de *G. gulo*. A filogenia morfológica de Characinae contendo todas as espécies de *Galeocharax* não resolve as relações, apresentando uma politomia (Mattox & Toledo-Piza 2012). Nenhuma filogenia molecular entre espécies desses gêneros, nem estudos aplicando técnicas de filogenômica foram realizadas em Characinae até presente momento.

DNA Barcode

Estudos baseados em identificação molecular usando a metodologia DNA barcode (Hebert *et al.,* 2003) tem revelado números subestimados de espécies antes não reconhecidas e vem se mostrando resolutivos para questões taxonômicas em diferentes grupos de peixes neotropicais, tais como Characidae (Pereira *et al.,* 2011; Silva *et al.* ,2013, Barreto *et al*., 2018), Lebiasinidae (Benzaquem *et al.,* 2015), Loricariidae (Costa-Silva *et al.,* 2015; de Borba *et al.,* 2019; Fagundes *et al*., 2020), Serrasalmidae (Machado *et al.,* 2018; Mateussi *et al.,* 2020; Ota *et al.,* 2020) e Curimatidae (Melo *et al.,* 2016b). O uso dessa técnica pode então ser promissor para Characinae podendo possivelmente revelar uma diversidade ainda não reconhecida, uma vez que não há estudos de identificação molecular para Characinae e os estudos recentes têm identificado novas espécies (Menezes & Lucena 2014; Lucena, Antonetti & Lucena 2018; Guimarães, Brito, Ferreira & Ottoni 2018; Lucena

& Lucena 2019) principalmente em *Roeboides* e *Phenacogaster* que existem espécies de ampla distribuição e em *Phenacogaster* que apresenta o complexo *P. pectinata* e outras espécies ainda não descritas (Lucena & Malabarba 2010). Com isso, o DNA barcoding foi utilizado no presente estudo de Charcinae a fim de elucidar as unidades taxonômicas que os compõem e no direcionamento dos estudos filogenômicos.

Elementos ultraconservados (UCEs)

Os elementos ultraconservados (*ultraconserved elements* - UCEs) são regiões do genoma extremamente conservadas e assim compartilhadas entre grupos pertencentes a linhagens muito distintas, como, por exemplo, aves e humanos (Bejerano *et al.,* 2004). Eles foram descritos por Bejerano *et al.* (2004), que encontraram 481 segmentos maiores de 200 pares de bases que eram absolutamente (100%) conservados em regiões ortólogas de humanos, ratos e camundongos e altamente conservados nos genomas de galinhas e cães (95–99%, respectivamente). Estudos posteriores mostraram que os UCEs também estão presentes em diversos outros organismos, como outros vertebrados, insetos e fungos (Siepel *et al.,* 2005; Faircloth *et al.,* 2012).

O papel dos UCEs no genoma ainda não está esclarecido (Dermitzakis *et al.*, 2005), embora os UCEs tenham sido associados com regulação gênica ou desenvolvimento (Sandelin *et al.,* 2004; Woolfe *et al.,* 2004) e se tem assumido que os UCEs são importantes pela sua natureza extremamente conservada entre grupos muito distantes filogeneticamente. Os UCEs são identificados nos organismos pelo alinhamento de vários genomas e por regiões desse alinhamento com áreas com conservação de sequências muito altas (95–100%) e filtradas utilizando critérios específicos como o comprimento das sequências (e.g. Bejerano *et al.,* 2004). Suas sequências conservadas permitem uma fácil identificação e alinhamento entre genomas e a premissa de contínua variabilidade nas sequências que flanqueiam cada UCE sugere que eles podem ser um tipo

de “fóssil molecular”, retendo um sinal de história evolutiva em diversas escalas de tempo, dependendo da distância da região central dos UCEs (Faircloth *et al.,* 2012).

Faircloth *et al.* (2012) introduziram os UCEs como uma nova classe de marcadores moleculares em estudos filogenéticos através do enriquecimento de bibliotecas genômicas contendo centenas ou milhares de loci, utilizando sequenciamento de nova geração (Faircloth *et al.,* 2012). Como as sequências de UCEs são altamente conservadas elas são utilizadas para o anelamento de sondas (probes), a partir das quais as sequências flanqueadoras dos UCEs são lidas. A presença de regiões com diferentes níveis de variabilidade tem tornado esta técnica muito promissora no campo da sistemática filogenética (Pennisi 2013). Além disso, os UCEs têm sido utilizados em muitos níveis de comparação entre organismos, de populações até grandes grupos (McCormack *et al.*, 2012, 2013; Crawford *et al.,* 2012; Smith *et al.,* 2014; Starrett *et al.,* 2016).

Em peixes o primeiro estudo foi realizado por Faircloth *et al.* (2013). Nesse estudo foram sequenciados 500 loci de cerca de 30 espécies de Actinopterygii e as filogenias obtidas mostram nós altamente resolvidos em todos níveis (recentes e antigos). Os resultados suportaram as relações entre *Amia* e *Lepisosteus* (Holostei) e revelaram que os Elopomorpha e depois os Osteoglossomorpha são as primeiras linhagens a divergir entre as linhagens dos teleósteos. Em sequência, outros trabalhos foram realizados mostrando os UCEs como promissor e excelente marcador molecular para estudos filogenéticos: Amarsipidae (Harrington *et al.,* 2016), Characiformes (Chakrabarty *et al.*, 2017); Acanthomorpha (Alfaro *et al.,* 2018), Loricariidae (Roxo *et al.,* 2019), Trichomycteridae (Ochoa *et al*., 2020), Serrasalmidae (Mateussi *et al*., 2020) e Heptapteridae (Silva *et al.*, 2021).

Em comparação com os estudos baseados em marcadores moleculares multi-locus o uso dos UCEs tem se mostrado bastante eficiente. A razão para este rápido crescimento está ligada a diferentes características de sua abordagem, como a obtenção de dados de eventos de divergência

recente e antiga (Crawford *et al.*, 2012; Harvey *et al.*, 2016; Manthey *et al.*, 2016) e alto custo-benefício em função do tempo e baixo custo dada a grande quantidade de dados gerados.

JUSTIFICATIVA

Várias questões no âmbito da sistemática de Characinae ainda permanecem não resolvidas como levantadas acima e aqui sumarizadas: 1) a posição de *Acanthocharax* ainda é incerta, devido à falta de amostragem para evidência molecular; 2) a posição de *Priocharax* ainda não foi testada utilizando dados moleculares; 3) a monofilia de *Cynopotamus* foi rejeitada pelas hipóteses moleculares de Oliveira *et al.* (2011), Tagliacollo *et al*. (2012) e Melo *et al*. (2016) e de *Roeboides* em Betancur-R *et al.* (2019); 4) as relações interespecíficas em cada um dos gêneros necessitam de uma maior amostragem de táxons terminais; 5) a diversidade molecular de espécies não é conhecida. A Tabela 1 mostra o número de táxons analisados pelos principais estudos morfológicos e moleculares de Characidae (Mirande 2018), morfológico da subfamília Characinae (Mattox & Toledo-Piza 2012) e molecular de Characidae (Oliveira *et al.*, 2011). Pode-se notar uma necessidade de maior amostragem de espécies de Characinae para geração de hipóteses filogenéticas mais robustas. Além disso, vários problemas taxonômicos ainda existentes em vários gêneros de Characinae podem ser esclarecidos com a identificação molecular de espécies. Finalmente, os recentes avanços na sistemática molecular com a utilização de supermatrizes em escala genômica tem proporcionado uma ótima oportunidade para obter filogenias mais robustas, e sua consequente utilização em estudos macroevolutivos.

Tabela 1. Número de espécies analisadas por gênero de Characinae em estudos anteriores e em nosso estudo.

Gênero	Espécies válidas	Oliveira *et al*., 2011	Mattox & Toledo-Piza, 2012	Mirande, 2018	Este estudo
Acanthocharax	1	0	1	0	**1**
Acestrocephalus	8	1	4	1	**3**
Charax	17	1	4	2	**10**
Cynopotamus	12	2	5	1	**9**
Galeocharax	3	1	3	1	**3**
Phenacogaster	23	1	4	3	**13**
Roeboides	21	1	6	2	**18**
Total	**85**	7	27	10	**57**

OBJETIVOS

Com base nas hipóteses prévias de relacionamentos interespecíficos e intergenéricos da subfamília Characinae, e desta com outras subfamílias de Characidae, o presente projeto teve como objetivos:

1) Elaborar e testar as hipóteses de relações filogenéticas entre os gêneros da subfamília Characinae e desta com os demais membros de Characidae;

2) Elaborar e testar as hipóteses de relações filogenéticas entre as espécies de cada um dos gêneros de Characinae;

3) Realizar a reconstrução ancestral da dieta de Characinae e testar se as mudanças na dieta influenciaram a diversificação morfológica do tamanho e forma do corpo;

4) Gerar sequências utilizando a técnica de DNA *barcode* para permitir uma identificação molecular dos peixes da subfamília Characinae;

5) Identificar linhagens/ espécies descritas ou não descritas.

BIBLIOGRAFIA

Alfaro, M. E., Faircloth, B. C., Harrington, R. C., Sorenson, L., Friedman, M., Thacker, C. E., ... & Near, T. J. (2018). Explosive diversification of marine fishes at the Cretaceous–Palaeogene boundary. Nature Ecology & Evolution, 2(4), 688-696.

Antonetti, D. A., Lucena, C. A. S., & Lucena, Z. M. S. (2018). Two new species of *Phenacogaster* (Actinopterygii: Characidae) from the rio Tocantins drainage, Brazil. Verlag Dr. Friedrich Pfeil.

Barreto, S. B., Silva, A. T., Batalha-Filho, H., Affonso, P. R. A., & Zanata, A. M. (2018). Integrative approach reveals a new species of *Nematocharax* (Teleostei: Characidae). Journal of fish biology, 93(6), 1151-1162.

Benzaquem, D. C., Oliveira, C., da Silva Batista, J., Zuanon, J., & Porto, J. I. R. (2015). DNA barcoding in pencilfishes (Lebiasinidae: Nannostomus) reveals cryptic diversity across the Brazilian Amazon. PloS one, 10(2), e0112217.

Betancur-R, R., Arcila, D., Vari, R. P., Hughes, L. C., Oliveira, C., Sabaj, M. H., & Ortí, G. (2019). Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: The monophyly of characiform fishes. Evolution, 73(2), 329-345.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W. J., Mattick, J. S., & Haussler, D. (2004). Ultraconserved elements in the human genome. Science, 304(5675), 1321-1325.

Costa-Silva, G. J., Rodriguez, M. S., Roxo, F. F., Foresti, F., & Oliveira, C. (2015). Using different methods to access the difficult task of delimiting species in a complex Neotropical hyperdiverse group. PLoS One, 10(9), e0135075.

Dermitzakis E.T., Reymond A. & Antonarakis S.E. (2005). Conserved non-genic sequences — an unexpected feature of mammalian genomes. Nature Reviews Genetics 6, 151–157.

Chakrabarty, P., Faircloth, B. C., Alda, F., Ludt, W. B., Mcmahan, C. D., Near, T. J., ... & Sorenson, L. (2017). Phylogenomic systematics of ostariophysan fishes: ultraconserved elements support the surprising non-monophyly of Characiformes. Systematic Biology, 66(6), 881-895.

Crawford, N. G., Faircloth, B. C., McCormack, J. E., Brumfield, R. T., Winker, K., & Glenn, T. C. (2012). More than 1000 ultraconserved elements provide evidence that turtles are the sister group of archosaurs. Biology letters, 8(5), 783-786.

Crawford, N.G., Parham, J.F., Sellas, A.B., Faircloth, B.C., Glenn, T.C., Papenfuss, T.J., Henderson, J.B., Hansen, M.H. & Simison, W.B. (2015) A phylogenomic analysis of turtles. Molecular Phylogenetics and Evolution 83, 250–257.

de Borba, R. S., Mariotto, S., Centofante, L., Henrique Zawadzki, C., & Pasquali Parise-Maltempi, P. (2019). Molecular discrimination of *Ancistrus* lineages (Siluriformes: Loricariidae) using barcode DNA tool. Mitochondrial DNA Part A, 30(4), 602-608.

Eigenmann, C. H. (1910). Catalogue of the fresh-water fishes of tropical and south temperate America, p. 375–511.

Eigenmann, C. H. (1912). The freshwater fishes of British Guiana, including a study of the ecological grouping of species and the relation of the fauna of the plateau to that of the lowlands (No. 114). Carnegie Institute.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T. & Glenn T.C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61, 717–726.

Faircloth B.C., Sorenson L., Santini F. & Alfaro M.E. (2013). A Phylogenomic Perspective on the Radiation of Ray-Finned Fishes Based upon Targeted Sequencing of Ultraconserved Elements (UCEs). PLoS ONE 8, e65923.

Fagundes, P. C., Pereira, E. H., & Reis, R. E. (2020). Iterative taxonomic study of *Pareiorhaphis hystrix* (Siluriformes, Loricariidae) suggests a single, yet phenotypically variable, species in south Brazil. PloS one, 15(9), e0237160.

Fricke, R., Eschmeyer, W. N. & R. van der Laan. (2022). Eschmeyer's catalog of fishes: genera, species, references. Available at http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (last accessed 12 Jan 2022).

Géry, J. (1966). *Hoplocharax goethei*, a new genus and species of South American characoid fishes, with a review of the sub-tribe Heterocharacini. Ichthyologica, 38, 281-296.

Géry, J. (1972). Corrected and supplemented descriptions of certain characoid fishes described by Henry W. Fowler, with revisions of several of their genera. Studies on Neotropical Fauna and Environment, 7(1), 1-35.

Géry J. (1977). Characoids of the world. Neptune City, NJ: THF Publications.

Giovannetti, V., Toledo-Piza, M., & Menezes, N. A. (2017). Taxonomic revision of *Galeocharax* (Characiformes: Characidae: Characinae). Neotropical Ichthyology, 15(1).

Guimarães, E. C., OTTONI, F. P., Ferreira, B. R. A., & BRITO, P. S. D. (2018). A new species of *Charax* (Ostariophysi, Characiformes, Characidae) from northeastern Brazil.

Harrington, R. C., Faircloth, B. C., Eytan, R. I., Smith, W. L., Near, T. J., Alfaro, M. E., & Friedman, M. (2016). Phylogenomic analysis of carangimorph fishes reveals flatfish asymmetry arose in a blink of the evolutionary eye. BMC Evolutionary Biology, 16(1), 224.

Harvey, M. G., Smith, B. T., Glenn, T. C., Faircloth, B. C., & Brumfield, R. T. (2016). Sequence capture versus restriction site associated DNA sequencing for shallow systematics. Systematic biology, 65(5), 910-924.

Hercos, A. P., de Queiroz, H. L., & de Almeida, H. L. (2009). Peixes ornamentais do Amanã. Instituto de Desenvolvimento Sustentável Mamirauá.

Hebert, P. D., Ratnasingham, S., & De Waard, J. R. (2003). Barcoding animal life: *cytochrome c oxidase subunit 1* divergences among closely related species. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(suppl_1), S96-S99.

Javonillo, R., Malabarba, L. R., Weitzman, S. H., & Burns, J. R. (2010). Relationships among major lineages of characid fishes (Teleostei: Ostariophysi: Characiformes), based on molecular sequence data. Molecular Phylogenetics and Evolution, 54(2), 498-511.

Lucena, C. D. (1998). Relações filogenéticas e definição do gênero *Roeboides* Günther (Ostariophysi: Characiformes: Characidae). Comunicações do Museu de Ciências e Tecnologia da PUCRS, série Zoologia, 11, 19-59.

Lucena, C. D. (2000). Revisão taxonômica e filogenia das espécies transandinas do gênero *Roeboides* Günther (Teleostei: Ostariophysi: Characiformes). Comunicações do Museu de Ciências e Tecnologia da PUCRS, série Zoologia, 13, 3-63.

Lucena, C. A. S. (2003). Revisão taxonômica e relações filogenéticas das espécies de *Roeboides* grupo-*microlepis* (Ostariophysi, Characiformes, Characidae). Iheringia. Série Zoologia, 93(3), 283-308.

Lucena C. A. S., Menezes N. A. (2003). Subfamily Characinae. In: Reis RE, Kullander SO, Ferraris CJ Jr, orgs. Checklist of the freshwater fishes of South and Central America. Porto Alegre, RS: EDIPUCRS, 200–208.

Lucena, C. A. S. D. (2007). Revisão taxonômica das espécies do gênero *Roeboides* grupo-*affinis* (Ostariophysi, Characiformes, Characidae). Iheringia. Série Zoologia, 97(2), 117-136.

Lucena, Z. M. S. D., & Malabarba, L. R. (2010). Description of nine new species of *Phenacogaster* (Ostariophysi: Characiformes: Characidae) and notes on the other species of genus. Zoologia (Curitiba), 27(2), 263-304.

Lucena, Z. M. S. D., & Lucena, C. A. S. D. (2019). A new glass tetra species of *Phenacogaster* from the rio Salitre, rio São Francisco drainage, Brazil (Characiformes: Characidae). Neotropical Ichthyology, 17.

Machado, V. N., Collins, R. A., Ota, R. P., Andrade, M. C., Farias, I. P., & Hrbek, T. (2018). One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. Scientific reports, 8(1), 1-12.

Manthey, J. D., Campillo, L. C., Burns, K. J., & Moyle, R. G. (2016). Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (Aves, Genus: Piranga). Systematic biology, 65(4), 640-650.

Matamoros, W. A., Chakrabarty, P., Angulo, A., Garita-Alvarado, C. A., & McMahan, C. D. (2013). A new species of *Roeboides* (Teleostei: Characidae) from Costa Rica and Panama, with a key to the middle American species of the genus. Neotropical ichthyology, 11, 285-290.

Mateussi, N. T., Melo, B. F., & Oliveira, C. (2020). Molecular delimitation and taxonomic revision of the wimple piranha *Catoprion* (Characiformes: Serrasalmidae) with the description of a new species. Journal of Fish Biology, 97(3), 668-685.

Mateussi, N. T., Melo, B. F., Ota, R. P., Roxo, F. F., Ochoa, L. E., Foresti, F., & Oliveira, C. (2020). Phylogenomics of the Neotropical fish family Serrasalmidae with a novel intrafamilial classification (Teleostei: Characiformes). Molecular Phylogenetics and Evolution, 153, 106945.

Mattox, G. M., & Toledo-Piza, M. (2012). Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). Zoological Journal of the Linnean Society, 165(4), 809-915.

Mattox, G. M., van der Sleen, P., Toledo-Piza, M. (2017). Subfamily Characinae—Characine Tetras. In: Van der Sleen P., Albert, J.S. editors. Field Guide to the Fishes of the Amazon, Orinoco, and Guianas. Princeton, New Jersey, Princeton University Press, p. 102-106.

McCormack, J. E., Faircloth, B. C., Crawford, N. G., Gowaty, P. A., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome research, 22(4), 746-754.

McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfield, R. T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular phylogenetics and evolution, 66(2), 526-538.

Melo, B. F., Benine, R. C., Silva, G. S., Avelino, G. S., & Oliveira, C. (2016). Molecular phylogeny of the Neotropical fish genus *Tetragonopterus* (Teleostei: Characiformes: Characidae). Molecular phylogenetics and evolution, 94, 709-717.

Melo, B. F., Ochoa, L. E., Vari, R. P., & Oliveira, C. (2016b). Cryptic species in the Neotropical fish genus *Curimatopsis* (Teleostei, Characiformes). Zoologica Scripta, 45(6), 650-658.

Menezes, N. A. (2006). Description of five new species of *Acestrocephalus* Eigenmann and redescription of *A. sardina* and *A. boehlkei* (Characiformes: Characidae). Neotropical Ichthyology, 4(4), 385-400.

Menezes, N. A. (2007). A new species of *Cynopotamus* Valenciennes, 1849 (Characiformes: Characidae) with a key to species of the genus. Zootaxa, 1635(1), 55-61.

Menezes, N. A., & Lucena, C. A. S. D. (2014). A taxonomic review of the species of *Charax* Scopoli, 1777 (Teleostei: Characidae: Characinae) with description of a new species from the rio Negro bearing superficial neuromasts on body scales, Amazon basin, Brazil. Neotropical Ichthyology, 12(2), 193-228.

Mirande, J. M. (2009). Weighted parsimony phylogeny of the family Characidae (Teleostei: Characiformes). Cladistics, 25(6), 574-613.

Mirande, J. M. (2010). Phylogeny of the family Characidae (Teleostei: Characiformes): from characters to taxonomy. Neotropical Ichthyology, 8(3), 385-568.

Mirande, J. M. (2019). Morphology, molecules and the phylogeny of Characidae (Teleostei, Characiformes). Cladistics, 35(3), 282-300.

Myers, G. S. (1960). The South American characid genera *Exodon*, *Gnathoplax*, and *Roeboexodon*, with notes on the ecology and taxonomy of characid fishes.

Ochoa, L. E., Datovo, A., DoNascimiento, C., Roxo, F. F., Sabaj, M. H., Chang, J., ... & Oliveira, C. (2020). Phylogenomic analysis of trichomycterid catfishes (Teleostei: Siluriformes) inferred from ultraconserved elements. Scientific reports, 10(1), 1-15.

Oliveira, C., Avelino, G. S., Abe, K. T., Mariguela, T. C., Benine, R. C., Ortí, G., ... & e Castro, R. M. C. (2011). Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. BMC evolutionary biology, 11(1), 275.

Ota, R. P., Machado, V. N., Andrade, M. C., Collins, R. A., Farias, I. P., & Hrbek, T. (2020). Integrative taxonomy reveals a new species of pacu (Characiformes: Serrasalmidae: Myloplus) from the Brazilian Amazon. Neotropical Ichthyology, 18(1).

Pennisi, E. (2013). Large-scale gene comparisons boost tree of life studies.

Pereira, L. H., Pazian, M. F., Hanner, R., Foresti, F., & Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Paraná Basin of Brazil. Mitochondrial Dna, 22(sup1), 87-96.

Sandelin A., Bailey P., Bruce S., Engström P.G., Klos J.M., Wasserman W.W., et al. (2004) Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. BMC Genomics 5, 1–9.

Sazima, I., & Machado, F. A. (1982). Habits and behavior of *Roeboides prognathus*, a scale-eating fish (Osteichthyes, Characoidei). Boletim de Zoologia da Universidade de São Paulo 7: 37–56.

Sazima, I. (1984). Scale-eating in characoids and other fishes. In Evolutionary ecology of neotropical freshwater fishes (pp. 9-23). Springer, Dordrecht.

Siepel, A., Bejerano, G., Pedersen, J. S., Hinrichs, A. S., Hou, M., Rosenbloom, K., ... & Haussler, D. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research, 15(8), 1034-1050.

Silva, G. S. C., Melo, B. F., Oliveira, C., & Benine, R. C. (2013). Morphological and molecular evidence for two new species of *Tetragonopterus* (Characiformes: Characidae) from central Brazil. Journal of Fish Biology, 82(5), 1613-1631.

Silva, G. S., Roxo, F. F., Melo, B. F., Ochoa, L. E., Bockmann, F. A., Sabaj, M. H., ... & Oliveira, C. (2021). Evolutionary history of Heptapteridae catfishes using ultraconserved elements (Teleostei, Siluriformes). Zoologica Scripta.

Smith, B.T., McCormack, J.E., Cuervo, A.M., Hickerson, M.J., Aleixo, A., Cadena, C.D., Pérez-Emán, J., Burney, C.W., Xie, X., Harvey, M.G., Faircloth, B.C., Glenn, T.C., Derryberry, E.P., Prejean, J., Fields, S. & Brumfield, R.T. (2014). The drivers of tropical speciation. Nature 515, 406–409.

Starrett, J., Derkarabetian, S., Hedin, M., Bryson Jr, R. W., McCormack, J. E., & Faircloth, B. C. (2017). High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. Molecular ecology resources, 17(4), 812-823.

Tagliacollo, V. A., Souza-Lima, R., Benine, R. C., & Oliveira, C. (2012). Molecular phylogeny of Aphyocharacinae (Characiformes, Characidae) with morphological diagnoses for the subfamily and recognized genera. Molecular phylogenetics and evolution, 64(2), 297-307.

Venere, P. C., & Garutti, V. (2011). Peixes do Cerrado: Parque Estadual da Serra Azul, rio Araguaia, MT. São Carlos: RiMa Editora, FAPEMAT.

Weitzman, S. H., & Vari, R. P. (1987). Two new species and a new genus of miniature characid fishes (Teleostei: Characiformes) from northern South America. Proceedings of the Biological Society of Washington, 100(3), 640-652.

Woolfe A., Goodson M., Goode D.K., Snell P., McEwen G.K., Vavouri T., et al. (2004). Highly Conserved Non-Coding Sequences Are Associated with Vertebrate Development. PLoS Biology 3, e7

Chapter 1

Phylogenomic analysis of the Neotropical fish subfamily Characinae using ultraconserved elements (Teleostei: Characidae)

Phylogenomic analysis of the Neotropical fish subfamily Characinae using ultraconserved elements (Teleostei: Characidae)

Camila S. Souza

Departamento de Biologia Estrutural e Funcional, Instituto de Biociências, Universidade Estadual Paulista, R. Prof. Dr. Antônio C. W. Zanin 250, 18618-689, Botucatu, SP, Brazil

Abstract

Characinae is one of the most species-rich subfamilies of Characidae and holds special taxonomic importance because it includes *Charax*, type-genus of Characidae and Characiformes. Currently, the monophyly and the hypotheses of intergeneric and interspecific relationships of Characinae are based on a few morphological and molecular studies but all with low species coverage. Given their diversity, taxonomic importance, and the lack of a taxon-dense phylogeny, we sought to buttress the systematic understanding of Characinae collecting DNA sequence data from ultraconserved elements (UCEs) of the genome from 98 specimens covering 57 species (67%) plus 17 characiforms as outgroups. We used maximum likelihood, Bayesian inference, and coalescent-based species tree approaches and the resulting phylogeny with 1,300 UCE loci (586,785 characters) reinforced the monophyly of the subfamily as well as of six genera: *Acestrocephalus*, *Charax*, *Cynopotamus*, *Galeocharax*, *Phenacogaster*, and *Roeboides*. The phylogeny reveals a novel hypothesis of intergeneric and interspecific relationships for the subfamily with *Phenacogaster* sister to all genera, and *Acanthocharax* sister to Cynopotamini (*Cynopotamus* (*Acestrocephalus Galeocharax*)) and Characini (*Charax Roeboides*). We propose a new tribe Acanthocharacini to allocate *Acanthocharax*, two subclades for *Phenacogaster*, two for *Cynopotamus*, three for *Charax*, and

reinforced the four subclades for *Roeboides* previously identified by morphological studies. Additionally, we generated a time-calibrated phylogeny for Characinae that suggested that the subfamily originated during the Miocene at around 19.4 million years ago. The results obtained here will contribute to the development of further research on the evolutionary processes modulating species diversification in Characinae.

Keywords: biodiversity, Characiformes, freshwater fish, Neotropics, Ostariophysi, systematics.

1. Introduction

The Neotropical ichthyofauna is mainly composed by non-cypriniform otophysan fishes (i.e., Characiformes, Siluriformes, and Gymnotiformes), which constitutes about 77% of the total species richness (Albert *et al*., 2011a, b). Among Neotropical otophysan groups, Characiformes represents one of the most ecomorphologically diverse orders of fishes with approximately 2,200 valid species (Fricke *et al.*, 2022). Within the order, the family Characidae, as proposed by Oliveira *et al*. (2011), is the most species-rich with a total of 1,228 species (Fricke *et al*., 2022). Due to its great diversity, Characidae includes some of the greatest taxonomic and systematic challenges among Neotropical fishes (Oliveira *et al*., 2011; Mirande, 2019), with the general morphology and anatomy of the species being highly conservative, and most members attaining maximum body sizes of 10 cm standard length (SL) or less (Froese & Pauly, 2021).

Characinae is the third most species-rich subfamily of Characidae and includes *Charax* Scopoli, 1777, type genus of Characiformes. The subfamily currently comprises 85 species (*sensu* Mattox & Toledo-Piza, 2012) distributed in rivers of the Neotropical region, including both sides of the Andes. In the east of the Andes, they occur from the northern coastal drainages of Venezuela to tributaries of the lower La Plata basin (Lucena & Menezes, 2003). Most members of the group are easily

recognized by their deep anterior bodies with a characteristic gibbosity (Figure 1), although this is absent in *Phenacogaster* Eigenmann, 1907 and *Acestrocephalus* Eigenmann, 1910 (Lucena & Menezes, 2003; Mattox & Toledo-Piza, 2012). They have adopted three distinct feeding strategies: carnivory, omnivory, and lepidophagy, although most species are carnivorous and feed mainly on insects and other fishes (Géry, 1977; Lucena & Menezes, 2003). *Roeboides* represents the only genus of Characinae with lepidophagous habits and possesses morphological specializations that aid in this unusual feeding style, such as external mammiliform teeth used to pluck scales from their prey (Sazima & Machado, 1982; Sazima, 1984).

The taxonomic history of Characinae has changed significantly during the two last decades, with its present composition defined by Lucena (1998), Lucena & Menezes (2003), Mirande (2009, 2010) and Mattox & Toledo-Piza (2012). The definition of Characinae proposed by Lucena & Menezes (2003) included 12 genera (*Acanthocharax* Eigenmann, 1912; *Acestrocephalus*; *Charax*; *Cynopotamus* Valenciennes, 1850; *Galeocharax* Fowler, 1910; *Gnathocharax* Fowler, 1913; *Heterocharax* Eigenmann, 1912; *Hoplocharax* Géry, 1966; *Lonchogenys* Myers, 1927; *Phenacogaster*; *Priocharax* Weitzman & Vari, 1987; *Roeboides*) based on the relatively deep body shape, presence of more than 20 conical teeth along the maxilla, presence of a conspicuous pseudotympanum anterior to the first pleural rib, and the retention of larval pectoral fin in specimens up to 41 mm SL. Conversely, the hypothesis proposed by Mirande (2009, 2010) based on morphological features included eight of the 12 genera (*sensu* Lucena & Menezes, 2003): *Acanthocharax*, *Acestrocephalus*, *Charax*, *Cynopotamus*, *Galeocharax*, *Phenacogaster*, *Priocharax*, and *Roeboides*, with the tentative inclusion of *Priocharax* based on Lucena & Menezes (2003) (Mirande, 2010: 492). He also proposed three additional genera, *Bryconexodon* Géry, 1980, *Exodon* Müller & Troschel, 1844, and *Roeboexodon* Géry, 1959 as belonging to Characinae forming a clade sister to all remaining genera of the subfamily. This three-genera group is composed of strictly lepidophagous characids and some of them had been previously proposed as closely related

to *Roeboides* in a pre-cladistic context (*e.g.*, Géry, 1964; Roberts, 1970). Later, with a broader set of taxa and characters, Mirande (2019:296) reinterpreted these results and assigned *Bryconexodon*, *Exodon* and *Roeboexodon* to their own subfamily, the Exodontinae.

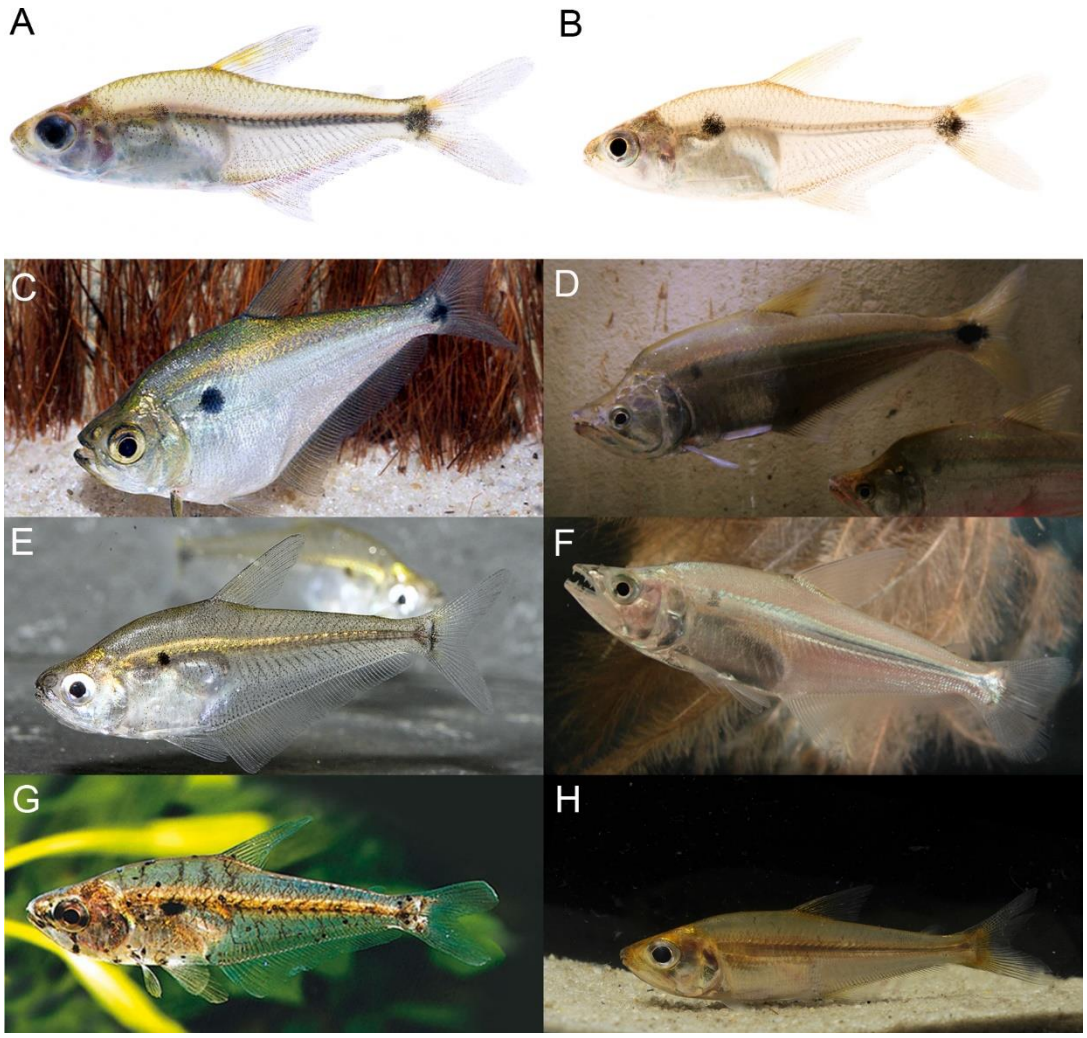


Figure 1. Representative specimens of Characinae. (A) *Phenacogaster* sp.; (B) *Phenacogaster eurytaenia*; (C) *Acanthocharax microlepis*; (D) *Cynopotamus atratoensis*; (E) *Roeboides descalvadensis*; (F) *Galeocharax humeralis*; (G) *Charax condei*; (H) *Acestrocephalus sardina*. Photographs by Martin Taylor (A, B), Johnny Jensen (C), Frank Schäefer (D, E), Frank Teigler (G), Pablo Giorgis (F), and Ralf Britz (H).

In parallel to these studies, Mattox & Toledo-Piza (2012) carried out a morphological study focusing on the subfamily Characinae to test its monophyly and the subunits proposed by Lucena (1998) and Lucena & Menezes (2003). The morphology-based results (Figure 2B) supported the monophyly of the seven genera in Characinae based on ten non-ambiguous synapomorphies. The authors arranged the genera in three tribes: Phenacogasterini as sister to Characini and Cynopotamini, with the intergeneric relationships (*Phenacogaster* ((*Charax Roeboides*) (*Acanthocharax* (*Cynopotamus* (*Acestrocephalus Galeocharax*))))). *Acanthocharax* was previously considered the sister-group to *Charax* and *Roeboides* (Lucena, 1998) but the former genus was found as sister to all other Cynopotamini more recently (Mattox & Toledo-Piza, 2012). Among the outgroups, *Tetragonopterus* has been consistently found as the sister group to Characinae based on both morphological and molecular data (Buckup, 1998; Javonillo *et al*., 2010; Oliveira *et al.*, 2011; Mattox & Toledo-Piza, 2012; Melo *et al.*, 2016; Betancur-R *et al*., 2019).

Weitzman & Vari (1987) described the enigmatic genus *Priocharax* and suggested that its morphology was similar to members of Characinae and Cynopotaminae. Lucena (1998) was the first to include *Priocharax* in a phylogenetic context, which resulted as the most basal lineage in the Characinae, leading to the subsequent classification of this genus in Characinae (Lucena & Menezes, 2003; Mirande, 2010). Mattox & Toledo-Piza (2012) were the first to propose that *Priocharax* was more related to the Heterocharacinae, a subset of former members of the Characinae given their own subfamily rank (Mirande, 2010; Mattox & Toledo-Piza, 2012). Oliveira *et al*. (2011) did not analyze *Priocharax* but showed that Heterocharacinae is nested within Acestrorhynchidae rather than to Characinae and therefore transferred Heterocharacinae into Acestrorhynchidae. Mirande (2019), based on literature information, suggested the placement of *Priocharax* in the subfamily Characinae. Betancur *et al.* (2019) found *Priocharax* as sister to *Roeboides* but, after tissue extraction and publication, an additional specimen of the same lot was examined by us and the identification as *Priocharax* was not confirmed. Additionally, the voucher specimen used by

Betancur *et al*. (2019) is no longer available for reidentification (LBP 17836; C. Souza, personal observation). Thus, the position of *Priocharax* in the phylogeny of Characidae remains unclear.

Other characiform studies also included representatives of Characinae using both molecular (Javonillo *et al*., 2010; Oliveira *et al*., 2011; Tagliacollo *et al*., 2012; Melo *et al.,* 2016; Betancur-R *et al*., 2019; Melo *et al.*, 2021), and total evidence analysis (Mirande, 2019). These studies agree about the relationships between Characinae and Tetragonopterinae (i.e., *Tetragonopterus*) (Javonillo *et al.*, 2010; Oliveira *et al.*, 2011; Tagliacollo *et al.*, 2012; Melo *et al.,* 2016; Mirande, 2019; Betancur-R. *et al*., 2019); however, the molecular studies did not corroborate the monophyly of *Cynopotamus* (Oliveira *et al*., 2011; Tagliacollo *et al*., 2012; Melo *et al.*, 2016). In addition, *Acanthocharax*, hypothesized as the sister group of either clades (*Charax Roeboides*) or (*Cynopotamus* (*Acestrocephalus Galeocharax*)) in morphological reconstructions (Lucena, 1998; Mattox & Toledo-Piza, 2012), had not been included in any molecular analysis.

The hypotheses of interspecific relationships among genera of the Characinae are also limited to morphological analyses and/or arrangements proposals based on taxonomic revisions and species descriptions (Lucena, 1998; Lucena, 2003, 2007; Menezes, 2006; Lucena & Malabarba, 2010; Menezes & Lucena, 2014; Giovannetti *et al*., 2017). These relationships remain unclear also due to the low number of taxa used in phylogenetic analyses (Oliveira *et al*., 2011; Mattox & Toledo-Piza, 2012; Mirande, 2019) (Figure 2). So far, the morphological phylogeny that covered interspecific relationships with the highest number of species of Characinae (i.e., 30%) is that of Mattox & Toledo-Piza (2012), although their study focused on testing the monophyly of the subfamily and the intergeneric relationships. In the field of molecular phylogenetics, ultraconserved elements (UCEs; Faircloth *et al*., 2012) have been used to explore the relationships within various animal groups, including fishes (Harrington *et al*., 2016; Chakrabarty *et al*., 2017; Alfaro *et al*., 2018; Roxo *et al*., 2019; Ochoa *et al.*, 2020; Mateussi *et al*., 2020; Melo *et al*., 2021). They constitute excellent markers

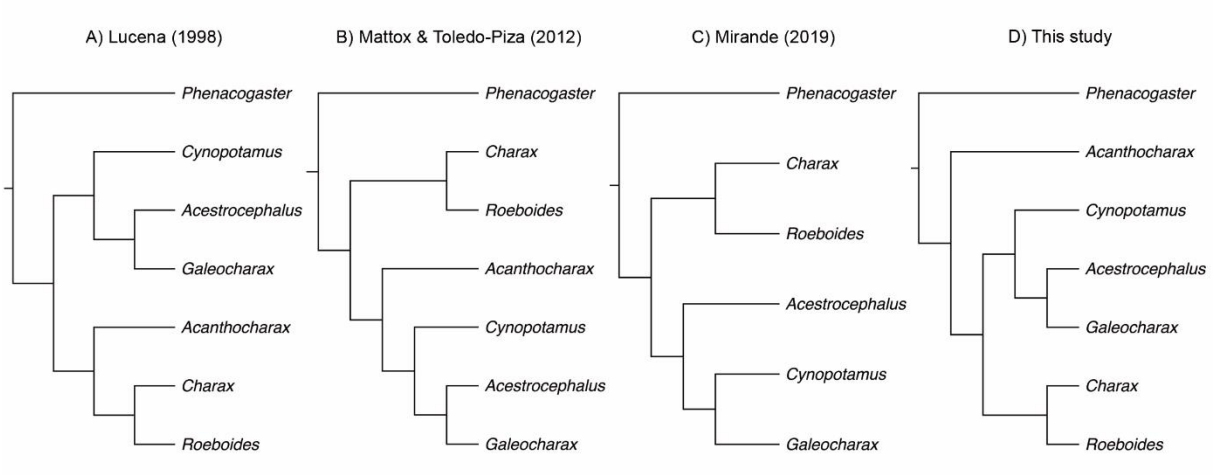


Figure 2. Intergeneric hypotheses of Characinae based on morphological (A- Lucena, 1998; B- Mattox & Toledo- Piza, 2012), total-evidence analysis (C- Mirande, 2019), and phylogenomic study (D- This study). Figures modified from original publications.

for phylogenetic studies due to their presence among a wide range of taxonomic groups, low degrees of ambiguity, and low saturation (Siepel *et al.*, 2005; Derti *et al.*, 2006; McCormack *et al.*, 2012; Alda *et al.*, 2021), despite well-known problems with incomplete lineage sorting (Alda *et al.*, 2019; Alda *et al.*, 2021). Here, we performed a phylogenomic study of Characinae using 57 out of 85 species (67%), including representatives of the seven genera currently assigned to the subfamily based on an UCE dataset to (i) test the monophyly of the genera, (ii) test previous phylogenetic hypotheses of intergeneric and interspecific relationships, and (iii) to better understand the evolution and biogeography of this subfamily.

2. Materials and Methods

2.1. Taxon sampling

Our study included 98 specimens from 57 species of Characinae (67%) with representatives of all seven genera. Numbers in parentheses represent sample size relative to the number of valid species:

Acanthocharax (1/1), *Acestrocephalus* (3/8), *Charax* (10/17), *Cynopotamus* (9/12), *Galeocharax* (3/3), *Phenacogaster* (13/23), and *Roeboides* (18/21). The outgroup included 16 species: one species of Acestrorhynchidae (*Acestrorhynchus microlepis* (Jardine, 1841)), one Cheirodontinae (*Protocheirodon pi* (Vari, 1978)), one Aphyocharacinae (*Aphyocharax pusillus* Günther, 1868), three Exodontinae (*Bryconexodon juruenae* Géry, 1980, *Exodon paradoxus* Müller & Troschel, 1844 and *Roeboexodon guyanensis* (Puyo, 1948)), two Spintherobolinae (*Amazonspinther dalmata* Bührnheim, Carvalho, Malabarba & Weitzman, 2008 and *Spintherobolus papilliferus* Eigenmann, 1911), two Stevardiinae (*Markiana nigripinnis* (Perugia, 1891) and *Planaltina myersi* (Böhlke, 1954), three Stethaprioninae (*Astyanax jordani* (Hubbs & Innes, 1936), *Paracheirodon axelrodi* (Schultz, 1956) and *Hyphessobrycon compressus* (Meek, 1904)), and three Tetragonopterinae (*Tetragonopterus georgiae* (Géry, 1965), *T. argenteus* Cuvier 1816, and *T. chalceus* Spix & Agassiz, 1829). *Acestrorhynchus microlepis* (Acestrorhynchidae) was used to root the trees. Separately, we used a 95% complete edge-trimmed matrix to include three species of *Priocharax* and only one specimen by species (85 taxa) to estimate a maximum likelihood (ML) tree and to time-calibrate the phylogeny to test the position and relationships of *Priocharax*. Voucher samples used in this study are deposited in the Laboratório de Biologia e Genética de Peixes, Universidade Estadual Paulista, Botucatu, Brazil (LBP), Coleção dos Recursos Genéticos, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil (INPA), Royal Ontario Museum, Toronto, Canada (ROM), Laboratório de Ictiologia, Museu de Ciências e Tecnologia, Porto Alegre, Brazil (MCP), Laboratório de Ictiologia e Pesca, Universidade Federal de Rondônia, Porto Velho, Brazil (UNIR), Laboratório de Genética e Biologia Molecular, Universidade Federal do Maranhão, São Luís, Brazil (LABGEN/UFMA), Coleção Ictiológica, Museu de Zoologia da Universidade de São Paulo, São Paulo, Brazil (MZUSP), and Smithsonian Tropical Research Institute, Panama City, Panama (STRI). The samples collected in this study are in agreement with Brazilian laws through SISBIO/MMA permit n. 3245 and procedures for collection, maintenance and analyses followed

the international guidelines for animal experiments through CEEAA IBB/UNESP protocol n. 304. The Supplementary Table 1 includes data from all ingroup and outgroup samples.

2.2. DNA extraction and sequencing

DNA was extracted from muscle, gills, or fin tissues with a DNeasy Tissue kit (Qiagen Inc.) following manufacturer's instructions. Then, 2 µl of each genomic DNA was quantified using fluorometry (Qubit, Life Technologies) to prepare the libraries using a concentration between 10–40 ng/µl. DNA libraries and sequencing were performed at Arbor Biosciences (AB; arborbiosci.com; Ann Arbor, MI, USA). Whole genomic DNA was first sheared with a QSonica Q800R instrument and selected to modal lengths of approximately 500 nt using a dual-step SPRI bead cleanup.

DNA libraries were prepared for the 117 specimens (98 ingroup taxa and 19 outgroup taxa) by modifying the Nextera (Epicentre Biotechnologies) library preparation protocol for solution-based target enrichment following Faircloth *et al.* (2012) and the number of PCR cycles following the recommendation of Faircloth *et al.* (2013). AB staff used the Nextera library preparation protocol of in vitro transposition followed by PCR to prune the DNA and attach sequencing adapters (Adey *et al.*, 2010). The Epicentre Nextera kit was used subsequentially to prepare transposase-mediated libraries with insert sizes averaging 100 bp (95% CI: 45 bp) following Adey *et al.* (2010). The libraries were enriched using a new probeset developed for ostariophysan fishes that captures sequence data for 2,708 UCE loci (Faircloth *et al.*, 2020). Then, the DNA was converted to Illumina sequencing libraries with a slightly modified version of the NEBNext(R) Ultra (TM) DNA Library Prep Kit for Illumina(R). After ligation of sequencing primers, libraries were amplified using KAPA HiFi HotStart ReadyMix (Kapa Biosystems) for six cycles using the manufacturer's recommended thermal profile and dual P5 and P7 indexed primers (Kircher *et al.*, 2012). After purification with

SPRI beads, libraries were quantified with the Quant-iT(TM) Picogreen(R) dsDNA Assay kit (ThermoFisher). AB staff then enriched pools comprising 100 ng each of eight libraries (800 ng total) using the MYbaits(R) Target Enrichment system (MYcroarray) following manual version 3.0. Sequencing was performed across two Illumina HiSeq paired-end 100 bp lanes using v4 chemistry.

2.3. Sequence data processing

After sequencing, adapter contamination, low quality bases, and sequences containing ambiguous base calls were trimmed using the Illumiprocessor software pipeline (Faircloth, 2013; https://github.com/faircloth-lab/illumiprocessor). After trimming, the assembled Illumina were read into contigs on a species-by-species basis using Velvet (Zerbino & Birney, 2008) on VelvetOptimiser (https://github.com/VictorianBioinformatics-Consortium/VelvetOptimiser). Following assembly, a custom Python program (match_contigs_to_probes.py), available in PHYLUCE (Faircloth, 2016) was used, integrating LASTZ (Harris, 2007) to align species specific contigs to our probe-UCE set. During the alignment, the latter program creates a relational data base of matches to UCE loci by taxon. We then used the get_match_counts.py program (also in PHYLUCE) to query the database and generate FASTA files for UCE loci that were identified across all taxa. A custom Python program (seqcap_align_2.py) was then used to align contigs using the MUSCLE alignment algorithm (Edgar, 2004) and to perform edge trimming.

2.4. Phylogenetic analyses

We performed phylogenetic analyses with 75% and 90% complementary matrices to test the effect of missing data: the 75% complete matrix contains loci present in at least 75% of taxa (i.e. 85 terminals or more), and the 90% complete matrix contains loci present in at least 90% of taxa (i.e.

103 terminals or more). We concatenated the matrices and performed analyses by ML in RAxML v8.019 (Stamatakis, 2014), and Bayesian inference (BI) in ExaBayes v1.4 (Aberer *et al*., 2014), and the coalescent-based species tree approach (AS) in ASTRAL-III (Zhang et al., 2018) using a 2×10 CPU, 256 GB Zungaro server at LBP-UNESP. For ML and BI analyses, the data was partitioned to account for variation in rates and patterns of molecular evolution among sites using PartitionUCE (Tagliacollo & Lanfear, 2018) and the substitution models estimated with hcluster in PartitionFinder v2 (Lanfear *et al*., 2014; 2016) through RAxML v8.0 (Stamatakis, 2006).

The ML analysis was performed using GTRCAT (Stamatakis, 2006) with 20 alternative runs of distinct parsimony starting trees in RAxML v8.019 (Stamatakis, 2014). The posteriori bootstrapping analysis was conducted using the autoMRE function in RAxML using the bootstopping criteria (Stamatakis, 2014; Pattengale *et al*., 2010). Bayesian inference was performed in ExaBayes v 1.597 (Aberer *et al*., 2014), with two independent runs, each with four chains (one cold and three heated) with 50,000,000 iterations and other parameters as default. Tree space was sampled every 100 generations to yield a total of 10,001 trees. We assessed convergence of the posterior distribution examining the ESS > 200 (effective sample size) and evaluating posterior trace distribution in Tracer v1.6.1 (Rambaut *et al*., 2014). We obtained the most likely set of trees from the posterior distribution of possible topologies using the consensus algorithm of ExaBayes with 10% burn-in. The coalescent-based species tree approach (AS) was inferred from individual gene trees using ASTRAL-III (Zhang *et al*., 2018). The individual gene trees used as input to ASTRAL-III were estimated by a ML analysis using a RAxML bootstrapped using the parameter –N = 2 and the GTRGAMMA model. Then, ASTRAL-III was used to infer species trees from the best gene trees, and to reconstruct the majority-rule consensus tree of the results.

2.5. Divergence time estimation

We estimated a time-calibrated phylogeny using an uncorrelated relaxed molecular clock (lognormal) using the BEAST v2 (Bouckaert et al., 2014). We used the 95% complete edge-trimmed matrix (91 UCE loci with 38,888bp) to estimate both the topology and node ages. We included a constraint on the root and two fossils as calibration points. First, a calibration point was assigned to the root of the tree (i.e., all taxa) to calibrate the node splitting Acestrorhynchidae and members of Characidae that, based on the recent time reconstruction of the entire order Characiformes (Melo *et al*., 2021), was estimated to have occurred in the Late Cretaceous at around 72 (83–60) millions of years ago (Ma) (normal distribution; offset = 72 Ma; mean = 0; standard deviation = 7.0). The second calibration point was based on the fossil †*Paleotetra entrecorregos*, UNG 2T-149, an articulated skeleton from the Entre-Córregos Formation in Minas Gerais, Brazil, dated to the Eocene-Oligocene boundary (Weiss *et al*., 2012). The fossil was hypothesized to be a stem Stevardiinae (Weiss *et al*., 2012) or a stem Characidae (Mirande, 2019). Given the natural uncertainty of the fossil and pending additional analysis with less missing data showing higher support for early nodes of Characidae, and considering that the origin of Characidae was estimated at around 65 Ma (Melo *et al*., 2021), we assigned it to the node stem of Stevardiinae following the original description and phylogenetic placement (Weiss *et al*., 2012) (log-normal distribution; offset = 33.9 Ma; mean = 5.0 Ma; standard deviation = 1.0). The third calibration point was *†Megacheirodon unicus*, MCP 3086-PV, an articulated skeleton from Oligocene-Miocene deposits of the Tremembé Formation in São Paulo, Brazil (Bührnheim *et al*., 2008). *†Megacheirodon* was hypothesized to be the sister clade to *Spintherobolus* and *Amazonspinther* (Bührnheim et al. 2008), thus we implemented it as a calibration of the node with both genera and other members of Characidae (log-normal distribution; offset = 23.8 Ma; mean = 1.0 Ma; standard deviation = 1.25). We also included a constraint prior for the monophyly of Characidae. The BEAST analyses were

conducted under a birth-death model for prior distributions and ran for 50 million generations sampling frequency at every 10,000th generation. We checked stationarity and sufficient mixing of parameters (ESS > 200) using Tracer v1.6 (Rambaut et al., 2014). A consensus tree was built using TreeAnnotator v1.8.2. All clade-age estimates are presented as the mean plus 95% highest posterior density (HPD) values.

3. Results and Discussion

3.1. Overall patterns and monophyly of Characinae

Sequencing and data filtering yielded an initial edge-trimmed aligned matrix comprising 2,527 UCE loci with a total of 990,049 base pairs (bp) for 114 specimens (98 Characinae and 16 outgroups) (Supplementary Table 2). We used two UCE matrices with 114 taxa to infer phylogenetic relationships: the 75% complete matrix was composed of 1,300 UCE loci and 586,785 bp, each of which contained sequence data for at least 85 taxa and the 90% complete matrix with 297 UCE loci and 137,232 bp each of which contained sequence data for at least 103 taxa. Phylogenies were resolved with high statistical support for most nodes regardless of matrix completeness (75% or 90%), or method of phylogenetic inference (ML, BI, or ASTRAL-III) (Figures 3–4; S1–S5).

The results of the ML trees and BI analyses of the edge-trimmed, 75% complete, partitioned matrices show identical topologies (Figures 3–4; S3). Overall, all phylogenetic inferences show the same topology to the intergeneric relationship, however ASTRAL-III presented the highest number of differences in interspecific relationships compared to the ML and the BI analyses. Details of the differences among each analysis can be observed in the resulted topologies (Figures 3–4; S1–S5).

Discrepancies in interspecific relationships presented by ASTRAL-III included *Galeocharax gulo* and species of *Phenacogaster* and *Roeboides* (Figures S4-S5) that show rapid and/or recent

diversification (Figure 5). Disagreements among trees could be caused by different factors such as incomplete lineage sorting (Pollard *et al*., 2006; Carstens & Knowles, 2007), mutation and recombination rate (Pollard *et al*., 2006, Rosser *et al.*, 2017), selective pressures (Malinsky *et al.*, 2015, Sun *et al*., 2015), hybridization and introgression (Toews & Brelsford, 2012, Denton *et al.*, 2014). Nevertheless, a significant explanation for different relationships found in ASTRAL-III can be attributable to deep coalescence processes, where multiples lineages tend to persist into the deeper portion of the species tree, that is common in species with rapid and/or recent diversification (Degnan & Rosenberg, 2006, 2009).

Despite the few discordances among topologies, we constructed our discussion based on the results of the ML tree of the edge-trimmed, 75% complete, partitioned matrix with 100% bootstrap for 76.6% of the nodes (Figures 3-4) and the topology that best concord with the morphology of the subfamily. The UCE phylogeny supports the monophyly of Characinae and of all non-monotypic characin genera: *Acestrocephalus, Galeocharax, Cynopotamus*, *Charax, Phenacogaster,* and *Roeboides* (Figures 3-4; ML = 100). As our taxon sampling contained only one specimen of *Acanthocharax*, we could not test the monophyly of the genus. The intergeneric relationships of Characinae and the interspecific relationships in *Cynopotamus* and *Galeocharax* are strongly supported in most of the nodes (Figures 3-4; S1-S5). Our reconstruction also shows the closer relationship between Characinae and *Tetragonopterus*, with Exodontinae as the immediate sister clade. This reconstruction matches recent molecular phylogenies of related taxa (e.g. Oliveira *et al.*, 2011; Melo *et al*., 2016, 2021). Furthermore, our results found *Priocharax* inside Stethaprioninae and the relationship between *Priocharax* and *Hyphessobrycon* and other characids rather than with Characinae (Figures 5 and S6).

The morphology-based phylogeny (Mattox & Toledo-Piza, 2012) supported the monophyly of Characinae based on ten non-ambiguous synapomorphies (details in Supplementary data), and the authors proposed three tribes in Characinae: Phenacogasterini (*Phenacogaster*), Characini (*Charax*

Roeboides) and Cynopotamini ((*Acanthocharax* (*Cynopotamus* (*Acestrocephalus Galeocharax*))). Herein, our phylogeny shows a slightly distinct arrangement: Phenacogasterini (*Phenacogaster*), Acanthocharacini new tribe (*Acanthocharax*), sister to Characini (*Charax Roeboides*) and Cynopotamini (*Cynopotamus* (*Acestrocephalus Galeocharax*)) (Figures 3-4). Given the new phylogenetic arrangement, Acanthocharacini Souza, Melo, Mattox & Oliveira, 2021 (ZooBank Nomenclatural Act: urn:lsid:zoobank.org:act:E32EDF2D-978C-4DB9-AD2C-8EAB9E1E5FE1), is herein proposed as a new tribe of Characinae.

3.1.1. *Acanthocharacini* Souza, Melo, Mattox & Oliveira, 2022

urn:lsid:zoobank.org:act:E32EDF2D-978C-4DB9-AD2C-8EAB9E1E5FE1

Included genus: *Acanthocharax* Eigenmann, 1912

Diagnosis: As for the genus (e.g., Eigenmann, 1912). Acanthocharacini is distinguished from all other tribes of Characinae by the presence of a spiniform projection on the preopercle (Eigenmann, 1912; Mattox & Toledo-Piza, 2012). The absence of predorsal scales were used previously to characterize the genus being useful to diagnose the tribe, albeit predorsal scales were independently lost in *Charax condei* and *C. stenopterus*.

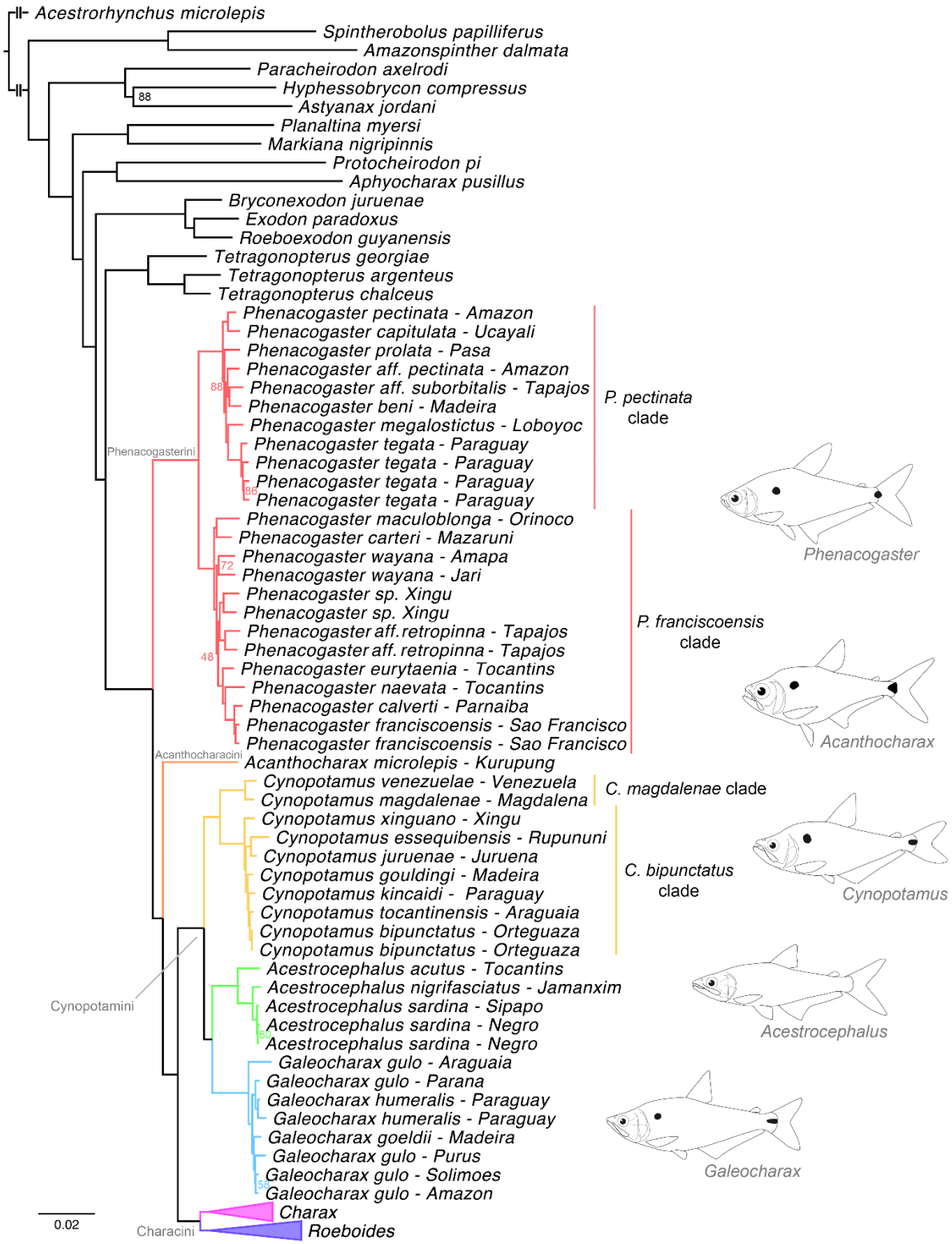


Figure 3. Phylogenetic relationships of Characinae based on a maximum likelihood analysis of the 75% complete matrix of ultraconserved elements (1,300 loci; 586,785 bp), highlighting the phylogenetic relationships in Phenacogasterini, Cynopotamini, and the position of Acanthocharacini. All nodes have bootstrap values ≥90% except where indicated.

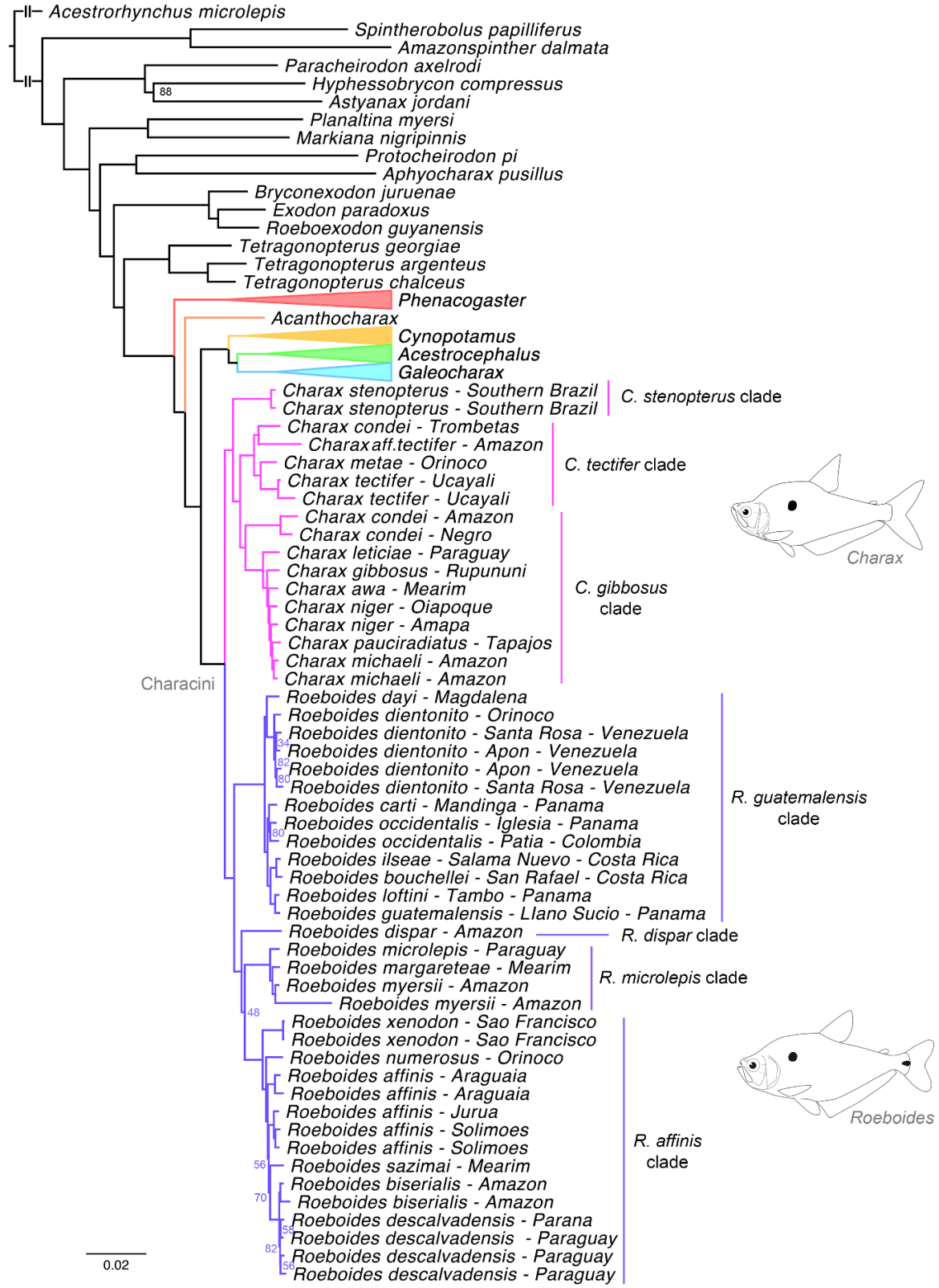


Figure 4. Phylogenetic relationships of Characinae based on a maximum likelihood analysis of the 75% complete matrix of ultraconserved elements (1,300 loci; 586,785 bp), highlighting the phylogenetic relationships in Characini. All nodes have bootstrap values ≥90% except where indicated.

3.2. Monophyly of genera and intergeneric relationships

Monophyly of *Phenacogaster*, the sole genus in the tribe Phenacogasterini, was supported herein, including 14 species (13 valid and one undescribed species) out of the 23 currently recognized in the genus (Fricke *et al.*, 2022). Morphologically, *Phenacogaster* has been diagnosed mainly by the presence of two longitudinal series of large, narrow, and elongate preventral scales and by the gap in the external tooth series of the premaxilla (Malabarba & Lucena, 1995; Lucena & Malabarba, 2010). The former character was interpreted as one of the five synapomorphies of *Phenacogaster* (Mattox & Toledo-Piza, 2012).

Monophyly of the clade comprising the other six genera (i.e., *Acanthocharax*, *Acestrocephalus*, *Charax*, *Cynopotamus*, *Galeocharax* and *Roeboides*) was supported herein with 100% support, as in previous hypotheses based on morphology (Figure 2; Lucena, 1998; Mattox & Toledo-Piza, 2012). This clade was strongly supported by the latter authors based on nine non-ambiguous synapomorphies, with one synapomorphy exclusive for this clade: lateral wings of mesethmoid poorly developed (Mattox & Toledo Piza, 2012).

Our phylogenetic hypothesis shows *Acanthocharax* as sister to the clade ((*Charax Roeboides*) (*Cynopotamus* (*Acestrocephalus Galeocharax*)) (Figure 3; ML = 100), which is congruent with the hypothesis based mainly on comparative myology (Howes, 1976). This result contrasts, however, to Lucena (1998), who proposed a closer relationship between *Acanthocharax* and the clade formed by *Charax* and *Roeboides* based on two synapomorphies. The more recent morphological phylogeny of the subfamily suggested the inclusion of *Acanthocharax* in the Cynopotamini based on five synapomorphies (Mattox & Toledo-Piza, 2012), four of which highly homoplastic in their analyses. The fifth synapomorphy is the shape of the lower pharyngeal toothplate which is more elongated posteriorly in *Acanthocharax*, *Cynopotamus*, *Acestrocephalus* and *Galeocharax*, but should be

reinterpreted either as convergent in the former genus and the Cynopotamini or as lost (*i.e.*, reversed to the shorter, more basal shape) in *Charax* and *Roeboides* based on our results.

The UCE phylogeny indicates *Cynopotamus* as sister to the clade with *Acestrocephalus* and *Galeocharax*, a relationship previously proposed by morphological studies (Menezes, 1976; Lucena, 1998; Mattox & Toledo-Piza, 2012). However, the relationships recovered herein do not support the total-evidence hypothesis that found *Acestrocephalus* as sister to a clade formed by *Cynopotamus* and *Galeocharax* (Mirande, 2019). Furthermore, *Acestrocephalus*, *Cynopotamus* and *Galeocharax* were individually corroborated as monophyletic assemblages (Figure 3; ML = 100), as shown in previous morphological studies (e.g., Menezes, 1976; 2006; Mattox & Toledo-Piza, 2012; Giovanetti *et al*., 2017). In the tribe Characini, *Charax* and *Roeboides* are resolved as sister genera in our molecular study and both are supported as monophyletic (Figure 4; ML = 100), a result also consistent with morphological studies (Lucena, 1998; 2000; Mirande, 2009; 2010; Mattox & Toledo-Piza, 2012).

3.3. Interspecific relationships

Hypotheses of interspecific relationships in Characinae were limited to morphological analyses or arrangements based on taxonomic revisions and species descriptions (Lucena, 2000, 2003, 2007; Menezes, 2006, 2007; Lucena & Malabarba, 2010; Menezes & Lucena, 2014; Giovannetti *et al*. 2017). *Phenacogaster*, the most species-rich genus, lacks formal descriptions for many species as well as taxonomic definitions of subgroups (Lucena & Malabarba, 2010). Based on 56% of its species diversity, our phylogeny consistently supports the presence of two main *Phenacogaster* clades (ML=100; Figure 4).

The first is the "*Phenacogaster pectinata* clade" (Figure 3), widely distributed across the Amazon and Paraguay basins with relatively strong node support (ML>86%). Species from the

Amazon (*P. beni* Eigenmann, 1911, *P. capitulata* Lucena & Malabarba, 2010, *P. megalostictus* Eigenmann, 1909, *P. pectinata* (Cope, 1870), *P. prolata* Lucena & Malabarba, 2010, *P.* aff. *pectinata* and *P.* aff. *suborbitalis*) are sister group to *P. tegata* (Eigenmann, 1911) from the Paraguay basin.

The second is the "*P. franciscoensis* clade", that includes *P. maculoblonga* Lucena & Malabarba, 2010, from the Orinoco basin sister to *P. carteri* (Norman, 1934) from Mazaruni River, with this subclade as sister to the remaining species in the *P. franciscoensis* clade. Within the latter group, *P. wayana* Le Bail & Lucena, 2010 from eastern Guiana Shield (Amapá and Jari rivers) is sister to two subclades: one with *P.* aff. *retropinna* Lucena & Malabarba, 2010 (Tapajós) sister to an undescribed species from the Xingu River, and another with successive less inclusive clades containing *P. eurytaenia* Antonetti, Lucena & Lucena, 2018 (Tocantins) sister to *P. naevata* Antonetti, Lucena & Lucena, 2018 (Tocantins) which is sister to a clade of *P. calverti* (Fowler, 1941) (Parnaíba) and *P. franciscoensis* Eigenmann, 1911 (São Francisco). Interestingly, the species from Tocantins, Parnaíba and São Francisco basins are related to each other forming a geographically-structured subclade. This result is similar to studies with other fish groups distributed throughout those basins such as *Otocinclus hasemani* and *O. xakriaba* (Schaefer, 1997), *Hisonotus* sp. 1, *Hisonotus* sp. 2 and *Parotocinclus* aff. *spilurus* (Roxo *et al.*, 2014), and *Prochilodus lacustris* and *P. brevis* (Melo *et al.*, 2018). The time-calibrated phylogeny suggests that "*Phenacogaster pectinata* clade" and "*P. franciscoensis* clade" diverged during the Miocene, at approximately 8.4 Ma (10.4–6.4 Ma 95% HPD; Figure 5).

Acanthocharax is the only monotypic genera of Characinae and restricted to its type species *A. microlepis* Eigenmann, 1912 with distribution through the Essequibo River and adjacent drainages of Guyana. In our analysis, it was represented by a single specimen from Kurupung River in Guyana and the phylogeny placed it as sister to the major clade containing Characini and Cynopotamini.

Our phylogenetic hypothesis supported the monophyly of *Cynopotamus* based on nine out of

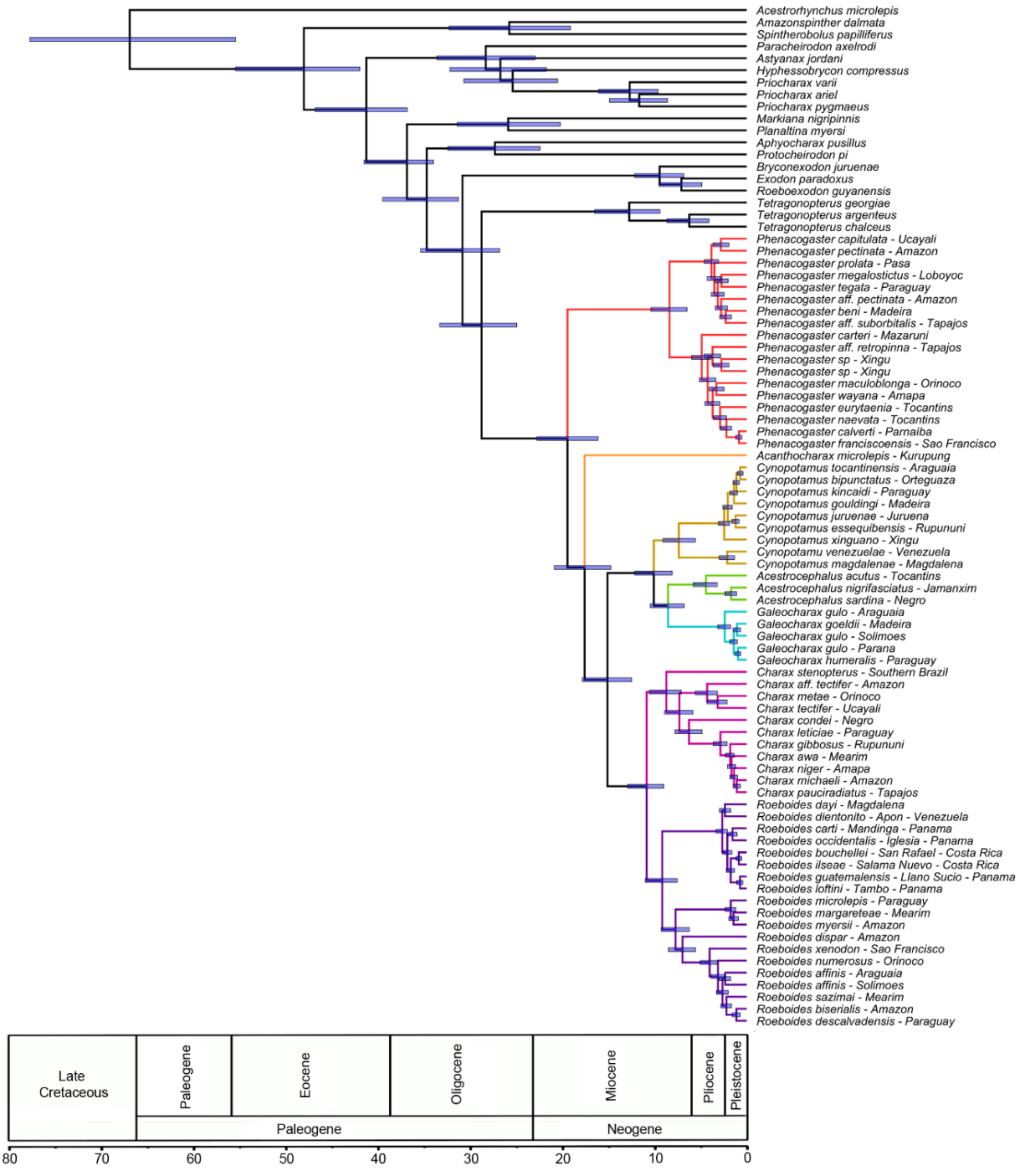


Figure 5. Time-calibrated phylogeny for Characinae based on a BEAST analysis of 91 UCE loci present for at least 95% of 85 specimens (66 Characinae and 19 outgroups). Node bars show the 95% highest posterior distribution of ages.

12 (75%) species of the genus. The species were grouped into two clades with high support (ML=100). The first is the "*C. magdalenae* clade" (Figure 3) that represents the first evidence of a close relationship among species from west of the Andes, with strong support for *C. venezuelae* (Schultz, 1944) from Santa Rosa River of Venezuela as sister to *C. magdalenae* (Steindachner, 1879) from Samaná River/Magdalena basin of Colombia. The second is the "*Cynopotamus bipunctatus* clade" formed by the following species from east of the Andes: *C. bipunctatus* Pellegrin, 1909, *C. essequibensis* Eigenmann, 1912, *C. gouldingi* Menezes, 1987, *C. juruenae* Menezes 1987, *C. kincaidi* (Schultz, 1950), *C. tocantinensis* Menezes, 1987, and *C. xinguano* Menezes 2007. According to timing estimates, the MRCA of *Cynopotamus* originated during the Miocene at approximately 10.0 Ma (12.5–8.0 Ma 95% HPD; Figure 5). The *C. magdalenae* clade (west of Andes) and *C. bipunctatus* clade (east of Andes) diverged between 9.0–5.5 Ma (95% HPD) which is coincident with the rise of the Eastern Cordillera (~12 Ma) and the rise of the Merida Andes (~8 Ma) with the isolation of the modern Maracaibo and Orinoco basins (Albert *et al.*, 2006).

Acestrocephalus is well supported as monophyletic in our analysis. Our phylogeny showed *A. acutus* Menezes, 2006 from the Tocantins River as sister to a clade formed by *A. nigrifasciatus* Menezes 2006 from Jamanxim/Tapajós and *A. sardina* (Fowler, 1913) from Sipapo and Negro rivers (Figure 3). Because our study included <50% of the species of *Acestrocephalus*, future studies with a greater number of species are necessary to better understand the relationships within the genus.

Galeocharax was recently revised with only three valid species [*G. goeldii* (Fowler, 1913), *G. gulo* (Cope, 1870) and *G. humeralis* (Valenciennes, 1834)] with *G. knerii* Steindachner, 1879 from the Upper Paraná river basin considered a junior synonym of *G. gulo* (type locality: Pebas, Peru, and widely distributed throughout Amazonas, Orinoco, Tocantins, and upper Paraná Rivers) (Giovannetti *et al.*, 2017). Here, we analyzed five samples of *G. gulo* from distinct drainages (Araguaia, Purus, Solimões, Amazon, and Paraná) and results did not support its monophyly (Figure 3; Figures S1-S5; ML=100). The topology found that the specimen from Araguaia River is sister to

the other two clades: one with *G. gulo* (Paraná; formerly *G. knerii*) and *G. humeralis* (Paraguay) and another with *G. goeldii* (Madeira) and *G. gulo* (Amazon drainages). Hence, our phylogeny does not support the proposal of synonymy of *G. knerii* and highlight that *G. gulo* from the Araguaia River might represent an undescribed species. Indeed, Giovannetti *et al*. (2017) discussed the morphological variation in populations from the Amazon, Paraná and Tocantins rivers, which would be interesting to be reinvestigated in light of our genomic data.

In *Charax*, molecular results strongly supported *C. stenopterus* from the coastal drainages of southern Brazil, the "*Charax stenopterus* clade", as a sister group of a large clade composed by species from the Amazon, Orinoco, Paraguay and Mearim river basins (Figures 4, S1-S5; ML=100). Within the large clade, two subclades emerge: "*C. tectifer* clade" and "*C. gibbosus* clade". We did not find monophyly for *C. condei* and *C. niger* (Figure 4; Figures S1-S5; ML=100). *Charax condei* (Géry & Knöppel, 1976) (Trombetas) and *C.* aff. *tectifer* (Amazon) were closer to a clade with *C. metae* Eigenmann 1922 (Orinoco) and *C. tectifer* (Cope, 1870) (Ucayali). This is a clade restricted to the proto-Orinoco-Amazonas, a previous fluvial configuration in northwestern South America, with species shared between the Orinoco and Amazon basin (Hoorn *et al*., 2010; Albert *et al.*, 2018). Within the *Charax gibbosus* clade, *C. condei* (Amazon and Negro) were related to a successive less inclusive clade with *C. leticiae* Lucena, 1987 (Paraguay), *C. gibbosus* (Linnaeus, 1758) (Rupununi), *C. awa* Guimarães, Brito, Ferreira & Ottoni, 2018 (Mearim), *C. niger* Lucena, 1989 (Oiapoque), *C. niger* (Amapá), *C. pauciradiatus* (Günther, 1864) (Tapajós) and *C. michaeli* Lucena, 1989 (Amazon). This biogeographic pattern approximates with the timing of the channelization of the Amazon River at approximately 10–5 Ma, and the subsequent connection and diversification of the fish faunas from western and eastern Amazon basin (Albert *et al.*, 2018; 2021).

In *Roeboides*, Lucena (1998) used osteological and meristic traits and proposed the division of *Roeboides* in four groups: *R. dispar* group, *R. guatemalensis* group, *R. microlepis* group, and the *R. affinis* group. In addition, Lucena (2000) considered *R. guatemalensis* and *R. microlepis* groups as

sister clades, with the *R. affinis* group as sister of the clade formed by these groups, and *R. dispar* as the sister to all other groups. Our results partially corroborated the composition of the four groups (herein termed clades) but did not support *R. guatemalensis* clade as sister to *R. microlepis* clade (Figure 4; Figures S1-S5). Our phylogenetic hypothesis showed the *R. guatemalensis* clade from west of the Andes as sister to a large monophyletic assemblage with *R. dispar* clade sister to *R. microlepis* and *R. affinis* clades (Figure 4). However, the relationship between the *R. microlepis* and *R. affinis* clades presented lower bootstrap support (ML=48%). Furthermore, our study did not support the monophyly of *R. affinis*, with one lineage of *R. affinis* (Araguaia) sister to a clade with *R. affinis* (Solimões and Juruá), *R. sazimai* (Mearim), *R. biserialis* (Amazon) and *R. descalvadensis* (Paraguay and Paraná). The time-calibrated tree estimated that the MRCA of *R. guatemalensis* clade (west-Andes) had its origin in the Pliocene (2.6 Ma, 3.3–2.0 Ma 95% HPD) and the last speciation event of this clade corresponds to species from Panama (0.7 Ma, 1.0–0.4 Ma 95% HPD) and Costa Rica (0.8 Ma, 1.1–0.5 Ma 95% HPD). This result suggests that the distribution and diversification of this clade into Central America are directly associated with the closure of the Isthmus of Panama at approximately 4–2.8 Ma (Bermingham & Martin, 1998; Coates & Stallard , 2013; McGirr *et al.*, 2021).

Overall, this study presents the most taxon-dense phylogeny for Characinae to date covering all genera and provides novel hypotheses for intergeneric and interspecific relationships. Limitations of our study include the relatively lower species coverage of *Acestrocephalus* (~38%), *Charax* (~58%) and *Phenacogaster* (~61%), the low support for the relationship between *Phenacogaster* aff. *retropinna* and *P.* sp. Xingu with others from the *P. franciscoensis* clade (48%), the low support for the relationship between the *Roeboides microlepis* clade and the *R. affinis* clade (48%), and the moderate support for the placement of *R. affinis* (56%) and *R. sazimai* (70%) within the *R. affinis* clade (Figs. 2–3). Finally, our results indicate various instances of relationships that are similar to previous morphological phylogenies (e.g. Lucena, 1998; Mattox & Toledo-Piza, 2012) and provide

an interesting framework for future investigation on the evolutionary processes modulating species diversification. Revisionary studies of species diversity and additional molecular data are also crucial to better understand the species diversity and evolution in Characinae.

Acknowledgements

We are grateful to Mary Burridge (ROM), María Fernando Castillo (STRI), Carlos Lucena (MCP), Carolina Doria (UFRO-ICT), Camila C. Ribas (INPA), Erick C. Guimarães (UFMA), and Michel D. Gianeti (MZUSP) for curatorial assistance and/or loan of tissues and vouchers. We are also thankful to LBP colleagues Luz E. Ochoa, Nadayca T. B. Mateussi, Fabio F. Roxo, Cristhian C. Conde-Saldaña and Eric. V. Ywamoto for advices and support with analyses. We also thank Martin Taylor, Johnny Jensen, Frank Schäefer, Frank Teigler, Pablo Giorgis, and Ralf Britz for providing photographs of Figure 1, and to Peter van der Sleen for fish illustrations in Figures 3-4.

Funding

This research was supported by the Brazilian agencies FAPESP proc. 17/06551-0 (CSS), 16/11313-8 and 18/23883-9 (BFM), 17/01970-4 (GMTM), 18/20610-1, 16/09204-6, 14/26508-3, CNPq proc. 200159/2020-8 (BFM), proc. 306054/2006-0 (CO), and the Stiassny AMNH Axelrod Research Curatorship of the American Museum of Natural History (BFM).

Bibliography

Aberer, A.J., Kobert, K., Stamatakis, A., 2014. ExaBayes: Massively parallel bayesian tree inference for the whole-genome era. Mol. Biol. Evol. 31, 2553–2556. https://doi.org/10.1093/molbev/msu236.

Adey, A., Morrison, H.G., Asan, Xun, X., Kitzman, J.O., Turner, E.H., Stackhouse, B., MacKenzie, A.P., Caruccio, N.C., Zhang, X., Shendure, J., 2010. Rapid, low-input, low-bias construction

of shotgun fragment libraries by high-density in vitro transposition. Genome Biol. 11. https://doi.org/10.1186/gb-2010-11-12-r119

Albert, J. S., Lovejoy, N. R., & Crampton, W. G., 2006. Miocene tectonism and the separation of cis-and trans-Andean river basins: Evidence from Neotropical fishes. Journal of South American Earth Sciences, 21(1-2), 14-27. https://doi.org/10.1016/j.jsames.2005.07.010

Albert, J.S., Bart, H.L., Reis, R.E., 2011a. Species richness and cladal diversity. In: Albert, J.S., Reis, R.E. (Eds.), Historical Biogeography of Neotropical Freshwater Fishes. University of California Press, Berkeley, Los Angeles, pp. 89–104.

Albert, J.S., Petry, P., Reis, R.E., 2011b. Major biogeographic and phylogenetic patterns. In: Albert, J.S., Reis, R.E. (Eds.), Historical Biogeography of Neotropical Freshwater Fishes. University of California Press, Berkeley, Los Angeles, pp. 21–57. https://doi.org/10.1525/california/9780520268685.003.0002

Albert, J. S., Val, P., & Hoorn, C., 2018. The changing course of the Amazon River in the Neogene: center stage for Neotropical diversification. Neotropical Ichthyol., 16 (3), e180033. https://doi.org/10.1590/1982-0224-20180033

Albert, J. S., Bernt, M. J., Fronk, A. H., Fontenelle, J. P., Kuznar, S. L., & Lovejoy, N. R., 2021. Late Neogene megariver captures and the Great Amazonian Biotic Interchange. Glob. Planet. Change, 205, 103554. https://doi.org/10.1016/j.gloplacha.2021.103554

Alda, F., Ludt, W.B., Elías, D.J., McMahan, C.D. and Chakrabarty, P., 2021. Comparing ultraconserved elements and exons for phylogenomic analyses of Middle American cichlids: when data agree to disagree. Genome Biol. Evol., 13(8), p.evab161. https://doi.org/10.1093/gbe/evab161

Alfaro, M.E., Faircloth, B.C., Harrington, R.C., Sorenson, L., Friedman, M., Thacker, C.E., Oliveros, C.H., Černý, D., Near, T.J., 2018. Explosive diversification of marine fishes at the Cretaceous-Palaeogene boundary. Ecol. Evol. Nat. https://doi.org/10.1038/ s41559-018-0494-6

Betancur-R, R., Arcila, D., Vari, R.P., Hughes, L.C., Oliveira, C., Sabaj, M.H., Ortí, G., 2019. Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: The monophyly of characiform fishes. Evolution 73, 329–345. https://doi.org/10.1111/evo.13649

Bermingham, E., & Martin, A. P., 1998. Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. Mol. Ecol. 7(4), 499-517. https://doi.org/10.1046/j.1365-294x.1998.00358.x

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., ... & Drummond, A. J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol., 10(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Buckup, P.A., 1998. Relationships of the Characidiinae and phylogeny of characiform fishes (Teleostei: Ostariophysi). In: Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M.S., Lucena, C.A.S. (Eds.), Phylogeny and classification of neotropical fishes. Edipucrs, Porto Alegre, pp. 123–144.

Bührnheim, C. M., Carvalho, T. P., Malabarba, L. R., & Weitzman, S. H., 2008. A new genus and species of characid fish from the Amazon basin: the recognition of a relictual lineage of characid fishes (Ostariophysi: Cheirodontinae: Cheirodontini). Neotropical Ichthyol., 6, 663-678. https://doi.org/10.1590/S1679-62252008000400016

Carvalho, T. P., & Albert, J. S., 2011. Eleven. The Amazon-Paraguay Divide. In Historical biogeography of Neotropical freshwater fishes (pp. 193-202). University of California Press.

Carstens, B. C., & Knowles, L. L., 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. Syst. Biol. 56(3), 400-411. https://doi.org/10.1080/10635150701405560

Chakrabarty, P., Faircloth, B.C., Alda, F., Ludt, W.B., Mcmahan, C.D., Near, T.J., et al., 2017. Phylogenomic systematics of ostariophysan fishes: Ultraconserved elements support the surprising non-monophyly of characiformes. Syst. Biol. 66 (6), 881–895. https://doi.org/10.1093/sysbio/syx038

Coates, A. G., & Stallard, R. F., 2013. How old is the Isthmus of Panama?. Bull. Mar. Sci., 89(4), 801-813. https://doi.org/10.5343/bms.2012.1076

Denton, R. D., Kenyon, L. J., Greenwald, K. R., & Gibbs, H. L., 2014. Evolutionary basis of mitonuclear discordance between sister species of mole salamanders (*Ambystoma* sp.). Mol. Eco., 23(11), 2811-2824. https://doi.org/10.1111/mec.12775

Degnan, J. H., & Rosenberg, N. A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2(5), e68. https://doi.org/10.1371/journal.pgen.0020068

Degnan, J. H., & Rosenberg, N. A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol., 24(6), 332-340. https://doi.org/10.1016/j.tree.2009.01.009

Derti, A., Roth, F.P., Church, G.M., Wu, C.T., 2006. Mammalian ultraconserved elements are strongly depleted among segmental duplications and copy number variants. Nat. Genet. 38, 1216-1220. https://doi.org/10.1038/ng1888

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nuc. Acid. Res. 32 (5), 1792–1797. https://doi.org/10.1093/nar/ gkh340

Eigenmann, C.H. 1912. The freshwater fishes of British Guiana, including a study of the ecological grouping of species, and the relation of the fauna of the plateau to that of the lowlands. Memoirs of the Carnegie Museum v. 5 (no. 1): i-xxii + 1-578, Pls. 1-103. https://www.biodiversitylibrary.org/page/7752475#page/35/mode/1up

Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61, 717–726. https://doi.org/10.1093/ sysbio/sys004

Faircloth, B.C., Sorenson, L., Santini, F., Alfaro, M.E., 2013. A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). PLoS ONE 8 (6), e65923. https://doi.org/10.1371/ journal.pone.0065923

Faircloth, B.C., 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. https://doi.org/10.1093/bioinformatics/btv646

Faircloth, B.C., Alda, F., Hoekzma, K., Burns, M.D., Oliveira, C., Albert, J.S., Melo, B.F., Ochoa, L.E., Roxo, F.F., Chakrabarty, P., Sidlauskas, B.L., Alfaro, M.E., 2020. A target enrichment bait set for studying relationships among ostariophysan fishes. Copeia 108, 47–60. https://doi.org/10.1643/CG-18-139

Fricke, R., Eschmeyer, W.N., Fong, J. D., 2021. Eschmeyer's catalog of fishes: genera/species by family/subfamily. (http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version accessed 24 03 2021.

Froese, R., Pauly, D., 2019. FishBase. World Wide Web Electronic Publications. Available from http://www.fishbase.org. Accessed 24 03 2021.

Géry, J., 1964. Poissons characoïdes nouveaux ou non signalés de l'Ilha do Bananal, Brésil. Vie et Milieu 17, 448-475

Géry, J., 1977. Characoids of the World. TFH Publications, Neptune City, New Jersey.

Giovannetti, V., Toledo-Piza, M., Menezes, N.A., 2017. Taxonomic revision of *Galeocharax* (Characiformes: Characidae: Characinae). Neotropical Ichthyol. 15. https://doi.org/10.1590/1982-0224-20160040

Harris, R., 2007. Improved pairwise alignment of genomic DNA. Ph.D. thesis. The Pennsylvania State University.

Harrington, R.C., Faircloth, B.C., Eytan, R.I., Smith, W.L., Near, T.J., Alfaro, M.E., Friedman, M., 2016. Phylogenomic analysis of carangimorph fishes reveals flatfish asymmetry arose in a blink

of the evolutionary eye. BMC Evol. Biol. 16, 1–14. https://doi.org/10.1186/s12862-016-0786-x

Hoorn, C., Wesselingh, F. P., Ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., ... & Antonelli, A., 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science, 330, 927-931. https://doi.org/10.1126/science.1194585

Howes, G.J., 1976. The cranial musculature and taxonomy of characoid fishes of the tribes Cynodontini and Characini. Bull. br. Mus. nat. Hist. Zool. 29, 201–248.

Javonillo, R., Malabarba, L.R., Weitzman, S.H., Burns, J.R., 2010. Relationships among major lineages of characid fishes (Teleostei: Ostariophysi: Characiformes), based on molecular sequence data. Mol. Phylogenet. Evol. 54, 498–511. https://doi.org/10.1016/j.ympev.2009.08.026

Kircher, M., Sawyer, S., Meyer, M., 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. 40 (1), e3. https:// doi.org/10.1093/nar/gkr771

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., Stamatakis, A., 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol. Biol. 14, 82. https://doi.org/10.1186/1471-2148-14-82

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773.

Lucena, C.A.S., 1998. Relações filogenéticas e definição do gênero *Roeboides* Günther (Ostariophysi: Characiformes: Characidae). Com. Museu Ciênc. PUCRS 11, 19–59.

Lucena, C.A.S., 2000. Revisão taxonômica e filogenia das espécies transandinas do gênero *Roeboides* Günther (Teleostei: Ostariophysi: Characiformes). Com. Museu Ciênc. PUCRS 13, 3–63.

Lucena, C.A.S., Menezes, N.A., 2003. Subfamily Characinae, in: Reis, R.E., Kullander, S.O., Ferraris, C.J.Jr. (Eds.), Checklist of the freshwater fishes of South and Central America. Edipucrs, Porto Alegre, pp. 200–208.

Lucena, C.A.S., 2007. Revisão taxonômica das espécies do gênero *Roeboides* grupo-affinis (Ostariophysi, Characiformes, Characidae). Iheringia 97, 117–136.

Lucena, Z.M.S., Malabarba, L.R., 2010. Descrição de nove espécies novas de *Phenacogaster* (Ostariophysi: Characiformes: Characidae) e comentários sobre as demais espécies do gênero. Zoologia 27, 263–304. http://dx.doi.org/10.1590/S1984-46702010000200014

Lundberg J.G., Marshall L.G., Guerrero J., Horton B., Malabarba M.C.S.L., Wesselingh F. 1998. The stage for Neotropical fish diversification: A history of tropical South American rivers. In: Malabarba L.R., Reis R.E., Vari R.P., Lucena Z.M.S., Lucena C.A.S., editors. Phylogeny and classification of neotropical fishes. Porto Alegre: Edipucrs. p. 13–48.

Malabarba, L.R., Lucena, Z.M.S., 1995. *Phenacogaster jancupa*, new species, with comments on the relationships and a new diagnosis of the genus (Ostariophysi: Characidae). Ichthyol. Explor. Freshw. 6, 337–344.

Malinsky, M., Challis, R. J., Tyers, A. M., Schiffels, S., Terai, Y., Ngatunga, B. P., ... & Turner, G. F., 2015. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. Science, 350(6267), 1493-1498. http://dx.doi.org/10.1126/science.aac9927

Matamoros, W.A., Chakrabarty, P., Angulo, A., Garita-Alvarado, C.A., McMahan, C.D. 2013. A new species of *Roeboides* (Teleostei: Characidae) from Costa Rica and Panama, with a key to the middle American species of the genus. Neotropical Ichthyol. 11, 285-290. https://doi.org/10.1590/S1679-62252013000200006

Mateussi, N.T., Melo, B.F., Ota, R.P., Roxo, F.F., Ochoa, L.E., Foresti, F., Oliveira, C. 2020. Phylogenomics of the Neotropical fish family Serrasalmidae with a novel intrafamilial classification (Teleostei: Characiformes). Mol. Phylogenet. Evol. 153, 106945. https://doi.org/10.1016/j.ympev.2020.106945

Mattox, G.M., Toledo-Piza, M., 2012. Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). Zool. J. Linnean Soc. 165, 809–915.

McCormack, J.E., Faircloth, B.C., Crawford, N.G., Gowaty, P.A., Brumfield, R.T., Glenn, T.C., 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Res. 22, 746–754. https://doi.org/10.1101/gr.125864.111

McGirr, R., Seton, M., & Williams, S., 2021. Kinematic and geodynamic evolution of the Isthmus of Panama region: Implications for Central American Seaway closure. Bulletin, 133(3-4), 867-884. https://doi.org/10.1130/B35595.1

Melo, B.F., Benine, R.C., Silva, G.S., Avelino, G.S., Oliveira, C., 2016. Molecular phylogeny of the Neotropical fish genus *Tetragonopterus* (Teleostei: Characiformes: Characidae). Mol. Phylogenet. Evol. 94, 709–717. https://doi.org/10.1016/j.ympev.2015.10.022

Melo, B. F., Dorini, B. F., Foresti, F., & Oliveira, C., 2018. Little divergence among mitochondrial lineages of *Prochilodus* (Teleostei, Characiformes). Frontiers in genetics, 9, 107. https://doi.org/10.3389/fgene.2018.00107

Melo, B. F., Sidlauskas, B. L., Near, T. J., Roxo, F. F., Ghezelayagh, A., Ochoa, L. E., ... & Oliveira, C. (2021). Accelerated diversification explains the exceptional species richness of tropical characoid fishes. Syst. Biol. https://doi.org/10.1093/sysbio/syab040

Menezes, N.A., 1976. On the Cynopotaminae, a new subfamily of Characidae (Osteichthyes, Ostariophysi, characoidei). Arq. Zool. 28, 1–91.

Menezes, N.A., 2006. Description of five new species of *Acestrocephalus* Eigenmann and redescription of *A. sardina* and *A. boehlkei* (Characiformes: Characidae). Neotropical Ichthyol. 4, 385–400. https://doi.org/10.1590/S1679-62252006000400002

Menezes, N.A., 2007. *Charax* goeldii Fowler, 1913 a valid species of *Galeocharax* Fowler (Characiformes: Characidae). Neotropical Ichthyol. 5, 21–24. https://doi.org/10.1590/S1679-62252007000100002

Menezes, N.A., Lucena, C.A.S., 2014. A taxonomic review of the species of *Charax* Scopoli, 1777 (Teleostei: Characidae: Characinae) with description of a new species from the rio Negro bearing superficial neuromasts on body scales, Amazon basin, Brazil. Neotropical Ichthyol. 12, 193–228. https://doi.org/10.1590/1982-0224-20130175

Mirande, J.M., 2009. Weighted parsimony phylogeny of the family Characidae (Teleostei: Characiformes). Cladistics 25, 1–40. https://doi.org/10.1111/j.1096-0031.2009. 00262.x

Mirande, J.M., 2010. Phylogeny of the family Characidae (Teleostei: Characiformes): from characters to taxonomy. Neotropical Ichthyol. 8, 385–568. https://doi.org/10. 1590/S1679-62252010000300001

Mirande, J.M., 2019. Morphology, molecules and the phylogeny of Characidae (Teleostei, Characiformes). Cladistics 35, 282–300. https://doi.org/10.1111/cla.12345

Montoya-Burgos, J. I., 2003. Historical biogeography of the catfish genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. Mol. Ecol. 12(7), 1855-1867. https://doi.org/10.1046/j.1365-294X.2003.01857.x

Ochoa, L.E., Datovo, A., DoNascimiento, C., Roxo, F.F., Sabaj, M.H., Chang, J., Melo, B.F., Silva, G.S.C., Foresti, F., Alfaro, M., Oliveira, C., 2020. Phylogenomic analysis of trichomycterid catfishes (Teleostei: Siluriformes) inferred from ultraconserved elements. Sci. Rep. 10, 2697. https://doi.org/10.1038/s41598-020-59519-w

Oliveira, C., Avelino, G.S., Abe, K.T., Mariguela, T.C., Benine, R.C., Ortí, G., Vari, R.P., Corrêa E Castro, R.M., 2011. Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. BMC Evol. Biol. 11, 275. https://doi.org/10.1186/1471- 2148-11-275

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R., Moret, B.M., Stamatakis, A., 2010. How many bootstrap replicates are necessary? J. Comput. Biol. 17, 337–354. https://doi.org/10.1089/cmb.2009.0179

Pollard, D. A., Iyer, V. N., Moses, A. M., & Eisen, M. B., 2006. Widespread discordance of gene trees with species tree in *Drosophila*: evidence for incomplete lineage sorting. PLoS genet. 2(10), e173. https://doi.org/10.1371/journal.pgen.0020173

Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer.

Ramirez, J. L., Birindelli, J. L., & Galetti Jr, P. M. 2017. A new genus of Anostomidae (Ostariophysi: Characiformes): diversity, phylogeny and biogeography based on cytogenetic, molecular and morphological data. Mol. Phylogenet. Evol. 107, 308-323. https://doi.org/10.1016/j.ympev.2016.11.012

Roberts, T.R., 1970. Scale-eating american characoid fishes: with special reference to *Probolodus heterostomus*. Proc. Calif. Acad. Sci. 38, 383–390.

Rosser, N. L., Thomas, L., Stankowski, S., Richards, Z. T., Kennington, W. J., & Johnson, M. S., 2017. Phylogenomics provides new insight into evolutionary relationships and genealogical discordance in the reef-building coral genus Acropora. Proc. R. Soc. Lond. B Biol. Sci., 284(1846), 20162182. https://doi.org/10.1098/rspb.2016.2182

Roxo, F. F., Albert, J. S., Silva, G. S., Zawadzki, C. H., Foresti, F., & Oliveira, C., 2014. Molecular phylogeny and biogeographic history of the armored Neotropical catfish subfamilies Hypoptopomatinae, Neoplecostominae and Otothyrinae (Siluriformes: Loricariidae). PLoS One, 9(8), e105564. https://doi.org/10.1371/journal.pone.0105564

Roxo, F.F., Ochoa, L.E., Sabaj, M.H., Lujan, N.K., Covain, R., Silva, G.S.C., Melo, B.F., Albert, J.S., Chang, J., Foresti, F., Alfaro, M.E., Oliveira, C., 2019. Phylogenomic reappraisal of the Neotropical catfish family Loricariidae (Teleostei: Siluriformes) using ultraconserved elements. Mol. Phylogenet. Evol. 148–165. https://doi.org/10.1016/j.ympev.2019.02.017

Sazima, I., Machado, F.A., 1982. Habits and behavior of *Roeboides prognathus*, a scale-eating fish (Osteichthyes, Characoidei). Bolm. Zool. Univ. S. Paulo 7, 35–56.

Sazima, I., 1984. Scale-eating in characoids and other fishes. In: Zaret, T.M. (Eds.) Evolutionary ecology of neotropical freshwater fishes. Springer, Dordrecht, pp. 9–23.

Schaefer, S. A., 1997. The Neotropical cascudinhos: systematics and biogeography of the *Otocinclus* catfishes (Siluriformes: Loricariidae). Proceedings of the Academy of Natural Sciences of Philadelphia, 148, 1-120.

Siepel A., Bejerano G., Pedersen J.S., Hinrichs A.S., Hou M., Rosenbloom K., Clawson H., Spieth J., Hillier L.W., Richards S., Weinstock G.M., Wilson R.K., Gibbs R.A., Kent W.J., Miller W., Haussler D., 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15, 1034–1050. http://www.genome.org/cgi/doi/10.1101/gr.3715005

Sivasundar, A., Bermingham, E., & Ortí, G., 2001. Population structure and biogeography of migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers. Mol. Ecol. 10(2), 407-417. https://doi.org/10.1046/j.1365-294X.2001.01194.x

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690. https://doi.org/10.1093/bioinformatics/btl446.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10. 1093/bioinformatics/btu033

Sun, J. T., Wang, M. M., Zhang, Y. K., Chapuis, M. P., Jiang, X. Y., Hu, G., ... & Hong, X. Y., 2015. Evidence for high dispersal ability and mito-nuclear discordance in the small brown planthopper, *Laodelphax striatellus*. Sci. Rep. 5(1), 1-10.

Tagliacollo, V.A., Souza-Lima, R., Benine, R.C., Oliveira, C., 2012. Molecular phylogeny of Aphyocharacinae (Characiformes, Characidae) with morphological diagnoses for the subfamily and recognized genera. Mol. Phylogenet. Evol. 64, 297-307. https://doi.org/10.1016/j.ympev.2012.04.007

Tagliacollo, V.A., Lanfear, R., 2018. Estimating improved partitioning schemes for ultraconserved elements. Mol. Biol. Evol. 35, 1798–1811. https://doi.org/10.1093/ molbev/msy069

Toews, D. P., & Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21(16), 3907-3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Weiss F.E., Malabarba L.R., Malabarba M.C., 2012. Phylogenetic relationships of *Paleotetra*, a new characiform fish (Ostariophysi) with two new species from the Eocene-Oligocene of south-eastern Brazil. J Syst. Palaeontol. 10:73–86. https://doi.org/10.1080/14772019.2011.565082

Weitzman, S.H., Vari, R.P., 1987. Two new species and a new genus of miniature characid fishes (Teleostei: Characiformes) from Northern South America. Proc. Biol. Soc. Wash. 100, 640–652.

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf. 19, 153. https://doi.org/10.1186/s12859-018-2129-y

Zerbino, D., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829. https://doi.org/10.1101/gr.074492.107

Supplementary information

Supplementary Table 1. Taxonomic sampling, voucher and museum number, and geographic information of the samples included in this study.

Family	Subfamily	*Genus*	Species	Voucher	Museum number	Drainage (River/Basin)	Country	Coordinates
Characidae	Characinae	*Acanthocharax*	*Acanthocharax microlepis*	T20425	ROM uncat	Kurupung	Guyana	N 06°13.324' W 060°09.068'
Characidae	Characinae	*Acestrocephalus*	*Acestrocephalus acutus*	96197	LBP 25321	Tocantins	Brazil	S 14°18'17.8" W 46°38'04.7"
Characidae	Characinae	*Acestrocephalus*	*Acestrocephalus nigrifasciatus*	MZict 003344	MZUSP 097494	Jamanxim	Brazil	S 7° 3' 52.0" W 55° 26' 28.0"
Characidae	Characinae	*Acestrocephalus*	*Acestrocephalus sardina*	T09194	ROM uncat	Sipapo	Guyana	N 06°48'13.5" W 73°16'23.1"
Characidae	Characinae	*Acestrocephalus*	*Acestrocephalus sardina*	33172	LBP 6876	Negro	Brazil	S 00°08.156' W 67°05.057'
Characidae	Characinae	*Acestrocephalus*	*Acestrocephalus sardina*	70161	LBP 17833	Negro	Brazil	S 00°30'5.3" W 64°49'12.2"
Acestrorhynchidae	Acestrorhynchinae	*Acestrorhynchus*	*Acestrorhynchus microlepis*	82946	LBP 21142	Amazon	Brazil	N 03°26'03.6" W 51°43'40.8"
Characidae	Spintherobolinae	*Amazonspinther*	*Amazonspinther dalmata*	46055	LBP 9309	Madeira	Brazil	S 8°11'46" W 63°51'48"
Characidae	Aphyocharacinae	*Aphyocharax*	*Aphyocharax pusillus*	22920	LBP 4046	Moa/Amazon	Brazil	S 7°37'20.0" W 72°47'42.2"
Characidae	Stethaprioninae	*Astyanax*	*Astyanax jordani*	24599	LBP 4527	-	Brazil	-
Characidae	Exodontinae	*Bryconexodon*	*Bryconexodon juruenae*	101796	LBP 19627	Apiacás	Brazil	S 10°19'50.5" W 56°59'02.2"
Characidae	Characinae	*Charax*	*Charax awa*	CICCAA 02152-2	UFMA uncat	Mearim	Brazil	S 3°40'48" W 45°19'51"
Characidae	Characinae	*Charax*	*Charax condei*	INPA-ICT 050072	CTGA 105925	Trombetas	Brazil	S 1°26'47.0" W 56°47'45.0"
Characidae	Characinae	*Charax*	*Charax condei*	74340	LBP 18290	Amazon	Brazil	S 00°26'10.0" W 64°57'05.8"
Characidae	Characinae	*Charax*	*Charax condei*	74420	LBP 18319	Negro	Brazil	S 00°30'5.3" W 64°49'12.2"
Characidae	Characinae	*Charax*	*Charax gibbosus*	T06353	ROM uncat	Rupununi	Guyana	-
Characidae	Characinae	*Charax*	*Charax leticiae*	21981	LBP 3730	Paraguay	Brazil	S 19°34'54.6' W 56°15'16.5"
Characidae	Characinae	*Charax*	*Charax metae*	61598	LBP 18653	Orinoco	Colombia	N 3°29'26.6" W 73°44'34.1"
Characidae	Characinae	*Charax*	*Charax michaeli*	86577	LBP 22517	Amazon	Colombia	S 04°11'45.6'' W 69°57'20.9"
Characidae	Characinae	*Charax*	*Charax michaeli*	87532	LBP 22517	Amazon	Colombia	S 04°11'45.6'' W 69°57'20.9"
Characidae	Characinae	*Charax*	*Charax niger*	82731	LBP 21086	Oiapoque	Brazil	N 03°48'47.6" W 51°48'31.6"
Characidae	Characinae	*Charax*	*Charax niger*	83211	LBP 21217	Amapá	Brazil	N 02°03'42.8" W 50°54'15.1"
Characidae	Characinae	*Charax*	*Charax pauciradiatus*	66967	LBP 16206	Tapajós	Brazil	S 04°28'11.2" W 56°17'01.1"
Characidae	Characinae	*Charax*	*Charax stenopterus*	20473	LBP 3338	Southern Brazil	Brazil	S 32°09'06.9" W 52°06'24.2"
Characidae	Characinae	*Charax*	*Charax stenopterus*	68399	LBP 17020	Southern Brazil	Brazil	S 32°05'24.1" W 52°15'09.7"
Characidae	Characinae	*Charax*	*Charax tectifer*	70280	LBP 17783	Ucayali	Peru	S 08°39'57.2" W 74°48'08.7"
Characidae	Characinae	*Charax*	*Charax tectifer*	70281	LBP 17783	Ucayali	Peru	S 08°39'57.2" W 74°48'08.7"
Characidae	Characinae	*Charax*	*Charax* aff. *tectifer*	86771	LBP 22793	Amazon	Brazil	S 04°11'25.5'' W 69°55'40.3"

Family	Subfamily	*Genus*	Species	Voucher	Museum number	Drainage (River/Basin)	Country	Coordinates
Characidae	Characinae	*Cynopotamus*	*Cynopotamus bipunctatus*	T24786	ROM uncat	Orteguaza	Colombia	-
Characidae	Characinae	*Cynopotamus*	*Cynopotamus bipunctatus*	91511	LBP 24313	Orteguaza	Colombia	N 1°31'09.3" W 75°32'19.0"
Characidae	Characinae	*Cynopotamus*	*Cynopotamus essequibensis*	T06254	ROM uncat	Rupununi	Guyana	-
Characidae	Characinae	*Cynopotamus*	*Cynopotamus gouldingi*	95300	LBP 25817	Madeira	Brazil	S 11°55'25.5" W 61°14'17.6"
Characidae	Characinae	*Cynopotamus*	*Cynopotamus juruenae*	MZict003991	MZUSP 095880	Juruena	Brazil	S 10°58' 30.0" W 55°44' 3.0"
Characidae	Characinae	*Cynopotamus*	*Cynopotamus kincaidi*	3567	LBP 19	Paraguay	Brazil	S 19°34.630' W 57°01.123'
Characidae	Characinae	*Cynopotamus*	*Cynopotamus magdalenae*	91521	LBP 24320	Magdalena	Colombia	-
Characidae	Characinae	*Cynopotamus*	*Cynopotamus tocantinensis*	36391	LBP 7769	Araguaia	Brazil	-
Characidae	Characinae	*Cynopotamus*	*Cynopotamus venezuelae*	29515	LBP 6132	Santa Rosa	Venezuela	N 09°38'53.8'' W 72°34'56.4"
Characidae	Characinae	*Cynopotamus*	*Cynopotamus xinguano*	96753	LBP 5967	Xingu	Brazil	S 13°50'55.7" W 53°15'26.7"
Characidae	Exodontinae	*Exodon*	*Exodon paradoxus*	44278	LBP 8851	Araguaia	Brazil	S 13°22'36.1' W 50°40'08.4"
Characidae	Characinae	*Galeocharax*	*Galeocharax goeldii*	UFRO-ICT 013389	UFRO uncat	Madeira	Brazil	S 08°48'30" W 63°36'53"
Characidae	Characinae	*Galeocharax*	*Galeocharax gulo*	16224	LBP 2391	Araguaia	Brazil	S 15°53'35.2" W 52°15'00.9"
Characidae	Characinae	*Galeocharax*	*Galeocharax gulo*	17088	LBP 2537	Purus	Brazil	S 08°50'54.8" W 68°41'23.7"
Characidae	Characinae	*Galeocharax*	*Galeocharax gulo*	86588	LBP 22568	Solimões	Brazil	S 04°19'28.0'' W 69°57'36.7"
Characidae	Characinae	*Galeocharax*	*Galeocharax gulo*	87955	LBP 22644	Amazon	Colombia	S 04°11'39.3'' W 69°58'06.9"
Characidae	Characinae	*Galeocharax*	*Galeocharax gulo*	96995	LBP 28889	Paraná	Brazil	-
Characidae	Characinae	*Galeocharax*	*Galeocharax humeralis*	55986	LBP 13453	Paraguay	Brazil	S 17°47'59.0" W 57°23'41.6"
Characidae	Characinae	*Galeocharax*	*Galeocharax humeralis*	56099	LBP 13483	Paraguay	Brazil	S 17°51'29.6" W 57°28'34.1"
Characidae	Stethaprioninae	*Hyphessobrycon*	*Hyphessobrycon compressus*	61599	LBP 19581	-	Belize	S 17°27'29.4" W 88°23'34.4"
Characidae	Stevardiinae	*Markiana*	*Markiana nigripinnis*	36284	LBP 7586	Paraguay	Brazil	S 16°11'39.5" W 55°48'25.1"
Characidae	Stethaprioninae	*Paracheirodon*	*Paracheirodon axelrodi*	24428	LBP 4472	Negro	Brazil	S 00°40'03.1" W 62°58'23.5"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster beni*	8084	UFRO uncat	Madeira	Brazil	S 08°41' W 63°51'
Characidae	Characinae	*Phenacogaster*	*Phenacogaster calverti*	27299	LBP 5582	Parnaíba	Brazil	S 09°09'51’ W 45°51'15'
Characidae	Characinae	*Phenacogaster*	*Phenacogaster capitulata*	72078	LBP 17802	Ucayali	Peru	S 08°35'44.2" W 74°48'04.3"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster carteri*	T21412	ROM uncat	Mazaruni	Guyana	-
Characidae	Characinae	*Phenacogaster*	*Phenacogaster eurytaenia*	93886	LBP 25604	Tocantins	Brazil	S 12°37'27.9" W 48°02'43.5"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster franciscoensis*	49188	LBP 10487	São Francisco	Brazil	S 17°14'32.3' W 46°28'03.8"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster franciscoensis*	49189	LBP 10487	São Francisco	Brazil	S 17°14'32.3' W 46°28'03.8"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster megalostictus*	T10318	ROM uncat	Loboyoc	Peru	-
Characidae	Characinae	*Phenacogaster*	*Phenacogaster naevata*	77592	LBP 19200	Tocantins	Brazil	S 11°03'14.3" W 48°34'22.0'
Characidae	Characinae	*Phenacogaster*	*Phenacogaster pectinata*	53617	LBP 12406	Amazon	Peru	S 04°09'04.2" W 73°28'25.7"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster* aff. *suborbitalis*	59025	LBP 14095	Tapajós	Brazil	S 04°55'58.8" W 56°51'51.6"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster* aff. *pectinata*	86591	LBP 22469	Amazon	Colombia	S 04°08'24.4'' W 69°56'53.4"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster prolata*	T09902	ROM uncat	Pasa	Venezuela	-
Characidae	Characinae	*Phenacogaster*	*Phenacogaster* aff. *retropinna*	66285	LBP 16035	Tapajós	Brazil	S 14°25'33.8" W 54°00'56.6"

Family	Subfamily	*Genus*	Species	Voucher	Museum number	Drainage (River/Basin)	Country	Coordinates
Characidae	Characinae	*Phenacogaster*	*Phenacogaster* aff. *retropinna*	81396	LBP 20843	Tapajós	Brazil	S 13°48'04.8" W 56°01'37.5"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster* sp.	66487	LBP 16061	Xingu	Brazil	S 14°38'24.0" W 53°55'38.0"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster* sp.	94285	LBP 25131	Xingu	Brazil	S 08°47'03.0" W 54°58'29.0"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster maculoblonga*	52572	LBP18683	Orinoco	Colombia	N 3°16'17.4" W 73°36'47.0"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster wayana*	101795	LBP 5402	Jari	Brazil	S 00°37’24’’ W 52°32’49’’
Characidae	Characinae	*Phenacogaster*	*Phenacogaster tegata*	24817	LBP 4683	Paraguay	Brazil	S 14°41'44" W 57°15"35"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster tegata*	28227	LBP 5795	Paraguay	Brazil	S 15°44'3.60" W 55°52'48.7"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster tegata*	36059	LBP 7641	Paraguay	Brazil	S 15°46'03.8" W 55°30'44.5"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster tegata*	49883	LBP 10785	Paraguay	Brazil	S 18°25'24.4" W 54°50'05.9"
Characidae	Characinae	*Phenacogaster*	*Phenacogaster wayana*	83213	LBP 21218	Amapá	Brazil	N 02°03'42.8" W 50°54'15.1"
Characidae	Stevardiinae	*Planaltina*	*Planaltina myersi*	75276	LBP 11680	Paraná	Brazil	S 17°56'39.5" W 47°58'09.6"
Characidae	-	*Priocharax*	*Priocharax ariel*	96382	LBP 17836	Negro	Brazil	S 00°25'55.6" W 65°01'16.3"
Characidae	-	*Priocharax*	*Priocharax pygmaeus*	96993	LBP 22464	Amazon	Brazil	S 04°08'24.4’’ W 69°56'53.4"
Characidae	-	*Priocharax*	*Priocharax varii*	96981	LBP 28495	Madeira	Brazil	S 08°52′53.5″ W 63°37′50.8″
Characidae	Cheirodontinae	*Protocheirodon*	*Protocheirodon pi*	49268	LBP 10565	Amazon	Brazil	S 10°03'28.6’’ W 67°51'25.6"
Characidae	Exodontinae	*Roeboexodon*	*Roeboexodon guyanensis*	59203	LBP 14159	Tapajós	Brazil	S 04°33'45.4" W 56°15'36.9"
Characidae	Characinae	*Roeboides*	*Roeboides affinis*	22929	LBP 4050	Juruá	Brazil	S 7°37'20.0" W 72°47'42.2"
Characidae	Characinae	*Roeboides*	*Roeboides affinis*	68750	LBP 17195	Araguaia	Brazil	S 15°10'23.2" W 51°09'27.1"
Characidae	Characinae	*Roeboides*	*Roeboides affinis*	68752	LBP 17195	Araguaia	Brazil	S 15°10'23.2" W 51°09'27.1"
Characidae	Characinae	*Roeboides*	*Roeboides affinis*	87169	LBP 22564	Solimões	Brazil	S 04°19'28.0’’ W 69°57'36.7"
Characidae	Characinae	*Roeboides*	*Roeboides affinis*	87171	LBP 22564	Solimões	Brazil	S 04°19'28.0’’ W 69°57'36.7"
Characidae	Characinae	*Roeboides*	*Roeboides biserialis*	49261	MCP 49261	Amazon	Brazil	S 02°11'29.4" W 54°51'28.0"
Characidae	Characinae	*Roeboides*	*Roeboides bouchellei*	STRI-2098	STRI uncat	San Rafael	Costa Rica	N 9°58'27.4" W 84°34'46.5"
Characidae	Characinae	*Roeboides*	*Roeboides carti*	STRI-324	STRI uncat	Mandinga	Panama	N 9°28'11.8" W 79°07'26.9"
Characidae	Characinae	*Roeboides*	*Roeboides dayi*	90745	LBP 24273	Magdalena	Colombia	N 5°11'47.2" W 74°45'52.6"
Characidae	Characinae	*Roeboides*	*Roeboides descalvadensis*	22254	LBP 3834	Paraguay	Brazil	S 19°34'17.3' W 56°14'44.8"
Characidae	Characinae	*Roeboides*	*Roeboides descalvadensis*	22783	LBP 3958	Paraguay	Brazil	S 15°54'03' W 56°01'17"
Characidae	Characinae	*Roeboides*	*Roeboides descalvadensis*	43169	LBP 9211	Paraná	Brazil	S 22°51'45.3" W 48°06'21.0"
Characidae	Characinae	*Roeboides*	*Roeboides descalvadensis*	56252	LBP 13526	Paraguay	Brazil	S 17°49'26.7" W 57°31'03.0"
Characidae	Characinae	*Roeboides*	*Roeboides biserialis*	72599	LBP 18034	Amazon	Brazil	S 03°05'57.7’" W58°27'18.0"
Characidae	Characinae	*Roeboides*	*Roeboides dientonito*	15633	LBP 2220	Orinoco	Venezuela	N 07º30'50.9" W 66º09'19.8"
Characidae	Characinae	*Roeboides*	*Roeboides dientonito*	29546	LBP 6106	Santa rosa	Venezuela	N 09°38'53.8' W 72°34'56.4"
Characidae	Characinae	*Roeboides*	*Roeboides dientonito*	29551	LBP 6106	Santa rosa	Venezuela	N 09°38'53.8' W 72°34'56.4"
Characidae	Characinae	*Roeboides*	*Roeboides dientonito*	29576	LBP 6114	Apón	Venezuela	N 10°01'42.0" W 72°25'58.0"
Characidae	Characinae	*Roeboides*	*Roeboides dientonito*	29578	LBP 6114	Apón	Venezuela	N 10°01'42.0" W 72°25'58.0"
Characidae	Characinae	*Roeboides*	*Roeboides dispar*	49266	LBP 10150	Amazon	Brazil	S 10°03'28.6’’ W 67°51'25.6"

Family	Subfamily	*Genus*	Species	Voucher	Museum number	Drainage (River/Basin)	Country	Coordinates
Characidae	Characinae	*Roeboides*	*Roeboides guatemalensis*	18531	LBP 2755	Llano Sucio	Panama	N 09°19’26.2" W 79°46'08.2"
Characidae	Characinae	*Roeboides*	*Roeboides ilseae*	STRI-2020	STRI uncat	Salama Nuevo	Costa Rica	N 8°54'15.3" W 83°26'21.6"
Characidae	Characinae	*Roeboides*	*Roeboides loftini*	AM-17	STRI uncat	Tambo	Panama	N 8°39'56.4" W 80°17'15.0"
Characidae	Characinae	*Roeboides*	*Roeboides margareteae*	CICCAA 02075-1	UFMA uncat	Mearim	Brazil	S 3°39'11" W 45°46'25"
Characidae	Characinae	*Roeboides*	*Roeboides microlepis*	26124	LBP 5098	Paraguay	Brazil	S 16°06'66" W 57°44'33"
Characidae	Characinae	*Roeboides*	*Roeboides myersii*	87569	LBP 22518	Amazon	Colombia	S 04°11'45.6’’ W 69°57'20.9"
Characidae	Characinae	*Roeboides*	*Roeboides myersii*	87570	LBP 22518	Amazon	Colombia	S 04°11'45.6’’ W 69°57'20.9"
Characidae	Characinae	*Roeboides*	*Roeboides numerosus*	47538	LBP 10217	Orinoco	Venezuela	N 07°37'24.4' W 66°24'48.0"
Characidae	Characinae	*Roeboides*	*Roeboides occidentalis*	STRI-3459	STRI uncat	Iglesia	Panama	N 8°25'23.0" W 78°00'05.0"
Characidae	Characinae	*Roeboides*	*Roeboides sazimai*	CICCAA 02111-2	UFMA uncat	Mearim	Brazil	S 3°43'48.0" W 45°35'07.0"
Characidae	Characinae	*Roeboides*	*Roeboides occidentalis*	91506	LBP 24307	Patía	Colombia	N 2°00'45.8" W 77°10'26.1"
Characidae	Characinae	*Roeboides*	*Roeboides xenodon*	48878	LBP 10387	São Francisco	Brazil	S 17°17'21.9" W 44°47'08.4"
Characidae	Characinae	*Roeboides*	*Roeboides xenodon*	49056	LBP 10444	São Francisco	Brazil	S 17°14'19.9' W 46°23'39.0"
Characidae	Spintherobolinae	*Spintherobolus*	*Spintherobolus papilliferus*	61600	LBP 20540	Southern Brazil	Brazil	S 23°41'44.3" W 46°03'10.7"
Characidae	Tetragonopterinae	*Tetragonopterus*	*Tetragonopterus argenteus*	22029	LBP 3758	Paraguay	Brazil	S 19°34'33.7' W 56°14'49.5"
Characidae	Tetragonopterinae	*Tetragonopterus*	*Tetragonopterus chalceus*	56694	LBP 13946	Tapajós	Brazil	S 04°16'49.5" W 59°59'26.1"
Characidae	Tetragonopterinae	*Tetragonopterus*	*Tetragonopterus georgiae*	82752	LBP 21093	Oiapoque	Brazil	N 03°48'47.6" W 51°48'31.6"

Supplementary Table 2. Values of the number of trimmed reads, number of contigs assembled, total base pair of contigs, number of UCE contigs, mean length of all contigs and their average coverage for samples.

Species	Voucher number	Museum number	Number of trimmed reads	Contigs assembled	Total bp contigs	UCE contigs	Total bp	Acession Number SRA (to be provider)
Acanthocharax microlepis	T20425	ROM uncat	1877491	95320	18986551	1711	1107139	
Acestrocephalus acutus	96197	LBP 25321	5552410	158811	36812329	1716	1217665	
Acestrocephalus nigrifasciatus	MZict 003344	MZUSP 097494	3612375	114793	30069817	1808	1475447	
Acestrocephalus sardina	T09194	ROM uncat	6206732	144214	33977853	1730	1230821	
Acestrocephalus sardina	33172	LBP 6876	2417235	112993	21301594	1825	1320513	
Acestrocephalus sardina	70161	LBP 17833	2691051	127793	22897321	1593	1045710	
Acestrorhynchus microlepis	82946	LBP 21142	2428153	208329	31788019	1532	790007	
Amazonspinther dalmata	46055	LBP 9309	1866742	65231	14504391	1811	1264008	
Aphyocharax pusillus	22920	LBP 4046	607315	24458	6172995	1658	1215382	
Astyanax jordani	24599	LBP 4527	1555975	52509	12011776	1807	1510645	
Bryconexodon juruenae	101796	LBP 19627	4378954	146490	33409085	1864	1423224	
Charax awa	CICCAA 02152-2	UFMA uncat	717160	44674	10564962	1769	1307419	
Charax condei	INPA-ICT 050072	CTGA 105925	2938292	99569	24091224	1773	1223686	
Charax condei	74340	LBP 18290	912116	32820	5737739	1532	455199	
Charax condei	74420	LBP 18319	302094	20522	3739870	1483	462733	
Charax gibbosus	T06353	ROM uncat	3863551	132103	31076638	1543	1058106	
Charax leticiae	21981	LBP 3730	4254867	204799	35725608	1732	1195472	
Charax metae	61598	LBP 18653	3137304	279869	44135637	1665	914222	
Charax michaeli	86577	LBP 22517	3261718	283321	43829983	1606	756482	
Charax michaeli	87532	LBP 22517	5675611	173555	34331596	1719	1147749	
Charax niger	82731	LBP 21086	1385494	75275	13677718	1764	1171830	
Charax niger	83211	LBP 21217	4557912	128268	30901571	1678	1268604	
Charax pauciradiatus	66967	LBP 16206	6158435	141253	27011692	1504	630112	
Charax stenopterus	20473	LBP 3338	2784253	42522	11856412	1706	1189117	
Charax stenopterus	68399	LBP 17020	1378262	139397	20749796	1564	768457	
Charax tectifer	70280	LBP 17783	5777570	180694	38375399	1458	811042	
Charax tectifer	70281	LBP 17783	12698337	234371	53402835	1327	702370	

Species	Voucher number	Museum number	Number of trimmed reads	Contigs assembled	Total bp contigs	UCE contigs	Total bp	Acession Number SRA (to be provider)
Charax aff. *tectifer*	86771	LBP 22793	4121063	100804	18048092	1315	460353	
Cynopotamus bipunctatus	T24786	ROM uncat	3214252	109779	27775636	1799	1421351	
Cynopotamus bipunctatus	91511	LBP 24313	3954389	115647	24531053	1787	1276430	
Cynopotamus essequibensis	T06254	ROM uncat	6113287	155873	35608392	1274	736085	
Cynopotamus gouldingi	95300	LBP 25817	485760	36067	8001093	1722	1076654	
Cynopotamus juruenae	MZict003991	MZUSP 095880	2991506	102245	27285738	1593	1257737	
Cynopotamus kincaidi	3567	LBP 19	1729753	27672	8987514	1636	1297777	
Cynopotamus magdalenae	91521	LBP 24320	3281790	100785	20840468	1822	1192517	
Cynopotamus tocantinensis	36391	LBP 7769	4883580	129089	30168041	1672	1097599	
Cynopotamus venezuelae	29515	LBP 6132	8962606	242195	52815265	1518	940255	
Cynopotamus xinguano	96753	LBP 5967	14008613	263409	62522749	1475	960240	
Exodon paradoxus	44278	LBP 8851	2020665	184324	28741642	1654	712453	
Galeocharax goeldii	UFRO-ICT 013389	UFRO uncat	7030141	172517	40963031	1616	1173506	
Galeocharax gulo	16224	LBP 2391	4842365	136004	32883994	1360	899939	
Galeocharax gulo	17088	LBP 2537	10436829	182754	32467282	1230	385610	
Galeocharax gulo	86588	LBP 22568	4370264	104420	24009441	1846	1322181	
Galeocharax gulo	87955	LBP 22644	2716879	118391	28804427	1781	1470720	
Galeocharax gulo	96995	LBP 28889	5037200	135499	32150722	1744	1191108	
Galeocharax humeralis	55986	LBP 13453	1873791	85516	16884682	1825	1221099	
Galeocharax humeralis	56099	LBP 13483	1375740	220092	34513913	1461	506036	
Hyphessobrycon compressus	61599	LBP 19581	3106427	84791	14097450	1527	648124	
Markiana nigripinnis	36284	LBP 7586	1909439	172838	27165057	1595	779487	
Paracheirodon axelrodi	24428	LBP 4472	2392042	174405	26452753	1594	669044	
Phenacogaster beni	8084	UFRO uncat	3260120	107240	25254542	1803	1262990	
Phenacogaster calverti	27299	LBP 5582	7883570	179795	42053472	1639	1106904	
Phenacogaster capitulata	72078	LBP 17802	3047061	102245	24956300	1797	1372484	
Phenacogaster carteri	T21412	ROM uncat	3143939	109238	25665793	1825	1302980	
Phenacogaster eurytaenia	93886	LBP 25604	2854466	108597	23953360	1806	1107557	
Phenacogaster franciscoensis	49188	LBP 10487	2481790	125655	21340828	1816	1166464	
Phenacogaster franciscoensis	49189	LBP 10487	2499520	108651	19670102	1845	1260484	

Species	Voucher number	Museum number	Number of trimmed reads	Contigs assembled	Total bp contigs	UCE contigs	Total bp	Acession Number SRA (to be provider)
Phenacogaster megalostictus	T10318	ROM uncat	3509774	113072	27268119	1792	1302191	
Phenacogaster naevata	77592	LBP 19200	8045271	182620	39689079	1374	705495	
Phenacogaster pectinata	53617	LBP 12406	5925538	213291	41349909	1731	903297	
Phenacogaster aff. *suborbitalis*	59025	LBP 14095	3579284	304366	46377043	1680	860080	
Phenacogaster aff. *pectinata*	86591	LBP 22469	3459207	109238	23448385	1841	1223561	
Phenacogaster prolata	T09902	ROM uncat	3464978	118062	26731124	1722	1141872	
Phenacogaster aff. *retropinna*	66285	LBP 16035	6232323	182370	39220228	1720	1031989	
Phenacogaster aff. *retropinna*	81396	LBP 20843	5930018	247971	48870009	1540	989193	
Phenacogaster sp.	66487	LBP 16061	2827616	97174	22932206	1736	1201838	
Phenacogaster sp.	94285	LBP 25131	3951380	140280	30431184	1741	1017071	
Phenacogaster maculoblonga	52572	LBP18683	9584014	106038	29368950	1216	607834	
Phenacogaster wayana	5402	LBP 21218	4982346	136784	31007922	1667	1072399	
Phenacogaster tegata	24817	LBP 4683	720041	37466	7491682	1697	1005205	
Phenacogaster tegata	28227	LBP 5795	1467856	146615	21990433	1570	740327	
Phenacogaster tegata	36059	LBP 7641	1626215	56500	13073073	1761	1234328	
Phenacogaster tegata	49883	LBP 10785	1948491	77179	16781004	1791	1247891	
Phenacogaster wayana	83213	LBP 21218	1902149	79151	18143566	1756	1208017	
Planaltina myersi	75276	LBP 11680	1900107	192673	27595692	1503	673157	
Priocharax ariel	96382	LBP 17836	3362630	254810	44215773	1379	593152	
Priocharax pygmaeus	96993	LBP 22464	2762984	216059	37673050	1422	657983	
Priocharax varii	96981	LBP 28495	4074576	311850	54327852	1358	587939	
Protocheirodon pi	49268	LBP 10565	2511759	217160	30825691	1489	633654	
Roeboexodon guyanensis	59203	LBP 14159	3500379	152110	31927493	1835	1500890	
Roeboides affinis	22929	LBP 4050	1908805	93295	19942005	1822	1310152	
Roeboides affinis	68750	LBP 17195	7007474	185491	41047041	1699	1126387	
Roeboides affinis	68752	LBP 17195	4088914	326901	57796795	1383	555590	
Roeboides affinis	87169	LBP 22564	3808120	161757	29751714	1786	1200150	
Roeboides affinis	87171	LBP 22564	2564075	117703	20969170	1751	1115520	
Roeboides biserialis	49261	MCP 49261	6119495	477137	85112389	1366	590953	
Roeboides bouchellei	STRI-2098	STRI uncat	3205245	112445	27835003	1801	1529458	

Species	Voucher number	Museum number	Number of trimmed reads	Contigs assembled	Total bp contigs	UCE contigs	Total bp	Acession Number SRA (to be provider)
Roeboides carti	STRI-324	STRI uncat	2817050	108138	26575429	1827	1541102	
Roeboides dayi	90745	LBP 24273	4152653	360897	54771948	1439	616306	
Roeboides descalvadensis	22254	LBP 3834	2104711	252451	39059175	1603	744857	
Roeboides descalvadensis	22783	LBP 3958	7867895	991655	161939966	1492	606454	
Roeboides descalvadensis	43169	LBP 9211	2188904	179817	29640291	1455	621781	
Roeboides descalvadensis	56252	LBP 13526	1331949	137199	22006306	1516	530136	
Roeboides biserialis	72599	LBP 18034	4433879	125964	29166836	1705	1169838	
Roeboides dientonito	15633	LBP 2220	5728762	183596	41705688	1757	1318600	
Roeboides dientonito	29546	LBP 6106	3296990	94378	20350591	1596	789722	
Roeboides dientonito	29551	LBP 6106	8459506	202654	46551334	1670	1103000	
Roeboides dientonito	29576	LBP 6114	3873021	207222	34366929	1807	1125183	
Roeboides dientonito	29578	LBP 6114	5749683	181241	37976916	1680	980272	
Roeboides dispar	49266	LBP 10150	6190339	154317	35191597	1630	982911	
Roeboides guatemalensis	18531	LBP 2755	1907664	70913	14676177	1592	700861	
Roeboides ilseae	STRI-2020	STRI uncat	3059068	109379	28934753	1829	1747762	
Roeboides loftini	AM-17	STRI uncat	3408292	119455	29961064	1800	1557738	
Roeboides margareteae	CICCAA 02075-1	UFMA uncat	3584061	126751	29855078	1781	1423784	
Roeboides microlepis	26124	LBP 5098	5463377	699610	110660529	1522	692241	
Roeboides myersii	87569	LBP 22518	8057964	190830	38091166	1147	419270	
Roeboides myersii	87570	LBP 22518	8791164	226724	49380739	1588	979023	
Roeboides numerosus	47538	LBP 10217	3475291	107390	25537215	1695	1248656	
Roeboides occidentalis	STRI-3459	STRI uncat	2294398	97500	24243839	1838	1628638	
Roeboides sazimai	CICCAA 02111-2	UFMA uncat	4569207	146914	33948020	1764	1323112	
Roeboides occidentalis	91506	LBP 24307	2081521	78699	16771035	1623	1093443	
Roeboides xenodon	48878	LBP 10387	1510325	29972	7368011	1500	939024	
Roeboides xenodon	49056	LBP 10444	2798665	77457	21347499	1838	1503391	
Spintherobolus papilliferus	61600	LBP 20540	2750620	234093	38151238	1565	825254	
Tetragonopterus argenteus	22029	LBP 3758	4499947	498451	71228641	1662	815787	
Tetragonopterus chalceus	56694	LBP 13946	2792781	119765	25543247	1840	1324641	
Tetragonopterus georgiae	82752	LBP 21093	4181681	154066	35554285	1859	1436975	

Supplementary data. List of morphological synapomorphies from Mattox & Toledo-Piza (2012) of the clades corroborated herein. Synapomorphies are organized by clades and are referred to by the number of the character in that study, and the state transition that represents the synapomorphy between parenthesis. Asterisks refer to exclusive synapomorphies of that clade.

Characinae (10 non-ambiguous synapomorphies):

4(0>2), absence of axilar scale on pelvic fin;

15(0>1), presence of a superficial lateral slit on *pars malaris* of *adductor mandibulae*;

16(0>1), presence of central raphe along anterior margin of *pars malaris* of *adductor mandibulae*;

31(0>1), *dilatator operculi* reaching at least the vertical through centre of orbit;

50(0>1), presence of superficial neuromasts associated with pores and striae on surface of infraorbital series;

74(0>1), toothed margin of maxilla longer than edentulous margin;

100(0>1), two anteriormost branchiostegal rays slender along entire length;

103(0>1), gill rakers on first branchial arch shaped as bony plates and covered with denticles;

128(0>1), presence of small notch on posteroventral margin of cleithrum;

141(2>0), absence of small ossifications associated with first proximal dorsal-fin radial.

Clade comprising *Acanthocharax*, *Acestrocephalus*, *Charax*, *Cynopotamus*, *Galeocharax* and *Roeboides* (nine non-ambiguous synapomorphies):

5(0>1), contralateral series of scales on the abdominal region forming a keel and bordering a gap along ventral mid-line between pelvic girdle and anus (reversed in *Acestrocephalus*);

12(0>1), dorsal extension of *pars rictalis* of the *adductor mandibulae* restricted ventrally;

37(0>1), lateral wings of mesethmoid poorly developed; *

38(0>1), the dorsal inflexion of the dorsal surface neurocranium resulting in large gibbosity (reversed in *Acestrocephalus* and *Galeocharax*);

90(0>1), presence of small and triangular posterodorsal projection of hyomandibula; 95 (0>1), space between posterior arms of dorsal hypohyal small forming a narrow slit;

97(0>2), articulation between anterior and posterior ceratohyals with strong interdigitating processes;

102(1>0), absence of bony lamella dorsal to fourth basibranchial, and lateral lamella of pelvic bone almost reaching anterior tip of rod-shaped element.

Characini (*Charax* and *Roeboides*) (five non-ambiguous synapomorphies):

47(0>1), supraoccipital spine well developed and collaborating with the great gibbosity;

60(0>2), absence of infraorbital 4 as an antogenous element (reversed to infraorbital 4 present but small in some species of *Roeboides*);

105 (1>0), presence of tooth plate on ventral surface of second pharyngobranchial (reversed in *Charax pauciradiatus*);

136 (0>1), anterior tip of pelvic bone located anterior to vertical through posterior margin of cleithrum; *

143 (0>1), first anal-fin pterygiophores with distinct posterodorsal slope (independently acquired in *Cynopotamus*).

Cynopotamini (*sensu* Menezes, 1976 as subfamily: *Acestrocephalus*, *Cynopotamus* and *Galeocharax*) (eight non-ambiguous synapomorphies):

1(1>2), presence of spinoid scales;

4(2>0), presence of a single axilar scale on pelvic fin;

11(0>2), dorsal limit of the pseudotympanum composed by the *lateralis superficialis* anteriorly and *obliquus superioris* posteriorly; *

16(1>0), absence of central raphe along anterior margin of *pars malaris* of the *adductor mandibulae* (independently acquired in *Charax condei*);

35(1>0), posterior margin of *levator operculi* situated posterior to vertical through posterior margin of *adductor operculi* (reversed in *Galeocharax knerii* [=*Galeocharax gulo*]);

41(1>0), absence of rhinosphenoid;

95(1>2), posterior arms of dorsal hypohyal fused to each other (independently acquired in *Charax pauciradiatus* and *C. stenopterus*);

98(0>1), interhyal with ventral tip slightly laterally flattened with small anterior projection.

Clade comprising *Acestrocephalus* and *Galeocharax* (six non-ambiguous synapomorphies):

38(1>0), dorsal surface of neurocranium flat (reversion of a synapomorphy of the clade comprising all characins except *Phenacogaster*);

63(0>1), anterior branch of laterosensory canal on infraorbital 6 elongate anteriorly;

77(1>0), presence of two tooth series on dentary;

105(1>0), presence of tooth plate on ventral surface of second pharyngobranchial;

128(1>0), posteroventral margin of cleithrum straight;

137(1>0), lateral lamella of pelvic bone short.

Phenacogasterini (*Phenacogaster*) (five non-ambiguous synapomorphies):

3(0>1), presence of two longitudinal series of large, narrow, and elongate preventral scales bent at the sides forming an angle with the side of the body; *

24(0>1), *segmentum mandibularis* of the *adductor mandibulae* dorsally pinnate;

68(1>0), anterior tip of nasal not reaching region between premaxilla and maxilla (independently acquired in *Roeboides dientonito* and *R. xenodon*);

89(0>1), posterior limit of metapterygoid-quadrate fenestra formed by metapterygoid, quadrate and symplectic;

120(0>1), presence of bony expansion on medial side of rib of fifth vertebra.

Acanthocharax microlepis (four non-ambiguous autapomorphies):

2(0>1), absence of pre-dorsal scales;

90(1>2), presence of large, approximately rounded posterodorsal projection of hyomandibula;

92(0>1), presence of spiniform projection on posterior margin of preopercle;

107(1>0), presence of two series of gill rakers on first ceratobranchial.

Acestrocephalus (six non-ambiguous synapomorphies):

5(1>0), contralateral series of longitudinal series of scales continuous on ventral surface, which is usually rounded and does not have keel or gap formed by scales (reversion of a synapomorphy of the clade comprising all characins except *Phenacogaster*);

17(0>1), dorsal origin of anterior portion of *pars mentalis* of the *adductor mandibulae* reaching articulation between hyomandibula and pterotic;

44(0>1), ascending process of parasphenoid reaching posterior margin of pterosphenoid; *

68(1>2), anterior tip of nasal reaching horizontal arm of premaxilla (independently acquired in *Roeboides affinis*);

90(1>2), presence of a large approximately round posterodorsal projection of hyomandibula.

108(0>1), presence of one series of gill rakers on second ceratobranchial;

Charax (four non-ambiguous synapomorphies):

12(1>2), dorsal extension of *pars rictalis* of the *adductor mandibulae* restricted ventrally to horizontal arm of preopercle;

41(1>0), absence of rhinosphenoid;

62(0>1), absence of anterodorsal branch of laterosensory canal on infraorbital 6;

128(1>2), presence of a large notch and projection on posteroventral margin of cleithrum. *

Cynopotamus (seven non-ambiguous synapomorphies):

39(0>1), epiphyseal bar with anterior and posterior lamellar expansions from frontals;

54(0>2), presence of a large and triangular projection on posteroventral portion of antorbital (reversed in *C. tocantinensis*);

64(0>2), presence of long process on infraorbital 6 at region of contact with infraorbital 5;

78(0>1), anterior portion of autopalatine separated by aperture delimiting medial rounded portion and straight lateral portion respectivelly in contact with neurocranium and maxilla;

85(0>1), presence of a process on the lateral margin of ectopterygoid (independently acquired in *Galeocharax gulo*);

121(0>1), rib of fifth vertebra longer and more robust than posterior ribs;

143(0>1), first anal-fin pterygiophores with distinct posterodorsal slope (acquired independently in Characini).

Galeocharax (five non-ambiguous synapomorphies):

7(0>1), scales on caudal fin covering lobes at least partially on median portion of fin;

30(0>1), limit of ventral portion of *levator arcus palatini* extending to anterior margin of metapterygoid;

55(0>1), presence of projection on anteroventral margin of infraorbital 1;

66(0>1), presence of a rudimentary lateral bony expansion on nasal;

127(1>2), presence of well-developed anteriorly directed projection along anteroventral margin of cleithrum (independently acquired in *Acestrocephalus sardina*, and in the clade comprising *Charax* and *Roeboides* as an ambiguous synapomorphy).

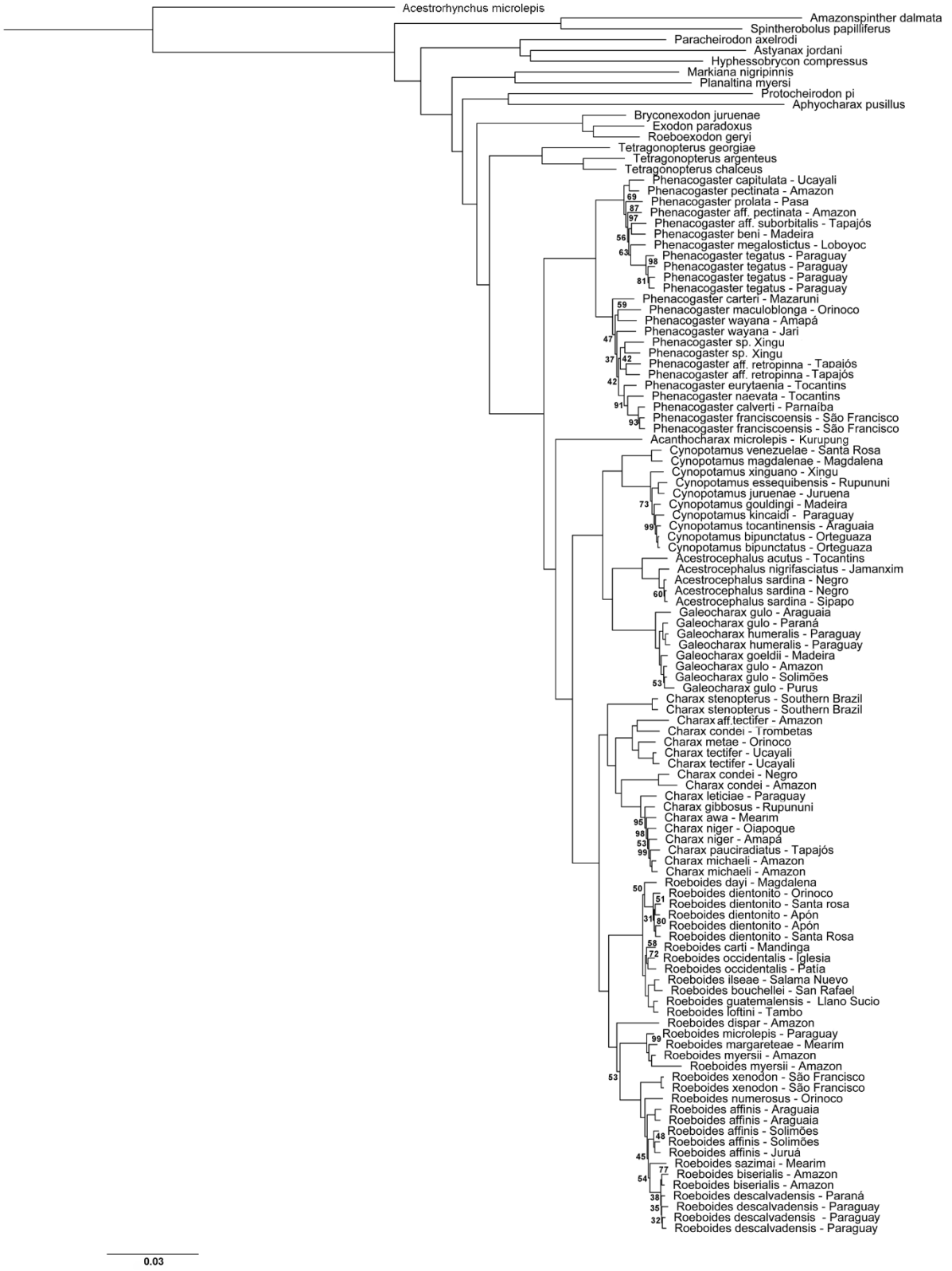
Roeboides (four non-ambiguous synapomorphies):

46(0>1), presence of projection on lateral tip of epiotic;

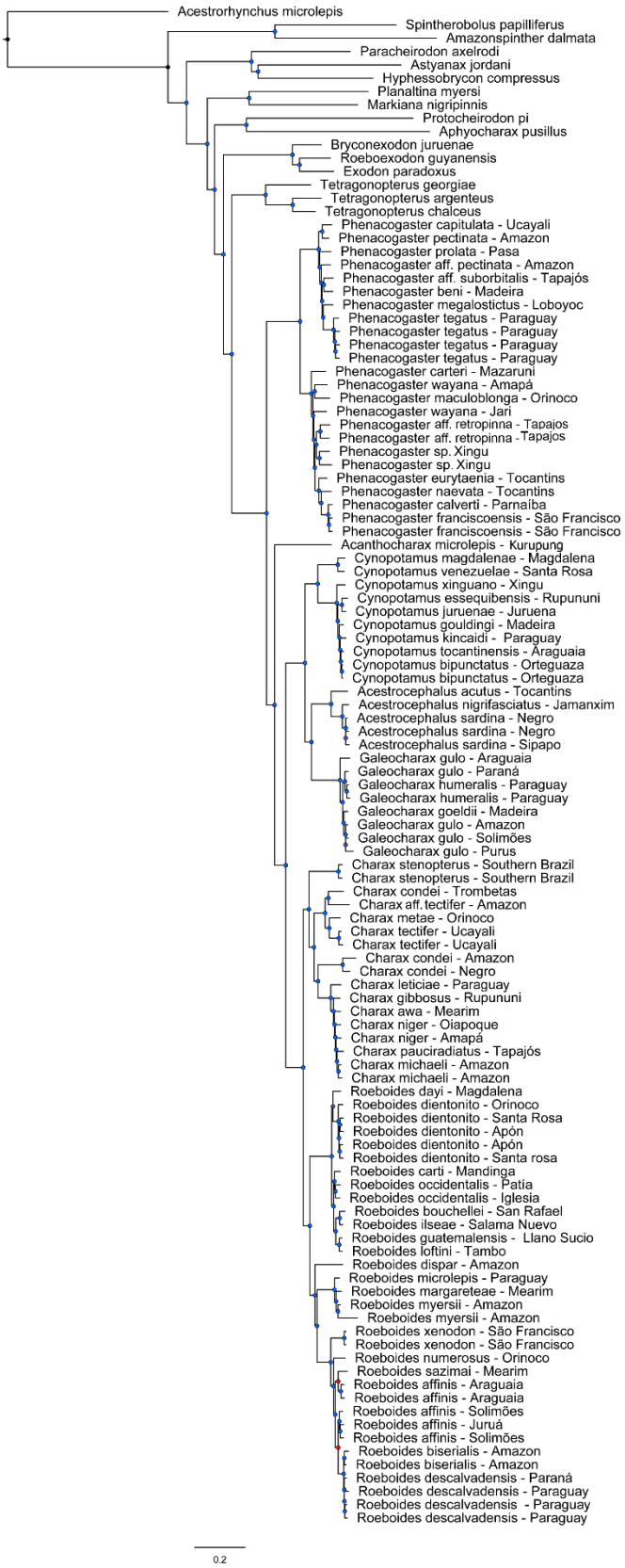
52(0>1), dorsal portion of antorbital positioned in space dorsal to lateral ethmoid and ventral to frontal; *

64(0>1), presence of short process on infraorbital 6 at region of contact with infraorbital 5, formed only by extension of laminar margin of bone (reversed in *Roeboides xenodon*); *

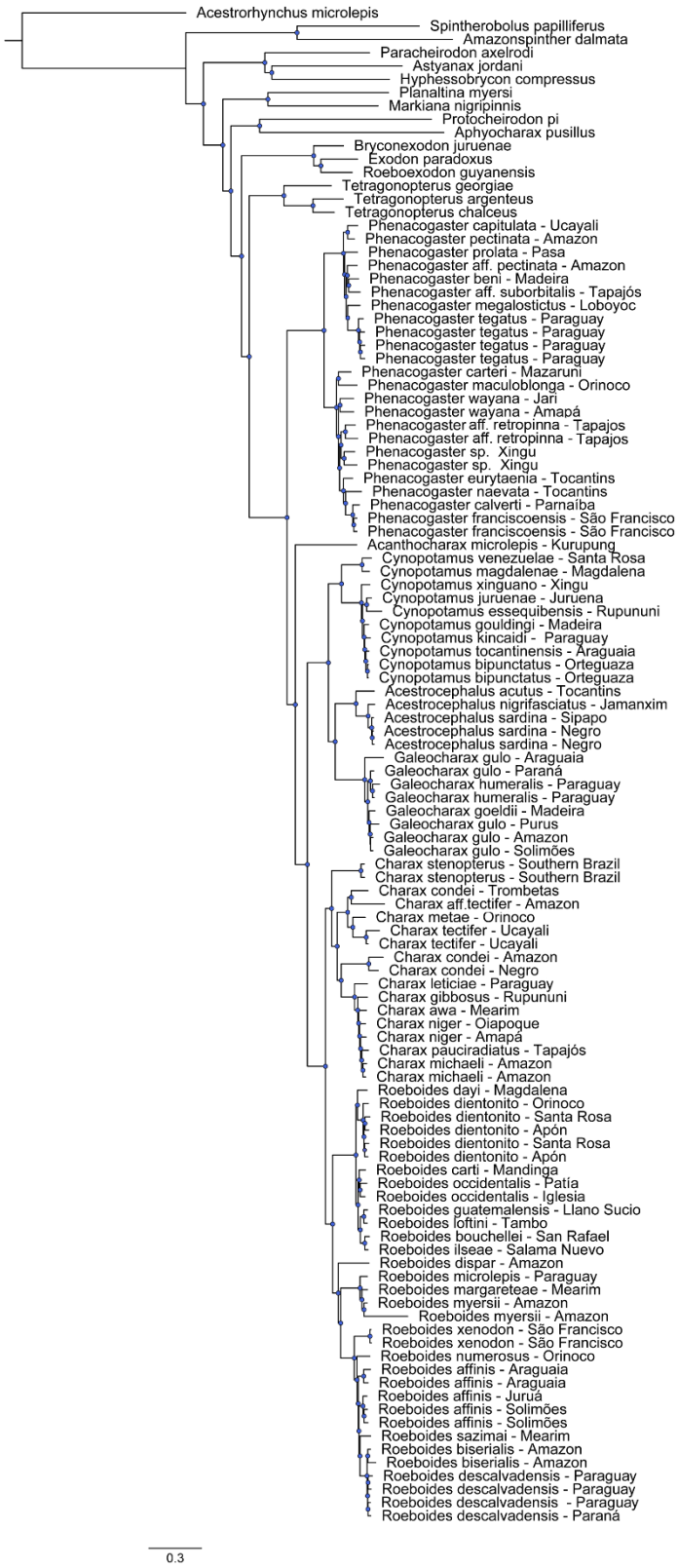
69(0>1), presence of mamilliform teeth outside mouth.



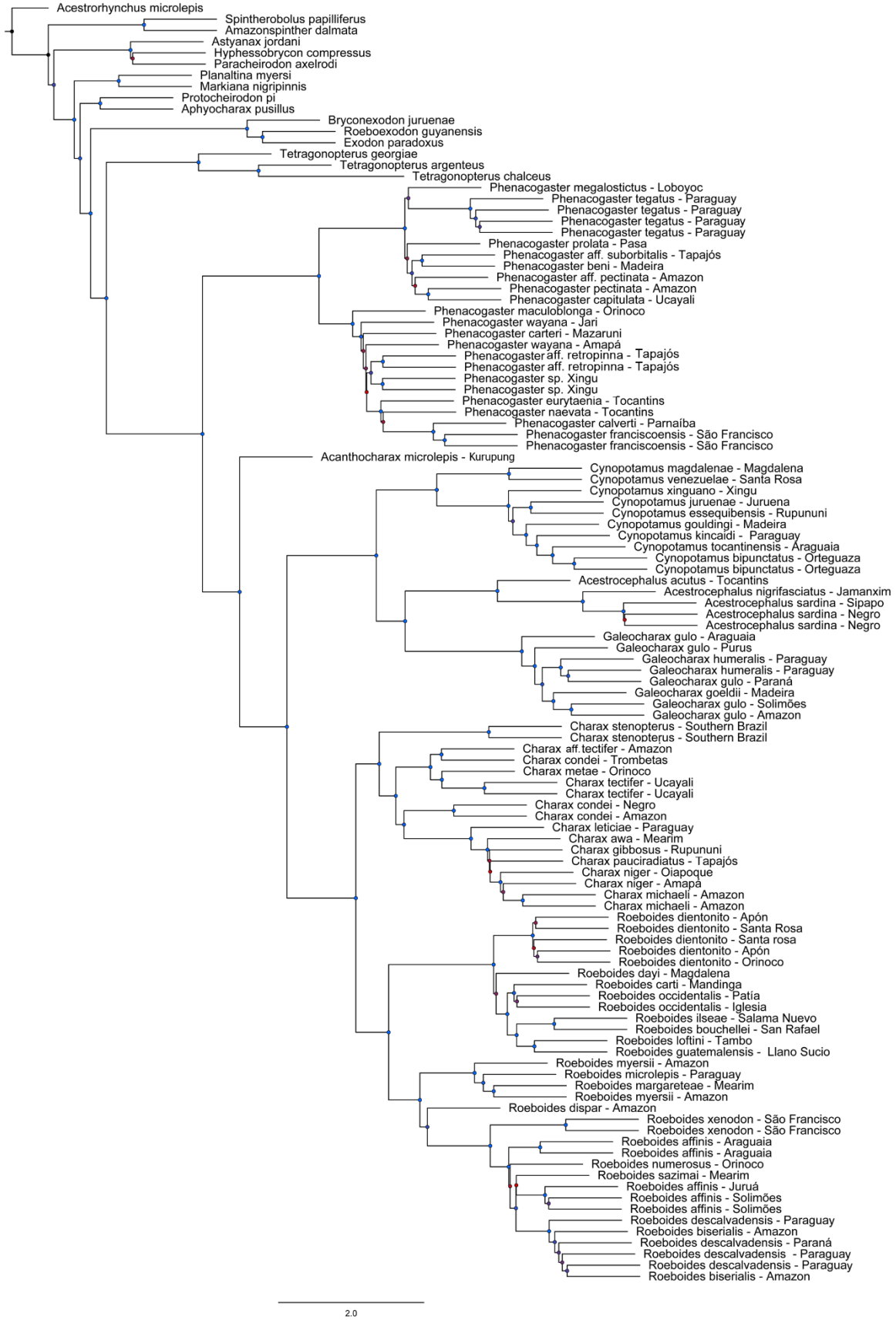
Supplementary Fig. 1. Maximum likelihood inference using 90% complete matrix. Only support values < 100% are shown.



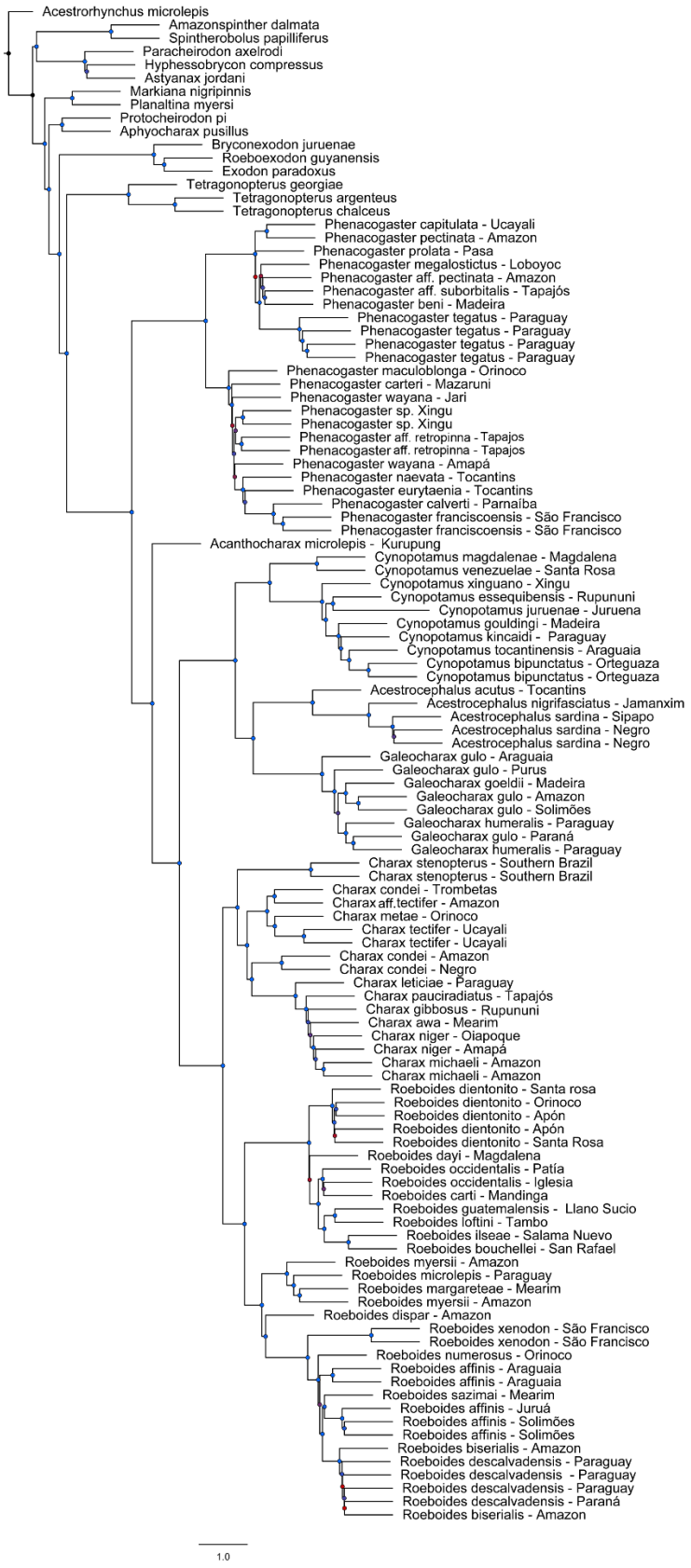
Supplementary Fig. 2. Phylogenetic hypothesis of Characinae based in Bayesian analysis using 90% complete matrix. Nodal support between 0.99-0.50 denoted by blue circle and support between 0.49-0.01 denoted by red circles.



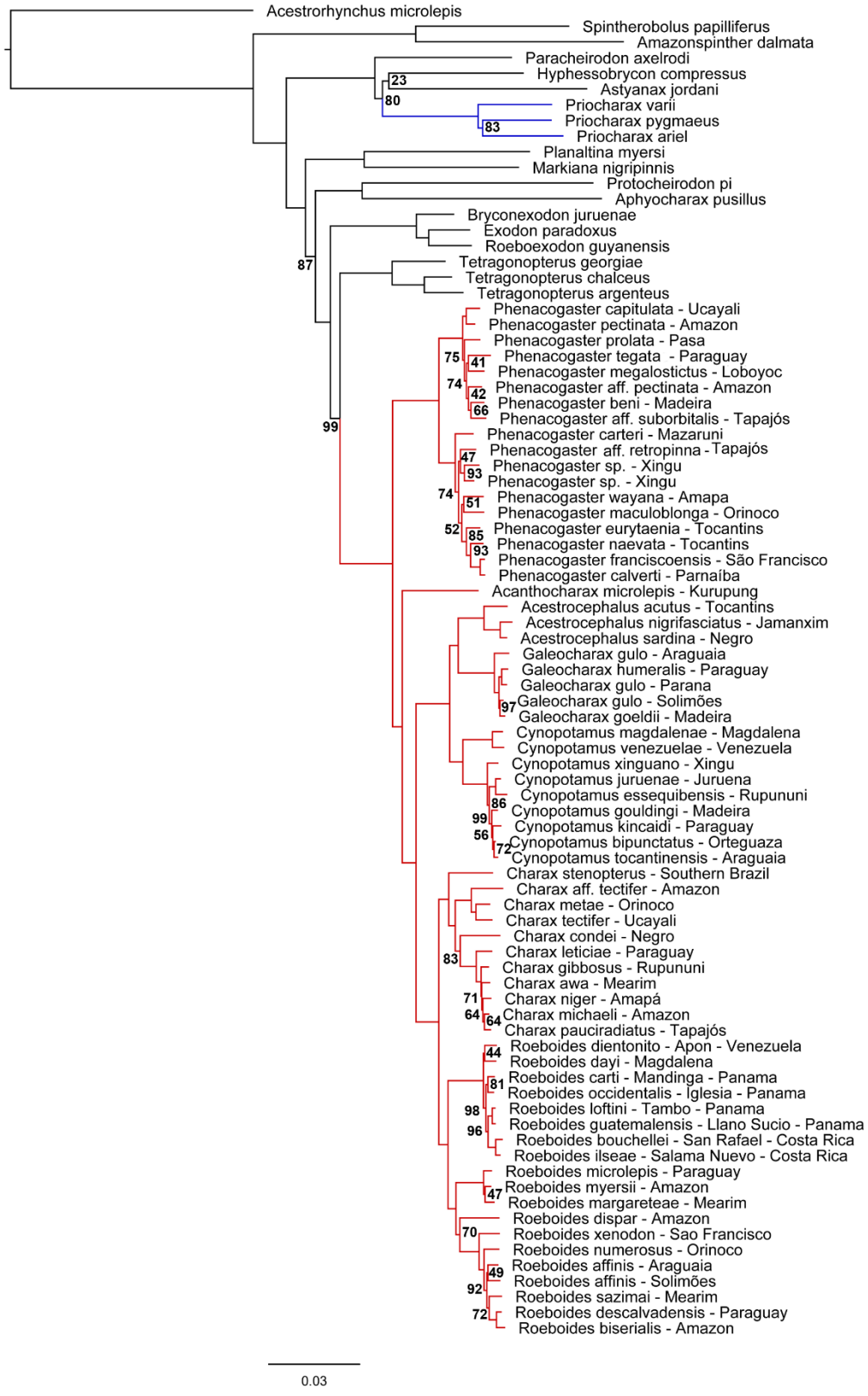
Supplementary Fig. 3. Phylogenetic hypothesis of Characinae based in Bayesian analysis using 75% complete matrix. Nodal support between 0.99-0.50 denoted by blue circle and support between 0.49-0.01 denoted by red circles.



Supplementary Fig. 4. Species tree inference from 90% complete matrix. Nodal support between 0.99-0.50 denoted by blue circle and support between 0.49-0.01 denoted by red circles.



Supplementary Fig. 5. Species tree inference from 75% complete matrix. Nodal support between 0.99-0.50 denoted by blue circle and support between 0.49-0.01 denoted by red circles.



Supplementary Fig. 6. Maximum likelihood inference using 95% complete matrix (only support values < 100% are shown). This result supports the relationship between *Priocharax* and *Hyphessobrycon* (Stethaprioninae) and other characids rather than with Characinae.

Chapter 2

Evolution and ecomorphology of the subfamily Characinae (Actinopterygii: Characiformes: Characidae)

Evolution and ecomorphology of the subfamily Characinae (Actinopterygii: Characiformes: Characidae)

Camila S. Souza

Departamento de Biologia Estrutural e Funcional, Instituto de Biociências, Universidade Estadual Paulista, R. Prof. Dr. Antônio C. W. Zanin 250, 18618-689, Botucatu, SP, Brazil

Abstract

The members of the species-rich characid subfamily Characinae occur widely throughout South and Central America. They are small to medium sized fishes, with the largest species not exceeding 240 mm in standard length. Most members of this group exhibit a deep anterior body with a characteristic gibbosity, though that feature is absent in *Phenacogaster* and *Acestrocephalus*. Though characins have been long-known to adopt three distinct feeding strategies: carnivory, omnivory and lepidophagy, no study has even reconstructed the evolutionary history of characin diet and its associated ecomorphology using modern methods. In this contribution, we link 2D and 3D morphometrics to a novel UCE phylogeny to reconstruct how diet and shape diversified across the phylogeny and determine whether diet predicts body and skull shape in this radiation of fishes. We generated phylomorphospaces using 153 lateral images of of 40 characin species and 147 micro-CT scans of skulls from 46 species of Characinae, tested for phylogenetic signal, reconstructed the evolutionary history of diet and shape, and used phylogenetic and non-phylogenetic ANOVAs to test for associations between diet, shape and size. All characters demonstrated significant phylogenetic signal. Omnivory is the probably the ancestral feeding state for Characinae, with carnivory also being a possibility. While nonphylogenetic ANOVA on 2D and 3D landmarks

indicated significant differences in body shape, body size, and skull shape among the three ecological groups, phylogenetic ANOVAs on size and the 2D landmarks were insignificant. Therefore, the body size and body shape differences among the three ecological groups could have been created by a random walk and are not necessarily the result of adaptation to different ecologies.

Keywords: Morphology, Diet, Body shape, Phylomorphospaces, ct-scan.

1. Introduction

Characiformes represents one of the most ecomorphologically diverse orders of fishes, with approximately 2,200 valid species (Fricke *et al*., 2022). Within this order the family Characidae is the most species-rich, with 1,238 species (Fricke *et al*., 2022). Perhaps because of its great diversity, Characidae includes some of the greatest taxonomic and systematic enigmas among Neotropical fishes (Javonillo *et al*., 2010; Mirande, 2010; Mirande, 2019; Oliveira *et al*., 2011). The morphology of the Characidae is highly conservative, with most members attaining small maximum body sizes of 10cm or less, and members of even the predatory genera rarely exceeding 20 cm (Reis *et al*., 2003; Mirande, 2018).

Characinae is the third most speciose subfamily of Characidae. It has great taxonomic relevance because it contains *Charax*, the type genus of the family and the order. Currently, all 85 species of Characinae are grouped into three tribes: Characini (*Charax* + *Roeboides*), Cynopotamini (*Acanthocharax* (*Cynopotamus* (*Acestrocephalus* + *Galeocharax*))), and Phenacogasterini (*Phenacogaster*) (*sensu* Mattox & Toledo-Piza, 2012; Fricke *et al*., 2022). Characins occur widely in rivers throughout the Neotropical region, including both sides of the Andes. In cis-Andean South America, they occur from the northern coastal drainages of Venezuela to tributaries of the Rio Paraguay and Rio Uruguay. Most members of the group are easily recognized by their deep anterior

bodies with a characteristic gibbosity, though this is absent in *Phenacogaster* and *Acestrocephalus* (Lucena & Menezes, 2003).

Members of Characinae have adopted three distinct feeding strategies: 41 carnivorous species, 23 omnivorous species and 21lepidophagous species (*sensu* Venere & Garutii, 2013; Mattox *et al*., 2017; Fricke *et al*., 2022). *Roeboides* represents the only characin genus e with lepidophagous feeding habits and possesses morphological specializations that aid this unusual feeding style, such as external mammiliform teeth (Sazima & Machado, 1983; Sazima, 1983). Several studies have examined ontogenetic changes in diet and such feeding morphologies within Characinae, mainly among species of *Roeboides* (Peterson & Winemiller, 1997; Peterson & McIntyre, 1998; Hahn *et al*., 2000; Novakowski *et al*., 2004; Bonato *et al*., 2017; Kolmann *et al.*, 2018b). However, these studies were based on a restricted taxonomic sampling and most were published before modern phylogenetic comparative methodologies became established. No study so far has examined the evolutionary history of dietary and ecomorphological diversification across the entire radiation of Characinae using such methods.

Herein, we apply phylogenetic ANOVA, tests of phylogenetic signal and ancestral state construction to conduct such an analysis in a phylomorphospace approach (Sidlauskas 2008). Such an approach provides a way of mapping the history of a clade's morphological diversification and inferring the magnitude and direction of shape change along any branch of the phylogeny. It has been used to investigate the correlation between shifts in morphology, habitat, and diet (Klingenberg & Ekau, 1996; Vidal-García & Keogh, 2017), patterns of body shape evolution along lineages (Benson & Choiniere, 2013; Serb *et al*., 2017), morphological divergence and convergence (Sakamoto & Ruta, 2012) and to visualize adaptive radiations (Arbour & López-Fernández, 2013). Though originally developed for use with 2D landmark data, recently 3D data generated through micro-CT scanning have been analyzed using phylomorphospaces (Bardua *et al*., 2019; Dollion *et al*., 2017; Evans *et al*., 2018; Kolmann *et al*., 2018a; Kolmann *et al*., 2018b; Vidal-García & Keogh,

2017). Micro-CT scanning produces high-resolution images of the internal anatomy of specimens while leaving them intact (Faulwetter *et al.*, 2013). The ambitious ScanAllFish project aims to scan every fish species in the world using the micro-CT scanning technique, thereby making images suitable for 3D phylomorphospace reconstruction widely available to the world's ichthyologists (Project: oVert: UW - CT Scan all Fishes: https://www.morphosource.org/Detail/ProjectDetail/Show/project_id/220).

The phylogenetic comparative analysis herein links 3D skull shape data generated as part of the ScanAllFish project to 2D body shape data, a dietary classification and a robust phylogeny to understand the morphological and trophic evolution of Characinae. In particular, we sought to reconstruct how diet evolved within Characinae, and determine whether dietary shifts influenced the morphological diversification of braincase shape, body shape, and body size in this subfamily of fishes.

2. Materials and Methods

2.1. Taxon sampling

We examined of 153 specimens acquired from museums (Table 1) representing 40 out of 85 species of Characinae (47%) plus the outgroup taxa *Tetragonopterus chalceus* and *T. argenteus*. Taxon sampling averaged 39% of the carnivorous species (1- *Acanthocharax microlepis*; 1- *Acestrocephalus*; 7- *Charax*; 5- *Cynopotamus*; 2- *Galeocharax*), 66% of the lepidophagus species (14-*Roeboides*); and 43% of the omnivorous species (10- *Phenacogaster*) (Table 1) within Characinae. The number of individuals measured per species ranged from 2 to 8 individuals.

Table 1. Species examined and their respective catalog numbers, number of specimens, maximum length (cm) and dietary classification.

Species	Catalog number	Specimens	Maximum length (cm)	Diet	Source
Acanthocharax microlepis	FMNH 53666	5	8.5	Carnivore	Mattox et al. 2017
Acestrocephalus sardina	ANSP192399/ AUM43599/AUM44032	8	13.5	Carnivore	Mattox et al. 2017
Charax pauciradiatus	LBP14117	4	10.1	Carnivore	Mattox et al. 2017
Charax gibbosus	ANSP135634	3	12.5	Carnivore	Mattox et al. 2017
Charax leticiae	LBP8771	4	10	Carnivore	Venere & Garutii, 2013
Charax metae	FMNH55145/LBP18734	5	11	Carnivore	Mattox et al. 2017
Charax michaeli	FMNH100336	4	13	Carnivore	Mattox et al. 2017
Charax stenopterus	LBP14529/LBP17020/LBP13020/LBP13169	4	9.4	Carnivore	Mattox et al. 2017
Charax tectifer	ANSP130523	4	10	Carnivore	Mattox et al. 2017
Cynopotamus bipunctatus	ANSP192522/ANSP198968/ANSP191144/ANSP128733	4	17.5	Carnivore	Venere & Garutii, 2013
Cynopotamus essequibensis	ANSP176793/ANSP189604/ANSP189604/ANSP175515	4	16	Carnivore	Mattox et al. 2017
Cynopotamus kincaidi	LBP19/LBP12638	2	17.4	Carnivore	Mattox et al. 2017
Cynopotamus magdalenae	USNM310473	2	40	Carnivore	Mattox et al. 2017
Cynopotamus gouldingi	ANSP 180805	2	16.5	Carnivore	Mattox et al. 2017
Galeocharax gulo	LBP2556/LBP13302	4	22	Carnivore	Mattox et al. 2017
Galeocharax humeralis	LBP13453	4	13.7	Carnivore	Mattox et al. 2017
Phenacogaster capitulata	LBP17802	4	3.5	Omnivore	Venere & Garutii, 2013
Phenacogaster eurytaenia	LBP25604	4	4.59	Omnivore	Venere & Garutii, 2013
Phenacogaster franciscoensis	LBP24038	4	4.2	Omnivore	Venere & Garutii, 2013
Phenacogaster megalostictus	LBP131826	3	3.7	Omnivore	Venere & Garutii, 2013
Phenacogaster naevata	LBP9385	4	3.45	Omnivore	Venere & Garutii, 2013
Phenacogaster pectinata	LBP22715	2	4.5	Omnivore	Venere & Garutii, 2013
Phenacogaster prolata	LBP18683	4	4.99	Omnivore	Venere & Garutii, 2013
Phenacogaster sp.	LBP16061	4	3.54	Omnivore	Venere & Garutii, 2013
Phenacogaster tegata	FMNH108487	8	3.2	Omnivore	Venere & Garutii, 2013
Phenacogaster wayana	ANSP189599	4	4.68	Omnivore	Venere & Garutii, 2013
Roeboides affinis	LBP22638	4	11	Lepidophage	Mattox et al. 2017
Roeboides bouchellei	FMNH128131	4	8.2	Lepidophage	Mattox et al. 2017
Roeboides carti	USNM311020	5	8.3	Lepidophage	Mattox et al. 2017
Roeboides dayi	FMNH85335	4	10.9	Lepidophage	Mattox et al. 2017
Roeboides descalvadensis	LBP5110	4	8.9	Lepidophage	Mattox et al. 2017
Roeboides dientonito	LBP6114	4	6.8	Lepidophage	Mattox et al. 2017
Roeboides dispar	ANSP180932	2	8.11	Lepidophage	Mattox et al. 2017
Roeboides guatemalensis	LBP2755	4	13	Lepidophage	Mattox et al. 2017
Roeboides ilseae	ANSP163737/ANSP164256	3	7.6	Lepidophage	Mattox et al. 2017

Species	Catalog number	Specimens	Maximum length (cm)	Diet	Source
Roeboides margareteae	ANSP95879/ANSP82289	2	18.47	Lepidophage	Mattox et al. 2017
Roeboides myersii	ANSP178154	3	15.77	Lepidophage	Mattox et al. 2017
Roeboides numerosus	ANSP196316	2	6.1	Lepidophage	Mattox et al. 2017
Roeboides occidentalis	ANSP99843	4	13	Lepidophage	Mattox et al. 2017
Roeboides xenodon	LBP10444	4	8.2	Lepidophage	Mattox et al. 2017
Tetragonopterus argenteus	OS 18365	3	12.7	Omnivore	Mattox et al. 2017
Tetragonopterus chalceus	OS 18616	4	11.2	Omnivore	Mattox et al. 2017

2.2. 2D landmarks

2.2.1. Body shape

We captured the body shape of each individual specimen and species by placing 2D landmarks on images obtained using a digital camera (Canon D90) equipped with a 60mm macro lens and following the phototank immersion method described in Sabaj Pérez (2009). We digitized each image using 15 landmarks and three curves (Fig. 1) located on lateral photographs using TPSDig 2 (Rohlf, 2017). The curves were decomposed into three series of semilandmarks in tpsUtil32 (Rohlf, 2015) to provide information along curves of the body that presented no obvious homologous landmarks that could be located in all species. The 50 total semilandmarks include 23 placed along the outline of the head between the tip of the snout and supraoccipital spine, 20 between the supraoccipital spine and the dorsal-fin origin and seven between the insertion and origin of the anal fin. We checked for possible digitization errors, and verified that we located the same number of landmarks for all samples by uniting data from all specimens of a species in tpsUtil32 (Rohlf, 2015) and then performing a Procrustes superimposition and relative warps analysis for each species in tpsRelw (Rohlf, 2003). After verifying that there were no large morphological outliers that represented digitization errors, we exported the consensus configuration for each species, and combined all the species consensus into a final TPS file in tpsUtil.

We constructed body shape morphospaces by relative warp analysis of that TPS file in tpsRelw, and also via principal components analysis in the geomorph package (v. 2.1.3) (Adams & Otárola-Castillo, 2013) in the R computing environment (R Core Development Team 2013). Semilandmarks were treated as such in all these analyses through the constrution of a "sliders" file, and a "links" file was constructed to aid visualization of landmark configurations. We also converted the TPS files into NTS format (another text file used for landmark coordinates and other variables by some programs).

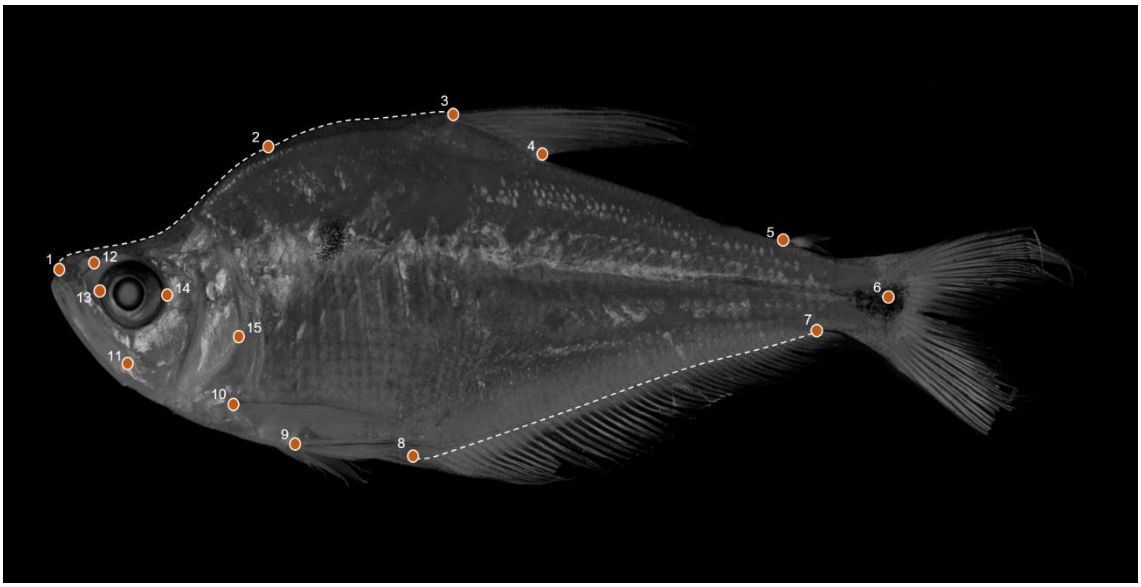


Figure 1: Lateral view of *Charax pauciradiatus* LBP14117, (110.99 mm SL). Landmarks represent: (1) anterior tip of the premaxilla; (2) posterior extreme of the supraoccipital spine; (3) dorsal-fin origin; (4) dorsal-fin insertion; (5) adipose-fin origin; (6) posterior extent of vertebral column and anterior of hypural plate; (7) anal-fin insertion; (8) anal-fin origin; (9) pelvic-fin origin; (10) pectoral-fin origin; (11) posterior extreme of maxilla; (12) posterior margin of posterior naris; (13) anterior margin of orbit along longest axis; (14) posterior margin of orbit along longest axis; (15) posterior extreme of opercle.

2.2.3. Body size

Maximum body size (MBS) of each species was obtained from data available in the literature (Reis *et al.*, 2003; Menezes, 2006, 2007; Lucena & Malabarba, 2010; Antonetti *et al.*, 2018). Those data appear in Table 1.

2.2.4. Ecological characterization

We identified and characterized the three distinct feeding strategies (carnivory, omnivory and lepidophagy) in the species of Characinae from data available in the literature, including species descriptions, field guides and checklists (Reis *et al.*, 2003, Menezes, 2006, 2007, Lucena & Malabarba, 2010, Venere & Garutii, 2013, Mattox *et al.*, 2017) (Table 1).

2.3. Phylogeny

We linked the morphometric and ecological data to a phylogeny of subfamily Characinae based on ultraconserved elements (maximum likelihood analysis of the 75% complete matrix of ultraconserved elements; 1,300 loci; 586,785 bp) (Souza et al., in review; Cap. 1). Using the drop.tip command from the APE package in R (Paradis & Schliep, 2019), we pruned out ingroup species for which ecomorphological data were not available. As outgroup taxa we retained *Tetragonopterus chalceus* and *T. argenteus*. We ultrametricized the phylogeny via penalized likelihood using the default settings of the chronos routine in APE.

2.3.1 Phylogenetic comparative methods

- *Ancestral state estimation*

Ancestral dietary shifts were determined using maximum likelihood (ML) and Bayesian inference using the ace command in APE, and the implementation of SIMMAP (Bollback, 2006) (*nsim*=100) in the R package phytools (Revell, 2012). To visualize the pattern of body size evolution we used the contMap and phenogram95 functions in phytools (Revell, 2012). These functions map a continuous character on to a phylogenetic tree by estimating ancestral states at each node using fastAnc (fast estimation of ML ancestral states).

- *Tests of phylogenetic signal*

To estimate the phylogenetic signal of the body size data we used K statistic of Blomberg *et al*. (2003) as implemented with the function *phylosig in* phytools (Revell, 2012). That statistic estimates the strength of phylogenetic signal in a dataset relative to that expected under a Brownian motion model of evolution. For the body shape data, we estimated phylogenetic signal using Kmult, the multivariate generalization of Bloomberg's K (Adams, 2014) as implemented in the *physignal* routine within the R package Geomorph (Adams & Otárola-Castillo, 2013). Kmult is a generalization of Blomberg's K found from the equivalency between statistical methods based on covariance matrices and those based on distance matrices (Adams, 2014). The advantage of this approach is that, in addition to inferring whether there is a small or large amount of signal present in the data, it provides a reference value for the degree to which the dataset departs from the Brownian motion model of evolution (Diniz-Filho *et al.,* 2012).

- *Phylomorphospace construction*

To visualize whether and how each ecological group and clade differs in body shape we constructed phylomorphospaces (Sidlauskas, 2008) using the "phylomorphospace" function in the R package "phytools" (Revell, 2012) and the *gm.prcomp* function in geomorph (Adams & Otárola-Castillo, 2013).

- *Phylogenetic and Nonphylogenetic ANOVA and MANOVA*

To test for significant morphological differences in body size and body shape among the three ecological groupings, we perfomed phylogenetic and nonphylogenetic ANOVA and MANOVA. Non-phylogenetic approaches used the *aov* function from the core R package for the body size data and the procD.lm routine in geomorph (Adams & Otárola-Castillo, 2013) for the body shape data. Phylogenetic ANOVA on the body size data used the *aov.phylo* routine in the geiger package (Harmon *et al*., 2008), and phylogenetic MANOVA on the body shape data used the phylogenetic least squares approach implemented by *procD.pgls* within geomorph (Adams & Otárola-Castillo, 2013).

2.4. 3D landmarks

2.4.1. Skull Shape

To test whether each ecological group (carnivory, omnivory and lepidophagy) differs in braincase morphology, we micro-CT scanned 147 specimens from museum or personal collections representing 45 species of Characinae (Supplementary Table 1), many of which are also represented in the body shape dataset. Because multiple individuals of a single species were needed to examine potential allometric shape changes in the the skull we reduced this dataset to just the species for which we were able to scan between 2 to 8 individuals (Table 2). The final CT dataset includes 108 specimens representing 26 species. Scanning was performed using the Bruker Skyscan 1173 at the Karel F. Liem Bio-Imaging Center at University of Washington Friday Harbor Labs (UW). We reconstructed the resulting image stacks using NRecon (Bruker microCT, Kontich, Belgium, 2016) and isolated individual fish from each batch in DataViewer 2.1 (Bruker, Kontich, Belgium, 2010). We converted these image stacks to DICOM file format for viewing and segmentation in the computer program Horos v2.0.1 (The Horos Project, 2015; http://www.horosproject.org/) and

CTVox 2.7 software (Bruker Corp., Billerica, MA). All microCT scans will be freely available for download from Morphosource (Boyer *et al.*, 2016; morphosource.org).

Table 2- Species, sample sizes and catalogue numbers. Note that many museum lots contain multiple individuals.

Species	Number of Individuals	Catalogue number
Acanthocharax microlepis	5	FMNH 53666
Acestrocephalus sardina	8	ANSP 192399/ AUM 44032/AUM 43599
Charax leticiae	4	LBP 8771
Charax metae	2	LBP 18734/ FMNH 55145
Charax michaeli	4	LBP 22517/ FMNH 100336
Charax pauciradiatus	4	LBP 14117
Charax stenopterus	4	LBP 17020/LBP 14529/LBP 13020/LBP 13169
Charax tectifer	3	ANSP 180523
Charax gibbosus	3	ANSP 135634
Cynopotamus essequibensis	3	ANSP 176793/ANSP 189604
Galeocharax humeralis	3	LBP13453
Phenacogaster sp	3	LBP 16061
Phenacogaster eurytaenia	4	LBP 25604
Phenacogaster franciscoensis	4	LBP 24038
Phenacogaster megalostictus	4	ANSP 131826
Phenacogaster naevata	4	LBP9385
Phenacogaster pectinata	3	LBP22715
Phenacogaster prolata	4	LBP 18683
Phenacogaster wayana	4	ANSP 189600/ ANSP 189599
Roeboides affinis	3	LBP22638
Roeboides descalvadensis	4	LBP5110
Roeboides dientonito	6	LBP6114/ ANSP 206220
Roeboides guatemalensis	3	LBP2755
Roeboides xenodon	4	LBP10444
Roeboides dayi	5	FMNH 85335
Roeboides myersii	3	ANSP 178154

2.4.2. Skull shape geometric morphometrics and effect of diet on skull shape

Specimens were digitized in three dimensions with 21 landmarks placed around the neurocranium, of which 16 were bilaterally symmetric and 5 fell along the sagittal plane (Table 3; Fig. 2) using 3D Slicer 4.10.2 (Kikinis *et al.*, 2014; Buser *et al.*, 2020). A principal components analysis (PCA) identified the principal axes of skull shape variation among the data. In the morphospace resulting from PCA, species were plotted using colors that correspond to their ecological group (carnivory, omnivory and lepidophagy). To test for significant skull shape variation among the three ecological groupings, we perfomed a non-phylogenetic Procrustes MANOVA. All analysis were performed using the geomorph (Adams & Otárola-Castillo, 2013) and morpho (Schlager *et al.*, 2019, R package v2.7) packages in R.

Table 3- Landmark definitions for the three-dimensional geometric morphometric analysis of skull shape.

Landmark	# Landmark description
1	Mesethmoid-anterior tip
2	Mesethmoid-dorsal left margin
3	Mesethmoid-dorsal right margin
4	Anterior margin of the frontal foramen (left side)
5	Posterior margin of the frontal foramen (right side)
6	Frontal-anteriormost segment of posterior fontanel
7	Anterior of supraoccipital crest
8	Supraoccipital-anterior right margin
9	Supraoccipital-anterior left margin
10	Tip of supraoccipital crest
11	Supraoccipital-posterior left margin
12	Supraoccipital-posterior right margin
13	Posterior of supraoccipital crest
14	Supracleithrum left margin
15	Supracleithrum right margin
16	Anterior limit between frontal and sphenotic (left side)

Landmark	# Landmark description
17	Anterior limit between frontal and sphenotic (right side)
18	Frontal-anterior left margin
19	Frontal-anterior right margin
20	Lateral ethmoid (left side)
21	Lateral ethmoid (right side)

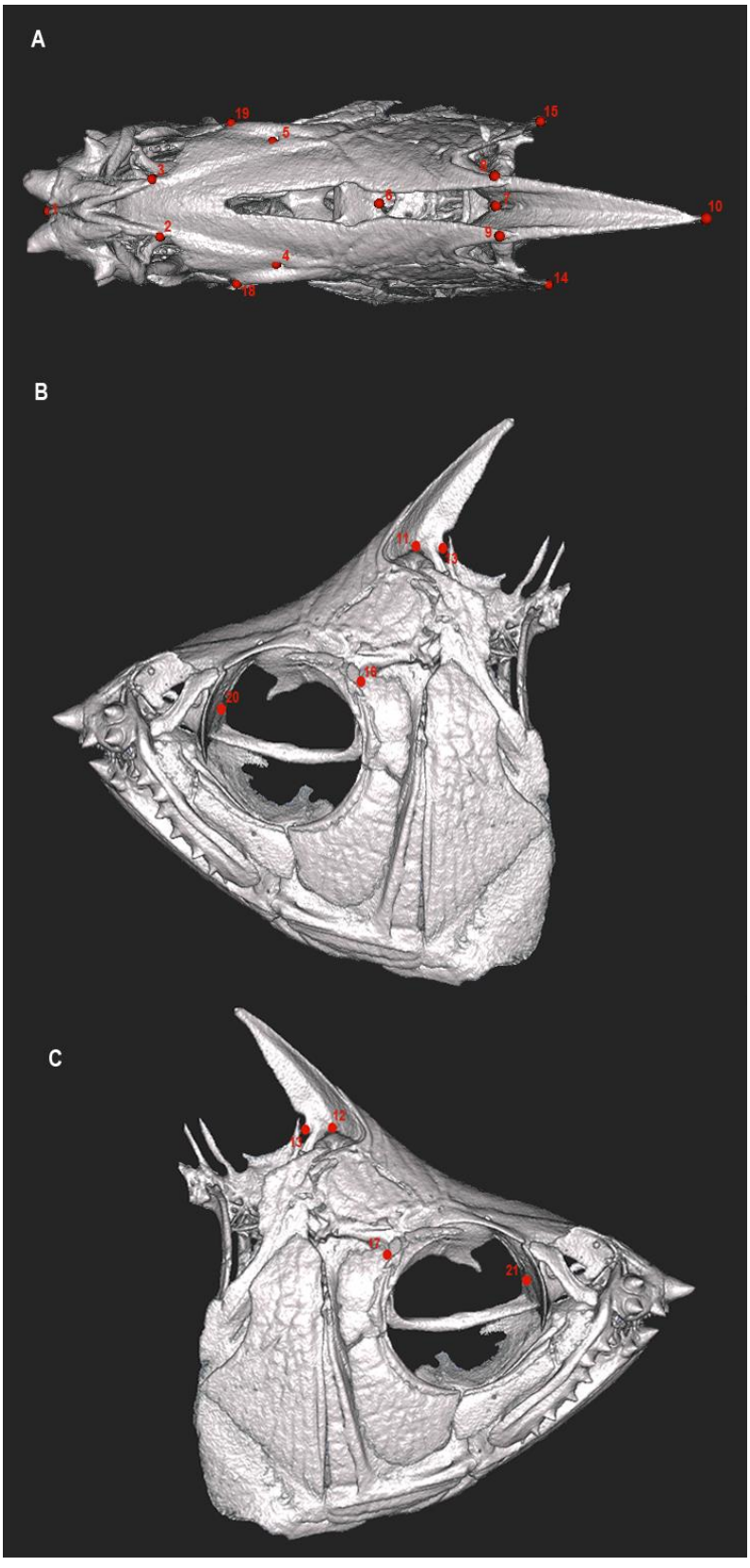


Figure 2: Three-dimensional skull surface rendering of *Roeboides numerosus* showing the location of 21 landmarks. A- dorsal view; B and C lateral views.

3. Results

3.1. 2D landmarks

Ancestral state estimation

The maximum likelihood (ML) and SIMMAP reconstructions inferred that general omnivory was probably the ancestral feeding state for Characinae (Fig. 3-4), with carnivory also being a possibility. Assuming that the most recent common ancestor of Characinae was indeed omnivorous, *Phenacogaster* was the only group to maintain that ancestral feeding state. A transition to carnivory occurred on the branch leading to the clade containing *Acanthocharax, Charax, Roeboides, Cynopotamus, Galeocharax* and *Acestrocephalus*. Lepidophagy was the last ancestral feeding state to evolve, and evolved along the branch leading to *Roeboides*, which are the only characin species that specialize on eating scales.

The reconstruction of trends in body size (Fig. 5) revealed an evolutionary transition toward smaller body sizes in the lineage leading to *Phenacogaster* and an overall increase in body sizes in the lineage leading to the clade containing *Cynopotamus, Acestrocephalus* and *Galeocharax.* That increase was followed by a transition back to smaller bodies in the lineage leading to *A. sardina* and a further increase in *Cynopotamus magdalenae*, which is substantially larger than any other member of the subfamily. An additional increase in body size occurred in the lineage leading to the species pair *Roeboides myersii* and *R. margareteae*.

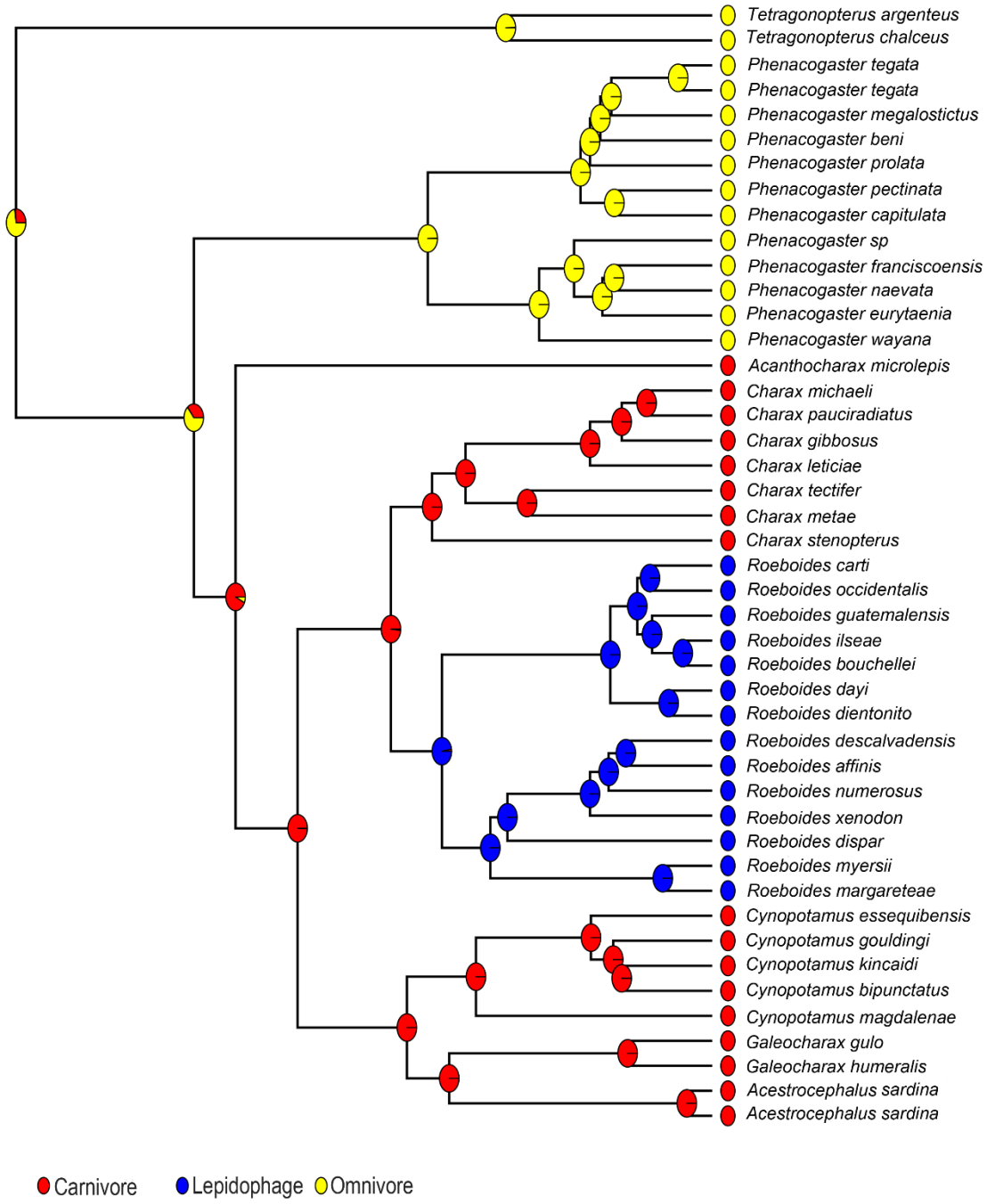


Figure 3. Reconstruction of the evolutionary history of diet of Characinae. Piecharts at nodes depict the relative probability of each diet in that ancestor.

Figure 4. SIMMAP cladogram illustrating ancestral state reconstruction of diet for the subfamily Charinae. Branch colors correspond to diet (Yellow: omnivore; Red: carnivore; Blue: lepidophage).

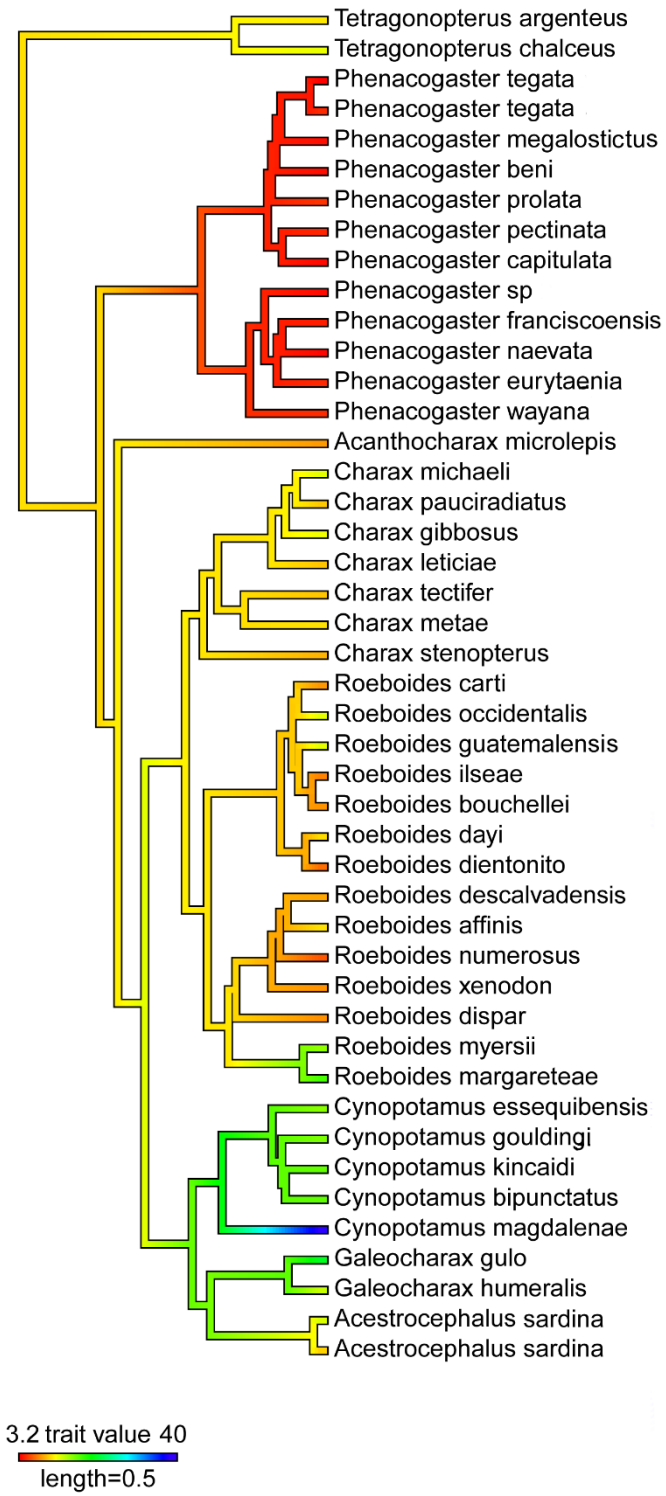


Figure 5. ContMap cladogram illustrating estimated ancestral maximum body size (MBS). Branch colors correspond to estimates of ln MBS (i.e., maximum standard body length measured from tip of snout to base of caudal fin), with red indicating smaller and green or blue indicating larger MBS.

Tests of phylogenetic signal

We estimated a K-statistic of 1.07 for the body shape data, matching the expectation of the Brownian motion model and representing a statistically significant level of phylogenetic signal (P=0.001) relative to the permuted null. The K statistic for the body size data of 0.848 also represents significant phylogenetic signal (P=0.002) and agrees closely with the Brownian expectation.

Body shape phylomorphospace

The first principal component (PC1) axis of the body shape phylomorphospace explained

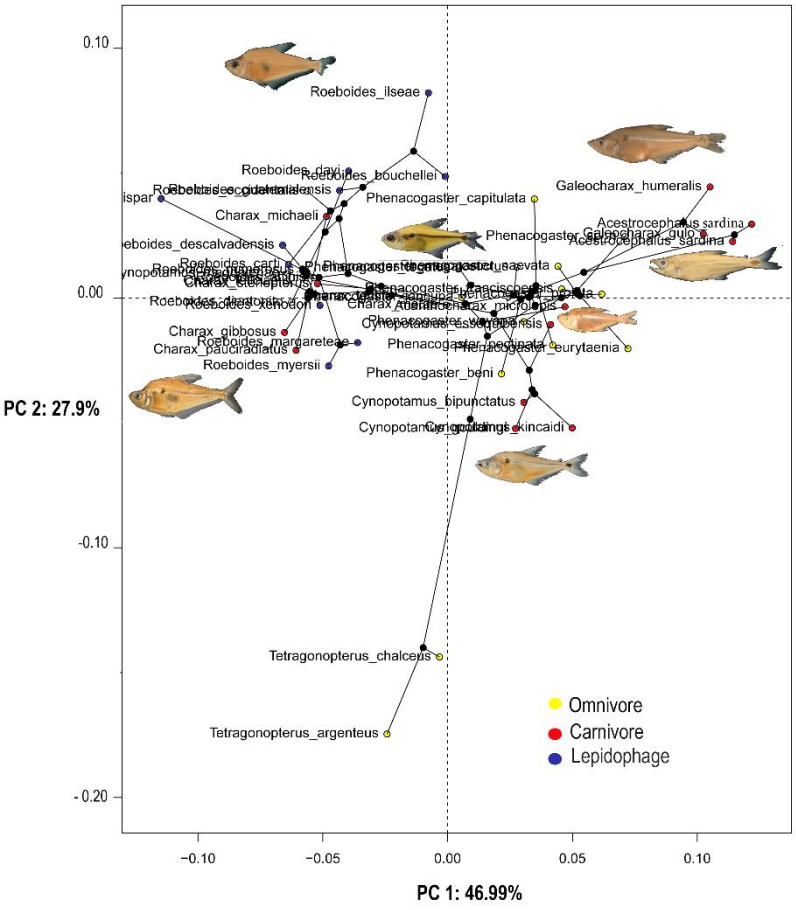


Figure 6.
Phylomorphospace plot for the subfamily Characinae in which terminal colors correspond to diet.

46.9% of variance in body shape among characin species, while the second explained an additional 27.9% (Fig. 6). PC1 illustrates variance between strongly concave dorsal profiles (negative values) and slightly concave or straight dorsal profiles (positive values). Variation along the PC2 axis describes the differences between elongate (positive values) and foreshortened body shapes (Fig. 6), and largely reflects the morphological differences between the outgroup (*Tetragonopterus*) and the ingroup species. Lepidophagous and omnivorous characins each have a relatively conserved morphology, and thus each trophic class occupies a confined region of morphospace. Lepidophages possess strongly concave dorsal profiles and omnivores slightly concave or straight dorsal profiles, and thus are well-separated along PC1. Unlike lepidophages and omnivores, the carnivorous characins possess three distinct morphologies in morphospace, one of which matches the lepidophagous morphology and another of which overlaps with the omnivorous morphology. A third carnivorous morphology lies at the positive extreme of PC1, and demonstrates a slightly concave dorsal shape and elongate body.

Phylogenetic and nonphylogenetic MANOVA and ANOVA

The nonphylogenetic MANOVA-shape ($F = 8.2599$, $p = 0.001$) and ANOVA- size ($F = 15.08$, $p = 1.16e-05$) was highly significant and showed that the dietary groupings differ in body shape and body size. However, the results of the phylogenetic MANOVA on shape ($F = 0.8269$, $p = 0.592$) and of the phylogenetic ANOVA on size ($F = 12.687$, $p = 0.3267$) were not significant. This result shows that body size and body shape differences among the three ecological groups are consistent with a random walk, and that the morphological similarity among species with similar ecologies may result from their close evolutionary relationship rather than from selection for an optimal phenotype.

3D landmarks

Skull shape geometric morphometrics and effect of diet on skull shape

The PCAs showed that the first two PC axes summarize the majority of shape variance (38.15% and 17.82%, respectively) (Fig. 7). Along PC1 axis, carnivores and lepidophages vary widely in skull shape, while omnivores occupy a confined portion of the diagram (Fig. 7).

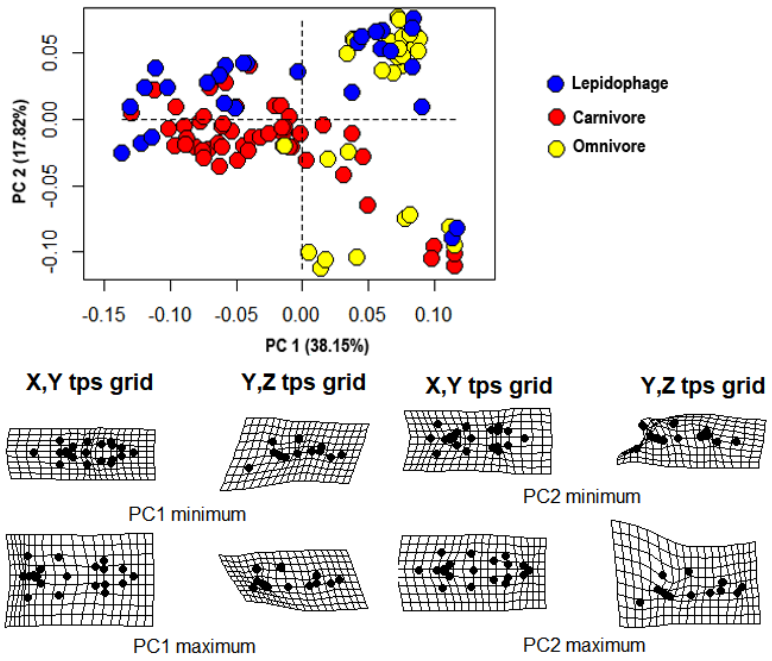


Figure 7. PCA values on skull shape variation based on species means, using the R package geomorph. Thin-plate spline (TPS) deformation grids are displayed to indicate variation on skull shape.

The variation is mainly associated with the anterior-posterior lengthening (negative values) and shortening (positive values) of the skull (Fig. 3). The second axis of variation (PC 2) corresponded to the width of the skull, with negative values representing wider skulls, and positve values represeting narrower skulls. All three trophic groups vary widely on PC2 (Fig. 7). The results of the

MANOVA showed that there was a significant difference in skull shape between at least two of the ecological groups ($F = 10.122$, $p = 0.001$, likely driven by the restriction of omnivores to positive or near-zero values of PC1).

4. Discussion

Diet appears to have evolved infrequently during the characin radiation. Assuming that the most recent common ancestor of Characinae was indeed omnivorous as suggested by both methods of ancestral state reconstruction, only characins of the genus *Phenacogaster* retained the ancestral diet. The prevalance of carnivory among the modern species most likely reflects a single transition away from omnivory, rather than multiple origins of that diet. Similarly, only a single transition to lepidophagy occurred within Characinae, in the lineage leading to *Roeboides*.

The body size evolution showed does not have an evolutionary tendency towards a general increase or decreases in size. Although, it is possible observed that the larger sizes prevail in most carnivorous fish. The hierarchy of increasing animal body size with increasing trophic level has been broadly accepted since Charles Elton's work in the early 1900s (Ou *et al*., 2017). Also, body size and trophic position are important and often ecologically linked traits in fishes (Bloom *et al.*, 2018). However, our results showed the ancestral state reconstruction of the diet based on a restricted phylogeny of Characinae with only one *Tetragonopterus* outgroup. In order to better conclude the evolution of the diet in Characinae is necessary to use on outgroup a wide range of Characiformes species.

Lepidophagous and omnivorous characins demonstrate distinctive morphologies, suggesting at first glance the existence of a link between diet and shape. Those distinctive morphologies drive the strong significance of the nonphylogenetic MANOVA/ANOVAs of body shape and size. However, phylogenetic versions of these analyses yielded insignifiant results. That lack of

significance reflects the strong phylogenetic signal in these data, and the scarcity of transitions among the ecological character states, and indicates that the data cannot reject the hypothesis that body size and body shape evolved as a random walk along the phylogeny. Such a random process can produce a wide variety of morphological patterns including unequal morphological diversity, apparent stasis, and clustering in morphospace (Raup & Gould 1974; Raup 1977; Foote 1996; Pie & Weitz 2005). Thus the morphological similarity among species with similar ecologies may result from their close evolutionary relationship rather than from selection for an optimal phenotype.

The substantial overlap in body shape and skull shape morphospace between species with different ecologies also suggests that even if a link between diet and morphology does exist within Characinae, it may be relatively weak. For example, the sister groups *Charax* and *Roeboides* share the same body shape, despite the former being carnivorous and the latter specializing on scales. Members of all three trophic groups overlap in skull shape, and while omnivores are absent from half of that morphospace, all omnivorous species share a fairly recent common ancestor and that restriction could easily also result from their shared ancestry. The influence of random processes may indeed provide the best explanation for the diversification of characin morphology.

Nevertheless, it is important to recognize that the inability to reject a random null hypothesis does not mean that the null hypothesis is true. Studies have demonstrated that body shapes can be influenced by many factors, such as ecological interactions, biomechanical constraints and natural selection (Wake & Larson, 1987; Gould, 2002; Adams & Nistri, 2010). The insignificant results herein may therefore reflect an ecological classification that does not capture the factors that determine characin morphology, particulary if diet is not primary. Alternatively, the morphological datasets may not fully capture the aspects of shape that respond most strongly to shifts in diet., such as differences in dentition or in jaw mechaniscs. Certainly, *Roeboides* species appear to have evolved specialized strategies and structures optimized for lepidophagy, such as the tooth and jaw characteristics that distinguish them from the carnivorous members of *Charax* species (carnivores)

(Kolmann *et al.* (2018b)), and a formal statistical analysis of those character systems would have likely revealed a more definitive separation among the ecological groupings than was apparent in the analysis of the general shapes of the body and skull.

Finally, the analysis herein may simply lack the statistical power needed to reject the random hypothesis, even if a true relationship between diet and body or skull shape does exist. Burns & Sidlauskas (2019) showed recently that half of all major shifts in characiform body shape coincide with changes in trophic niches, suggesting that dietary diversification sometimes drives characiform body evolution, or vice versa. Thus, that relationship may hold within Characinae, even if we lack the statistical power to detect it. Only a single transition exists among each dietary class within Characinae, and thus the dataset does not afford the ability to determine whether repeated transitons to the same ecology produce the same morphologies. Indeed, Burns & Sidlauskas (2019) in their a much larger analysis failed to recover a statistically significant signal of convergence in body shape among detrititivorous characiforms, despite a clear clustering of those species in morphospace and the presence of three independent origins of the ecology. The only significant convergence in that study occurred among predators, which represented one of the most frequently evolving ecologies among any exhibited within the order.

Thus, to fully unlock the reasons for the substantial morphological diversification of characins, future studies should measure those morphologies at a broader taxonomic and with greater ecological precision. As future studies permit more precise ecological characterization of each characiform species, and as the phylogeny of Characiformes expands to include more taxa, we anticipate substantial improvement in our ability to detect a link between ecology and shape within the subfamily Characinae, if such a link does in fact exist.

Acknowledgements

We are grateful to David C. Werneke (AUM), Susan Mochel (FMNH), Diane Pitassy (USNM), Tammes Menne (ZMUC), Mark Sabaj (ANSP) and Mariangeles Arce (ANSP) for curatorial assistance and/or loan of vouchers. We are also thankful to Thaddaeus Buser, Álvaro Cortés, e Jonathan Huie for advices and support with analyses.

Funding

This research was supported by the Brazilian agencies FAPESP proc. 17/06551-0 and 2019/04602-1 (CSS), 16/11313-8 and 18/23883-9 (BFM), 18/20610-1, 16/09204-6, 14/26508-3, CNPq proc. 200159/2020-8 (BFM), proc. 306054/2006-0 (CO), and the Stiassny AMNH Axelrod Research Curatorship of the American Museum of Natural History (BFM).

Bibliography

Adams, D. C., & Otárola-Castillo, E. (2013). geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods in Ecology and Evolution, 4(4), 393-399.

Adams, D. C. (2014). A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. Systematic biology, 63(5), 685-697.

Adams, D. C., & Nistri, A. (2010). Ontogenetic convergence and evolution of foot morphology in European cave salamanders (Family: Plethodontidae). BMC Evolutionary Biology, 10(1), 216.

Antonetti, D.A., Lucena, C. A. S., Lucena, Z. M. S. (2018). Two new species of *Phenacogaster* (Actinopterygii: Characidae) from the rio Tocantins drainage, Brazil. Ichthyological Exploration of Freshwaters, 28(3):271-80.

Arbour, J. H., & López-Fernández, H. (2013). Ecological variation in South American geophagine cichlids arose during an early burst of adaptive morphological and functional evolution. Proceedings of the Royal Society B: Biological Sciences, 280(1763), 20130849.

Bardua, C., Wilkinson, M., Gower, D. J., Sherratt, E., & Goswami, A. (2019). Morphological evolution and modularity of the caecilian skull. BMC evolutionary biology, 19(1), 30.

Benson, R. B., & Choiniere, J. N. (2013). Rates of dinosaur limb evolution provide evidence for exceptional radiation in Mesozoic birds. Proceedings of the Royal Society B: Biological Sciences, 280(1768), 20131780.

Bollback, J. P. (2006). SIMMAP: stochastic character mapping of discrete traits on phylogenies. BMC bioinformatics, 7(1), 88.

Bonato, K. O., Burress, E. D., & Fialho, C. B. (2017). Dietary differentiation in relation to mouth and tooth morphology of a neotropical characid fish community. Zoologischer Anzeiger, 267, 31-40.

Boyer, D. M., Gunnell, G. F., Kaufman, S., & McGeary, T. M. (2016). Morphosource: Archiving and sharing 3-d digital specimen data. The Paleontological Society Papers, 22, 157-181.Bloom, D. D., Burns, M. D., & Schriever, T. A. (2018). Evolution of body size and trophic position in migratory fishes: a phylogenetic comparative analysis of Clupeiformes (anchovies, herring, shad and allies). Biological Journal of the Linnean Society, 125(2), 302-314.

Bloom, D. D., Burns, M. D., & Schriever, T. A. (2018). Evolution of body size and trophic position in migratory fishes: a phylogenetic comparative analysis of Clupeiformes (anchovies, herring, shad and allies). Biological Journal of the Linnean Society, 125(2), 302-314.

Blomberg, S. P., Garland Jr, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution, 57(4), 717-745.

Burns, M. D., & Sidlauskas, B. L. (2019). Ancient and contingent body shape diversification in a hyperdiverse continental fish radiation. Evolution.

Buser, T. J., Boyd, O. F., Cortés, Á., Donatelli, C. M., Kolmann, M. A., Luparell, J. L., ... & Summers, A. P. (2020). The natural historian's guide to the CT galaxy: step-by-step instructions for preparing and analyzing computed tomographic (CT) data using cross-platform, open access software. Integrative Organismal Biology.

Diniz-Filho, J. A. F., Santos, T., Rangel, T. F., & Bini, L. M. (2012). A comparison of metrics for estimating phylogenetic signal under alternative evolutionary models. Genetics and molecular biology, 35(3), 673-679.

Dollion, A. Y., Measey, G. J., Cornette, R., Carne, L., Tolley, K. A., da Silva, J. M., ... & Herrel, A. (2017). Does diet drive the evolution of head shape and bite force in chameleons of the genus Bradypodion?. Functional ecology, 31(3), 671-684.

Evans, K. M., Bernt, M. J., Kolmann, M. A., Ford, K. L., & Albert, J. S. (2018). Why the long face? Static allometry in the sexually dimorphic phenotypes of Neotropical electric fishes. Zoological Journal of the Linnean Society.

Faircloth, B. C., Alda, F., Hoekzema, K., Burns, M. D., Oliveira, C., Albert, J. S., ... & Sidlauskas, B. L. (2020). A target enrichment bait set for studying relationships among ostariophysan fishes. Copeia, 108(1), 47-60.

Faulwetter, S., Vasileiadou, A., Kouratoras, M., Dailianis, T., & Arvanitidis, C. (2013). Micro-computed tomography: Introducing new dimensions to taxonomy. ZooKeys, (263), 1.

Foote, M. (1996). Models of morphological diversification: Pp. 62–86 in D. Jablonski, D. Erwin, and J. Lipps, eds. Evolutionary paleobiology.

Fricke, R., Eschmeyer, W. N. & R. van der Laan (2022). Eschmeyer's catalog of fishes: genera, species, references. Available at http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (last accessed 15 Jan 2022).

Gery, J. (1977). Characoids of the world. TFH Public. Inc. Neptune City. 672p.

Gould, S. J. (2002). The structure of evolutionary theory. Harvard University Press.

Hahn, N. S., Pavanelli, C. S., & Okada, E. K. (2000). Dental development and ontogenetic diet shifts of *Roeboides paranensis* Pignalberi (Osteichthyes, Characinae) in pools of the upper Rio Paraná floodplain (state of Paraná, Brazil). Revista Brasileira de Biologia, 60(1), 93-99.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2008). GEIGER: investigating evolutionary radiations. Bioinformatics, 24(1), 129-131.

Kikinis, R., Pieper, S. D., & Vosburgh, K. G. (2014). 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In Intraoperative imaging and image-guided therapy (pp. 277-289). Springer, New York, NY.

Klingenberg, C. P., & Ekau, W. (1996). A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). Biological Journal of the Linnean Society, 59(2), 143-177.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. Molecular ecology resources, 11(2), 353-357.

Kolmann, M. A., Urban, P., & Summers, A. P. (2018a). Structure and function of the armored keel in piranhas, pacus, and their allies. The Anatomical Record, 303(1), 30-43.

Kolmann, M. A., Huie, J. M., Evans, K., & Summers, A. P. (2018b). Specialized specialists and the narrow niche fallacy: a tale of scale-feeding fishes. Royal Society open science, 5(1), 171581.

Javonillo, R., Malabarba, L. R., Weitzman, S. H., & Burns, J. R. (2010). Relationships among major lineages of characid fishes (Teleostei: Ostariophysi: Characiformes), based on molecular sequence data. Molecular Phylogenetics and Evolution, 54(2), 498-511.

Lucena, C. A. S., & Menezes, N. A. (2003). Subfamily Characinae. In: Reis RE, Kullander SO, Ferraris CJ Jr, orgs. Checklist of the freshwater fishes of South and Central America. Porto Alegre, RS: EDIPUCRS, 200–208.

Lucena, Z. M. S. D., & Malabarba, L. R. (2010). Descrição de nove espécies novas de *Phenacogaster* (Ostariophysi: Characiformes: Characidae) e comentários sobre as demais espécies do gênero. Zoologia (Curitiba), 27(2), 263-304.

Mattox, G. M., & Toledo-Piza, M. (2012). Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). Zoological Journal of the Linnean Society, 165(4), 809-915.

Mattox, G. M., Van Der Sleen, P., & Toledo-Piza, M. (2017). Characinae – Characine Tetras. In van der Sleen, P., & Albert, J. S. (Eds.), Field guide to the fishes of the Amazon, Orinoco, and Guianas (Vol. 115) (pp 102-105). Princeton University Press.

Menezes, N. A. (2006). Description of five new species of *Acestrocephalus* Eigenmann and redescription of *A. sardina* and *A. boehlkei* (Characiformes: Characidae). Neotropical Ichthyology, 4(4), 385-400.

Menezes, N. A. (2007). A new species of *Cynopotamus* Valenciennes, 1849 (Characiformes: Characidae) with a key to species of the genus. Zootaxa, 1635(1), 55-61.

Mirande, J. M. (2010). Phylogeny of the family Characidae (Teleostei: Characiformes): from characters to taxonomy. Neotropical Ichthyology, 8(3), 385-568.

Mirande, J. M. (2019). Morphology, molecules and the phylogeny of Characidae (Teleostei, Characiformes). Cladistics, 35(3), 282-300.

Novakowski, G. C., Fugi, R., & Hahn, N. S. (2004). Diet and dental development of three species of *Roeboides* (Characiformes: Characidae). Neotropical Ichthyology, 2(3), 157-162.

Oliveira, C., Avelino, G. S., Abe, K. T., Mariguela, T. C., Benine, R. C., Ortí, G., ... & e Castro, R. M. C. (2011). Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. BMC evolutionary biology, 11(1), 275.

Ou, C., Montaña, C. G., & Winemiller, K. O. (2017). Body size–trophic position relationships among fishes of the lower Mekong basin. Royal Society open science, 4(1), 160645.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics, 35(3), 526-528.

Pérez, M. H. S. (2009). Photographic atlas of fishes of the Guiana Shield. Bulletin of the Biological Society of Washington, 17(1), 52-59.

Peterson, C. C., & Winemiller, K. O. (1997). Ontogenic diet shifts and scale-eating in *Roeboides dayi*, a Neotropical characid. Environmental Biology of Fishes, 49(1), 111-118.

Peterson, C. C., & McIntyre, P. (1998). Ontogenetic diet shifts in *Roeboides affinis* with morphological comparisons. Environmental Biology of Fishes, 53(1), 105-110.

Pie, M. R., & Weitz, J. S. (2005). A null model of morphospace occupation. The American Naturalist, 166(1), E1-E13.

Raup, D. M., & Gould, S. J. (1974). Stochastic simulation and evolution of morphology-towards a nomothetic paleontology.

Raup, D. M. (1977). Probabilistic models in evolutionary paleobiology: A random walk through the fossil record produces some surprising results. American Scientist, 65(1), 50-57.

Reis, R. E., Kullander, S. O., & Ferraris, C. J. (2003). Check list of the freshwater fishes of South and Central America. Edipucrs.

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution, 3(2), 217-223.

Rohlf, F. J. (2003). tpsRelw, relative warps analysis. version 1.36 Department of Ecology and Evolution. State University of New York at Stony Brook. software.

Rohlf, F. J. (2015). The tps series of software. Hystrix, 26(1).

Rohlf, F. J. (2017). TpsDig, Version 2.22, 2015.

Sakamoto, M., & Ruta, M. (2012). Convergence and divergence in the evolution of cat skulls: temporal and spatial patterns of morphological diversity. PLoS One, 7(7), e39752.

Sazima, I., & Machado, F. A. (1983). Hábitos e comportamento de *Roeboides prognathus*, um peixe lepidófago (Osteichthyes, Characoidei). Boletim de Zoologia, 7(7), 37-56.

Sazima, I. (1983). Scale-eating in characoids and other fishes. Environmental Biology of Fishes, 9(2), 87-101.

S. Schlager, G. Jefferis, D. Ian, M.S. Schlager Package 'Morpho' (2019). http://cran.r-project.org/web/packages/Morpho/Morpho.pdf

Serb, J. M., Sherratt, E., Alejandrino, A., & Adams, D. C. (2017). Phylogenetic convergence and multiple shell shape optima for gliding scallops (Bivalvia: Pectinidae). Journal of evolutionary biology, 30(9), 1736-1747.

Sidlauskas, B. (2008). Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. Evolution: International Journal of Organic Evolution, 62(12), 3135-3156.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312-1313.

Team, R. C. (2013). R: A language and environment for statistical computing.

Venere, P. C., & Garutti, V. (2011). Peixes do Cerrado: Parque Estadual da Serra Azul, rio Araguaia, MT. São Carlos: RiMa Editora, FAPEMAT. (pp 71)

Vidal-García, M., & Keogh, J. S. (2017). Phylogenetic conservatism in skulls and evolutionary lability in limbs–morphological evolution across an ancient frog radiation is shaped by diet, locomotion and burrowing. BMC Evolutionary Biology, 17(1), 165.

Wake, D. B., & Larson, A. (1987). Multidimensional analysis of an evolving lineage. Science, 238(4823), 42-48.

Supplementary information

Supplementary Table 1. Characin samples that were CT-scanned.

Catalogue number	Species	Name scan
FMNH 53666	*Acanthocharax microlepis*	scan_acantho_1
FMNH 53666	*Acanthocharax microlepis*	scan_acantho_1
FMNH 53666	*Acanthocharax microlepis*	scan_acantho_1
FMNH 53666	*Acanthocharax microlepis*	scan_acantho_1
FMNH 53666	*Acanthocharax microlepis*	scan_acantho_1
AUM44032	*Acestrocephalus sardina*	Acestrocephalus can 1b
AUM44032	*Acestrocephalus sardina*	Acestrocephalus can 1b
AUM43599	*Acestrocephalus sardina*	Acestrocephalus can 2
AUM43599	*Acestrocephalus sardina*	Acestrocephalus can 2
AUM 43599	*Acestrocephalus sardina*	Acestrocephalus can 3b
ANSP 192399	*Acestrocephalus sardina*	scan_cynopotamus_8
ANSP 192399	*Acestrocephalus sardina*	scan_cynopotamus_8
ANSP 192399	*Acestrocephalus sardina*	scan_cynopotamus_8
ANSP 180523	*Charax tectifer*	scan_charax_1
ANSP 180523	*Charax tectifer*	scan_charax_1
ANSP 180523	*Charax tectifer*	scan_charax_1
ANSP 180523	*Charax tectifer*	scan_charax_1
ANSP 135634	*Charax gibbosus*	scan_cynopotamus_4
ANSP 135634	*Charax gibbosus*	scan_cynopotamus_4
ANSP 135634	*Charax gibbosus*	scan_roebo_5
LBP8771	*Charax leticiae*	Charax can 1
LBP8771	*Charax leticiae*	Charax can 1
LBP8771	*Charax leticiae*	Charax can 1
LBP8771	*Charax leticiae*	Charax can 1
LBP18734	*Charax metae*	Charax can 3
LBP18734	*Charax metae*	Charax can 3
FMNH 55145	*Charax metae*	scan_charax_14
FMNH 55145	*Charax metae*	scan_charax_14
FMNH 55145	*Charax metae*	scan_charax_14
FMNH 55145	*Charax metae*	scan_charax_14
FMNH 55145	*Charax metae*	scan_charax_14
LBP22517	*Charax michaeli*	Charax can 3
LBP22517	*Charax michaeli*	Charax can 3
FMNH100336	*Charax michaeli*	scan_roebo_9
FMNH100336	*Charax michaeli*	scan_roebo_9
FMNH100336	*Charax michaeli*	scan_roebo_9
FMNH100336	*Charax michaeli*	scan_roebo_9
ANSP 101413	*Charax microlepis*	scan_roebo_3

Catalogue number	Species	Name scan
LBP14117	*Charax pauciradiatus*	Charax can 2c
LBP14117	*Charax pauciradiatus*	Charax can 2c
LBP14117	*Charax pauciradiatus*	Charax can 2c
LBP14117	*Charax pauciradiatus*	Charax can 2c
LBP 17020	*Charax stenopterus*	Charax can 4
LBP 14529	*Charax stenopterus*	Charax can 4
LBP13020	*Charax stenopterus*	Roeboides can 5
LBP13169	*Charax stenopterus*	Roeboides can 6
ANSP 192403	*Cynopotamus bipunctatus*	scan_cynopotamus_8
ANSP 176793	*Cynopotamus essequibensis*	scan_cynopotamus_4
ANSP 189604	*Cynopotamus essequibensis*	scan_cynopotamus_4
ANSP 189604	*Cynopotamus essequibensis*	scan_roebo_5
USNM310473	*Cynopotamus magdalenae*	Cynopotamus can 2
LBP 2556	*Galeocharax gulo*	Acestrocephalus can 2
LBP 2556	*Galeocharax gulo*	Acestrocephalus can 3b
LBP13453	*Galeocharax humeralis*	Acestrocephalus can 1b
LBP13453	*Galeocharax humeralis*	Acestrocephalus can 1b
LBP13453	*Galeocharax humeralis*	Galeocharax can 1b
LBP13453	*Galeocharax humeralis*	Galeocharax can 1b
FMNH 54598	*Phenacogaster beni*	scan_roebo_3
LBP25604	*Phenacogaster eurytaenia*	Camilla_Can_APS_
LBP25604	*Phenacogaster eurytaenia*	Camilla_Can_APS_
LBP25604	*Phenacogaster eurytaenia*	Camilla_Can_APS_
LBP25604	*Phenacogaster eurytaenia*	Camilla_Can_APS_
LBP24038	*Phenacogaster franciscoensis*	Phenacogaster can4
LBP24038	*Phenacogaster franciscoensis*	Phenacogaster can4
LBP24038	*Phenacogaster franciscoensis*	Phenacogaster can4
LBP24038	*Phenacogaster franciscoensis*	Phenacogaster can4
LBP4683	*Phenacogaster jancupa*	scan_phenacogaster_10
LBP4683	*Phenacogaster jancupa*	scan_phenacogaster_10
LBP17082	*Phenacogaster maculoblonga*	Phenacogaster can5
LBP17082	*Phenacogaster maculoblonga*	Phenacogaster can5
LBP17082	*Phenacogaster maculoblonga*	Phenacogaster can5
LBP17082	*Phenacogaster maculoblonga*	Phenacogaster can5
ANSP 131826	*Phenacogaster megalostictus*	scan_phenacogaster_12
ANSP 131826	*Phenacogaster megalostictus*	scan_phenacogaster_12
ANSP 131826	*Phenacogaster megalostictus*	scan_phenacogaster_12
ANSP 131826	*Phenacogaster megalostictus*	scan_phenacogaster_12
ANSP 190567	*Phenacogaster microstictus*	scan_phenacogaster_10
ANSP 190567	*Phenacogaster microstictus*	scan_phenacogaster_10
LBP9385	*Phenacogaster naevata*	Phenacogaster can7
LBP9385	*Phenacogaster naevata*	Phenacogaster can7
LBP9385	*Phenacogaster naevata*	Phenacogaster can7

Catalogue number	Species	Name scan
LBP9385	*Phenacogaster naevata*	Phenacogaster can7
LBP22715	*Phenacogaster pectinata*	Phenacogaster can3
LBP22715	*Phenacogaster pectinata*	Phenacogaster can3
LBP22715	*Phenacogaster pectinata*	Phenacogaster can3
LBP 5795	*Phenacogaster prolata*	scan_phenacogaster_13
LBP 5795	*Phenacogaster prolata*	scan_phenacogaster_13
LBP 18683	*Phenacogaster prolata*	scan_phenacogaster_13
LBP 18683	*Phenacogaster prolata*	scan_phenacogaster_13
LBP 18683	*Phenacogaster prolata*	scan_phenacogaster_13
LBP 18683	*Phenacogaster prolata*	scan_phenacogaster_13
LBP16061	*Phenacogaster* sp.	Phenacogaster can7
LBP 16061	*Phenacogaster* sp.	Phenacogaster can1c
LBP 16061	*Phenacogaster* sp.	Phenacogaster can1c
LBP 16061	*Phenacogaster* sp.	Phenacogaster can1c
LBP 16061	*Phenacogaster* sp.	Phenacogaster can1c
FMNH108487	*Phenacogaster tegata*	scan_phenacogaster_12
FMNH108487	*Phenacogaster tegata*	scan_phenacogaster_12
ANSP189600	*Phenacogaster wayana*	scan_phenacogaster_11a
ANSP189600	*Phenacogaster wayana*	scan_phenacogaster_11a
ANSP189599	*Phenacogaster wayana*	scan_phenacogaster_11a
ANSP189599	*Phenacogaster wayana*	scan_phenacogaster_11a
LBP22638	*Roeboides affinis*	Roeboies scan 1
LBP22638	*Roeboides affinis*	Roeboies scan 1
LBP22638	*Roeboides affinis*	Roeboies scan 1
LBP22638	*Roeboides affinis*	Roeboies scan 2
FMNH 128131	*Roeboides bouchellei*	scan_roebo_2
USNM311020	*Roeboides carti*	Roeboides can 3
USNM311020	*Roeboides carti*	Roeboides can 3
USNM311020	*Roeboides carti*	Roeboides can 4
USNM311020	*Roeboides carti*	Roeboides can 4
FMNH 85335	*Roeboides dayi*	scan_roebo_6
FMNH 85335	*Roeboides dayi*	scan_roebo_6
FMNH 85335	*Roeboides dayi*	scan_roebo_6
FMNH 85335	*Roeboides dayi*	scan_roebo_6
FMNH 85335	*Roeboides dayi*	scan_roebo_6
LBP5110	*Roeboides descalvadensis*	Roeboides can 5
LBP5110	*Roeboides descalvadensis*	Roeboides can 6
LBP5110	*Roeboides descalvadensis*	Roeboides can 7
LBP5110	*Roeboides descalvadensis*	Roeboides can 7
ANSP 206220	*Roeboides dientonito*	scan_roebo_7
ANSP 206220	*Roeboides dientonito*	scan_roebo_7
ANSP 206220	*Roeboides dientonito*	scan_roebo_7
LBP6114	*Roeboides dientonito*	Roeboides can 3

Catalogue number	Species	Name scan
LBP6114	*Roeboides dientonito*	Roeboides can 3
LBP6114	*Roeboides dientonito*	Roeboides can 4
ANSP180932	*Roeboides dispar*	scan_roebo_9
ANSP180932	*Roeboides dispar*	scan_roebo_9
LBP2755	*Roeboides guatemalensis*	Roeboies scan 1
LBP2755	*Roeboides guatemalensis*	Roeboies scan 2
LBP2755	*Roeboides guatemalensis*	Roeboies scan 2
LBP2755	*Roeboides guatemalensis*	Roeboies scan 2
ANSP 164256	*Roeboides ilsea*	scan_roebo_3
ANSP 120320	*Roeboides magdalenae*	scan_roebo_3
ANSP 95879	*Roeboides margareteae*	scan_roeboides_15
ZMUC P241391	*Roeboides microlepis*	Roeboides can 8b
ANSP 178154	*Roeboides myersii*	scan_cynopotamus_4
ANSP 178154	*Roeboides myersii*	scan_roebo_5
ANSP 178154	*Roeboides myersii*	scan_roebo_5
ANSP 196316	*Roeboides numerosus*	scan_charax_1
ANSP 196316	*Roeboides numerosus*	scan_roebo_3
ANSP 99843	*Roeboides occidentalis*	scan_roebo_7
ANSP 99843	*Roeboides occidentalis*	scan_roebo_7
LBP10444	*Roeboides xenodon*	Roeboides can 5
LBP10444	*Roeboides xenodon*	Roeboides can 5
LBP10444	*Roeboides xenodon*	Roeboides can 6
LBP10444	*Roeboides xenodon*	Roeboides can 6

Chapter 3

Molecular identification and description of a new species of *Phenacogaster* (Characidae: Characinae)

Molecular identification and description of a new species of *Phenacogaster* (Characidae: Characinae)

Camila S. Souza

Departamento de Biologia Estrutural e Funcional, Instituto de Biociências, Universidade Estadual Paulista, R. Prof. Dr. Antônio C. W. Zanin 250, 18618-689, Botucatu, SP, Brazil

Abstract

Phenacogoster is the most species-rich genus of Characinae containing 23 valid species widely distributed in riverine systems of cis-Andean South America. However, little has been advanced in the molecular studies for *Phenacogoster*. Here, we used partial sequences of the mitochondrial gene *cytochrome c oxidase subunit I* (COI) and species delimitation methods to the identification of 13 valid species of *Phenacogaster* and to assess the intraspecific genetic diversity. Additionally, we recognize and describe a new species from the Xingu and Araguaia basins.

Keywords: Phenacogasterii, Neotropical fish, Taxonomy, COI, biodiversity

1. Introduction

The Neotropical fish subfamily Characinae contains 85 valid species and 7 genera (*Acanthocharax* Eigenmann 1912, *Acestrocephalus* Eigenmann 1910, *Charax* Scopoli 1777, *Cynopotamus* Valenciennes 1850, *Galeocharax* Fowler 1910, *Phenacogaster* Eigenmann 1907, and *Roeboides* Günther 1864) (Mattox & Toledo-Piza, 2012). *Phenacogaster* is the most species-rich genus of

Characinae, with 23 species widely distributed in riverine systems of cis-Andean South America (Fricke *et al.,* 2022). They are small fishes that reach up to 6 cm standard length and popularly known as “lambaris”, “glass tetras”, “mojaritas”, and “yaya” (Lucena & Malabarba, 2010).

Phenacogaster is morphologically distinguished from other Characinae genera by dorsal profile without gibbosity, preventral region flat and covered by two longitudinal series of elongate and imbricated scales that form a zig-zag pattern in ventral view (Mattox *et al.,* 2017). The latest taxonomic revision of the genus recognized nine new species, almost doubling the number of species then recognized for the genus, and provided an identification key for the species (Lucena & Malabarba, 2010). The authors proposed the *P. pectinata* complex with *P. pectinata* (Cope 1870), *P. microstictus* Eigenmann 1909, *P. beni* Eigenmann 1911 and *P. suborbitalis* (Ahl 1936) (Lucena & Malabarba, 2010). However, several undescribed species are recognized in this group and the difficulty to diagnose them is associated by the broad distribution and lack of consistent morphological differences (Géry, 1972; Lucena & Malabarba, 2010). After that revision, only three species were described: *P. naevata* and *P. eurytaenia* from the Tocantins basin (Antonetti *et al*., 2018) and *P. julliae* from the São Francisco basin (Lucena & Lucena, 2019).

The use of molecular information can help the species delimitation (Barreto *et al*., 2018; Mateussi *et al.*, 2020; Ota *et al.,* 2020) and indicate potential hidden diversity (Pires *et al.,* 2017; Machado *et al.,* 2018; Ramirez *et al.,* 2020). In this context, and considering the lack of molecular studies for *Phenacogaster*, we used partial sequences of the mitochondrial gene *cytochrome c oxidase subunit I* (COI) and species delimitation methods to the identification of 13 valid species of *Phenacogaster* and to assess the intraspecific genetic diversity. Additionally, we recognize and describe a new species from the Xingu and Araguaia basins.

2. Materials and Methods

Taxon sampling

We used specimens of *Phenacogaster* collected recently or obtained from ichthyological collections, and morphologically identified with the help of identification keys (e.g. Lucena & Malabarba, 2010). All fishes were collected following the Brazilian laws through SISBIO/MMA permit n. 3,245 and procedures for collection, maintenance and analyses followed the international guidelines for animal experiments through CEEAA IBB/UNESP protocol n. 304.

Molecular analysis

DNA Extraction, Amplification and Sequencing

One hundred and one taxa were included in the molecular analyses (Supplementary Table 1). Ninety-eight sequences were generated herein, and three sequences were obtained from Genbank (1- *Tetragonopterus carvalhoi* Melo, Benine, Mariguela & Oliveira 2011 and 2- *Phenacogaster wayana* Le Bail & Lucena 2010). DNA was extracted from samples of muscle, liver or gills, using the extraction method of the Ivanova *et al.* (2006). Partial sequences of the *cytochrome c oxidase subunit I* (*COI*) gene were amplified by polymerase chain reaction (PCR) with primers FishF1, FishR1 and FishF2, FishR2 (Ward *et al.*, 2005) or L6252-Asn and H7271-COXI (Melo *et al*., 2011). PCR amplifications were performed in a total volume of 12.5 µl that included 1.25 µl of 10X buffer, 0.25 µl of $MgCl_2$ (50 mM), 0.2 µl dNTPs (2 mM), 0.5 µl of each primer (5 mM), 0.1 µl of PHT Taq DNA polymerase (Phoneutria), 1.0 µl of genomic DNA (200 ng) and 8.7 µl ddH_2O. PCR conditions consisted of an initial denaturation (5 min at 94°C), followed by 30 cycles of chain denaturation (50s at 94°C), primer hybridization (45s at 50-54°C) and nucleotide extension (1 min at 68°C), and

a final extension (10 min at 68°C). All PCR products were checked on 1% agarose gels and then purified with ExoSap-IT (USB Corporation) following the manufacturer's instructions. The purified PCR products were submitted to sequencing reactions using BigDye Terminator v 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and purified again by ethanol precipitation. Products were loaded onto an ABI 3130 DNA Analyzer automatic sequencer (Applied Biosystems).

Data analysis

Sequences were assembled using Geneious v7.1.9 (Kearse *et al.*, 2012) and subsequently aligned with Muscle (Edgar, 2004) under default parameters. The method of Xia *et al.* (2003) were used to evaluate the occurrence of substitution saturation in DAMBE v5.3.38 (Xia, 2013). Nucleotide variation, substitution patterns and the best-fit model of nucleotide evolution were estimated in MEGA v10 (Kumar *et al.*, 2018).

The maximum-likelihood (ML) and neighbor-joining (NJ) trees were performed with 1000 bootstrap replicates, using the K2P+G model (Kimura 1980) in MEGA v10. Two distinct species delimitation methods were performed: Automatic Barcode Gap Discovery analysis (ABGD; Puillandre *et al.*, 2012) available in the ABGD webserver (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html) using Kimura (K80; 1.0) evolutionary model with X = 1.0, Pmin = 0.001 and Pmax = 0.1 and Poisson Tree Process (PTP; Zhang *et al.*, 2013) using the best maximum-likelihood (ML) tree, 100,000 generations, and other parameters at default in the bPTP webserver (http://species.h-its.org/ptp/). The values of genetic distance were calculated in MEGA v10 using the K2P model+G and groups were ordered based on morphological identification. ABGD, bPTP and genetic distance analysis were performed without the outgroup *Tetragonopterus carvalhoi*.

Morphological analysis

Morphometric and meristic data were taken on the left side of each specimen whenever possible following Fink & Weitzman (1974) and Lucena & Malabarba (2010). Measurements were taken point to point with a precision of 0.1 mm using a digital caliper. Counts of vertebrae and supraneurals were obtained from cleared and stained specimens (c&s) prepared with the protocol from Taylor & Van Dyke (1985). All measurements are expressed as percents of standard length (SL), except subunits of the head that are expressed as percents of head length (HL). In each description the frequency of each count is provided in parentheses, and the count of the holotype is indicated by an asterisk.

3. Results

Molecular identification

Partial sequences of the *COI* gene were obtained for 101 specimens representing 13 out of 23 valid species of *Phenacogaster* (56.2%). The matrix contained 678 pb (190 pb variable sites) and nucleotide composition of 24.6% adenine, 27.4% citosine, 18% guanine and 30% thymine. DAMBE indicated no saturation for either transitions or transversions in both asymmetrical (Iss.cAsym) and symmetrical (Iss.cSym) topologies. ML and NJ trees recovered almost the same topology, and both supported the recognition of the new species (Figure 1; Supplementary Fig. 1). ABGD delimited 16 species and recognized *P. pectinata* and *P. capitulata* as a single species (Fig. 1; Supplementary Fig. 2). Conversely, bPTP delimited a disproportionate number of 38 species, but also identified *P. pectinata* and *P. capitulata* as a single species (Fig. 1; Supplementary Fig 3). The overall mean of

K2P genetic distances was 0.077 ± 0.008. Interspecific genetic distances ranged from 0.026 ± 0.007 to 0.141 ± 0.020 and intraspecific distances ranged from 0.000 ± 0.000 to 0.006 ± 0.004 (Table 1).

Table 1- Pairwise K2P genetic distances among species of *Phenacogaster*. Intraspecific genetic variations are highlighted in bold in the last column. Numbers below diagonal are values of interspecific distances and numbers above diagonal are respective values of standard deviation.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**Intraspecific genetic distances**
1- ***P. microstictus***		0.007	0.013	0.011	0.011	0.011	0.013	0.011	0.016	0.017	0.017	0.012	0.015	0.014	0.016	0.015	0.015	**-**
2- ***P. prolata***	0.026		0.011	0.010	0.010	0.009	0.012	0.010	0.017	0.017	0.016	0.013	0.015	0.013	0.017	0.014	0.015	**0**
3- ***P. pectinata***	0.079	0.069		0.007	0.011	0.010	0.013	0.011	0.018	0.020	0.019	0.016	0.017	0.016	0.018	0.018	0.019	**0.001±0.003**
4- ***P. capitulata***	0.051	0.043	0.027		0.010	0.011	0.015	0.011	0.022	0.023	0.021	0.016	0.018	0.018	0.020	0.017	0.017	**0**
5- ***P. beni***	0.058	0.052	0.065	0.046		0.006	0.007	0.008	0.016	0.018	0.018	0.013	0.015	0.014	0.017	0.015	0.015	**0**
6- ***P.*** **aff.** ***suborbitalis***	0.062	0.052	0.057	0.048	0.030		0.007	0.008	0.016	0.017	0.018	0.014	0.014	0.013	0.016	0.014	0.016	**0.003±0.002**
7- ***P.*** **aff.** ***beni***	0.073	0.070	0.072	0.067	0.029	0.028		0.008	0.018	0.018	0.019	0.016	0.017	0.016	0.017	0.018	0.021	**0**
8- ***P. tegata***	0.058	0.053	0.069	0.047	0.039	0.035	0.029		0.016	0.016	0.017	0.013	0.015	0.013	0.016	0.014	0.015	**0.003±0.001**
9- ***P. wayana***	0.102	0.111	0.129	0.125	0.111	0.113	0.118	0.110		0.014	0.016	0.014	0.012	0.012	0.016	0.013	0.015	**0**
10- ***P. retropinna***	0.110	0.111	0.141	0.131	0.127	0.120	0.120	0.110	0.080		0.009	0.010	0.011	0.010	0.012	0.010	0.011	**0.005±0.002**
11- ***P.*** **sp. n**	0.114	0.112	0.139	0.117	0.131	0.126	0.134	0.120	0.109	0.045		0.010	0.013	0.011	0.013	0.012	0.014	**0.006±0.003**
12- ***P. maculoblonga***	0.071	0.079	0.114	0.092	0.090	0.091	0.099	0.081	0.086	0.053	0.057		0.009	0.009	0.011	0.009	0.010	**0**
13- ***P. calverti***	0.100	0.103	0.123	0.103	0.105	0.103	0.112	0.101	0.067	0.058	0.077	0.051		0.008	0.011	0.008	0.011	**0**
14- ***P.*** **aff.** ***retropinna***	0.086	0.085	0.106	0.103	0.088	0.084	0.098	0.082	0.068	0.051	0.067	0.045	0.042		0.012	0.008	0.010	**0.006±0.002**
15- ***P. franciscoensis***	0.113	0.118	0.133	0.124	0.125	0.118	0.119	0.115	0.106	0.073	0.085	0.068	0.068	0.070		0.010	0.008	**0.010±0.003**
16- ***P. eurytaenia***	0.093	0.090	0.124	0.107	0.099	0.098	0.118	0.094	0.073	0.050	0.068	0.050	0.041	0.038	0.051		0.007	**0**
17- ***P. naevata***	0.096	0.095	0.131	0.108	0.103	0.107	0.135	0.099	0.088	0.058	0.073	0.051	0.060	0.053	0.032	0.030		**0**

Figure 1. ML tree of species of *Phenacogaster* based on the *COI* gene (678 pb). Vertical bars at right represent the number of species obtained by the ABGD and bPTP analyses. Numbers near nodes represent bootstrap support (values <70% are not shown). Codes before tip names represent tissue numbers or deposited sequences in NCBI.

Morphological identification

***Phenacogaster* sp. n. (Fig. 2, Tab. 2)**

Figure 2. *Phenacogaster* sp. n., MZUSP 120058, holotype, 26.7 mm SL, holotype, Brazil, Pará, Novo Progresso, Amazon basin, Rio Xingu.

Holotype. MZUSP 120058, 26.7 mm SL, Brazil, Pará, Novo Progresso, Amazon basin, rio Xingu (Riacho em estrada de terra perpendicular à BR163 entre a FAB e Cachoeira da Serra, provável afluente do Curuá), 08°29’59” S 54°58’06.1” W, 07 Aug 2015, F. Dagosta, M.M.F. Marinho, P. Camelier, V. Giovanetti.

Paratypes: All from Brazil. MZUSP 120058, 28, 20.4-34.9 mm SL, same data as holotype. MZUSP 97621, 49, 21.6-33.6 mm SL, 4 (c&s), Pará, Altamira, Amazon basin, Rio Xingu, Rio Curuá-Iriri, 08°15’17” S 55°06’40” W, 27 Oct 2007, JLO Birindelli, L Souza, AL Netto-Ferreira, MH Sabaj, NK Lujan. LBP 25217 (tissue number: 94032), 1, 30.58 mm SL, Pará, Altamira, Amazon basin, Rio Xingu, Rio Treze de Maio, 08°39’06.9” S 55°02’09.1” W, 24 Sep 2017, AC Dias, CS Souza, C Souto, N Flausino Jr, R Devidé. LBP 30738, 1, 38.06 mm SL, Mato Grosso, Primavera do Leste, Amazon basin, Rio Xingu, Córrego Xavante, 14°38’24” S 53°55’38” W, 23 Aug 2021, CS Souza, L Reia, GSC. Silva, EV Ywamoto. LBP 16061, 9, 22.9-35.4 mm SL, Mato Grosso, Primavera do

Leste, Amazon basin, Rio Xingu, Córrego Xavante, 14°38'24" S 53°55'38" W, 05 Aug 2012, C Oliveira, M Taylor, GHC Silva, JHM Martinez. LBP15807, 2, 21.58-28.19 mm SL, Mato Grosso, Querência, Amazon basin, Rio Xingu, Afluente rio Feio, 12°33'20.5" S 52°16'16.1" W, 31 Jul 2012, C Oliveira, M Taylor, GHC Silva, JHM Martinez. LBP 15835 (tissue number: 64949), 2, 19.3- 26.9 mm SL, Mato Grosso, Querência, Amazon basin, Rio Xingu, Rio Feio, 12°31'55.7" S 52°20'29.8" W, 31 Jul 2012, C Oliveira, M Taylor, GHC Silva, JHM Martinez. MZUSP 97708, 8, 27.3-32.4 mm SL, Mato Grosso, Santo Antonio do Leste, Araguaia basin, rio Suspiro, 14° 52'30.0" S 54° 5' 0.0" W, 18 Jan 2002, N Menezes, O Oyakawa, GM Guazzelli, R Quevedo. MZUSP 118678, 8, 26.9-29.6, Mato Grosso, Mato Grosso, Santo Antonio do Leste, Araguaia basin, rio Suspiro. 14° 52' 30.92" S 54° 05'1.47" W, 17 Nov 2014, F Dagosta, O Ohara, V Giovannetti.

Diagnosis. *Phenacogaster* sp. n. is distinguished from all congeners except *P. tegata* (Eigenmann 1911), *P. carteri* (Norman 1934), *P. napoatilis* (Lucena & Malabarba 2010), and *P. capitulata* (Lucena & Malabarba 2010) by the presence of incomplete lateral line. It differs from *P. tegata* by the presence of round humeral blotch near pseudotympanum and distant to vertical line of dorsal-fin origin (vs. humeral blotch horizontally-elongated and posteriorly distant from pseudotympanum, close to vertical line of dorsal-fin origin). The new species also differs from *P. carteri*, *P. napoatilis*, and *P. capitulata* by the presence of humeral blotch rounded in males and females (vs. presence or absence humeral blotch in only females). *Phenacogaster* sp. n. can be distinguished from *P. retropinna* Lucena & Malabarba, 2010, from Xingu and Araguaia basins by presence of incomplete lateral line (vs. complete), anal-fin origin at vertical line through base of first or second dorsal-fin branched rays (vs. anal-fin origin at vertical through fifth or sixth dorsal-fin branched rays) and supplemented by overlapped ranges of body depth (29.4–36.2% SL vs 25.1–32.8% SL). Finally, *Phenacogaster* sp. n. can be diagnosed from *P. ojitata* Lucena & Malabarba, 2010, from Xingu basin by presence of incomplete lateral line (vs. complete), larger orbital diameter (34–42.9% HL

vs 33–36.9% HL), and caudal peduncle blotch slightly extending over middle of caudal-fin rays (vs. caudal peduncle blotch diamond-shaped and further extending over middle caudal-fin rays).

Description. Morphometric data in Table 2. Body compressed and relatively slender. Dorsal profile convex from anterior tip of upper jaw to dorsal-fin origin; slightly straight from dorsal-fin base to adipose-fin originand slightly concave from that point to base of dorsal procurrent caudal-fin rays. Ventral profile convex from tip of lower jaw to anal-fin origin. Body profile along anal-fin base straight. Ventral profile of caudal peduncle straight to slightly concave. Preventral area flattened. Mouth terminal with lower jaw slightly shorter than upper jaw; posterior tip of maxilla reaching midpoint of second infraorbital. Pseudotympanum extending from region of first rib to anterior border of second rib. Premaxilla with two tooth rows. Outer row with 6(1) or 8(1) total teeth, divided in medial and lateral sections by gap; medial section with 3(2) tricuspid teeth; lateral section with 3(1) or 5(1). Inner row with 6(1) or 9(1) total teeth, 4(1) or 5(1) tricuspid teeth followed by 2(1) or 4 (1) conical teeth. Maxilla with 22(2), 23(1), or 29(1) conical teeth. Dentary with 16(1), 18(2), or 19 (1) teeth, in single row, with 4(1), 5(1), 6(1) or 7(1) tricuspid teeth followed by 11(2), 12(1) or 14(1) conical teeth (Fig. 3).

Dorsal-fin rays ii,8*(10). Anal-fin rays iv,30(1), 31(1), 33(1) or 34*(1). Pectoral fin rays i,11*(3) or i,12(2). Pelvic-fin rays i,7*(5); its tip reaching beyond anal-fin origin. Lateral line incomplete. Lateral line longitudinal to 30(1), 32(8), 33(2), 34*(35), or 36(4). Pored scales 7(4), 8(9), 9*(14), 10(6), 11(9), 12(4), 13(1), 14 (4), or 16(1); some specimens with perforated scales 2(2) or 5(3) before the last scales non-perforated. Scale series between lateral line and dorsal-fin origin 5(6), 6*(42), or 7(5). Scale series between lateral line and anal-fin origin 4(24), 5*(24), or 6(7). Gill rakers on upper limb of first gill arch 4(14) or 5(5); gill rakers on lower limb 7(19). Total vertebrae 35(3) or 36(1): precaudal 15(4), caudal 20(3) or 21 (1). Supraneurals 4(4).

Table 2. Morphometric data of *Phenacogaster* sp. n. (n= 39, holotype included). SD = standard deviation. Range includes holotype.

	Holotype	n	Range	Mean	SD
Standard length (SL) (mm)	26.7	39	24.1 - 38.06	29.5	-
Percents of standard length					
Greatest body depth	31.5	39	29.4 - 36.2	32.5	1.7
Snout to dorsal-fin origin	53.3	39	50.6 - 55.3	52.9	1.1
Snout to pectoral-fin origin	26.7	39	26.6 - 31.5	28.7	1.3
Snout to pelvic-fin origin	42.7	39	39.1 - 45.3	41.8	1.4
Snout to anal-fin origin	53.8	39	51.3 - 58.9	55.6	1.9
Dorsal-fin origin to hypural joint	51.5	39	47.7 - 54.3	51	1.8
Dorsal-fin origin to anal-fin origin	31.2	39	29.3 - 36.9	32.8	1.7
Dorsal-fin origin to pelvic-fin origin	32.3	39	31.9 - 39.9	34.6	1.7
Dorsal-fin origin to pectoral-fin origin	38.9	39	36 - 41.3	38.4	1.4
Caudal-peduncle depth	9.1	39	8.6 - 11.4	9.7	0.7
Pectoral-fin length	16.6	39	16.5 - 22.8	20.3	1.7
	Holotype	**n**	**Range**	**Mean**	**SD**
Pelvic-fin length	15.2	39	14.1 - 20.4	17.4	1.5
Head length	28	39	24 - 29.8	27.3	1.2
Percents of head length					
Snout length	26.5	39	23.4 - 34.1	27.4	2.2
Orbital diameter	36.2	39	34 - 42.9	37.4	2.1
Interorbital width	27	39	24.4 – 32.3	28	2.2

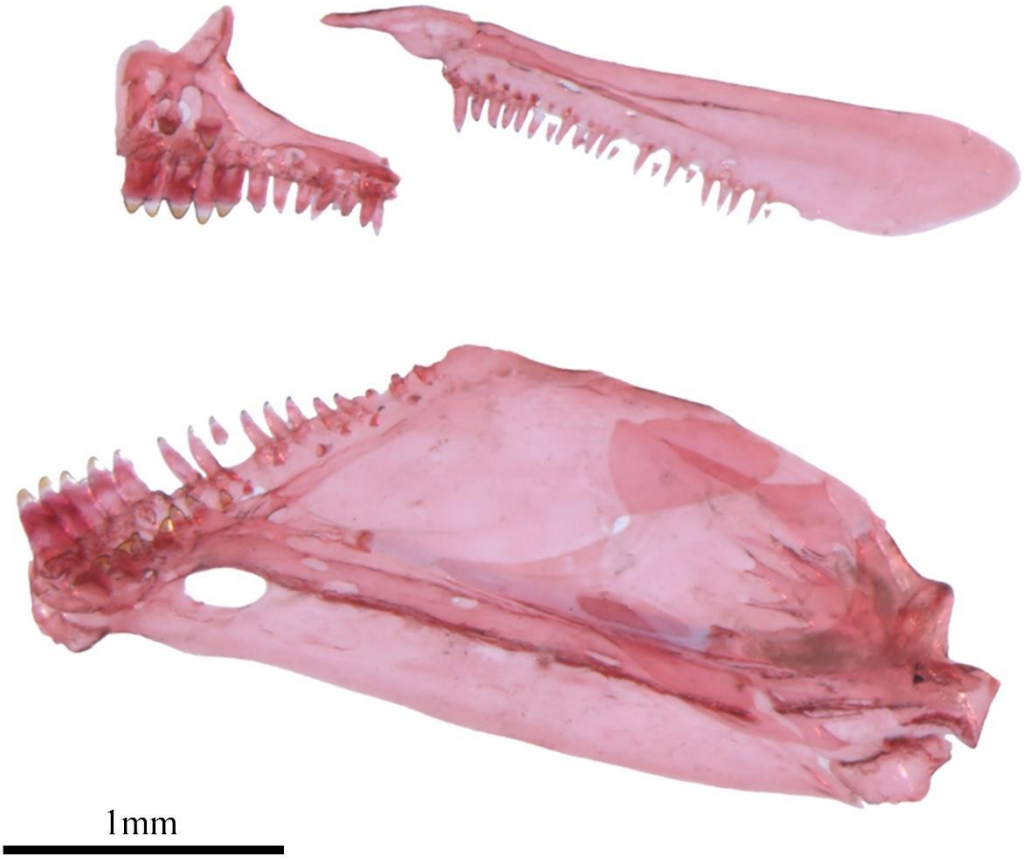


Figure 3. *Phenacogaster* sp. n., MZUSP 97621, 23.81 mm SL, paratype: maxilla (top right), premaxilla (top left), and lower jaw (bottom). Lateral view, left side.

Color in alcohol. Overall ground coloration yellowish. Dorsolateral region of body with chromatophores along margins of scales. Ventrolateral region less pigmented. Humeral blotch rounded or oval immediately to second to five lateral-line scales, extending from the rear of pseudotympanum, covering about three to five scale rows vertically. Caudal peduncle with diamond-shaped blotch of dark pigmentation. Anal, pelvic, pectoral, and dorsal fins hyaline scattered with small dark chromatophores. Adipose fin hyaline (Fig. 1).

Color in life. Overall ground coloration yellowish to golden. Dorsolateral region of body with chromatophores along margins of scales. Ventrolateral region less pigmented. Humeral blotch rounded or oval posteriorly with bright yellow to orange blotch. Diamond-shaped caudal blotch over

middle portions of caudal peduncle, slightly extending posteriorly to distal portion of medial caudal-fin rays. All fins orange to yellowish coloration, more intense at anterior half of caudal-fin lobes. Posterior tip of caudal and dorsal fins scattered by small dark chromatophores.

Sexual dimorphism. Mature males with hooks on pelvic and anal fin rays.

Distribution. *Phenacogaster* sp.n. is known from the upper Xingu and upper Araguaia rivers, Amazon basin, Brazil (Fig. 4).

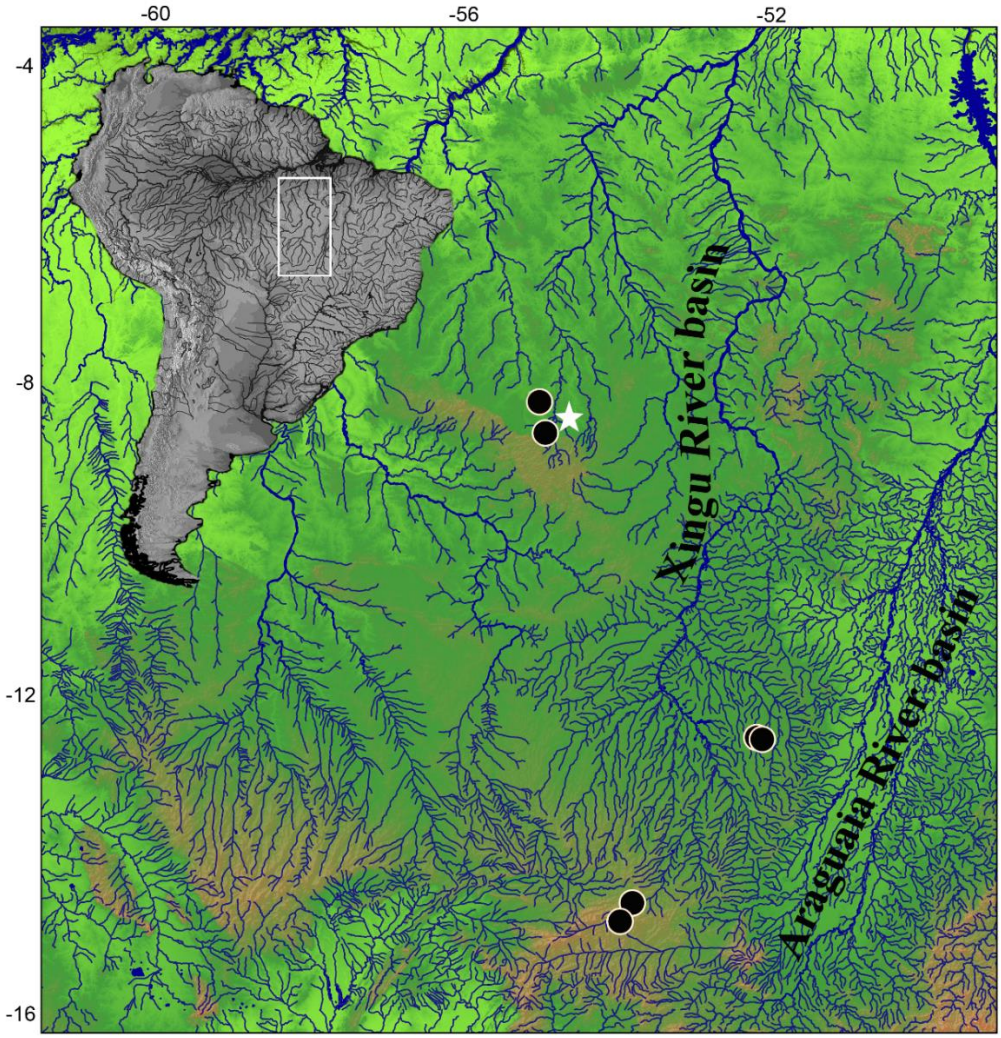


Figure 4. Map showing the distribution of *Phenacogaster* sp. n. in the upper Xingu and Araguaia river basins. Type locality = white star.

Conservation status. *Phenacogaster* sp. n. is widely distributed across the upper Xingu and Araguaia river basins, and there are no significant threats affecting its populations. Therefore, we suggest the category Least Concern (LC) according to the International Union for Conservation of Nature criteria (IUCN Standards and Petitions Subcommittee, 2014).

Comparative material examined. ***Phenacogaster retropinna***: MZUSP81267 (17, 32.9 - 39,2 mm SL), Brazil, Amazonas, Amazon River basin, rio Negro. 0º16'22.0" N 69º 54'3.0" W, 7 Nov 2002, Lima et al. MZUSP99771 (14, 32- 40.3 mm SL), Brazil, Mato Grosso, Aripuanã, Madeira River basin, rio Aripuanã, Balneário Primavera, a jusante do salto de Dardanelos. 10º09'54" S 59º26'55" W, 12 Dec 2004, F. A. Machado, C.M.C Leite, N.E. Silva & R. Rosa. MZUSP 30550 (12, 18.5-30.5 mm SL), Brazil, Mato Grosso, Gaúcho do Norte, Xingu River basin, rio Xingu, encontro dos rios Culuene e Sete de Setembro (canal). 12º 56' 0.0" S 52º49'0.0" W, 23 Jul 1984, M Goulding, Portugal & Carvalho. LBP 25926 (2, 33.61- 36.49 mm SL), Brazil, Mato Grosso, Paranatinga, Xingu River basin, Rio Culuene. 13°50'50.8" S 53°15'40.2" W, 24 Jan 2018, NF Junior, N Estevão, FA Machado. LBP 15676 (105, 21.5-43.3 mm SL), Brazil, Mato Grosso, Ribeirão Cascalheira, Xingu River basin, Córrego do Gato. 13°09'13.6" S 51°55'18.7" W, 30 Jul 2012, C Oliveira, M Taylor, GHC Silva, JHM Martinez. ***P. capitulata***: LBP 17802 (6, 27.38–31.97 mm SL), Peru, Pucallpa, Coronel Portillo, Ucayali River basin. 08°35'44.2" S 74°48'04.3" W, 18 Jun 2013, R Britzke. ***P. tegata***: MZUSP 35889 (5, 26.9-37.9 mm SL), Brazil, Mato Grosso, Itiquira, Paraguay River basin, rio Piquiri, faz. Santo Antônio do Paraíso. 17º 12'0.0" S 54º 9" 0.0" W, 3/x/1979, J.H.B. Medeiros & J. C. Oliveira. MZUSP 96694 (10, 24.8-27.7 mm SL), Brazil, Mato Grosso, Barão do Melgaço, Paraguay River basin, rio Mutum.16º19'30"S 55º49'59"W, F.A. Machado, F.C.T. Lima et al. LBP 7641 (6, 35.8–39.1 mm SL), Brazil, Mato Grosso, Santo Antonio do Leverger, Paraguay River basin. 15°46'03.8" S 55°30'44.5" W, 01 Mar 2009, M Mehanna, PA Campos. LBP 7606 (16, 21.7–31.8 mm SL), Brazil, Mato Grosso, Barão de Melgaço, Paraguay River basin, Lagoa Marginal rio Cuiabá. 16°11'39.5" S 55°48'25.1" W, 29 Jan 2021, C Oliveira, GA Lopez, R Britzke. ***P. ojitata***:

MZUSP 30551 (36.3 mm SL), Brazil, Pará, rio Curuá, Serra do Cachimbo, rodovia Santarém-Cuiabá, poço de cachoeira. 09º22'0.0" S 54º52'0.0" W, 15 Aug 1984, M Goulding. MZUSP 100922 (28.4 - 33.5 mm SL; 2 d&c, 31.2- 30.5 mm SL), Brazil, Pará, rio Curuá, Serra do Cachimbo, rodovia Santarém- Cuiabá, poço de cachoeira. 09º22'0.0" S 54º52'0.0" W, 15 Aug 1984, M Goulding. MZUSP 97588 (30.5-48.8 mm SL), Brazil, Pará, Altamira, Xingu basin, rio Curuá, na ponte da BR163. 08°53'54" S 55°59'20" W, 29 Oct 2007, J. Birindelli, L. Sousa, A. Netto-Ferreira, M. Sabaj-Perez, N. Lujan. ***P. napoatilis***: MZUSP 38667 (9, 21.5-35 mm SL), Equador, Napo, Napo River basin, rio Jatuncocha, 2km acima da Laguna Jatuncocha, 1º3'0.00" S 75º31'4.0" W, 26 Oct 1981, D. Stewart & M. Ibarra.

4. Discussion

The present study shows the results from the first molecular analysis of the genus *Phenacogaster*. The molecular methods delimited 38 and 16 species based on the application of the bPTP and ABGD, respectively. The discrepancy in the results obtained among the analysis may be related to the variety of algorithms and implementations of each method (Luo *et al.*, 2018; Dellicour & Flot, 2018). Also, the population sizes, number of species involved, and speciation rates can affect the delimitation (Ahrens *et al.*, 2016). Despite that, we conservatively recognize only 13 valid species of *Phenacogaster* and one new species (described herein). Preliminary morphological analyses indicate that *P.* aff. *suborbitalis* (Tapajós), *P.* aff. *beni* (Madeira), and *P.* aff. *retropinna* (Arinos/Tapajós) may represent cryptic or yet unrecognized species but a broad morphological examination is needed to confirm such conclusions.

Although *Phenacogaster* is morphologically distinct of other Characinae genera, the species identification is rather complex and based mainly on color variation (Lucena & Malabarba, 2010). Widely distributed and undescribed species were recently grouped in the proposed *P. pectinata*

complex (*P. pectinata, P. microstictus, P. beni*), by presenting a set of characters such as humeral blotch present in females, humeral blotch absent or restricted to a few chromatophores in males, and the origin of the anal fin vertically that projects at the vertical line of the dorsal-fin origin, and the complete lateral line. Here, both topologies (Fig. 1; Sup. 1) partially corroborate the composition of the *P. pectinata* complex. The analyses indicate *P.* aff. *suborbitalis, P.* aff. *beni, P. capitulata, P. tegata*, and *P. prolata* forming a large clade (Fig. 1). Additionally, the phylogeny of the subfamily Characinae based on genomic-wide data supports this clade (Souza *et al.*, in review; Cap. 1).

Phenacogaster sp. n., described in the present study, is known from the upper Xingu and upper Araguaia river basins of the Brazilian Shield (Fig. 4). Lucena & Malabarba (2010) described two species with distribution in the Xingu river basin, *P. retropinna* from Negro, Madeira, Xingu and Araguaia rivers, and *P. ojitata* restricted to the Rio Curuá, a left tributary of the Xingu river basin. Here, molecular and morphological evidence support that *Phenacogaster* sp. n. represents a distinct species from *P. retropinna*. The mitochondrial data analysis showed a high genetic divergence between these two species (0.045±0.009; Tab. 1). Unfortunately, no tissue of *P. ojitata* were available for molecular analyses. However, *Phenacogaster* sp. n. can be easily distinguished from *P. ojitata* by the presence of incomplete lateral line (vs. complete) and supplemented by larger orbital diameter (34- 42.9% HL vs 33-36.9% HL). Furthermore, *Phenacogaster* sp. n. can be distinguished from other *Phenacogaster* species with incomplete lateral line by presence of humeral blotch (vs. absence of humeral blotch in *P. carteri*); the presence of humeral blotch in males and females (vs. absence humeral blotch in males in *P. napoatilis* and *P. capitulata*); humeral blotch in males and females located near pseudotympanum and distant from dorsal-fin origin (vs. humeral blotch in males and females distant from the pseudotympanum and near dorsal-fin origin in *P. tegata*).

The incomplete lateral line is a character found in four out of 23 valid species of *Phenacogaster* (Lucena & Malabarba, 2010) and independently acquired along the phylogeny (Souza et al., in

review; Cap. 1). Incompletely pored lateral line or even absence in many characids are result of loss of terminal stages in the developmental sequence (Marinho *et al.*, 2021). *Cyphocharax punctatus* (Vari & Nijssen 1986), *C. boiadeiro* Melo 2017, and *Moenkhausia andrica* Reia, Oliveira & Benine 2021 has an increase of the number of pored scales associated with the ontogenetic development of specimens (Vari, 1992; Melo, 2017, Reia *et al.*, 2021). We observed young and adults of *Phenacogaster* sp. n with incomplete lateral line with only five of 39 individuals with discontinuous lateral line scale. A comparable pattern of specimens with incomplete and discontinuous lateral line scale was not yet reported to species of *Phenacogaster*, but observed in characids such as *Hyphessobrycon cachimbensis* Travassos, 1964 and *Diapoma obi* (Casciotta, Almirón, Piálek & Rícan, 2012) (Marinho *et al.*, 2021).

This is the first molecular study of *Phenacogaster* and although the data cover 56% of the described species, the results support the recognition of the new species from the upper Xingu and Araguaia basins. Additionally, the study suggests the occurrence of an underestimated diversity in the genus and represents an important step to better delimit and recognize the species diversity in *Phenacogaster*.

Acknowledgements

We are grateful to Mary Burridge (ROM), María Fernando Castillo (STRI), Carlos Lucena (MCP), Carolina Doria (UFRO-ICT), Camila C. Ribas (INPA), Erick C. Guimarães (UFMA), and Michel D. Gianeti (MZUSP) for curatorial assistance and/or loan of tissues and vouchers.

Funding

This research was supported by the Brazilian agencies FAPESP proc. 17/06551-0 (CSS), 16/11313-8 and 18/23883-9 (BFM), 18/20610-1, 16/09204-6, 14/26508-3 and CNPq proc. 306054/2006-0 (CO).

Bibliography

Ahrens, D., Fujisawa, T., Krammer, H. J., Eberle, J., Fabrizi, S., & Vogler, A. P. (2016). Rarity and incomplete sampling in DNA-based species delimitation. Systematic Biology, 65(3), 478-494.

Antonetti, D. A., Lucena, C. A. S., & Lucena, Z. M. S. (2018). Two new species of *Phenacogaster* (Actinopterygii: Characidae) from the rio Tocantins drainage, Brazil. Verlag Dr. Friedrich Pfeil.

Barreto, S. B., Silva, A. T., Batalha-Filho, H., Affonso, P. R. A., & Zanata, A. M. (2018). Integrative approach reveals a new species of *Nematocharax* (Teleostei: Characidae). Journal of fish biology, 93(6), 1151-1162.

Dellicour, S., & Flot, J. F. (2018). The hitchhiker's guide to single-locus species delimitation. Molecular Ecology Resources, 18(6), 1234-1246.

Fink, W. L., & Weitzman, S. H. (1974). The so-called cheirodontin fishes of Central America with descriptions of two new species (Pisces: Characidae). Smithsonian Contributions to Zoology.

Fricke, R., Eschmeyer, W. N. & R. van der Laan (2020). Eschmeyer's catalog of fishes: genera, species, references. Available at http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (last accessed 30 April 2021).

Géry, J. (1972). Corrected and supplemented descriptions of certain characoid fishes described by Henry W. Fowler, with revisions of several of their genera. Studies on Neotropical Fauna and Environment, 7(1), 1-35.

Ivanova, N. V., Dewaard, J. R., & Hebert, P. D. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular ecology notes, 6(4), 998-1002.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647-1649.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of molecular evolution, 16(2), 111-120.

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(6), 1547-1549.

Lucena, Z. M. S. D., & Malabarba, L. R. (2010). Descrição de nove espécies novas de *Phenacogaster* (Ostariophysi: Characiformes: Characidae) e comentários sobre as demais espécies do gênero. Zoologia (Curitiba), 27, 263-304.

Lucena, Z. M. S. D., & Lucena, C. A. S. D. (2019). A new glass tetra species of *Phenacogaster* from the rio Salitre, rio São Francisco drainage, Brazil (Characiformes: Characidae). Neotropical Ichthyology, 17.

Luo, A., Ling, C., Ho, S. Y., & Zhu, C. D. (2018). Comparison of methods for molecular species delimitation across a range of speciation scenarios. Systematic Biology, 67(5), 830-846.

Marinho, M. M. F., Ohara, W. M., & Dagosta, F. C. P. (2021). A new species of *Moenkhausia* (Characiformes: Characidae) from the rio Madeira basin, Brazil, with comments on the evolution and development of the trunk lateral line system in characids. Neotropical Ichthyology, 19.

Mateussi, N. T. B., Melo, B. F., & Oliveira, C. (2020). Molecular delimitation and taxonomic revision of the wimple piranha *Catoprion* (Characiformes: Serrasalmidae) with the description of a new species. Journal of Fish Biology, 97(3), 668-685.

Mattox, G. M., & Toledo-Piza, M. (2012). Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). Zoological Journal of the Linnean Society, 165(4), 809-915.

Mattox et al. (2017). Subfamily Characinae. In: van der Sleen, P., & Albert, J. S. (Eds.). (2017). Field guide to the fishes of the Amazon, Orinoco, and Guianas (Vol. 115). Princeton University Press.

Machado, V. N., Collins, R. A., Ota, R. P., Andrade, M. C., Farias, I. P., & Hrbek, T. (2018). One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. Scientific reports, 8(1), 1-12.

Melo, B. F., Ochoa, L. E., Vari, R. P., & Oliveira, C. (2016). Cryptic species in the Neotropical fish genus *Curimatopsis* (Teleostei, Characiformes). Zoologica Scripta, 45(6), 650-658.

Ota, R. P., Machado, V. N., Andrade, M. C., Collins, R. A., Farias, I. P., & Hrbek, T. (2020). Integrative taxonomy reveals a new species of pacu (Characiformes: Serrasalmidae: *Myloplus*) from the Brazilian Amazon. Neotropical Ichthyology, 18.

Pires, A. A., Ramirez, J. L., Galetti, P. M., Troy, W. P., & Freitas, P. D. (2017). Molecular analysis reveals hidden diversity in *Zungaro* (Siluriformes: Pimelodidade): a genus of giant South American catfish. Genetica, 145(3), 335-340.

Puillandre, N., Lambert, A., Brouillet, S., & ACHAZ, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular ecology, 21(8), 1864-1877.

Ramirez, J. L., Rosas-Puchuri, U., Cañedo, R. M., Alfaro-Shigueto, J., Ayon, P., Zelada-Mázmela, E., ... & Velez-Zuazo, X. (2020). DNA barcoding in the Southeast Pacific marine realm: Low coverage and geographic representation despite high diversity. PloS one, 15(12), e0244323.

Xia, X. (2013). DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Molecular biology and evolution, 30(7), 1720-1728.

Xia, X., Xie, Z., Salemi, M., Chen, L., & Wang, Y. (2003). An index of substitution saturation and its application. Molecular phylogenetics and evolution, 26(1), 1-7.

Taylor, W. R., & Van Dyke, G. C. (1985). Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage study. Cybium (Paris), 9(2), 107-119.

Vari, R. P. (1992). Systematics of the neotropical characiform genus *Cyphocharax* Fowler (Pisces: Ostariophysi). Smithsonian Contributions to Zoology.

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29(22), 2869-2876.

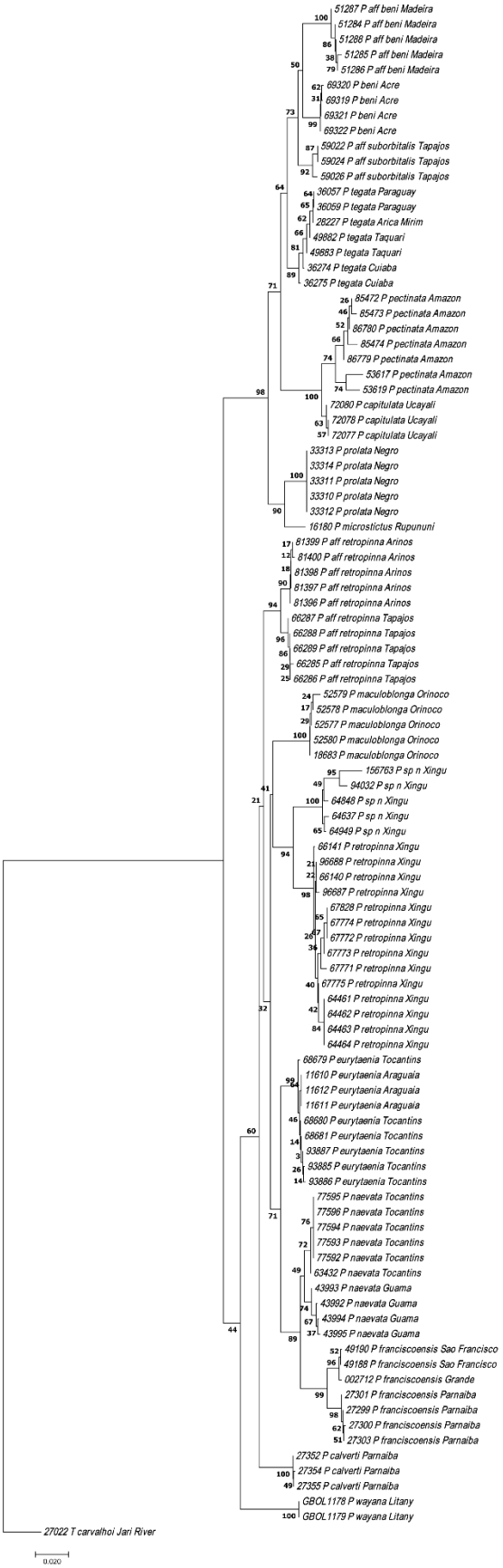
Ward, R. D., Hanner, R., & Hebert, P. D. (2009). The campaign to DNA barcode all fishes, FISH-BOL. Journal of fish biology, 74(2), 329-356.

Supplementary information

Supplementary Table 1. List of the all specimens used in species delimitation analyses.

Museum number	Voucher	Species	Drainage	Country	Coordinates	Genbank number
LBP 19200	77592, 77593, 77594, 77595, 77596	*P. naevata*	Tocantins River	Brazil	11°03'14.3"S 48°34'22.0"W	
LBP 15353	63432	*P. naevata*	Tocantins River	Brazil	10°12'43.8"S 40°27'09.3"W	
LBP 9385	43992, 43993, 43994, 43995	*P. naevata*	Guamá River	Brazil	01°34'00.5"S 47°09'51.4"W	
LBP 10487	49188, 49190	*P. franciscoensis*	São Francisco River	Brazil	17°14'32.3'S 46°28'03.8"W	
MZUSP 114504	002712	*P. franciscoensis*	Grande River	Brazil	12°17'57.7"S 45°00'56.8"W	
LBP 5582	27299, 27300, 27301, 27303	*P. franciscoensis*	Parnaíba River	Brazil	09°09'51"S 45°51'15"W	
LBP 1507	11610, 11611, 11612	*P. eurytaenia*	Araguaia River	Brazil	15°52'40.4" 52°18'15.5"W	
LBP 17173	68679, 68680, 68681	*P. eurytaenia*	Tocantins River	Brazil	14°58'37.4"S 48°40'42.1"W	
LBP 25604	93885, 93886, 93887	*P. eurytaenia*	Tocantins River	Brazil	12°37'27.9"S 48°02'43.5"W	
LBP 15966	66140, 66141	*P. retropinna*	Xingu River	Brazil	13°29'41.8"S 53°04'57.7"W	
LBP 15676	64461, 64462, 64463, 64464	*P. retropinna*	Xingu River	Brazil	13°09'13.6"S 51°55'18.7"W	
LBP 25926	96687, 96688	*P. retropinna*	Xingu River	Brazil	13°50'50.8"S 53°15'40.2"W	
LBP 16691	67828	*P. retropinna*	Xingu River	Brazil	03°24'19.8"S 52°05'48.0"W	
LBP 16670	67771, 67772, 67773, 67774, 67775	*P. retropinna*	Xingu River	Brazil	03°30'14.3"S 52°02'19.9"W	
LBP 18683	52577, 52578, 52579, 52580,18683	*P. maculoblonga*	Orinoco River	Colômbia	3°16'17.4"N 73°36'47.0"W	
LBP 15734	64637	*Phenacogaster* sp. n	Xingu River	Brazil	12°53'04.3"S 52°02'00.3"W	
LBP 25217	94032	*Phenacogaster* sp. n	Xingu River	Brazil	08°39'06.9"S 55°02'09.1"W	
LBP 15807	64848, 64849	*Phenacogaster* sp. n	Xingu River	Brazil	12°33'20.5"S 52°16'16.1"W	
LBP 15676	156763	*Phenacogaster* sp. n	Xingu River	Brazil	13°09'13.6"S 51°55'18.7"W	
LBP 5612	27352, 27354, 27355	*P. calverti*	Parnaiba River	Brazil	7°30'55.0"S 46°04'53.0"W	
LBP 20843	81396, 81397, 81398, 81399, 81400	*P.* aff. *retropinna*	Arinos River	Brazil	13°48'04.8"S 56°01'37.5"W	
LBP 16035	66285, 66286, 66287, 66288, 66289	*P.*aff. *retropinna*	Xingu River	Brazil	14°25'33.8"S 54°00'56.6"W	
MHNG 2759.079	GBOL1178, GBOL1179	*P. wayana*	Litany River	French Guiana	2°56'16.7"N 54°10'20.6"W	MZ051714, MZ051920
LBP 6931	33310, 33311, 33312, 33313, 33314	*P. prolata*	Negro River	Brazil	0°04'66.5"N 66°48'54.6"W	
ROM uncat	16180	*P. microstictus*	Essequibo River	Guyana	2°09'33.5"N 59°17'33.5"W	

Museum number	Voucher	Species	Drainage	Country	Coordinates	Genbank number
LBP 22680	85472, 85473, 85474	*P. pectinata*	Amazon River	Brazil	4°12'02.7"S 69°55'35.3"W	
LBP 22418	86779, 86780	*P. pectinata*	Amazon River	Colômbia	4°07'33.8"S 70°00'28.9"W	
LBP 12406	53617, 53619	*P. pectinata*	Amazon River	Peru	4°09'04.2"S 73°28'25.7"W	
LBP 17802	72077, 72078, 72080	*P. capitulata*	Ucayali River	Peru	8°35'44.2"S 74°48'04.3"W	
LBP 5795	28227	*P. tegata*	Aricá-Mirim River	Brazil	15°44'03.6"S 55°52'48.7"W	
LBP 10785	49882, 49883	*P. tegata*	Paraguay River	Brazil	18°25'24.4"S 54°50'05.9"W	
LBP 7606	36274, 36275	*P. tegata*	Paraguay River	Brazil	16°11'39.5"S 55°48'25.1"W	
LBP 7641	36057, 36059	*P. tegata*	Paraguay River	Brazil	15°46'03.8"S 55°30'44.5"W	
LBP 14095	59022, 59024, 59026	*P.* aff. *suborbitalis*	Tapajós River	Brazil	4°55'58.8"S 56°51'51.6"W	
LBP 16865	69319, 69320, 69321, 69322,	*P. beni*	Acre River	Brazil	10°04'44.3"S 67°32'33.9"W	
LBP 12024	51284, 51285, 51286, 51287, 51288	*P.* aff. *beni*	Madeira River	Brazil	7°56'06.6"S 63°27'21.2"W	
LBP 5376	27022	*T. carvalhoi*	Jari River	Brazil	0°33'51.0"S 52°34'45.0"W	HM070393.1



Supplementary Fig. 1. NJ tree of species of *Phenacogaster*, based on the COI gene.


16/12/2021 10:37 abgd web

Initial Partition with prior maximal distance P=7.74e-03 ; Barcode gap distance = 0.021
Distance K80 Kimura MinSlope=1.000000
Download (left click and save) or see below the tree file corresponding to this partition: click here

Group[1 | n: 1 ;id: 16180_P_microstictus_Rupununi
Group[2 | n: 5 ;id: 33312_P_prolata_Negro 33310_P_prolata_Negro 33311_P_prolata_Negro 33313_P_prolata_Negro 33314_P_prolata_Negro
Group[3 | n: 10 ;id: 86780_P_pectinata_Amazon 85472_P_pectinata_Amazon 85473_P_pectinata_Amazon 86779_P_pectinata_Amazon 85474_P_pectinata_Amazon 53617_P_pectinata_Amazon 53619_P_pectinata_Amazon 72078_P_capitulata_Ucayali 72077_P_capitulata_Ucayali 72080_P_capitulata_Ucayali
Group[4 | n: 4 ;id: 69320_P_beni_Acre 69322_P_beni_Acre 69321_P_beni_Acre 69319_P_beni_Acre
Group[5 | n: 3 ;id: 59026_P_aff_suborbitalis_Tapajos 59022_P_aff_suborbitalis_Tapajos 59024_P_aff_suborbitalis_Tapajos
Group[6 | n: 5 ;id: 51285_P_aff_beni_Madeira 51284_P_aff_beni_Madeira 51287_P_aff_beni_Madeira 51288_P_aff_beni_Madeira 51286_P_aff_beni_Madeira
Group[7 | n: 7 ;id: 36275_P_tegata_Cuiaba 36274_P_tegata_Cuiaba 36057_P_tegata_Paraguay 36059_P_tegata_Paraguay 28227_P_tegata_Arica_Mirim 49883_P_tegata_Taquari 49882_P_tegata_Taquari
Group[8 | n: 2 ;id: GBOL1178_P_wayana_Litany GBOL1179_P_wayana_Litany
Group[9 | n: 14 ;id: 67771_P_retropinna_Xingu 67773_P_retropinna_Xingu 67828_P_retropinna_Xingu 67772_P_retropinna_Xingu 67774_P_retropinna_Xingu 64461_P_retropinna_Xingu 64462_P_retropinna_Xingu 64463_P_retropinna_Xingu 64464_P_retropinna_Xingu 66141_P_retropinna_Xingu 96687_P_retropinna_Xingu 67775_P_retropinna_Xingu 96688_P_retropinna_Xingu 66140_P_retropinna_Xingu
Group[10 | n: 5 ;id: 156763_P_sp_n_Xingu 94032_P_sp_n_Xingu 64637_P_sp_n_Xingu 64848_P_sp_n_Xingu 64949_P_sp_n_Xingu
Group[11 | n: 5 ;id: 52579_P_maculoblonga_Orinoco 18683_P_maculoblonga_Orinoco 52580_P_maculoblonga_Orinoco 52577_P_maculoblonga_Orinoco 52578_P_maculoblonga_Orinoco
Group[12 | n: 3 ;id: 27354_P_calverti_Parnaiba 27352_P_calverti_Parnaiba 27355_P_calverti_Parnaiba
Group[13 | n: 10 ;id: 81396_P_aff_retropinna_Arinos 81397_P_aff_retropinna_Arinos 81398_P_aff_retropinna_Arinos 81399_P_aff_retropinna_Arinos 81400_P_aff_retropinna_Arinos 66285_P_aff_retropinna_Tapajos 66287_P_aff_retropinna_Tapajos 66286_P_aff_retropinna_Tapajos 66288_P_aff_retropinna_Tapajos 66289_P_aff_retropinna_Tapajos
Group[14 | n: 7 ;id: 27301_P_franciscoensis_Parnaiba 27299_P_franciscoensis_Parnaiba 27300_P_franciscoensis_Parnaiba 27303_P_franciscoensis_Parnaiba 002712_P_franciscoensis_Grande 49190_P_franciscoensis_Sao_Francisco 49188_P_franciscoensis_Sao_Francisco
Group[15 | n: 9 ;id: 68679_P_eurytaenia_Tocantins 93887_P_eurytaenia_Tocantins 68680_P_eurytaenia_Tocantins 68681_P_eurytaenia_Tocantins 93885_P_eurytaenia_Tocantins 11611_P_eurytaenia_Araguaia 11610_P_eurytaenia_Araguaia 11612_P_eurytaenia_Araguaia 93886_P_eurytaenia_Tocantins
Group[16 | n: 10 ;id: 77592_P_naevata_Tocantins 77593_P_naevata_Tocantins 77594_P_naevata_Tocantins 77595_P_naevata_Tocantins 77596_P_naevata_Tocantins 63432_P_naevata_Tocantins 43992_P_naevata_Guama 43994_P_naevata_Guama 43993_P_naevata_Guama 43995_P_naevata_Guama

https://bioinfo.mnhn.fr/abi/public/abgd/temp/14888.1787332702/groupe5.init.html 1/1

Supplementary Fig. 2. ABGD dataset reporting the number of groups delimitated to *Phenacogaster*.

```
# Max likilhood partition
Species 1 (support = 0.474)
     GBOL1178_P_wayana_Litany,GBOL1179_P_wayana_Litany

Species 2 (support = 0.127)

96687_P_retropinna_Xingu,67771_P_retropinna_Xingu,67773_P_retropinna_Xingu,67828_P_retropinna_Xingu,67772_P_retropinna_Xingu,67774_P_retropinna_Xingu,64461_P_retropinna_Xingu,64462
_P_retropinna_Xingu,64463_P_retropinna_Xingu,64464_P_retropinna_Xingu,66141_P_retropinna_Xingu,67775_P_retropinna_Xingu,96688_P_retropinna_Xingu,66140_P_retropinna_Xingu

Species 3 (support = 0.202)

68681_P_eurytaenia_Tocantins,93885_P_eurytaenia_Tocantins,68680_P_eurytaenia_Tocantins,68679_P_eurytaenia_Tocantins,93887_P_eurytaenia_Tocantins,93886_P_eurytaenia_Tocantins,11611_
P_eurytaenia_Araguaia,11610_P_eurytaenia_Araguaia,11612_P_eurytaenia_Araguaia

Species 4 (support = 0.081)

85472_P_pectinata_Amazon,85473_P_pectinata_Amazon,86780_P_pectinata_Amazon,86779_P_pectinata_Amazon,85474_P_pectinata_Amazon,53617_P_pectinata_Amazon,53619_P_pectinata_Amazon,72077
_P_capitulata_Ucayali,72078_P_capitulata_Ucayali,72080_P_capitulata_Ucayali

Species 5 (support = 0.358)
     43993_P_naevata_Guama,43995_P_naevata_Guama,43994_P_naevata_Guama,43992_P_naevata_Guama

Species 6 (support = 0.299)
     002712_P_franciscoensis_Grande,49188_P_franciscoensis_Sao_Francisco,49190_P_franciscoensis_Sao_Francisco

Species 7 (support = 0.602)
     36275_P_tegata_Cuiaba,49883_P_tegata_Taquari

Species 8 (support = 0.939)
     16180_P_microstictus_Rupununi

Species 9 (support = 0.404)
     59026_P_aff_suborbitalis_Tapajos,59022_P_aff_suborbitalis_Tapajos,59024_P_aff_suborbitalis_Tapajos

Species 10 (support = 0.349)
     69320_P_beni_Acre,69322_P_beni_Acre,69321_P_beni_Acre,69319_P_beni_Acre

Species 11 (support = 0.721)
     36274_P_tegata_Cuiaba

Species 12 (support = 0.543)
     36057_P_tegata_Paraguay

Species 13 (support = 0.357)
     36059_P_tegata_Paraguay

Species 14 (support = 0.333)
     156763_P_sp_n_Xingu,94032_P_sp_n_Xingu,64637_P_sp_n_Xingu

Species 15 (support = 0.407)
     64848_P_sp_n_Xingu,64949_P_sp_n_Xingu

Species 16 (support = 0.752)
     33312_P_prolata_Negro

Species 17 (support = 0.569)
     33310_P_prolata_Negro

Species 18 (support = 0.378)
     33311_P_prolata_Negro

Species 19 (support = 0.714)
     27299_P_franciscoensis_Parnaiba

Species 20 (support = 0.792)
     18683_P_maculoblonga_Orinoco

Species 21 (support = 0.651)
     27355_P_calverti_Parnaiba

Species 22 (support = 0.160)
     77596_P_naevata_Tocantins,63432_P_naevata_Tocantins,77595_P_naevata_Tocantins,77594_P_naevata_Tocantins,77592_P_naevata_Tocantins

Species 23 (support = 0.831)
     77593_P_naevata_Tocantins

Species 24 (support = 0.598)
     52579_P_maculoblonga_Orinoco

Species 25 (support = 0.169)
     28227_P_tegata_Arica_Mirim

Species 26 (support = 0.169)
     49882_P_tegata_Taquari

Species 27 (support = 0.229)
     51288_P_aff_beni_Madeira,51286_P_aff_beni_Madeira,51287_P_aff_beni_Madeira,51285_P_aff_beni_Madeira

Species 28 (support = 0.734)
     51284_P_aff_beni_Madeira

Species 29 (support = 0.331)
     27354_P_calverti_Parnaiba

Species 30 (support = 0.331)
     27352_P_calverti_Parnaiba

Species 31 (support = 0.184)
     81399_P_aff_retropinna_Arinos,81400_P_aff_retropinna_Arinos,81398_P_aff_retropinna_Arinos,81397_P_aff_retropinna_Arinos,81396_P_aff_retropinna_Arinos

Species 32 (support = 0.250)
     66285_P_aff_retropinna_Tapajos,66287_P_aff_retropinna_Tapajos,66286_P_aff_retropinna_Tapajos,66288_P_aff_retropinna_Tapajos,66289_P_aff_retropinna_Tapajos

Species 33 (support = 0.202)
     52577_P_maculoblonga_Orinoco,52578_P_maculoblonga_Orinoco

Species 34 (support = 0.400)
     52580_P_maculoblonga_Orinoco

Species 35 (support = 0.468)
     27301_P_franciscoensis_Parnaiba

Species 36 (support = 0.238)
     27300_P_franciscoensis_Parnaiba,27303_P_franciscoensis_Parnaiba

Species 37 (support = 0.178)
     33313_P_prolata_Negro

Species 38 (support = 0.178)
     33314_P_prolata_Negro
```


Supplementary Fig. 3. Bayesian Poisson Tree Processes (bPTP) delimitation of species of *Phenacogaster*.

Chapter 4

Molecular species delimitation of the Cynopotamini genera *Acestrocephalus*, *Cynopotamus*, and *Galeocharax* (Teleostei: Characidae: Characinae)

Molecular species delimitation of the Cynopotamini genera *Acestrocephalus, Cynopotamus,* and *Galeocharax* (Teleostei: Characidae: Characinae)

Camila S. Souza

Departamento de Biologia Estrutural e Funcional, Instituto de Biociências, Universidade Estadual Paulista, R. Prof. Dr. Antônio C. W. Zanin 250, 18618-689 Botucatu SP, Brazil

Abstract

Molecular investigations using DNA barcoding have improved our knowledge of the Neotropical ichthyofauna. Here we analyzed 46 *cytochrome c oxidase subunit I* partial sequence representing of 12 species of Cynopotamini (37% *Acestrocephalus*, 50% *Cynopotamus* and 100% *Galeocharax*) and used species delimitation methods to test the morphological hypotheses of the species limits and the presence of multiple biological units. The analyses identified different numbers of genetic lineages (ABGD = 18 lineages and bPTP = 10 lineages) and both methods indicated potential candidate species to genus *Galeocharax*.

Keywords: Fish, Diversity, Systematics, DNA barcode, COI

1. Introduction

The characiform Characinae comprises 85 valid species in eight genera broadly distributed throughout river basins of the Neotropical region (*sensu* Mattox & Toledo-Piza, 2012). Species of the subfamily are easily recognized by their deep anterior bodies with a characteristic gibbosity,

except *Phenacogaster* and *Acestrocephalus* (Lucena & Menezes, 2003), and by the presence of superficial neuromasts over the head (Mattox *et al.*, 2017). The subfamily is monophyletic and arranged in three tribes: 1) Characini with *Charax* Scopoli, 1977 and *Roeboides* Günther, 1864; 2) Cynopotamini with *Acanthocharax* Eigenmann, 1912, *Acestrocephalus* Eigenmann, 1910, *Cynopotamus* Valenciennes in Cuvier & Valenciennes, 1850, and *Galeocharax* Fowler, 1910; 3) Phenacogasterini with *Phenacogaster* Eigenmann, 1907 (*sensu* Mattox & Toledo-Piza, 2012). More recently, based in phylogenomic data, Souza *et al.* (in review; Cap. 1) proposed a new tribe, Acanthocharacini, to allocated *Acanthocharax*.

Cynopotamini is supported by eight non-ambiguous synapomorphies, in which the presence of spinoid scales is one of the clearest features for the three included genera (Mattox & Toledo-Piza, 2012). *Galeocharax* is monophyletic and supported by five synapomorphies (Mattox & Toledo-Piza, 2012) and externally identified by the presence of scales on the caudal fin, two series of dentary teeth, and the lateral expansion of the tubular portion of the nasal bone (Mattox *et al.*, 2017). Giovannetti *et al.* (2017) revised *Galeocharax* and recognized three valid species: *Galeocharax humeralis* (Valenciennes, 1834) from the Rio Paraguai and the lower Rio Paraná, *Galeocharax gulo* (Cope, 1870) from the Amazon, Orinoco, Oyapock, Araguaia, Tocantins, and upper Paraná river basins, and *Galeocharax goeldii* (Fowler, 1913) from the Rio Madeira basin River basin. The authors also proposed *G. knerii* (Steindachner, 1879) from the upper rio Paraná basin, as a junior synonym of *G. gulo*.

The second genus is *Acestrocephalus*, characterized by contralateral series of longitudinal series of scales continuous on ventral surface, usually rounded and not forming a keel or gap of scales, the ascending process of parasphenoid reaching posterior margin of pterosphenoid, and anterior tip of nasal reaching horizontal arm of premaxilla near base of tooth series (Mattox & Toledo-Piza, 2012). *Acestrocephalus* is distributed from northern to central South America and is currently represented by eight species: *A. acutus* Menezes, 2006, *A. anomalus* (Steindachner, 1880), *A. boehlkei* Menezes,

1977, *A. maculosus* Menezes, 2006, *A. nigrifasciatus* Menezes, 2006, *A. pallidus* Menezes, 2006, *A. sardina* (Fowler, 1913), and *A. stigmatus* Menezes, 2006.

The third genus in Cynopotamini is *Cynopotamus*, characterized externally by a relatively large body sizes, one tooth series on dentary, spinous scales and predorsal scales present (Mattox *et al.*, 2017) and internally by presence an epiphyseal bar with anterior and posterior lamellar expansions from frontals, anterior margin of palatine separated by aperture delimiting medial rounded portion and straight lateral portion in contact with neurocranium and maxilla, respectively, presence of process on lateral margin of ectopterygoid (Mattox Toledo-Piza, 2012). *Cynopotamus* has 13 species: *C. amazonum* (Günther, 1868), *C. argenteus* (Valenciennes, 1837), *C. atratoensis* (Eigenmann, 1907), *C. bipunctatus* Pellegrin, 1909, *C. essequibensis* Eigenmann, 1912, *C. gouldingi* (Menezes, 1987), *C. juruenae* Menezes, 1987, *C. kincaidi* (Schultz, 1950), *C. magdalenae* (Steindachner 1879), *C. tocantinensis* Menezes, 1987, *C. venezuelae* (Schultz, 1944), *C. xinguano* Menezes, 2007 distributed in South America.

Although the tribe Cynopotamini possesses a high morphological diversity, it has never been subject of a molecular investigation aiming to investigate species boundaries. Recent molecular investigations in the Neotropical ichthyofauna have found the presence of multiple genetic lineages within species and have improved our current knowledge in terms of biological diversity (Pereira *et al.*, 2013; Gomes *et al.*, 2015; Díaz *et al.*, 2016; Rossini *et al.*, 2016; Ramirez *et al.*, 2017; Machado *et al.*, 2018; Mateussi *et al.*, 2019; Jennings *et al.*, 2019). Here, we generated mitochondrial data for 12 species of the tribe and used species delimitation analyses to test the morphological hypotheses of the species limits for every analyzed species and the presence of multiple biological units.

2. Materials and Methods

Taxon sampling

Taxon sampling was represented by 12 valid species of Cynopotamini and were collected or obtained from ichthyological collections, and morphologically identified by identification keys (Giovannetti *et al*. 2017; Menezes, 2006, 2007). All fishes were collected following the Brazilian laws through SISBIO/MMA permit n. 3,245 and procedures for collection, maintenance and analyses followed the international guidelines for animal experiments through CEEAA IBB/UNESP protocol n. 304. Voucher specimens were fixed in 95% ethanol or 10% formalin and transferred to 70% ethanol for permanent storage, and posteriorly deposited in the collection of the Laboratório de Biologia e Genética de Peixes, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, Brazil (LBP).

DNA Extraction, Amplification and Sequencing

DNA was extracted from muscles, liver or gills, using the extraction method of the Ivanova *et al*. (2006). Partial sequences of the *cytochrome c oxidase subunit I* (*COI*) gene were amplified by polymerase chain reaction (PCR) with primers FishF1, FishR1 and FishF2, FishR2 (Ward *et al*., 2005) or L6252-Asn and H7271-COXI (Melo *et al*., 2011). PCR amplifications were performed in a total volume of 12.5 µl that included 1.25 µl of 10X buffer, 0.25 µl of $MgCl_2$ (50 mM), 0.2 µl dNTPs (2 mM), 0.5 µl of each primer (5 mM), 0.1 µl of PHT Taq DNA polymerase (Phoneutria), 1.0 µl of genomic DNA (200 ng) and 8.7 µl ddH_2O. PCR conditions consisted of an initial denaturation (5 min at 94°C), followed by 30 cycles of chain denaturation (50s at 94°C), primer hybridization (45s at 50-54°C) and nucleotide extension (1 min at 68°C), and the final extension (10

min at 68°C). All PCR products were checked on 1% agarose gels and purified with ExoSap-IT (USB Corporation) following the manufacturer's instructions. The purified PCR products were submitted to sequencing reactions using BigDye Terminator v 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and purified again by ethanol precipitation. Products were loaded onto an ABI 3130 DNA Analyzer automatic sequencer (Applied Biosystems).

Molecular analysis

Sequences of each forward and reverse strands were assembled using Geneious v7.1.9 (Kearse *et al.*, 2012) and contigs were subsequently aligned with Muscle algorithm under default parameters (Edgar, 2004). Substitution saturation was evaluated using the method of Xia *et al*. (2003) in DAMBE v5.3.38 (Xia, 2013). Nucleotide variation, substitution patterns and the best-fit model of nucleotide evolution were estimated in MEGA v10 (Kumar *et al*., 2018). The values of genetic distance were calculated in MEGA v10 using the TrN+G model (Tamura, 1992), as suggested by the best-fit model test in MEGA. Groups were ordered based on preliminary topologies.

The maximum likelihood (ML) analysis was performed under RAxML HPC-PTHREADS-SSE3 (Stamatakis, 2006) using five random parsimony trees with the GTR+G model (Stamatakis *et al*., 2008) on 2x 40 CPU 128GB *Brycon* server at LBP/UNESP. We also estimated a neighbor-joining tree (NJ) with 1000 bootstrap replicates, using the TrN+G (Tamura, 1992) in MEGA v10 (Kumar *et al*., 2018). Two species delimitation methods were performed in this study: 1) Automatic Barcode Gap Discovery analysis (ABGD; Puillandre *et al.* , 2012) available in the ABGD webserver (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html) using Kimura (K80; 2.0), evolutionary model with X = 1.0, Pmin = 0.001 and Pmax = 0.1; 2) Bayesian Poisson Tree Process (bPTP; Zhang *et al*., 2013) using the best maximum-likelihood (ML) tree, 100,000 generations, and other parameters at default in the bPTP webserver (http://species.h-its.org/ptp/). We used DnaSP v5

(Librado & Rozas, 2009) with two reduced matrices excluding sites with missing data to estimate the diversity, number, and distribution of haplotypes. Thereafter, we generated two haplotype networks using the median-joining analysis (Bandelt, Forster, & Röhl, 1999) available in PopART 1.7 (Leigh & Bryant, 2015): one for *Acestrocephalus* plus *Cynopotamus*, and the other for *Galeocharax.*

3. Results

A total of 46 partial sequences of *COI* were analyzed: *Acestrocephalus sardina* (2 specimens), *A. nigrifasciatus* (1), *A. acutus* (1), *Galeocharax gulo* (12), *G. goeldii* (2), *G. humeralis* (8), *Cynopotamus argenteus* (5), *C. juruenae* (2), *C. essequibensis* (4), *C. xinguano* (1), *C. gouldingi* (2), *C. goeldii* (2). Four additional sequences were used as outgroups: *Acanthocharax microlepis* (1), *Phenacogaster calverti* (1), *P. franciscoensis* (1), *Tetragonopterus carvalhoi* (1) (Supplementary table 1). The corresponding number of species for each genus was *Acestrocephalus* (37.5%), *Cynopotamus* (50%), and *Galeocharax* (100%). The final matrix had 540 base pairs (bp) with a total 182 variable sites. The nucleotide composition was 25.8% adenine (A), 24.8% cytosine (C), 16.6% guanine (G) and 32.7% thymine (T). DAMBE indicated no saturation in either transitions or transversions in both asymmetrical (Iss.cAsym) and symmetrical (Iss.cSym) topologies.

For our dataset, ML and NJ tree presented the same topology, and ABGD and bPTP delimited different numbers of genetic lineages (Figure 1; Supplementary Figs. 1-2-3). ABGD indicated 18 genetic lineages and revealed three lineages within *Galeocharax gulo* (*G. gulo* from Solimões River; *G. gulo* from Araguaia River; *G. gulo* from Paraná River) (Fig. 1). Differently, bPTP delimited 10 genetic lineages not differentiating species of *Galeocharax* and *Cynopotamus* (Fig. 1). The values

of interspecific distances among the lineages of *Galeocharax gulo* ranged from 0.017±0.005 to 0.049±0.010 and intraspecific distances ranged from 0 to 0.002±0.001 (Table 1).

Population genetic analyses for the reduced matrix of *Acestrocephalus* plus *Cynopotamus* resulted in a total of 10 haplotypes with haplotype diversity (Hd) of 0.9000 and *Galeocharax* with 13 haplotypes and Hd = 0.8893.

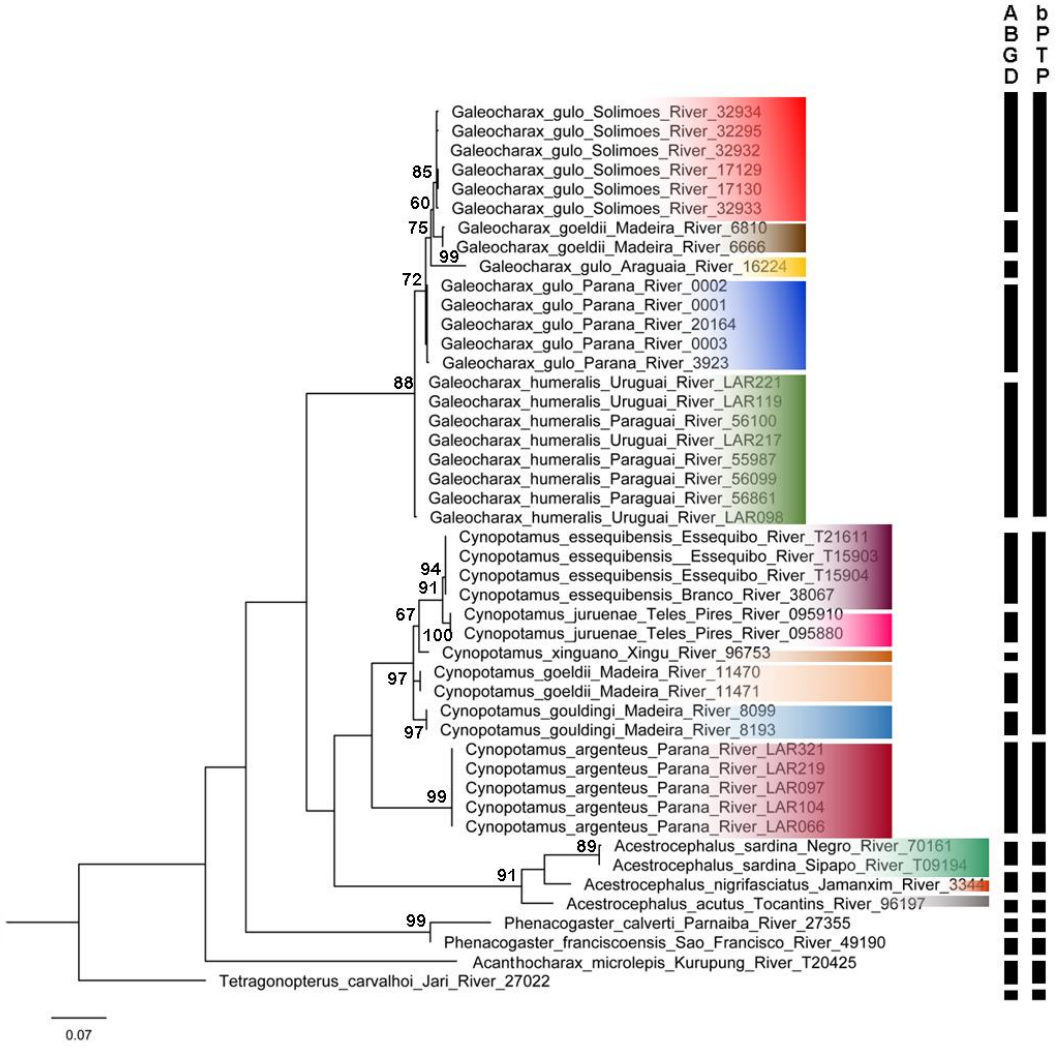


Figure 1. ML tree of species of *Cynopotamus*, *Acestrocephalus* and *Galeocharax*, based on the *COI* gene (543 pb). Vertical bars at right represent the number of species obtained by the ABGD and bPTP analyses. Numbers near nodes represent bootstrap support (Values <50% are not shown). Numbers of the specimen are after tip names.

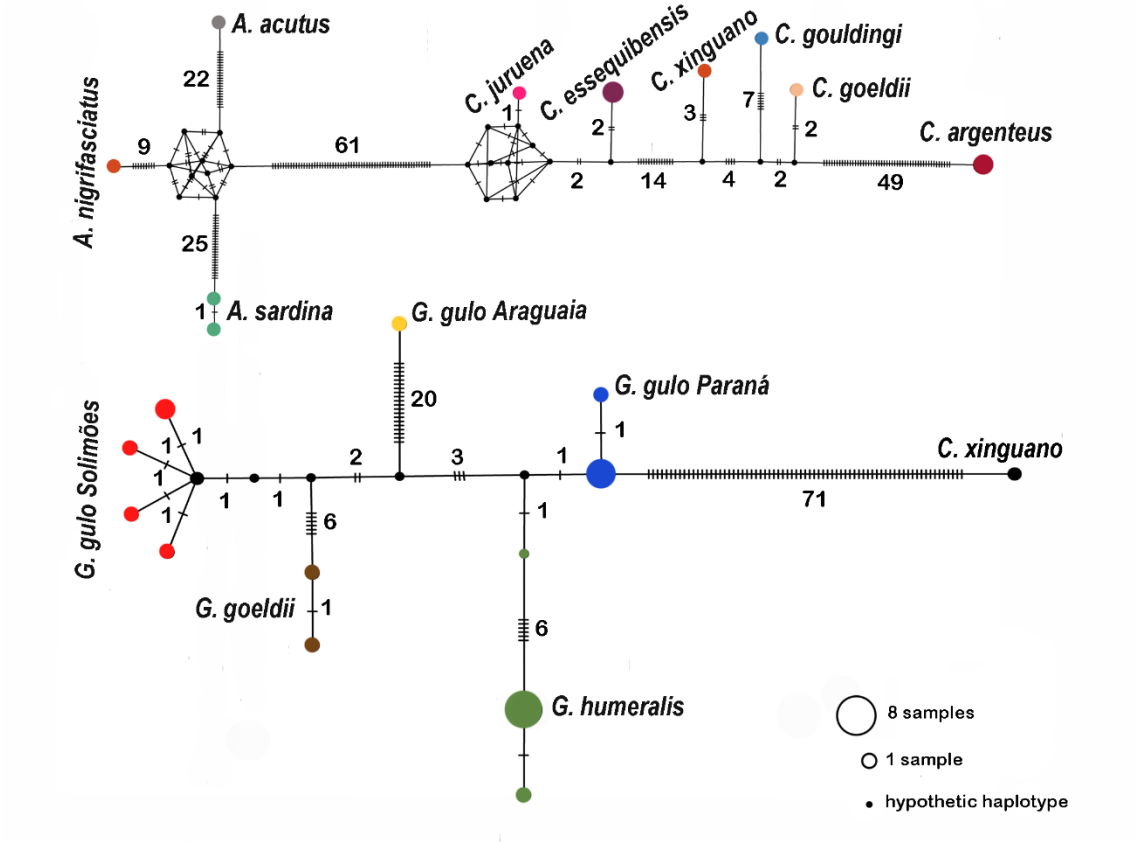


Figure 2. Haplotype network analyses showing the distribution of the distinct haplotypes of species of *Cynopotamus*, *Acestrocephalus* and lineages and species of *Galeocharax*.

Table 1. Genetic distances among species of *Acanthocharax, Cynopotamus*, *Acestrocephalus* and *Galeocharax.* Intraspecific genetic variations are highlighted in bold in the last column. Numbers below diagonal are values of interspecific distances and numbers above diagonal are respective values of standard deviation.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**Intraspecific genetic distances**
1- *Acanthocharax microlepis*		0.036	0.039	0.040	0.038	0.037	0.034	0.036	0.034	0.037	0.034	0.033	0.035	0.033	0.034	-
2- *Acestrocephalus sardina*	0.245		0.014	0.019	0.029	0.021	0.022	0.022	0.023	0.021	0.024	0.025	0.025	0.025	0.028	**0.001±0.001**
3- *A. nigrifasciatus*	0.250	0.088		0.013	0.026	0.020	0.024	0.022	0.022	0.021	0.024	0.023	0.024	0.024	0.028	-
4- *A. acutus*	0.274	0.129	0.081		0.026	0.020	0.022	0.023	0.023	0.020	0.022	0.022	0.023	0.022	0.026	-
5- *Cynopotamus argenteus*	0.235	0.253	0.218	0.207		0.021	0.020	0.018	0.019	0.020	0.023	0.023	0.022	0.024	0.026	**0**
6- *C. juruenae*	0.238	0.195	0.193	0.181	0.153		0.010	0.011	0.011	0.005	0.022	0.021	0.022	0.021	0.024	-
7- *C. xinguano*	0.219	0.203	0.223	0.188	0.144	0.051		0.007	0.008	0.009	0.022	0.021	0.022	0.021	0.025	-
8- *C. goeldii*	0.232	0.211	0.215	0.195	0.126	0.055	0.027		0.007	0.010	0.023	0.022	0.022	0.021	0.025	**0**
9- *C. gouldingi*	0.219	0.211	0.215	0.188	0.138	0.058	0.034	0.025		0.010	0.024	0.023	0.023	0.022	0.027	-
10- *C. essequibensis*	0.240	0.196	0.196	0.177	0.151	0.013	0.043	0.049	0.051		0.022	0.021	0.021	0.021	0.023	-
11- *Galeocharax goeldii*	0.225	0.213	0.201	0.187	0.196	0.189	0.190	0.193	0.201	0.183		0.008	0.005	0.007	0.011	**0.001±0.001**
12- *G. humeralis*	0.210	0.216	0.193	0.180	0.186	0.173	0.176	0.180	0.187	0.174	0.033		0.007	0.005	0.011	**0.004±0.004**
13- *G. gulo - Solimões River*	0.224	0.226	0.203	0.189	0.185	0.184	0.184	0.187	0.194	0.178	0.018	0.025		0.005	0.010	**0.002±0.001**
14- *G. gulo - Paraná River*	0.211	0.219	0.199	0.179	0.195	0.176	0.169	0.173	0.178	0.170	0.025	0.016	0.017		0.010	**0.001±0.001**
15- *G. gulo - Araguaia River*	0.210	0.248	0.238	0.223	0.221	0.189	0.207	0.210	0.207	0.183	0.054	0.063	0.050	0.049		-

The haplotype distribution of the matrix with *Acestrocephalus* and *Cynopotamus* was: one haplotype for *A. nigrifasciatus, A. acutus, C. argenteus, C. juruenae, C. essequibensis C. xinguano* and two haplotypes for *A. sardina* and *C. gouldingi* (Fig. 2). The matrix of *Galeocharax* has one haplotype for *Cynopotamus xinguano* (outgroup) and 12 haplotypes for *Galeocharax*. *G. gulo* from Araguaia River presented one haplotype, *G. gulo* from Paraná River, G. *humeralis* and *G. goeldii* presented two haplotypes each, and *G. gulo* from Solimões River presented five haplotypes (Fig. 2). Analyses indicated lack of haplotype sharing among lineages of *Galeocharax gulo*.

4. Discussion

The present study provides the first molecular data for species identification and delimitation of 12 species for the Cynopotamini genera *Acestrocephalus, Cynopotamus,* and *Galeocharax*. The ABGD and bPTP presented distinct results for the numbers of lineages/species (Fig. 1). However, the ML, NJ and ABGD results identified the presence of three genetic lineages within the widely distributed *Galeocharax gulo.* The only interspecific genetic divergence values lower than 2% was between *G. gulo* from the Paraná river and *G. gulo* from the Solimões river (0.017±0.005) and between *G. gulo* from the Paraná and *G. humeralis* from Paraguai and Uruguay basins (0.016±0.005). In addition, we did not find shared haplotypes among species of *Galeocharax* or among lineages of *Galeocharax gulo*, suggesting a strong genetic structure for each species and lineages.

Giovanetti *et al.* (2017) examined morphological data for a high number of specimens of *Galeocharax* from the Amazonas, Tocantins, and Upper Paraná River basins. Their results revealed overlaps in the diagnostic characters of both *G. gulo* and *G. knerii* (deeper body and presence of fewer teeth on the posterior row of the dentary (Menezes, 1976)), and all the meristic and morphometric characters. The authors proposed *G. knerii* from the Upper Paraná River basin as a junior synonym of *G. gulo* from the Amazon River basin. However, Mattox & Toledo-Piza (2012)

found differences in myological characters, in which specimens from the upper Upper Paraná basin had the posterior margin of the *adductor operculi* at the same vertical line than the *levator operculi*, while the *Galeocharax* from the Amazonas River basin had the posterior margin of the *levator operculi* extending posteriorly through the margin of the *adductor operculi*. Although Mattox & Toledo-Piza (2012) analyzed fewer specimens, this would indicate a clear morphological diagnosis for *G. knerii*. Recently, our phylogenomic reconstruction of the entire subfamily Characinae (Souza *et al*. in review; Cap.1) supported the monophyly of three lineages of *G. gulo* (*G. gulo* Araguaia River, *G. gulo* Paraná River, and *G. gulo* Amazon River). The molecular results found here and the phylogenomic analysis (Souza *et al*. in review) allied to previous morphological data (Menezes 1976; Mattox & Toledo-Piza 2012) does not support the proposal of synonymy of *G. knerii* for the Paraná River.

The lineage of *Galeocharax* from the Araguaia River showed the higher interspecific distances among the lineages from Paraná River and Solimões River (>4%). Giovanetti *et al*. (2017) did not examine samples from the Araguaia River but from Tocantins basin and found a tendency towards lower values in the number of branched anal-fin rays and total number of vertebrae when compared to specimens from the remainder of the geographic distribution of the species. Because the Araguaia basin has been recognized as an area with high endemicity of freshwater fishes (Albert & Reis 2011; Lima & Moreira, 2003; Britski and Lima 2008; Hrbek *et al.*, 2014; Antonetti *et al*., 2018), further research with additional morphological and molecular data is encouraged to test the hypothesis that this lineage represents an undescribed species.

Additionally, some questions still remain to be explored because of the gap in our taxon sampling for the Tocantins, Xingu and Orinoco basins. Besides that, considering that Taphorn (1992: 185) suggested that specimens of *Galeocharax* from the Apure River, Venezuela, could represent an undescribed species and the recent taxonomic revisions of *Acestrocephalus* and

Cynopotamus (Menezes, 2006, 2007) resulted in the descriptions of various new species, the diversity of *Galeocharax* is possibly underestimated.

In *Acestrocephalus* and *Cynopotamus*, we did not find genetically distinct lineages. The overall genetic divergence among species of *Acestrocephalus* (8.1% - 12.9%) and *Cynopotamus* (2.5% - 15.3%) is higher than in *Galeocharax* (1.7% - 4.9%) indicating an ancient diversification of *Acestrocephalus* and *Cynopotamus*, and a recent diversification of *Galeocharax*. It is notorious that different groups have different evolutionary rates (Krieger & Fuerst 2002) and the lower interspecific genetic divergence observed among some Neotropical fishes could be related to recent high levels of diversification (Hubert *et al.*, 2007; Carvalho *et al.*, 2011; Pereira *et al.*, 2011, 2013; Ramirez & Galetti, 2015; Ramirez *et al.*, 2017; Shimabukuro-Dias *et al.*, 2017).

Our analyses were performed with 37.5% of the species of *Acestrocephalus*, 50% of *Cynopotamus*, and 100% of *Galeocharax*, which represent the first array of data to permit the molecular identification of many species of these three genera. Additionally, this study is the first to provide barcode sequences to identify the rare *Acanthocharax microlepis* (used as outgroup). The lack of available material for genetic studies and limitations to collect Cynopotamini species are challenges for future studies. Despite these limitations, this is the most complete molecular study for Cynopotamini and results contribute to a precise species identification that are fundamental for conservation purposes.

Acknowledgements

We are grateful to Mary Burridge (ROM), María Fernando Castillo (STRI), Carlos Lucena (MCP), Carolina Doria (UFRO-ICT), Camila C. Ribas (INPA), Erick C. Guimarães (UFMA), and Michel D. Gianeti (MZUSP) for curatorial assistance and/or loan of tissues and vouchers. We are also thankful to Luz E. Ochoa, Rafaela Ota and George M. T. Mattox for advices.

Funding

This research was supported by the Brazilian agencies FAPESP proc. 17/06551-0 (CSS), 16/11313-8 and 18/23883-9 (BFM), 18/20610-1, 16/09204-6, 14/26508-3 and CNPq proc. 306054/2006-0 (CO).

Bibliography

Albert, J. S., & Reis, R. (Eds.). (2011). Historical biogeography of Neotropical freshwater fishes. University of California Press.

Antonetti, D. A., Lucena, C. A. S., & Lucena, Z. M. S. (2018). Two new species of *Phenacogaster* (Actinopterygii: Characidae) from the rio Tocantins drainage, Brazil. Ichthyological exploration of freshwaters, 28(3), 271-280.

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution, 16(1), 37-48.

Britski, H. A., & Lima, F. C. (2008). A new species of *Hemigrammus* from the upper Rio Tapajós basin in Brazil (Teleostei: Characiformes: Characidae). Copeia, 2008(3), 565-569.

de Carvalho, D. C., Oliveira, D. A., Pompeu, P. S., Leal, C. G., Oliveira, C., & Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. Mitochondrial Dna, 22(sup1), 80-86.

Díaz, J., Villanova, G. V., Brancolini, F., del Pazo, F., Posner, V. M., Grimberg, A., & Arranz, S. E. (2016). First DNA barcode reference library for the identification of South American freshwater fish from the lower Paraná river. Plos one, 11(7), e0157419.

Fricke, R., Eschmeyer, W. N. & R. van der Laan (2020). Eschmeyer's catalog of fishes: genera, species, references. Available at

http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (last accessed 18 July 2020).

Giovannetti, V., Toledo-Piza, M., & Menezes, N. A. (2017). Taxonomic revision of *Galeocharax* (Characiformes: Characidae: Characinae). Neotropical ichthyology, 15(1).

Gomes, L. C., Pessali, T. C., Sales, N. G., Pompeu, P. S., & Carvalho, D. C. (2015). Integrative taxonomy detects cryptic and overlooked fish species in a neotropical river basin. Genetica, 143(5), 581-588.

Hubert, N., Duponchelle, F., Nunez, J., GARCIA-DAVILA, C. A. R. M. E. N., Paugy, D., & RENNO, J. F. (2007). Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the diversification of the Neotropical ichthyofauna. Molecular Ecology, 16(10), 2115-2136.

Hrbek, T., da Silva, V. M. F., Dutra, N., Gravena, W., Martin, A. R., & Farias, I. P. (2014). A new species of river dolphin from Brazil or: how little do we know our biodiversity. Plos one, 9(1), e83623.

Ivanova, N. V., Dewaard, J. R., & Hebert, P. D. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular ecology notes, 6(4), 998-1002.

Jennings, W. B., Ruschi, P. A., Ferraro, G., Quijada, C. C., Silva-Malanski, A. C. G., Prosdocimi, F., & Buckup, P. A. (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(2), 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(6), 1547-1549.

Leigh, J. W., & Bryant, D. (2015). POPART: full-feature software for haplotype network construction. Methods in ecology and evolution, 6(9), 1110-1116.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25(11), 1451-1452.

Lima, F. C., & Moreira, C. R. (2003). Three new species of *Hyphessobrycon* (Characiformes: Characidae) from the upper rio Araguaia basin in Brazil. Neotropical Ichthyology, 1(1), 21-33.

Lucena C. A. S., Menezes N. A. (2003). Subfamily Characinae. In: Reis RE, Kullander SO, Ferraris CJ Jr, orgs. Checklist of the freshwater fishes of South and Central America. Porto Alegre, RS: EDIPUCRS, 200–208.

Machado, V. N., Collins, R. A., Ota, R. P., Andrade, M. C., Farias, I. P., & Hrbek, T. (2018). One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. Scientific reports, 8(1), 1-12.

Mateussi, N. T., Melo, B. F., Foresti, F., & Oliveira, C. (2019). Molecular data reveal multiple lineages in piranhas of the genus *Pygocentrus* (Teleostei, Characiformes). Genes, 10(5), 371.

Mattox, G. M., & Toledo-Piza, M. (2012). Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). Zoological Journal of the Linnean Society, 165(4), 809-915.

Mattox, G. M., van der Sleen, P., Toledo-Piza, M. (2017). Subfamily Characinae—Characine Tetras. In: Van der Sleen P., Albert, J.S. editors. Field Guide to the Fishes of the Amazon, Orinoco, and Guianas. Princeton, New Jersey, Princeton University Press, p. 102-106.

Melo, B. F., Benine, R. C., Silva, G. S., Avelino, G. S., & Oliveira, C. (2016). Molecular phylogeny of the Neotropical fish genus *Tetragonopterus* (Teleostei: Characiformes: Characidae). Molecular phylogenetics and evolution, 94, 709-717.

Menezes, N. A., & NA, M. (1976). On the Cynopotaminae, a new subfamily of Characidae (Osteichthyes, Ostariophysi, Characoidei). Arquivos de Zoologia, São Paulo 28: 1–91.

Pereira, L. H., Pazian, M. F., Hanner, R., Foresti, F., & Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Paraná Basin of Brazil. Mitochondrial Dna, 22(sup1), 87-96.

Pereira, L. H., Hanner, R., Foresti, F., & Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna?. BMC genetics, 14(1), 20.

Puillandre, N., Lambert, A., Brouillet, S., & ACHAZ, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular ecology, 21(8), 1864-1877.

Ramirez, J. L., & Galetti Jr, P. M. (2015). DNA barcode and evolutionary relationship within *Laemolyta* Cope 1872 (Characiformes: Anostomidae) through molecular analyses. Molecular phylogenetics and evolution, 93, 77-82.

Ramirez, J. L., Birindelli, J. L., Carvalho, D. C., Affonso, P. R., Venere, P. C., Ortega, H., ... & Galetti Jr, P. M. (2017). Revealing hidden diversity of the underestimated neotropical ichthyofauna: DNA barcoding in the recently described genus *Megaleporinus* (Characiformes: Anostomidae). Frontiers in genetics, 8, 149.

Rossini, B. C., Oliveira, C. A. M., Melo, F. A. G. D., Bertaco, V. D. A., Astarloa, J. M. D. D., Rosso, J. J., ... & Oliveira, C. (2016). Highlighting *Astyanax* species diversity through DNA barcoding. Plos one, 11(12), e0167203.

Stamatakis, A., Hoover, P., & Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. Systematic biology, 57(5), 758-771.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312-1313.

Shimabukuro-Dias, C. K., Costa Silva, G. J. D., Ashikaga, F. Y., Foresti, F., & Oliveira, C. (2017). Molecular identification of the fish fauna from the pantanal flood plain area in Brazil. Mitochondrial Dna Part A, 28(4), 588-592.

Smith, M. A., Fisher, B. L., & Hebert, P. D. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philosophical Transactions of the Royal Society B: Biological sciences, 360(1462), 1825-1834.

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+ C-content biases. Molecular biology and evolution, 9(4), 678-687.

Taphorn DC. The characiform fishes of the Apure river drainage, Venezuela. Biollania, Edición Especial (Monografias Cientificas del Museo de Ciencias Naturales). 1992;4:1-537

Xia, X., Xie, Z., Salemi, M., Chen, L., & Wang, Y. (2003). An index of substitution saturation and its application. Molecular phylogenetics and evolution, 26(1), 1-7.

Xia, X. (2013). DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Molecular biology and evolution, 30(7), 1720-1728.

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29(22), 2869-2876.

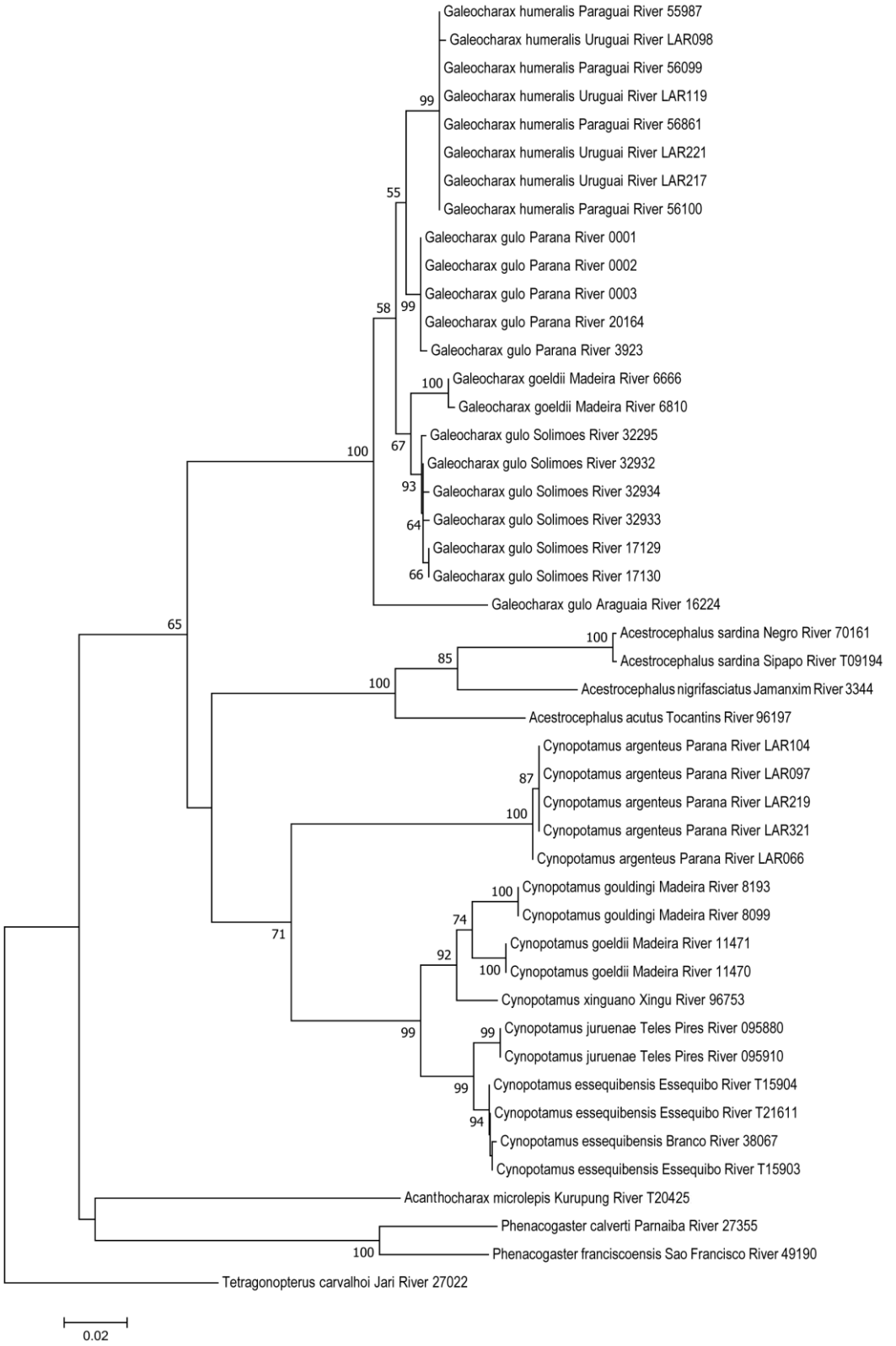
Ward, R. D., Hanner, R., & Hebert, P. D. (2009). The campaign to DNA barcode all fishes, FISH-BOL. Journal of fish biology, 74(2), 329-356.

Supplementary information

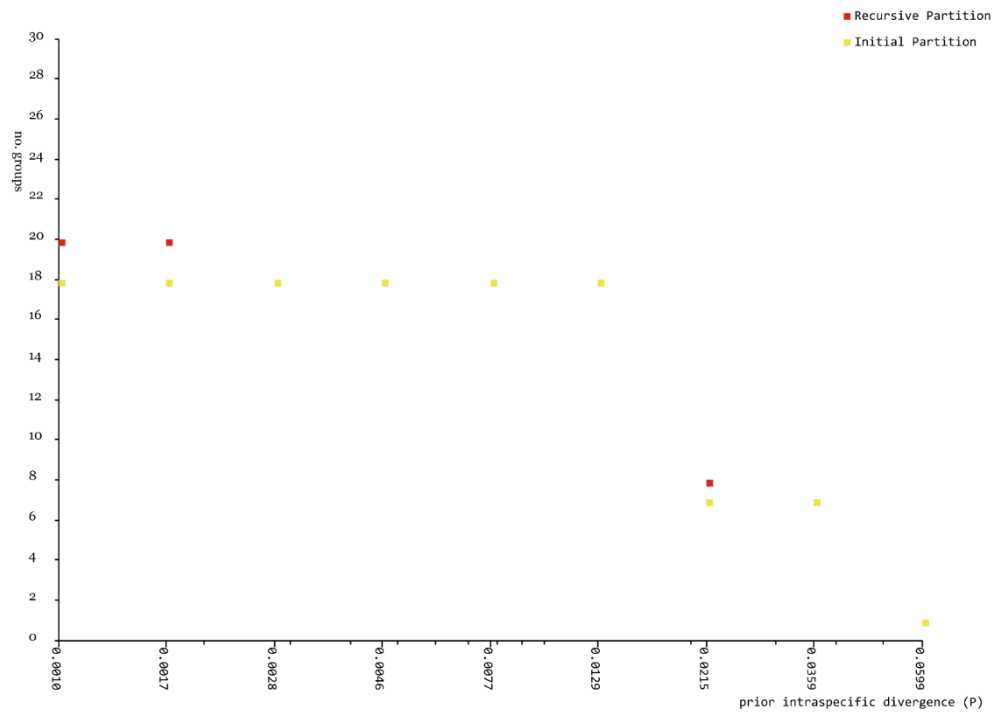
Supplementary table 1. List of the all specimens used in species delimitation analyses. 10 sequences were obtained from Genbank (KU288813, KU288844, KU288839, KU288920, KU288999, KU288840, KU288918, KU288857, KU288921, HM070393).

Genus	Species	Drainage	Country	Voucher	Tissue number	Museum number	GenBank n.
Acanthocharax	*A. microlepis*	Kurupung River	Guiana	T20425	20425	ROM	
Acestrocephalus	*A. acutus*	Tocantins River	Brazil	96197	96197	LPB 25321	
Acestrocephalus	*A. sardina*	Sipapo River	Venezuela	T09194	T09194	ROM	
Acestrocephalus	*A. sardina*	Negro River	Brazil	70161	70161	17833	
Acestrocephalus	*A. nigrifasciatus*	Jamanxim River	Brazil	MUZUSP003344	3344	97469	
Cynopotamus	*C. argenteus*	Paraná River	Argentina	LAR066	LAR066	MAG ICT 0066	KU288813
Cynopotamus	*C. argenteus*	Paraná River	Argentina	LAR104	LAR104		KU288844
Cynopotamus	*C. argenteus*	Paraná River	Argentina	LAR097	LAR097		KU288839
Cynopotamus	*C. argenteus*	Paraná River	Argentina	LAR219	LAR219	MAG ICT 0208	KU288920
Cynopotamus	*C. argenteus*	Paraná River	Argentina	LAR321	LAR321	MG-ZV-P-00211	KU288999
Cynopotamus	*C. gouldingi*	Madeira River	Brazil	UFRO-ICT 017897	8193	UFRO	
Cynopotamus	*C. gouldingi*	Madeira River	Brazil	8099 UFRO	8099	UFRO	
Cynopotamus	*C. goeldii*	Madeira River	Brazil	UFRO	11471	UFRO	
Cynopotamus	*C. goeldii*	Madeira River	Brazil	UFRO-ICT 023656	11470	UFRO	
Cynopotamus	*C. xinguano*	Xingu River	Brazil	96753	96753	LBP 25967	
Cynopotamus	*C. juruenae*	Teles Pires River	Brazil	MUZUSP003991	95880	95880	
Cynopotamus	*C. juruenae*	Teles Pires River	Brazil	MUZUSP004013	95910	95910	
Cynopotamus	*C. essequibensis*	Essequibo River	Guyana	T15904	T15904	ROM	
Cynopotamus	*C. essequibensis*	Essequibo River	Guyana	T21611	T21611	ROM	
Cynopotamus	*C. essequibensis*	Branco River	Brazil	38067	38067	LBP 8158	
Cynopotamus	*C. essequibensis*	Essequibo River	Guyana	T15903	T15903	ROM	
Galeocharax	*G. humeralis*	Paraguai River	Brazil	55987	55987	LBP 13453	
Galeocharax	*G. humeralis*	Uruguai River	Brazil	LAR098	LAR098		KU288840
Galeocharax	*G. humeralis*	Uruguai River	Brazil	LAR217	LAR217	MAG ICT 0206	KU288918
Galeocharax	*G. humeralis*	Uruguai River	Brazil	LAR119	LAR119		KU288857
Galeocharax	*G. humeralis*	Uruguai River	Brazil	LAR221	LAR221	MAG ICT 0210	KU288921
Galeocharax	*G. humeralis*	Paraguai River	Brazil	56099	56099	LBP 13483	
Galeocharax	*G. humeralis*	Paraguai River	Brazil	56100	56100	LBP 13483	
Galeocharax	*G. humeralis*	Paraguai River	Brazil	56861	56861	LBP 13694	
Galeocharax	*G. gulo*	Parana River	Brazil	20164	20164	LBP 3496	
Galeocharax	*G. gulo*	Parana River	Brazil	96994	1	LBP 28889	
Galeocharax	*G. gulo*	Parana River	Brazil	96695	2	LBP 28889	
Galeocharax	*G. gulo*	Parana River	Brazil	96996	3	LBP 28889	
Galeocharax	*G. gulo*	Parana River	Brazil	3923	3923	LBP 76	
Galeocharax	*G. goeldii*	Madeira River	Brazil	UFRO-ICT 013389	6666	UFRO	

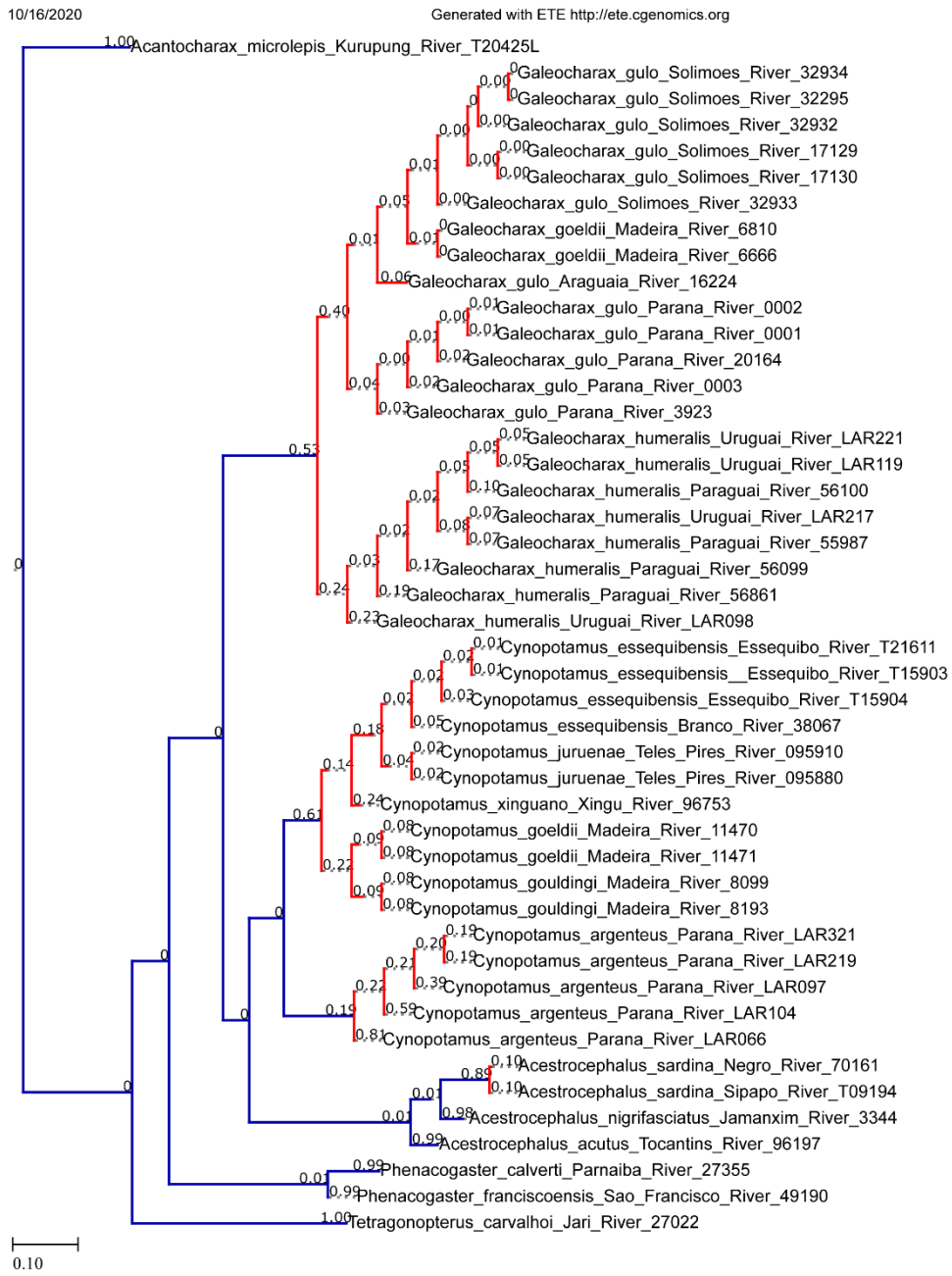
Genus	Species	Drainage	Country	Voucher	Tissue number	Museum number	GenBank n.
Galeocharax	*G. goeldii*	Madeira River	Brazil	UFRO-ICT 006854	6810	UFRO	
Galeocharax	*G. gulo*	Solimões River	Brazil	INPA	P 32932	INPA	
Galeocharax	*G. gulo*	Solimões River	Brazil	INPA	P 32933	INPA	
Galeocharax	*G. gulo*	Solimões River	Brazil	INPA-ICT 055530	P 32295	INPA	
Galeocharax	*G. gulo*	Solimões River	Brazil	INPA	P 32934	INPA	
Galeocharax	*G. gulo*	Solimões River	Brazil	17129	17129	LBP 2556	
Galeocharax	*G. gulo*	Solimões River	Brazil	17130	17130	LBP 2556	
Galeocharax	*G. gulo*	Araguaia River	Brazil	16224	16224	LBP 2391	
Phenacogaster	*P. franciscoensis*	Parnaíba River	Brazil	49190	49190	LBP 10487	
Phenacogaster	*P. calverti*	Parnaiba River	Brazil	27355	27355	LBP 5612	
Tetragonopterus	*T. carvalhoi*	Jari River	Brazil	27022	27022	LBP 5376	HM070393



Supplementary Fig. 1. NJ tree of species of *Acestrocephalus*, *Cynopotamus* and *Galeocharax*, based on the COI gene (543 pb). Values < 50% are not shown.



Supplementary Fig. 2. Automatic partition of the dataset reporting the number of groups inside the initial and recursive partitions.



Supplementary Fig. 3. Bayesian Poisson Tree Processes (bPTP) delimitation tests of species using the maximum likelihood tree.

Chapter 5

Molecular identification of species of the tribe Characini (Teleostei: Characidae: Characinae)

Molecular identification of species of the tribe Characini (Teleostei: Characidae: Characinae)

Camila S. Souza

Departamento de Biologia Estrutural e Funcional, Instituto de Biociências, Universidade Estadual Paulista, R. Prof. Dr. Antônio C. W. Zanin 250, 18618-689 Botucatu SP, Brazil

Abstract

DNA barcoding has successfully been used in fish species identification and improved our knowledge of the Neotropical ichthyofauna. Characini comprised species of *Charax* and *Roeboides* distributed by neotropical region. Here we assessed 65% of valid species of these genera using a DNA barcode approach to generate a reference library and evaluate the species diversity in *Charax* and *Roeboides*. The results of ABGD and bPTP, in general, were congruent with the morphological identification but delimited two genetic lineages for each of the following species: *C. leticiae*, *C. tectifer*, and *R. affinis*, revealing a previously unknown genetic diversity. The molecular findings can be used along with morphological traits to better understand each species. Additionally, considering the biological importance of these species, precise species identification is quite desirable and fundamental for the conservation of the whole biodiversity.

Keywords: Fish, Diversity, Systematics, DNA barcode, COI

1. Introduction

Within the characid subfamily Characinae, the tribe Characini contains species of *Charax* Scopoli, 1777 and *Roeboides* Günther, 1864. This tribe is monophyletic and supported by five synapomorphies with one exclusive: anterior tip of pelvic bone located anterior to vertical through posterior margin of cleithrum (Mattox & Toledo-Piza, 2012). Currently, *Charax* is represented by 17 valid species widely distributed in the East-Andean rivers of South America, whereas *Roeboides* is represented by 21 valid species distributed in rivers of the South and Central America in both sides of the Andes (*sensu* Mattox *et al*., 2018; Fricke *et al*., 2022). These two genera contain small to medium-sized species (maximum size about 130 mm SL in *Charax* and 190 mm SL in *Roeboides*) with deep bodies. *Charax* is generally carnivorous while species of *Roeboides* are lepidophagous with morphological specializations such as presence of external mammiliform teeth (Sazima & Machado, 1983; Sazima, 1983; Lucena, 2007; Menezes & Lucena, 2014).

Charax is diagnosed by dorsal extension of section A1 of adductor mandibulae along preopercle restricted to horizontal arm of preopercle, absence of rhinosphenoid, absence of anterodorsal branch of laterosensory canal on infraorbital 6, presence of large notch and projection on posteroventral margin of cleithrum (Mattox & Toledo-Piza, 2012). All the systematic studies of *Charax* are based on morphological data (e.g., Eigenmann, 1912, 1922; Géry & Knöppel 1976; Lucena 1987, 1989; Menezes & Lucena, 2014; Guimarães *et al.*, 2018). The last studies provided more extensive geographic distribution for various species, including the description of *C. delimai* from Rio Negro, the synonymization of *C. unimaculatus* in *C. michaeli* (Menezes & Lucena, 2014) and the description of *C. awa*, the first species from coastal river basins of northeastern Brazil (Guimarães *et al.*, 2018).

Roeboides is diagnosed by six synapomorphies, three of which are related to the arrangement of mamilliform teeth on premaxilla and dentary (Lucena, 1998). Based on osteological

synapomorphies especially those related to jaws as well as meristic traits, the species of *Roeboides* are organized in four groups: the *R. dispar* group (*R. dispar* Lucena, 2001), the *R. guatemalensis* group (*R. guatemalensis* (Günther, 1864); *R. dayi* (Steindachner, 1878); *R. occidentalis* Meek & Hildebrand, 1916; *R. bouchellei* Fowler, 1923; *R. dientonito* Schultz, 1944; *R. ilsea* Bussing, 1986; *R. carti* Lucena, 2000), the *R. microlepis* group (*R. microlepis* (Reinhardt, 1851); *R. myersii* Gill, 1870; *R. araguaito* Lucena, 2003; *R. margaretae* Lucena, 2003), and the *R. affinis* group (*R. xenodon* (Reinhardt, 1851); *R. affinis* (Günther, 1868); *R. biserialis* (Garman, 1890); *R. descalvadensis* Fowler, 1932; *R. numerous* Lucena, 2000; *R. oligistos* Lucena, 2000; *R. sazimai* Lucena, 2007) (Lucena, 1998, 2000, 2003, 2007, 2011). Based on a preliminary molecular analysis and morphological features, Bussing (1998) provided evidence of an undescribed species of *Roeboides* from Costa Rica and Panama. Later on, Matamoros *et al*. (2013) described *R. bussingi*, the last and more recent description of a species of *Roeboides*.

The diversity of *Charax* and *Roeboides* had little effort in the last decade and molecular studies are not available. Biological identification through molecular tools has been successfully used to recognize the Neotropical freshwater fish fauna and flagging new species (e.g., Mattox *et al*., 2020; Ochoa *et al*., 2020; García-Melo *et al*., 2019; Machado *et al*., 2018; Mateussi *et al*., 2017, 2019; Melo *et al*., 2016; Pereira *et al*., 2013). Moreover, the geographical distribution of *Charax* and *Roeboides* is highly complex, with some species occurring in restricted basins (e.g. *Roeboides xenodon*, *Roeboides numerosus*, *Charax condei* (Géry & Knöppel, 1976), *Charax metae* Eigenmann, 1922)), and others having broad distributions occurring even in multiple hydrographic basins (e.g. *Roeboides affinis*, *Roeboides descalvadensis*, and *C. leticiae* Lucena, 1987).

Currently, 38 valid species are reported for Characini and probably this diversity is underestimated. In addition, only 34.2% of barcode sequences have been generated from described species. In this context, we used *DNA barcoding* and species delimitation methods for 172

specimens of Characini (65% of valid species) aiming to evaluate the species diversity in *Charax* and *Roeboides*.

2. Materials and Methods

Taxon sampling

Fishes representing nominal species of *Charax* and *Roeboides* were collected or obtained from ichthyological collections, and morphologically identified with the help of identification keys (Menezes & Lucena, 2014, Lucena, 2003, 2007). All fishes were collected following the Brazilian laws through SISBIO/MMA permit n. 3,245 and procedures for collection, maintenance and analyses followed the international guidelines for animal experiments through CEEAA IBB/UNESP protocol n. 304. Voucher specimens were fixed in 95% ethanol or 10% formalin and transferred to 70% ethanol for permanent storage, and posteriorly deposited in the collection of the Laboratório de Biologia e Genética de Peixes, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, São Paulo, Brazil (LBP). We used a total of 172 specimens representing 8 species of *Charax* and 17 species of *Roeboides*. One species of *Tetragonopterus* and two species of *Phenacogaster* were used as outgroup based on previous phylogenies (Oliveira *et al.*, 2011; Melo *et al*., 2021). Thirteen out of 172 sequences were obtained from GenBank (Supplementary Table 1).

DNA Extraction, amplification and sequencing

DNA was extracted from samples of muscle, liver or gills, using the extraction method of the Ivanova *et al*. (2006). Partial sequences of the *cytochrome c oxidase subunit I* (*COI*) gene were amplified by polymerase chain reaction (PCR) with primers FishF1, FishR1 and FishF2, FishR2

(Ward *et al*., 2005) or L6252-Asn and H7271-COXI (Melo *et al*., 2011). PCR amplifications were performed in a total volume of 12.5 µl that included 1.25 µl of 10X buffer, 0.25 µl of $MgCl_2$ (50 mM), 0.2 µl dNTPs (2 mM), 0.5 µl of each primer (5 mM), 0.1 µl of PHT Taq DNA polymerase (Phoneutria), 1.0 µl of genomic DNA (200 ng) and 8.7 µl ddH_2O. PCR conditions consisted of an initial denaturation (5 min at 94°C), followed by 30 cycles of chain denaturation (50s at 94°C), primer hybridization (45s at 50-54°C) and nucleotide extension (1 min at 68°C), and a final extension (10 min at 68°C). All PCR products were checked on 1% agarose gels and then purified with ExoSap-IT (USB Corporation) following the manufacturer's instructions. The purified PCR products were submitted to sequencing reactions using BigDye Terminator v 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and purified again by ethanol precipitation. Products were loaded onto an ABI 3130 DNA Analyzer automatic sequencer (Applied Biosystems).

Data analysis

Sequences were assembled using Geneious v7.1.9 (Kearse *et al.*, 2012) and subsequently aligned with Muscle (Edgar, 2004) under default parameters. To evaluate the occurrence of substitution saturation, was used the method of Xia *et al*. (2003) in DAMBE v5.3.38 (Xia, 2013). Nucleotide variation, substitution patterns and the best-fit model of nucleotide evolution were estimated in MEGA v10 (Kumar *et al*., 2018). *Charax* and *Roeboides* data were independently analyzed.
The maximum-likelihood (ML) tree was obtained using a random starting tree with GTRCAT model (Stamatakis *et al*., 2008) through RAxML HPC2 on XSEDE v8.2.12 (Stamatakis, 2014) as implemented on the CIPRES webserver (Miller *et al*., 2010). All other parameters were left at default.

Two distinct species delimitation methods were performed: 1) Automatic Barcode Gap Discovery analysis (ABGD; Puillandre *et al.*, 2012) available in the ABGD webserver

(https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html) using Kimura (K80; 2.0) evolutionary model with X = 1.0, Pmin = 0.001 and Pmax = 0.1; 2) Poisson Tree Process (PTP; Zhang *et al*., 2013) using the best maximum-likelihood (ML) tree, 100,000 generations, and other parameters at default in the bPTP webserver (http://species.h-its.org/ptp/). We used DnaSP v5 (Librado & Rozas, 2009) with two reduced matrices excluding sites with missing data to estimate the haplotype diversity, haplotype number, and haplotype distribution. Thereafter, we generated haplotype networks using the median-joining analysis (Bandelt, Forster, & Röhl, 1999) available in PopART 1.7 (Leigh & Bryant, 2015). The values of genetic distance were calculated in MEGA v10 using the Tamura 3-parameter model (Tamura, 1992) in MEGA. Groups were ordered based on ABGD and bPTP results, except for *R. biserialis* that was based on the ML tree and the morphological identification.

3. Results

Barcode sequences were obtained for 172 specimens representing 8 out of 17 species of *Charax*, 17 out of 21 species of *Roeboides* and three species of outgroup (1- *Phenacogaster franciscoensis*. 1- *P. calverti*. 1- *Tetragonopterus carvalhoi*). The matrix of *Charax* had 42 sequences with 513 pb (253 variable sites) and nucleotide composition of 24.3%(A), 26.9%(C), 17.6%(G), and 31.2%(T). The matrix of *Roeboides* had 133 sequences with 513 pb (212 variable sites) and nucleotide composition of 24.1%(A), 26.4%(C), 18.4%(G), and 31.1%(T). DAMBE indicated no saturation in either transitions or transversions in both asymmetrical (Iss.cAsym) and symmetrical (Iss.cSym) topologies. ABGD and bPTP did not discriminate some species, such as except ABGD (*C. pauciradiatus* and *C. niger),* (*R. biserialis* and *R. descalvadensis*), (*R. myersii*, *R. microlepis* and *R. margaretae*) and (*R. ilseae* and *R. bouchellei*) (Fig. 1-2).

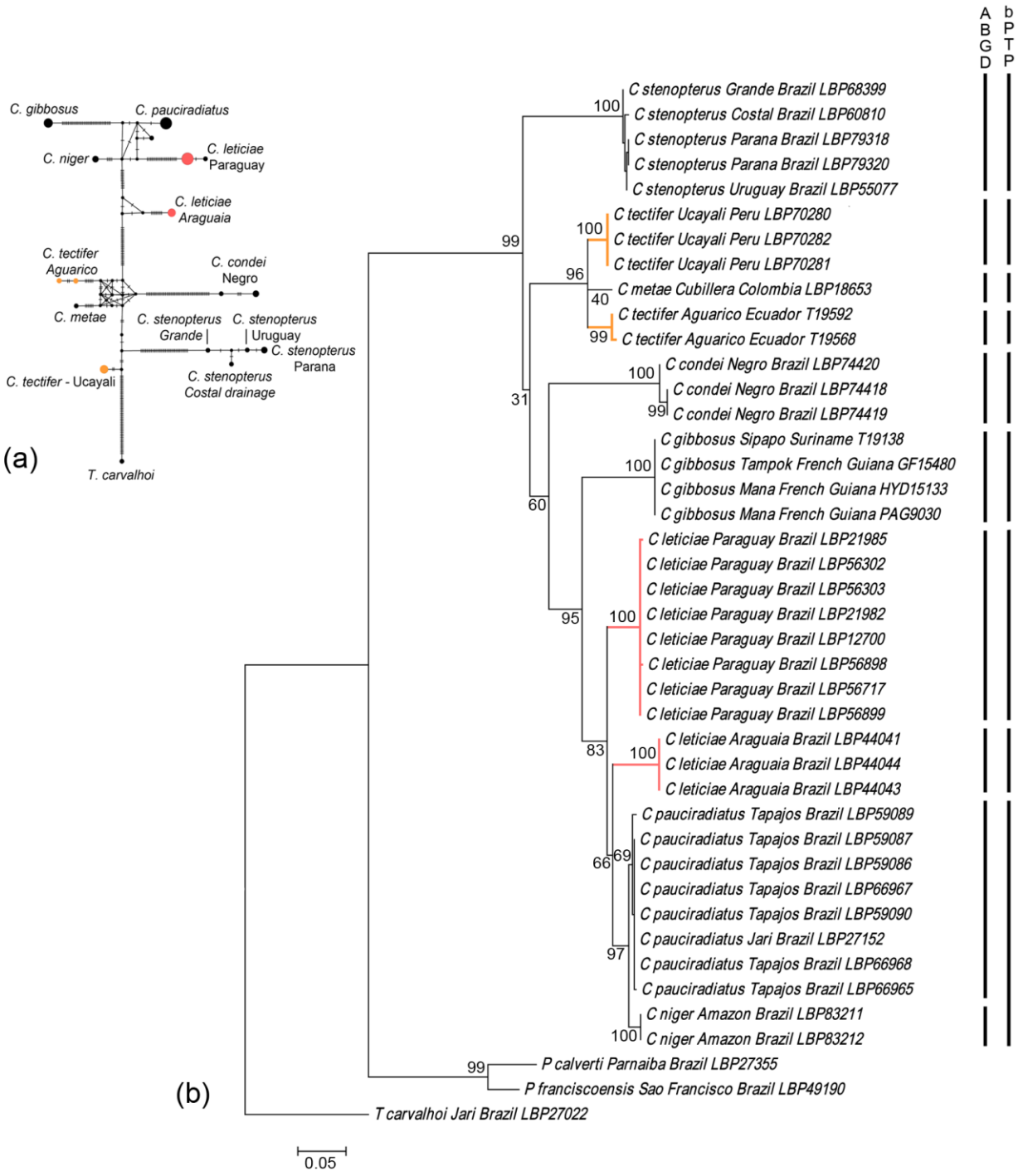


Figure 1. a) Haplotype network analyses showing the distribution of the distinct haplotypes of species and lineages of *Charax*. b) ML tree of species of *Charax* based on the *COI* gene (513 pb). Vertical bars at right represent the number of species obtained by the ABGD and bPTP analyses. Numbers near nodes represent bootstrap support. Numbers of the specimen are after tip names.

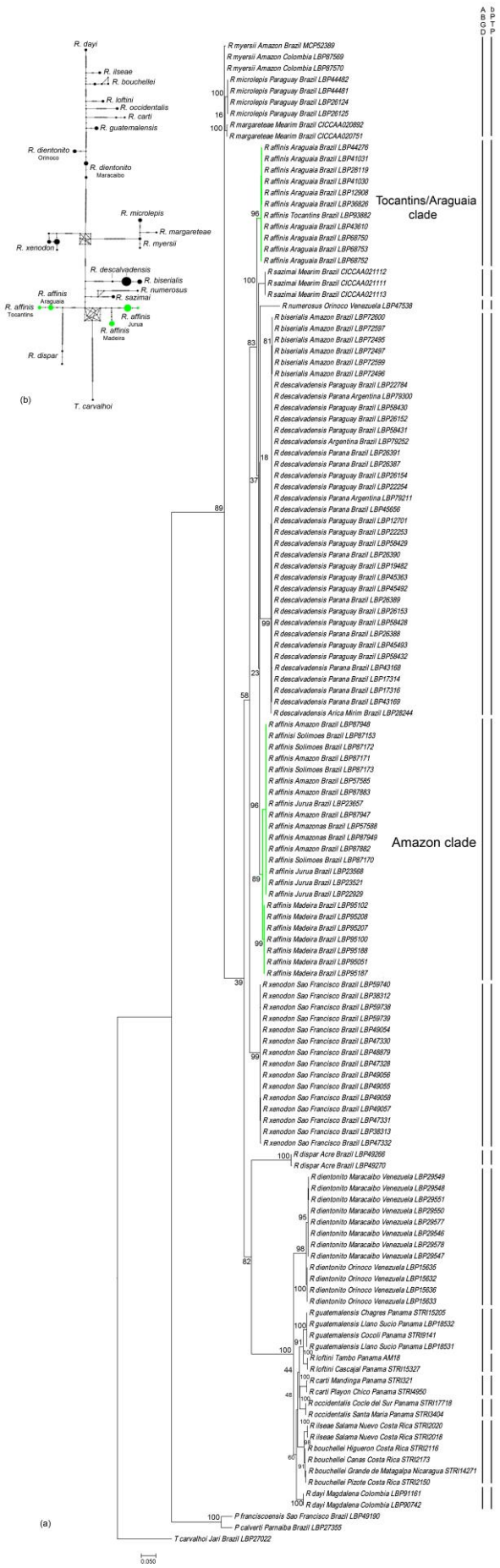


Figure 2. a) Haplotype network analyses showing the distribution of the distinct haplotypes of species and lineages of *Roeboides*. b) ML tree of species of *Roeboides* based on the *COI* gene (513 pb). Vertical bars at right represent the number of species obtained by the ABGD and bPTP analyses. Numbers near nodes represent bootstrap support. Numbers of the specimen are after tip names.

However, ABGD and bPTP delimited two genetic lineages for each of the following species: *C. leticiae*, *C. tectifer*, and *R. affinis* (Figures 1–2; Supplementary Figs. 1-4). The value of interspecific genetic distance between the two lineages of *C. leticiae* was 0.083 ± 0.021 (Araguaia and Paraguay basins), between the lineages of *C. tectifer* was 0.047 ± 0.010 (Aguarico and Ucayali rivers) and between the lineages of *R. affinis* was 0.040 ± 0.009 (Tocantins/Araguaia and Amazon basins) (Tables 1-2).

Table 1. Pairwise K2P genetic distances among species of *Charax*. Intraspecific genetic distances are highlighted in bold in the last column. Numbers below diagonal are values of interspecific distances and numbers above diagonal are respective values of standard deviation.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Intraspecific genetic variations**
1- *C. condei*		0.026	0.027	0.024	0.024	0.025	0.023	0.027	0.022	0.024	**0.005±0.002**
2- *C. stenopterus*	0.195		0.023	0.022	0.023	0.022	0.022	0.022	0.025	0.023	**0.005±0.002**
3- *C. metae*	0.177	0.139		0.024	0.025	0.024	0.021	0.022	0.010	0.010	-
4- *C. gibbosus*	0.179	0.164	0.141		0.018	0.018	0.019	0.020	0.022	0.023	**0.004±0.002**
5- *C. niger*	0.166	0.157	0.143	0.112		0.006	0.013	0.014	0.023	0.022	**0**
6- *C. pauciradiatus*	0.172	0.154	0.143	0.114	0.019		0.012	0.013	0.023	0.022	**0.002±0.001**
7- *C. leticiae* Paraguay	0.158	0.153	0.129	0.123	0.066	0.059		0.015	0.022	0.021	**0**
8- *C. leticiae* Araguaia	0.183	0.144	0.130	0.120	0.073	0.068	0.083		0.022	0.021	**0**
9- *C. tectifer* Aguarico	0.152	0.174	0.041	0.152	0.154	0.155	0.144	0.138		0.010	**0.004±0.003**
10- *C. tectifer* Ucayali	0.168	0.153	0.042	0.162	0.143	0.144	0.141	0.134	0.047		**0**

Table 3. Pairwise K2P genetic distances among species of *Roeboides*. Intraspecific genetic distances are highlighted in bold in the last column. Numbers below diagonal are values of interspecific distances and numbers above diagonal are respective values of standard deviation.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**Intraspecific genetic distances**
1- *R. dientonito*		0.015	0.017	0.015	0.015	0.014	0.016	0.016	0.029	0.028	0.029	0.028	0.027	0.025	0.026	0.026	0.025	0.023	0.026	**0.006±0.002**
2- *R. dayi*	0.087		0.012	0.012	0.014	0.014	0.013	0.012	0.029	0.026	0.026	0.025	0.025	0.026	0.030	0.030	0.026	0.024	0.028	**0**
3- *R. ilseae*	0.099	0.058		0.006	0.013	0.012	0.012	0.013	0.033	0.028	0.026	0.027	0.029	0.026	0.032	0.032	0.030	0.026	0.031	**0**
4- *R. bouchellei*	0.085	0.057	0.018		0.010	0.010	0.011	0.013	0.030	0.027	0.025	0.026	0.030	0.026	0.032	0.032	0.030	0.026	0.030	**0.002±0.002**
5- *R. loftini*	0.087	0.070	0.066	0.048		0.008	0.010	0.011	0.029	0.029	0.027	0.028	0.027	0.024	0.030	0.030	0.026	0.025	0.027	**0**
6- *R. guatemalensis*	0.077	0.067	0.061	0.049	0.029		0.011	0.012	0.029	0.030	0.029	0.030	0.028	0.023	0.030	0.030	0.028	0.025	0.026	**0**
7- *R. occidentalis*	0.088	0.060	0.052	0.049	0.047	0.047		0.011	0.032	0.027	0.025	0.026	0.027	0.026	0.032	0.032	0.029	0.026	0.031	**0**
8- *R. carti*	0.094	0.057	0.067	0.063	0.057	0.058	0.048		0.033	0.032	0.031	0.030	0.028	0.029	0.029	0.029	0.026	0.027	0.030	**0.002±0.002**
9- *R. dispar*	0.215	0.213	0.232	0.216	0.216	0.215	0.227	0.239		0.025	0.026	0.025	0.027	0.023	0.024	0.024	0.023	0.020	0.022	**0.002±0.002**
10- *R. microlepis*	0.204	0.182	0.202	0.196	0.211	0.220	0.198	0.234	0.175		0.006	0.005	0.019	0.019	0.023	0.023	0.023	0.019	0.021	**0**
11- *R. margareteae*	0.206	0.172	0.183	0.176	0.191	0.207	0.179	0.214	0.183	0.019		0.004	0.020	0.019	0.024	0.024	0.023	0.019	0.021	**0.002±0.002**
12- *R. myersii*	0.210	0.174	0.194	0.188	0.202	0.218	0.189	0.224	0.177	0.013	0.012		0.019	0.018	0.023	0.022	0.022	0.018	0.019	**0.003±0.002**
13- *R. numerosus*	0.193	0.175	0.197	0.205	0.197	0.198	0.195	0.207	0.196	0.122	0.127	0.122		0.015	0.018	0.018	0.016	0.017	0.016	-
14- *R. xenodon*	0.183	0.180	0.187	0.188	0.178	0.172	0.187	0.218	0.166	0.120	0.120	0.116	0.092		0.017	0.017	0.015	0.015	0.013	**0.001±0.001**
15- *R. biserialis*	0.187	0.213	0.225	0.224	0.215	0.219	0.225	0.213	0.169	0.153	0.158	0.154	0.104	0.103		0.002	0.013	0.013	0.012	**0**
16- *R. descalvadensis*	0.184	0.211	0.225	0.224	0.215	0.216	0.222	0.210	0.169	0.151	0.156	0.152	0.103	0.101	0.002		0.013	0.012	0.012	**0**
17- *R. sazimai*	0.185	0.187	0.206	0.205	0.189	0.208	0.204	0.195	0.165	0.153	0.148	0.142	0.093	0.082	0.065	0.063		0.011	0.010	**0**
18- *R. affinis.* Araguaia	0.170	0.175	0.183	0.183	0.187	0.188	0.190	0.200	0.143	0.127	0.122	0.117	0.097	0.078	0.064	0.062	0.045		0.009	**0.001±0.001**
19- *R. affinis.* Amazon	0.191	0.198	0.210	0.204	0.196	0.194	0.216	0.217	0.164	0.146	0.141	0.136	0.091	0.075	0.062	0.060	0.049	0.040		**0.008±0.003**

The reduced matrix of *Charax* resulted in a total of 18 haplotypes with haplotype diversity (Hd) of 0.9269 and the matrix of *Roeboides* resulted in 33 haplotypes and 0.9230 Hd. The haplotype distribution of *Charax* was: one haplotype for *C. metae, C. gibbosus, C. niger, C tectifer* (Ucayali clade), *C. leticiae* (Araguaia clade), two haplotypes for *C. condei, C. tectifer* (Aguarico clade), *C. leticiae* (Paraguay clade) and *C. pauciradiatus*, and four haplotypes for *C. stenopterus* (Fig. 1a). For the matrix of *Roeboides*, the results were: one haplotype for each of *R. biserialis, R. dayi, R. ilseae, R. loftini, R. occidentalis, R. microlepis, R. margareteae, R. numerous, R. sazimai, R. affinis* (Madeira clade) and *C. stenopterus* (outgroup), two haplotypes for *R. bouchellei*, *R. carti*, *R. descalvadensis*, *R. dispar*, *R. myersii*, and *R. affinis* (Tocantins/Araguaia clade), *R. affinis* (Amazon clade), three haplotypes for *R. dientonito* and *R. guatemalensis*, and four haplotypes for *R. xenodon* (Fig. 2b).

4. Discussion

Species delimitation analyses were effective to identify the most species of *Charax* and *Roeboides* used in this study and to identify the composition of the four groups of *Roeboides* (Lucena, 1998). Additionally, our results revealed a different number of genetic lineages of *C. leticiae, C. tectifer,* and *R. affinis*. Specifically, *C. leticiae, C. tectifer,* and *R. affinis* were represented here by two genetic lineages showing a high degree of genetic distance.

Charax leticiae had two distinct genetic lineages without shared haplotype (Paraguay and Araguaia rivers; 0.081±0.013) and *C. tectifer* also had two lineages recognized by all methods and without shared haplotypes (Ucayali and Aguarico rivers - Peru; 0.047±0.010) (Fig. 1, Table 1). *Charax leticiae* is broadly distributed in the Tocantins, Araguaia, Paraguay, and Aripuanã river basins whereas *C. tectifer* is known from Ucayali and Tambo river basin, Peru, tributaries of the Amazon basin in Colombia, and the Napo basin in Ecuador (Menezes & Lucena, 2014).

Additional genetic units likely exist in the other drainages, but we were unable to obtain samples for molecular analyses. Previous morphological studies did not find significant statistical differences in meristic and morphometric data among *C. leticiae* from Tocantins/Araguaia and Paraguay basins and among *C. tectifer* samples from Colombia, Ecuador, and Peru (Menezes & Lucena, 2014). Our results suggest that these lineages may represent cryptic species: biological units classified as a single nominal species and morphologically indistinguishable (Mayr, 1942). However, since we did not perform any morphological analysis and the samples used in our study is from different localities compared to those by Menezes & Lucena (2014), this cannot be tested at this moment. Further morphological investigation in *C. leticiae* and *C. tectifer* are needed to confirm if the Paraguay and Aguarico lineages represent undescribed species.

Roeboides affinis is recognized as a species with a wide distribution that can be found in Paraguay, Paraná, Uruguay, Amazon, Orinoco, Tocantins, Araguaia river basins, and Guyana and Suriname. Species delimitation analyses recognized two lineages to specimens of *R. affinis* (*R. affinis* – Tocantins/Araguaia clade and *R. affinis* Amazon clade; 0.040±0.009) (Fig. 2; Table 2). Furthermore, the haplotype network shows a lack of any shared haplotype between the basins and corroborates the singularity of each lineage. A review study of *R. affinis* showed differences in morphometric, meristic, and morphological data, but some data overlap and no pattern with geographical correlation exists (Lucena, 2007).

Molecular studies in several Neotropical fish groups have shown that widely distributed nominal species isolated in different hydrographic systems can sometimes represent species complexes (Pereira *et al.,* 2011; Bagley *et al*., 2015; Costa-Silva *et al.,* 2015; Melo *et al.,* 2016; Souza *et al.,* 2018; Mateussi *et al.,* 2019). A species complex is generally understood as a group of sibling or cryptic species that need a critical revision to clarify the taxa involved and the diagnostic traits (Sigovini *et al*., 2016). The combination of morphological and our molecular data provides evidence that *R. affinis* represents a species complex, however since did not perform any

morphological analysis and our molecular data has a limited sampling, we did not disregard that *R. affinis* may contain different taxonomic entities and our molecular data might be important to conduct future taxonomic studies for this species.

Although our molecular analyses identified lineages for *C. leticiae, C. tectifer* and *R. affinis*, they did not discriminate some species that showed low genetic divergence between them (<2%) (*R. biserialis* and *R. descalvadensis*), (*R. myersii, R. microlepis* and *R. margaretae*) and (*R. ilseae* and *R. bouchellei*) (Fig. 2; Table 2). Despite this, no evidence of shared haplotypes among the species were observed, and interestingly, these species showed a close relationship confirmed by morphological (Lucena, 1998) and phylogenomic data (Souza *et al.* in review; Cap.1), suggesting a recent and rapid speciation for these species. A similar example was found in other Neotropical fish groups, such as Parodontidae (Bellafronte *et al.,* 2013), *Rineloricaria* (Costa-Silva *et al.,* 2015) and *Astyanax* (Rossini *et al.*, 2016). From a different point of view, other biological processes may also explain these results, such as introgression, incomplete lineage sorting and hybridization leading to taxonomic incongruences between morphological and molecular data (Wendt *et al.,* 2019).

This is the first genetic study investigating the biodiversity within *Charax* and *Roeboides*, and in general, the results were congruent with the morphological identification. However, further morphological and molecular studies are needed to better define the taxonomic status of *C. leticiae*, *C. tectifer* and *R. affinis*. Furthermore, additional studies using a greater number of samples from additional localities are necessary to understand the species limits that presented low interspecific genetic divergence.

Acknowledgements

We are grateful to Mary Burridge (ROM), María Fernando Castillo (STRI), Carlos Lucena (MCP), Carolina Doria (UFRO-ICT), Camila C. Ribas (INPA), Erick C. Guimarães (UFMA), and Michel

D. Gianeti (MZUSP) for curatorial assistance and/or loan of tissues and vouchers. We are also thankful to Luz E. Ochoa and George M. T. Mattox for advices.

Funding

This research was supported by the Brazilian agencies FAPESP proc. 17/06551-0 (CSS), 16/11313-8 and 18/23883-9 (BFM), 18/20610-1, 16/09204-6, 14/26508-3 and CNPq proc. 306054/2006-0 (CO).

Bibliography

Bagley, J. C., Alda, F., Breitman, M. F., Bermingham, E., van den Berghe, E. P., & Johnson, J. B. (2015). Assessing species boundaries using multilocus species delimitation in a morphologically conserved group of neotropical freshwater fishes, the *Poecilia sphenops* species complex (Poeciliidae). PLoS One, 10(4), e0121139.

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution, 16(1), 37-48.

Bellafronte, E., Mariguela, T. C., Pereira, L. H. G., Oliveira, C., & Moreira-Filho, O. (2013). DNA barcode of Parodontidae species from the La Plata river basin-applying new data to clarify taxonomic problems. Neotropical Ichthyology, 11(3), 497-506.

Cope, E. D. (1870). Contribution to the ichthyology of the Marañon. Proceedings of the American Philosophical Society of Philadelphia, 11: 559-570.

Cope, E. D. (1894). On three new genera of Characinidae. American Naturalist, 28: 67.

Costa-Silva, G. J., Rodriguez, M. S., Roxo, F. F., Foresti, F., & Oliveira, C. (2015). Using different methods to access the difficult task of delimiting species in a complex Neotropical hyperdiverse group. PLoS One, 10(9), e0135075.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research, 32(5), 1792-1797.

Eigenmann, C. H. (1912). The freshwater fishes of British Guiana, including a study of the ecological grouping of species, and the relation of the fauna of the plateau to that of the lowlands. Memoirs of the Carnegie Museum, 5: 1-578.

Eigenmann, C. H. (1922). Peces colombianos de las cordilleras y de los llanos al oriente de Bogota. Boletin de la Sociedad Colombiana de Ciencias Naturales, 66: 159-168.

Fricke, R., Eschmeyer, W. N. & R. van der Laan (2020). Eschmeyer's catalog of fishes: genera, species, references. Available at http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (last accessed 30 April 2021).

García-Melo, J. E., Oliveira, C., Da Costa Silva, G. J., Ochoa-Orrego, L. E., Garcia Pereira, L. H., & Maldonado-Ocampo, J. A. (2019). Species delimitation of neotropical Characins (Stevardiinae): Implications for taxonomy of complex groups. PloS one, 14(6), e0216786.

Géry, J. & H. A. Knoppel, (1976). Un characin translucide nouveau: *Asiphonichthys condei* n. sp. (Cypriniforme, Characoidei, Characidae). Revue française d'Aquariologie Herpétologie, 3: 47-54.

Guimarães, Erick C., Pâmella S. De Brito, Beldo R. A. Ferreira & Felipe P. Ottoni. 2018. A new species of *Charax* (Ostariophysi, Characiformes, Characidae) from northeastern Brazil. Zoosystematics and Evolution 94(1): 83-93.

Günther, A. (1864). Catalogue of the fishes in the British Museum. Catalogue of the Physostomi, containing the families Siluridae, Characinidae, Haplochitonidae, Sternoptychidae, Scopelidae, Stomiatidae in the collection of the British Museum. Catalogue of the fishes in the British Museum, 5: 1-455.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647-1649.

Kner, R. (1858). Zur Familie der Characinen. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch-Naturwissenschaftliche Classe, 32: 163-168.

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(6), 1547-1549.

Leigh, J. W., & Bryant, D. (2015). POPART: full-feature software for haplotype network construction. Methods in ecology and evolution, 6(9), 1110-1116.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25(11), 1451-1452.

Linnaeus, C. (1758). Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. Editio decima, reformata. Holmiae 1: 1-824.

Lucena, C. A. S. (1987). Revisão e redefinição do gênero neotropical *Charax* Scopoli, 1777 com a descrição de quatro espécies novas (Pisces; Characiformes; Characidae). Comunicações do Museu de Cências da PUCRS, Porto Alegre, 40: 5-124.

Lucena, C. A. S. (1989). Trois nouvelles espèces du genre Charax Scopoli, 1777 pour la region Nord du Brésil (Characiformes, Characidae, Characinae). Revue francaise d'Aquariologie Herpétologie, 15: 97-104.

Lucena, C. D. (1998). Relações filogenéticas e definição do gênero *Roeboides* Günther (Ostariophysi: Characiformes: Characidae). Comunicações do Museu de Ciências e Tecnologia da PUCRS, série Zoologia, 11, 19-59.

Lucena CAS (2000) Revisão taxonômica e filogenia das espécies transandinas do gênero *Roeboides* Günther (Teleostei: Ostariophysi: Characiformes). Comunicações Museu Ciências Tecnologia PUCRS, Série Zoologia 13: 3–63.

Lucena, C. A. S. (2003). Revisão taxonômica e relações filogenéticas das espécies de *Roeboides* grupo-*microlepis* (Ostariophysi, Characiformes, Characidae). Iheringia. Série Zoologia, 93(3), 283-308.

Lucena, C. A. S. D. (2007). Revisão taxonômica das espécies do gênero *Roeboides* grupo-*affinis* (Ostariophysi, Characiformes, Characidae). Iheringia. Série Zoologia, 97(2), 117-136.

Lucena, C. A. S. (2011). A new species of *Roeboides* from Panama (Characiformes: Characidae). Revista de Biología Tropical, 4: 1663-1667.

Machado, V. N., Collins, R. A., Ota, R. P., Andrade, M. C., Farias, I. P., & Hrbek, T. (2018). One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. Scientific reports, 8(1), 1-12.

Matamoros, W. A., Chakrabarty, P., Angulo, A., Garita-Alvarado, C. A., & McMahan, C. D. (2013). A new species of *Roeboides* (Teleostei: Characidae) from Costa Rica and Panama, with a key to the middle American species of the genus. Neotropical ichthyology, 11(2), 285-290.

Mateussi, N. T., Pavanelli, C. S., & Oliveira, C. (2017). Molecular identification of cryptic diversity in species of cis-Andean *Mylossoma* (Characiformes: Serrasalmidae). Mitochondrial Dna Part A, 28(5), 778-780.

Mattox, G. M., & Toledo-Piza, M. (2012). Phylogenetic study of the Characinae (Teleostei: Characiformes: Characidae). Zoological Journal of the Linnean Society, 165(4), 809-915.

Mattox et al. (2017). Subfamily Characinae. In: van der Sleen, P., & Albert, J. S. (Eds.). (2017). Field guide to the fishes of the Amazon, Orinoco, and Guianas (Vol. 115). Princeton University Press.

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

Melo, B. F., Ochoa, L. E., Vari, R. P., & Oliveira, C. (2016). Cryptic species in the Neotropical fish genus *Curimatopsis* (Teleostei, Characiformes). Zoologica Scripta, 45(6), 650-658.

Menezes, N. A., & Lucena, C. A. S. D. (2014). A taxonomic review of the species of *Charax* Scopoli, 1777 (Teleostei: Characidae: Characinae) with description of a new species from the rio Negro bearing superficial neuromasts on body scales, Amazon basin, Brazil. Neotropical Ichthyology, 12(2), 193-228.

Miller, M.A., Pfeiffer, W., & Schwartz, T. (2010) "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1 - 8.

Ochoa, L. E., Melo, B. F., García-Melo, J. E., Maldonado-Ocampo, J. A., Souza, C. S., Albornoz-Garzón, J. G., ... & Oliveira, C. (2020). Species delimitation reveals an underestimated diversity of Andean catfishes of the family Astroblepidae (Teleostei: Siluriformes). Neotropical Ichthyology, 18(4).

Pereira, L. H., Hanner, R., Foresti, F., & Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna?. BMC genetics, 14(1), 1-14.

Puillandre, N., Lambert, A., Brouillet, S., & ACHAZ, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular ecology, 21(8), 1864-1877.

Rossini, B. C., Oliveira, C. A. M., Melo, F. A. G. D., Bertaco, V. D. A., Astarloa, J. M. D. D., Rosso, J. J., ... & Oliveira, C. (2016). Highlighting *Astyanax* species diversity through DNA barcoding. PLoS One, 11(12), e0167203.

Sazima, I., & Machado, F. A. (1983). Hábitos e comportamento de *Roeboides prognathus*, um peixe lepidófago (Osteichthyes, Characoidei). Boletim de Zoologia, 7(7), 37-56.

Sigovini, M., Keppel, E., & Tagliapietra, D. (2016). Open Nomenclature in the biodiversity era. Methods in Ecology and Evolution, 7(10), 1217-1225.

Stamatakis, A., Hoover, P., & Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. Systematic biology, 57(5), 758-771.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312-1313.

Xia, X. (2013). DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. Molecular biology and evolution, 30(7), 1720-1728.

Xia, X., Xie, Z., Salemi, M., Chen, L., & Wang, Y. (2003). An index of substitution saturation and its application. Molecular phylogenetics and evolution, 26(1), 1-7.

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics, 29(22), 2869-2876.

Ward, R. D., Hanner, R., & Hebert, P. D. (2009). The campaign to DNA barcode all fishes, FISH-BOL. Journal of fish biology, 74(2), 329-356.

Wendt, E. W., Silva, P. C., Malabarba, L. R., & Carvalho, T. P. (2019). Phylogenetic relationships and historical biogeography of *Oligosarcus* (Teleostei: Characidae): Examining riverine landscape evolution in southeastern South America. Molecular phylogenetics and evolution, 140, 106604.

Supplementary information

Supplementary Table 1. List of the all specimens used in species delimitation analyses.

Genus	Species	Drainage	Country	Coordinates	Voucher	Catalog number	GenBank n.
Charax	*C. condei*	Negro	Brazil	S 00°30'5.3" W 64°49'12.2"	LBP74418	LBP 18319	
Charax	*C. condei*	Negro	Brazil	S 00°30'5.3" W 64°49'12.2"	LBP74419	LBP 18319	
Charax	*C. condei*	Negro	Brazil	S 00°30'5.3" W 64°49'12.2"	LBP74420	LBP 18319	
Charax	*C. gibbosus*	Mana	French Guiana	N 5°38'57.8" W 53°48'57.2"	HYD15133	MHNG uncat	
Charax	*C. gibbosus*	Mana	French Guiana	N 5°38'54.1" W 53°45'55.4"	PAG9030	MHNG uncat	
Charax	*C. gibbosus*	Corantijn	Suriname	-	T19138	ROM uncat	
Charax	*C. gibbosus*	Tampok	French Guiana	3°23'31.8"N 53°49'52.0"W	GF15480	MHNG uncat	
Charax	*C. leticiae*	Paraguay	Brazil	S 18°25'42.5" W 54°50'02.8"	LBP12700	LBP 1480	
Charax	*C. leticiae*	Araguaia	Brazil	S 13°19'22.8' W 50°37'20.7"	LBP44041	LBP 8771	
Charax	*C. leticiae*	Araguaia	Brazil	S 13°19'22.8' W 50°37'20.7"	LBP44043	LBP 8771	
Charax	*C. leticiae*	Araguaia	Brazil	S 13°19'22.8' W 50°37'20.7"	LBP44044	LBP 8771	
Charax	*C. leticiae*	Araguaia	Brazil	S 19°34'54.6' W 56°15'16.5"	LBP21982	LBP 3730	
Charax	*C. leticiae*	Araguaia	Brazil	S 19°34'54.6' W 56°15'16.5"	LBP21985	LBP 3730	
Charax	*C. leticiae*	Araguaia	Brazil	S 17°49'26.7" W 57°31'03.0"	LBP56302	LBP 13542	
Charax	*C. leticiae*	Araguaia	Brazil	S 17°49'26.7" W 57°31'03.0"	LBP56303	LBP 13542	
Charax	*C. leticiae*	Araguaia	Brazil	S 17°49'29.8" W 57°30'42.3"	LBP56717	LBP 13658	
Charax	*C. leticiae*	Araguaia	Brazil	S 17°47'33.7" W 57°33'26.7"	LBP56898	LBP 13706	
Charax	*C. leticiae*	Araguaia	Brazil	S 17°47'33.7" W 57°33'26.7"	LBP56899	LBP 13706	
Charax	*C. metae*	Cubillera	Colombia	N 3°29'26.6" W 73°44'34.1"	LBP61598	LBP 18653	
Charax	*C. niger*	Amazon	Brazil	N 2°03'42.8" W 50°54'15.1"	LBP83211	LBP 21217	

Genus	Species	Drainage	Country	Coordinates	Voucher	Catalog number	GenBank n.
Charax	*C. niger*	Amazon	Brazil	N 2°03'42.8" W 50°54'15.1"	LBP83211	LBP 21217	
Charax	*C. pauciradiatus*	Jari	Brazil	S 00°56'00" W 52°32'30"	LBP27152	LBP 5426	
Charax	*C. pauciradiatus*	Tapajós	Brazil	S 04°52'14.1" W 56°51'19.0"	LBP59086	LBP 14117	
Charax	*C. pauciradiatus*	Tapajós	Brazil	S 04°52'14.1" W 56°51'19.0"	LBP59087	LBP 14117	
Charax	*C. pauciradiatus*	Tapajós	Brazil	S 04°52'14.1" W 56°51'19.0"	LBP59089	LBP 14117	
Charax	*C. pauciradiatus*	Tapajós	Brazil	S 04°52'14.1" W 56°51'19.0"	LBP59090	LBP 14117	
Charax	*C. pauciradiatus*	Tapajós	Brazil	S 04°28'11.2" W 56°17'01.1"	LBP66965	LBP 16206	
Charax	*C. pauciradiatus*	Tapajós	Brazil	S 04°28'11.2" W 56°17'01.1"	LBP66967	LBP 16206	
Charax	*C. pauciradiatus*	Tapajós	Brazil	S 04°28'11.2" W 56°17'01.1"	LBP66968	LBP 16206	
Charax	*C. stenopterus*	Costal	Brazil	S 33°41'22.6" W 53°26'22.3"	LBP60810	LBP 14529	
Charax	*C. stenopterus*	Grande	Brazil	S 32°05'24.1" W 52°15'09.7"	LBP68399	LBP 17020	
Charax	*C. stenopterus*	Parana	Brazil	S 31°38'26.4" W 60°38'10.9"	LBP79318	LBP 19942	
Charax	*C. stenopterus*	Parana	Brazil	S 31°38'26.4" W 60°38'10.9"	LBP79320	LBP 19942	
Charax	*C. stenopterus*	Uruguay	Brazil	S 29°30'42.8" W 56°43'09.9"	LBP55077	LBP 13169	
Charax	*C. tectifer*	Aguarico	Ecuador	N 0°00'30.4" W 77°23'40.9"	T19568	ROM uncat	
Charax	*C. tectifer*	Aguarico	Ecuador	N 0°01'43.8" W 77°23'36.2"	T19592	ROM uncat	
Charax	*C. tectifer*	Ucayali	Peru	S 08°39'57.2" W 74°48'08.7"	LBP70280	LBP 17783	
Charax	*C. tectifer*	Ucayali	Peru	S 08°39'57.2" W 74°48'08.7"	LBP70281	LBP 17783	
Charax	*C. tectifer*	Ucayali	Peru	S 08°39'57.2" W 74°48'08.7"	LBP70282	LBP 17783	
Phenacogaster	*P. calverti*	Parnaiba	Brazil	S 07°30'55’ W 46°04'53'	LBP27355	LBP 5612	
Phenacogaster	*P. franciscoensis*	São Francisco	Brazil	S 17°14'32.3' W 46°28'03.8"	LBP49190	LBP 10487	
Roeboides	*R. affinis*	Amazon	Brazil	-	LBP57585	LBP 14776	
Roeboides	*R. affinis*	Amazon	Brazil	S 04°19'28.0'’ W 69°57'36.7"	LBP87171	LBP 22564	
Roeboides	*R. affinis*	Amazon	Brazil	S 04°14'04.6'’ W 69°57'42.2"	LBP87882	LBP 22607	
Roeboides	*R. affinis*	Amazon	Brazil	S 04°14'04.6'’ W 69°57'42.2"	LBP87883	LBP 22607	
Roeboides	*R. affinis*	Amazon	Brazil	S 04°11'39.3'’ W 69°58'06.9"	LBP87947	LBP 22638	

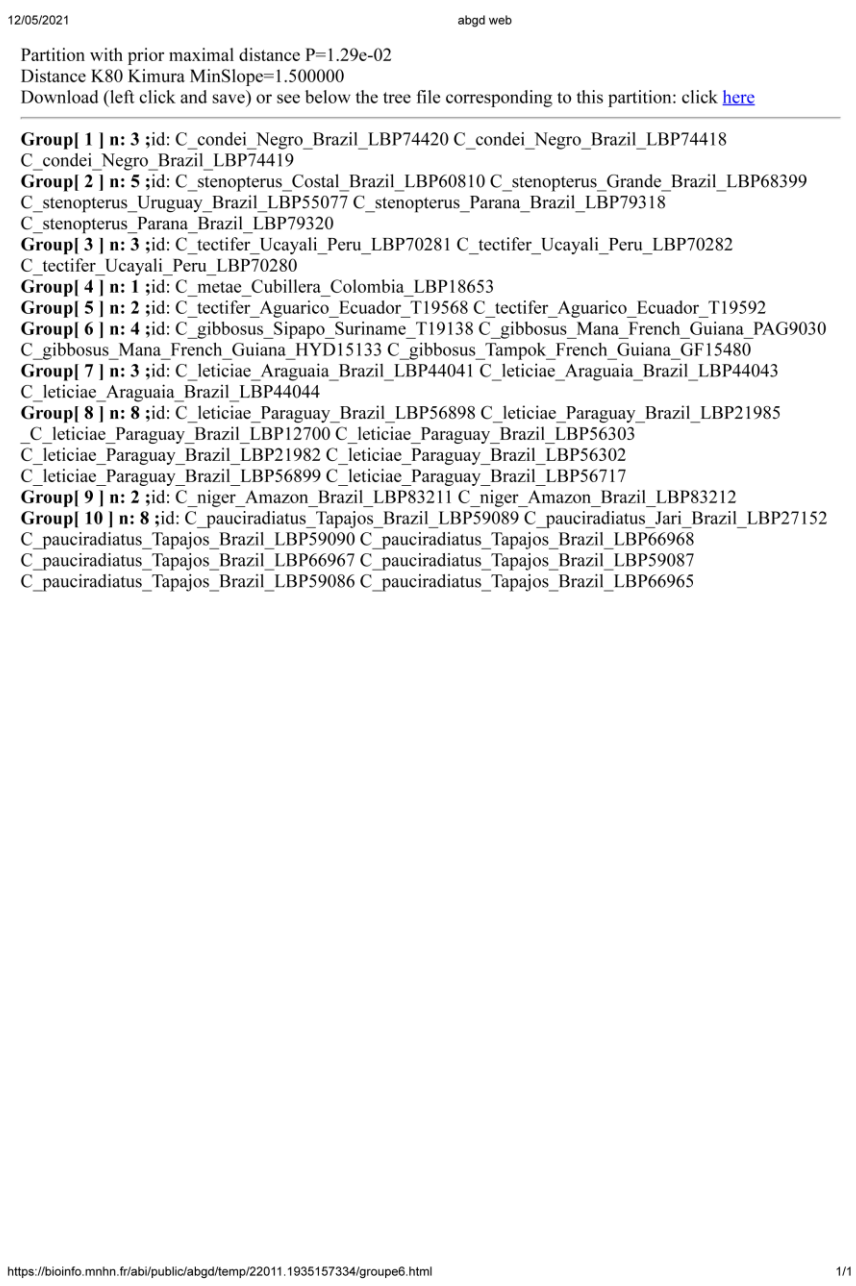
Genus	Species	Drainage	Country	Coordinates	Voucher	Catalog number	GenBank n.
Roeboides	*R. affinis*	Amazon	Brazil	S 04°11'39.3'' W 69°58'06.9"	LBP87948	LBP 22638	
Roeboides	*R. affinis*	Amazon	Brazil	-	LBP57588	LBP 14776	
Roeboides	*R. affinis*	Amazon	Brazil	S 04°11'39.3'' W 69°58'06.9"	LBP87949	LBP 22638	
Roeboides	*R. affinis*	Araguaia	Brazil	S 15°53'35.6" W 52°15'01.0"	LBP12908	LBP 1837	
Roeboides	*R. affinis*	Araguaia	Brazil	S 15°53'31.5" W 52°15'02.0"	LBP28119	LBP 5753	
Roeboides	*R. affinis*	Araguaia	Brazil	-	LBP36826	LBP 7828	
Roeboides	*R. affinis*	Araguaia	Brazil	S 13°18'37.3" W 50°36'47.6"	LBP41030	LBP 12795	
Roeboides	*R. affinis*	Araguaia	Brazil	S 13°18'37.3" W 50°36'47.6"	LBP41031	LBP 12795	
Roeboides	*R. affinis*	Araguaia	Brazil	S 13°19'00.0" W 50°37'00.0"	LBP43610	LBP 12724	
Roeboides	*R. affinis*	Araguaia	Brazil	S 13°22'36.1' W 50°40'08.4"	LBP44276	LBP 8850	
Roeboides	*R. affinis*	Araguaia	Brazil	S 15°10'23.2" W 51°09'27.1"	LBP68750	LBP 17195	
Roeboides	*R. affinis*	Araguaia	Brazil	S 15°10'23.2" W 51°09'27.1"	LBP68752	LBP 17195	
Roeboides	*R. affinis*	Araguaia	Brazil	S 15°10'23.2" W 51°09'27.1"	LBP68753	LBP 17195	
Roeboides	*R. affinis*	Juruá	Brazil	S 7°37'20.0" W 72°47'42.2"	LBP22929	LBP 4050	
Roeboides	*R. affinis*	Juruá	Brazil	S 7°34'28.8' W 72°55'24.9"	LBP23521	LBP 4088	
Roeboides	*R. affinis*	Juruá	Brazil	-	LBP23657	LBP 4135	
Roeboides	*R. affinis*	Juruá	Brazil	-	LBP23658	LBP 4135	
Roeboides	*R. affinis*	Madeira	Brazil	S 11°54'45.2" W 61°13'36.7"	LBP95051	LBP 24629	
Roeboides	*R. affinis*	Madeira	Brazil	S 11°54'13.0" W 61°14'08.1"	LBP95100	LBP 25381	
Roeboides	*R. affinis*	Madeira	Brazil	S 11°54'13.0" W 61°14'08.1"	LBP95102	LBP 25381	
Roeboides	*R. affinis*	Madeira	Brazil	S 11°55'25.5" W 61°14'17.6"	LBP95187	LBP 25783	
Roeboides	*R. affinis*	Madeira	Brazil	S 11°55'25.5" W 61°14'17.6"	LBP95188	LBP 25783	
Roeboides	*R. affinis*	Madeira	Brazil	S 11°54'11.1" W 61°14'12.5"	LBP95207	LBP 25405	
Roeboides	*R. affinis*	Madeira	Brazil	S 11°54'11.1" W 61°14'12.5"	LBP95208	LBP 25405	
Roeboides	*R. affinis*	Solimões	Brazil	S 04°19'28.0'' W 69°57'36.7"	LBP87170	LBP 22564	
Roeboides	*R. affinis*	Solimões	Brazil	S 04°19'28.0'' W 69°57'36.7"	LBP87172	LBP 22564	

Genus	Species	Drainage	Country	Coordinates	Voucher	Catalog number	GenBank n.
Roeboides	*R. affinis*	Solimões	Brazil	S 04°19'28.0'' W 69°57'36.7"	LBP87173	LBP 22564	
Roeboides	*R. affinis*	Tocantins	Brazil	S 12°37'27.9" W 48°02'43.5"	LBP93882	LBP 25603	
Roeboides	*R. affinis*	Solimões	Brazil	S 04°19'28.0'' W 69°57'36.7"	LBP87153	LBP 22563	
Roeboides	*R. bouchellei*	Canas	Costa Rica	N 10°20'53.9" W 85°10'07.7"	STRI2173	STRI-00426	MG496227.1
Roeboides	*R. bouchellei*	Grande de Matagalpa	Nicaragua	N 13°12'29.9" W 84°56'59.6"	STRI14271	STRI-05975	MG496231.1
Roeboides	*R. bouchellei*	Higueron	Costa Rica	N 10°20'33.7" W 85°04'33.2"	STRI2116	STRI-00424	MG496228.1
Roeboides	*R. bouchellei*	Pizote	Costa Rica	N 10°54'30.2" W 85°12'40.7"	STRI2150	STRI-00428	MG496230.1
Roeboides	*R. carti*	Mandinga	Panama	N 9°28'11.8" W 79°07'27.1"	STRI321	STRI-00416	MG937234.1
Roeboides	*R. carti*	Playon Chico	Panama	N 9°15'49.8" W 78°13'19.2"	STRI4950	STRI-00402	MG937233.1
Roeboides	*R. dayi*	Magdalena	Colombia	N 4°42'04.0" W 76°51'10.1"	LBP90742	LBP 24360	
Roeboides	*R. dayi*	Magdalena	Colombia	N 5°11'47.2" W 74°45'52.6"	LBP91161	LBP 24273	
Roeboides	*R. biserialis*	Amazon	Brazil	S 03°07'07.0" W 58°2714.7"	LBP72495	LBP 18001	
Roeboides	*R. biserialis*	Amazon	Brazil	S 03°07'07.0" W 58°2714.7"	LBP72496	LBP 18001	
Roeboides	*R. biserialis*	Amazon	Brazil	S 03°07'07.0" W 58°2714.7"	LBP72497	LBP 18001	
Roeboides	*R. biserialis*	Amazon	Brazil	S 03°05'57.7" W 58°2718.0"	LBP72597	LBP 18034	
Roeboides	*R. biserialis*	Amazon	Brazil	S 03°05'57.7" W 58°2718.0"	LBP72599	LBP 18034	
Roeboides	*R. biserialis*	Amazon	Brazil	S 03°05'57.7" W 58°2718.0"	LBP72600	LBP 18034	
Roeboides	*R. descalvadensis*	Parana	Argentina	S 31°29'39.6" W 60°26'47.0"	LBP79252	LBP 19912	
Roeboides	*R. descalvadensis*	Aricá Mirim	Brazil	S 15°44'3.60" W 55°52'48.7"	LBP28244	LBP 5801	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 18°25'42.5" W 54°50'02.8"	LBP12701	LBP 7207	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 14°40'32.8" W 56°13'14.0"	LBP19482	LBP 3232	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 19°34'17.3' W 56°14'44.8"	LBP22253	LBP 3834	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 19°34'17.3' W 56°14'44.8"	LBP22254	LBP 3834	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 15°54'03' W 56°01'17"	LBP22784	LBP 3958	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 16°06'66" W 57°44'33"	LBP26152	LBP 5110	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 16°06'66" W 57°44'33"	LBP26153	LBP 5110	

Genus	Species	Drainage	Country	Coordinates	Voucher	Catalog number	GenBank n.
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 16°06'66" W 57°44'33"	LBP26154	LBP 5110	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 19°47'03.9' W 56°49'19.8"	LBP45363	LBP 9846	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 19°26'02.1' W 57°03'08.9"	LBP45492	LBP 9886	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 19°26'02.1' W 57°03'08.9"	LBP45493	LBP 9886	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 17°49'51.7" W 57°23'42.8"	LBP58428	LBP 14066	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 17°49'51.7" W 57°23'42.8"	LBP58429	LBP 14066	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 17°49'51.7" W 57°23'42.8"	LBP58430	LBP 14066	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 17°49'51.7" W 57°23'42.8"	LBP58431	LBP 14066	
Roeboides	*R. descalvadensis*	Paraguay	Brazil	S 17°49'51.7" W 57°23'42.8"	LBP58432	LBP 14066	
Roeboides	*R. descalvadensis*	Parana	Argentina	S 31°40'23.9" W 60°34'44.9"	LBP79211	LBP 19899	
Roeboides	*R. descalvadensis*	Parana	Argentina	S 31°30'36.4" W 60°28'12.6"	LBP79300	LBP 19934	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°43'03.2" W 53°17'27.6"	LBP17314	LBP 2619	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°43'03.2" W 53°17'27.6"	LBP17316	LBP 2619	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°47’29’’ W 53°20’58’’	LBP26387	LBP 5215	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°47’29’’ W 53°20’58’’	LBP26388	LBP 5215	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°47’29’’ W 53°20’58’’	LBP26389	LBP 5215	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°47’29’’ W 53°20’58’’	LBP26390	LBP 5215	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°47’29’’ W 53°20’58’’	LBP26391	LBP 5215	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°51'45.3" W 48°06'21.0"	LBP43168	LBP 9211	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°51'45.3" W 48°06'21.0"	LBP43169	LBP 9211	
Roeboides	*R. descalvadensis*	Parana	Brazil	S 22°02'52.3' W 53°41'38.1"	LBP45656	LBP 9642	
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 09°38'53.8' W 72°34'56.4"	LBP29546	LBP 6106	
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 09°38'53.8' W 72°34'56.4"	LBP29547	LBP 6106	
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 09°38'53.8' W 72°34'56.4"	LBP29548	LBP 6106	
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 09°38'53.8' W 72°34'56.4"	LBP29549	LBP 6106	
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 09°38'53.8' W 72°34'56.4"	LBP29550	LBP 6106	

Genus	Species	Drainage	Country	Coordinates	Voucher	Catalog number	GenBank n.
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 09°38'53.8' W 72°34'56.4"	LBP29551	LBP 6106	
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 10°01'42.0" W 72°25'58.0"	LBP29577	LBP 6114	
Roeboides	*R. dientonito*	Maracaibo	Venezuela	N 10°01'42.0" W 72°25'58.0"	LBP29578	LBP 6114	
Roeboides	*R. dientonito*	Orinoco	Venezuela	N 7º30'50.9" W 66º09'19.8"	LBP15632	LBP 2220	
Roeboides	*R. dientonito*	Orinoco	Venezuela	N 7º30'50.9" W 66º09'19.8"	LBP15633	LBP 2220	
Roeboides	*R. dientonito*	Orinoco	Venezuela	N 7º30'50.9" W 66º09'19.8"	LBP15635	LBP 2220	
Roeboides	*R. dientonito*	Orinoco	Venezuela	N 7º30'50.9" W 66º09'19.8"	LBP15636	LBP 2220	
Roeboides	*R. dispar*	Acre	Brazil	S 10°03'28.6" W 67°51'25.6"	LBP49266	LBP 10150	
Roeboides	*R. dispar*	Acre	Brazil	S 10°03'28.6" W 67°51'25.6"	LBP49270	LBP 10150	
Roeboides	*R. guatemalensis*	Chagres	Panama	N 9°21'36.3" W 79°19'20.3"	STRI15205	STRI-07236	MG937237.1
Roeboides	*R. guatemalensis*	Cocoli	Panama	N 8°58'49.9" W 79°37'21.7"	STRI9141	STRI-00407	MG937238.1
Roeboides	*R. guatemalensis*	Llano Sucio	Panama	N 09°19’26.2" W 79°46'08.2"	LBP18531	LBP 2755	
Roeboides	*R. guatemalensis*	Llano Sucio	Panama	N 09°19’26.2" W 79°46'08.2"	LBP18532	LBP 2755	
Roeboides	*R. ilseae*	Salama Nuevo	Costa Rica	N 8°54'15.3" W 83°26'21.6"	STRI2018	STRI-00431	MG496234.1
Roeboides	*R. ilseae*	Salama Nuevo	Costa Rica	N 8°54'15.3" W 83°26'21.6"	STRI2020	STRI-00431	
Roeboides	*R. loftini*	Cascajal	Panama	N 8°48'16.7" W 80°31'59.9"	STRI15327	STRI-05331	MG937241.1
Roeboides	*R. loftini*	Tambo	Panama	N 8°39'56.4" W 80°17'15.0"	AM18	STRI-00400	MG937240.1
Roeboides	*R. margareteae*	Mearim	Brazil	S 3°39'11" W 45°46'25"	CICCAA020751	UFMA uncat	
Roeboides	*R. margareteae*	Mearim	Brazil	S 3°43'48" W 45°35'7"	CICCAA020892	UFMA uncat	
Roeboides	*R. microlepis*	Paraguay	Brazil	S 16°06'66" W 57°44'33"	LBP26124	LBP 5098	
Roeboides	*R. microlepis*	Paraguay	Brazil	S 16°06'66" W 57°44'33"	LBP26125	LBP 5098	
Roeboides	*R. microlepis*	Paraguay	Brazil	S 16°05'14.8" W 57°42'17.8"	LBP44481	LBP 9248	
Roeboides	*R. microlepis*	Paraguay	Brazil	S 16°05'14.8" W 57°42'17.8"	LBP44482	LBP 9248	
Roeboides	*R. myersii*	Amazon	Brazil	S 2º14'15.6" W 54º 4' 35.7"	MCP52389	MCP 52389	
Roeboides	*R. myersii*	Amazon	Colombia	S 04°11'45.6'’ W 69°57'20.9"	LBP87569	LBP 22518	
Roeboides	*R. myersii*	Amazon	Colombia	S 04°11'45.6'’ W 69°57'20.9"	LBP87570	LBP 22518	

Genus	Species	Drainage	Country	Coordinates	Voucher	Catalog number	GenBank n.
Roeboides	*R. numerosus*	Orinoco	Venezuela	N 07°37'24.4' W 66°24'48.0"	LBP47538	LBP 10217	
Roeboides	*R. occidentalis*	Cocle del Sur	Panama	N 8°37'16.6" W 80°26'55.7"	STRI17718	STRI-07172	MG937245.1
Roeboides	*R. occidentalis*	Santa Maria	Panama	N 8°08'06.5" W 80°34'35.0"	STRI3404	STRI-00460	MG937247.1
Roeboides	*R. sazimai*	Mearim	Brazil	S 3°43'48.0" W 45°35'07.0"	CICCAA021111	UFMA uncat	
Roeboides	*R. sazimai*	Mearim	Brazil	S 3°43'48.0" W 45°35'07.0"	CICCAA021112	UFMA uncat	
Roeboides	*R. sazimai*	Mearim	Brazil	S 3°43'48.0" W 45°35'07.0"	CICCAA021113	UFMA uncat	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 15°19'24.2' W 43°39'52.5"	LBP38312	LBP 8267	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 15°19'24.2' W 43°39'52.5"	LBP38313	LBP 8267	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°13'33.7' W 44°48'27.9"	LBP47328	LBP 10324	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°13'33.7' W 44°48'27.9"	LBP47330	LBP 10324	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°13'33.7' W 44°48'27.9"	LBP47331	LBP 10324	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°13'33.7' W 44°48'27.9"	LBP47332	LBP 10324	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°17'21.9' W 44°47'08.4"	LBP48879	LBP 10387	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°14'19.9' W 46°23'39.0"	LBP49054	LBP 10444	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°14'19.9' W 46°23'39.0"	LBP49055	LBP 10444	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°14'19.9' W 46°23'39.0"	LBP49056	LBP 10444	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°14'19.9' W 46°23'39.0"	LBP49057	LBP 10444	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 17°14'19.9' W 46°23'39.0"	LBP49058	LBP 10444	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 09°55'28.4" W 37°07'22.5"	LBP59738	LBP 11559	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 09°55'28.4" W 37°07'22.5"	LBP59739	LBP 11559	
Roeboides	*R. xenodon*	São Francisco	Brazil	S 09°55'28.4" W 37°07'22.5"	LBP59740	LBP 11559	
Tetragonopterus	*T. carvalhoi*	Jari	Brazil	S 00°33"51" W 52°34'45"	LBP27022	LBP 5376	



12/05/2021 abgd web

Partition with prior maximal distance P=1.29e-02
Distance K80 Kimura MinSlope=1.500000
Download (left click and save) or see below the tree file corresponding to this partition: click here

---

Group[1] n: 3 ;id: C_condei_Negro_Brazil_LBP74420 C_condei_Negro_Brazil_LBP74418 C_condei_Negro_Brazil_LBP74419
Group[2] n: 5 ;id: C_stenopterus_Costal_Brazil_LBP60810 C_stenopterus_Grande_Brazil_LBP68399 C_stenopterus_Uruguay_Brazil_LBP55077 C_stenopterus_Parana_Brazil_LBP79318 C_stenopterus_Parana_Brazil_LBP79320
Group[3] n: 3 ;id: C_tectifer_Ucayali_Peru_LBP70281 C_tectifer_Ucayali_Peru_LBP70282 C_tectifer_Ucayali_Peru_LBP70280
Group[4] n: 1 ;id: C_metae_Cubillera_Colombia_LBP18653
Group[5] n: 2 ;id: C_tectifer_Aguarico_Ecuador_T19568 C_tectifer_Aguarico_Ecuador_T19592
Group[6] n: 4 ;id: C_gibbosus_Sipapo_Suriname_T19138 C_gibbosus_Mana_French_Guiana_PAG9030 C_gibbosus_Mana_French_Guiana_HYD15133 C_gibbosus_Tampok_French_Guiana_GF15480
Group[7] n: 3 ;id: C_leticiae_Araguaia_Brazil_LBP44041 C_leticiae_Araguaia_Brazil_LBP44043 C_leticiae_Araguaia_Brazil_LBP44044
Group[8] n: 8 ;id: C_leticiae_Paraguay_Brazil_LBP56898 C_leticiae_Paraguay_Brazil_LBP21985 _C_leticiae_Paraguay_Brazil_LBP12700 C_leticiae_Paraguay_Brazil_LBP56303 C_leticiae_Paraguay_Brazil_LBP21982 C_leticiae_Paraguay_Brazil_LBP56302 C_leticiae_Paraguay_Brazil_LBP56899 C_leticiae_Paraguay_Brazil_LBP56717
Group[9] n: 2 ;id: C_niger_Amazon_Brazil_LBP83211 C_niger_Amazon_Brazil_LBP83212
Group[10] n: 8 ;id: C_pauciradiatus_Tapajos_Brazil_LBP59089 C_pauciradiatus_Jari_Brazil_LBP27152 C_pauciradiatus_Tapajos_Brazil_LBP59090 C_pauciradiatus_Tapajos_Brazil_LBP66968 C_pauciradiatus_Tapajos_Brazil_LBP66967 C_pauciradiatus_Tapajos_Brazil_LBP59087 C_pauciradiatus_Tapajos_Brazil_LBP59086 C_pauciradiatus_Tapajos_Brazil_LBP66965

https://bioinfo.mnhn.fr/abi/public/abgd/temp/22011.1935157334/groupe6.html 1/1

Supplementary Fig. 1. ABGD dataset reporting the number of groups delimitated to *Charax*.

12/05/2021 abgd web

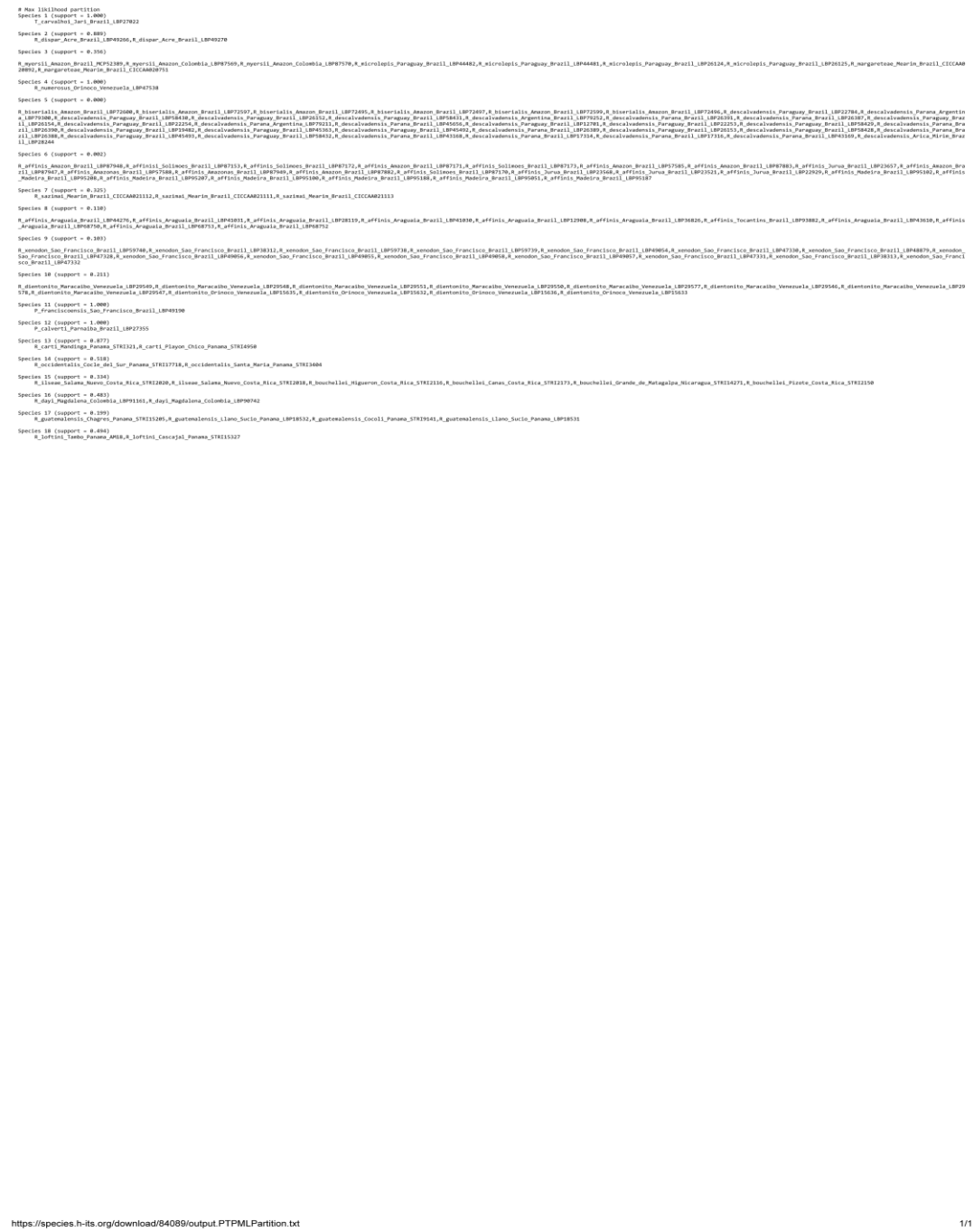
Initial Partition with prior maximal distance P=1.29e-02 ; Barcode gap distance = 0.032
Distance K80 Kimura MinSlope=1.500000
Download (left click and save) or see below the tree file corresponding to this partition: click here

Group[1] n: 12 ;id: R_dientonito_Orinoco_Venezuela_LBP15636 R_dientonito_Orinoco_Venezuela_LBP15632 R_dientonito_Orinoco_Venezuela_LBP15635 R_dientonito_Orinoco_Venezuela_LBP15633 R_dientonito_Maracaibo_Venezuela_LBP29548 R_dientonito_Maracaibo_Venezuela_LBP29549 R_dientonito_Maracaibo_Venezuela_LBP29551 R_dientonito_Maracaibo_Venezuela_LBP29547 R_dientonito_Maracaibo_Venezuela_LBP29546 R_dientonito_Maracaibo_Venezuela_LBP29550 R_dientonito_Maracaibo_Venezuela_LBP29577 R_dientonito_Maracaibo_Venezuela_LBP29578
Group[2] n: 2 ;id: R_dayi_Magdalena_Colombia_LBP90742 R_dayi_Magdalena_Colombia_LBP91161
Group[3] n: 6 ;id: R_ilseae_Salama_Nuevo_Costa_Rica_STRI2018 R_ilseae_Salama_Nuevo_Costa_Rica_STRI2020 R_bouchellei_Grande_de_Matagalpa_Nicaragua_STRI14271 R_bouchellei_Pizote_Costa_Rica_STRI2150 R_bouchellei_Higueron_Costa_Rica_STRI2116 R_bouchellei_Canas_Costa_Rica_STRI2173
Group[4] n: 6 ;id: R_loftini_Cascajal_Panama_STRI15327 R_loftini_Tambo_Panama_AM18 R_guatemalensis_Llano_Sucio_Panama_LBP18531 R_guatemalensis_Llano_Sucio_Panama_LBP18532 R_guatemalensis_Cocoli_Panama_STRI9141 R_guatemalensis_Chagres_Panama_STRI15205
Group[5] n: 2 ;id: R_occidentalis_Santa_Maria_Panama_STRI3404 R_occidentalis_Cocle_del_Sur_Panama_STRI17718
Group[6] n: 2 ;id: R_carti_Playon_Chico_Panama_STRI4950 R_carti_Mandinga_Panama_STRI321
Group[7] n: 2 ;id: R_dispar_Acre_Brazil_LBP49266 R_dispar_Acre_Brazil_LBP49270
Group[8] n: 9 ;id: R_microlepis_Paraguay_Brazil_LBP44481 R_microlepis_Paraguay_Brazil_LBP26125 R_microlepis_Paraguay_Brazil_LBP26124 R_microlepis_Paraguay_Brazil_LBP44482 R_margareteae_Mearim_Brazil_CICCAA020892 R_margareteae_Mearim_Brazil_CICCAA020751 R_myersii_Amazon_Colombia_LBP87570 R_myersii_Amazon_Colombia_LBP87569 R_myersii_Amazon_Brazil_MCP52389
Group[9] n: 1 ;id: R_numerosus_Orinoco_Venezuela_LBP47538
Group[10] n: 15 ;id: R_xenodon_Sao_Francisco_Brazil_LBP47332 R_xenodon_Sao_Francisco_Brazil_LBP38313 R_xenodon_Sao_Francisco_Brazil_LBP49057 R_xenodon_Sao_Francisco_Brazil_LBP49058 R_xenodon_Sao_Francisco_Brazil_LBP47331 R_xenodon_Sao_Francisco_Brazil_LBP59738 R_xenodon_Sao_Francisco_Brazil_LBP47330 R_xenodon_Sao_Francisco_Brazil_LBP49054 R_xenodon_Sao_Francisco_Brazil_LBP59739 R_xenodon_Sao_Francisco_Brazil_LBP49056 R_xenodon_Sao_Francisco_Brazil_LBP59740 R_xenodon_Sao_Francisco_Brazil_LBP48879 R_xenodon_Sao_Francisco_Brazil_LBP47328 R_xenodon_Sao_Francisco_Brazil_LBP49055 R_xenodon_Sao_Francisco_Brazil_LBP38312
Group[11] n: 36 ;id: R_biserialis_Amazon_Brazil_LBP72599 R_biserialis_Amazon_Brazil_LBP72497 R_biserialis_Amazon_Brazil_LBP72496 R_biserialis_Amazon_Brazil_LBP72495 R_biserialis_Amazon_Brazil_LBP72597 R_biserialis_Amazon_Brazil_LBP72600 R_descalvadensis_Arica_Mirim_Brazil_LBP28244 R_descalvadensis_Parana_Brazil_LBP43168 R_descalvadensis_Paraguay_Brazil_LBP58430 R_descalvadensis_Paraguay_Brazil_LBP58428 R_descalvadensis_Parana_Brazil_LBP43169 R_descalvadensis_Paraguay_Brazil_LBP22253 R_descalvadensis_Paraguay_Brazil_LBP58432 R_descalvadensis_Paraguay_Brazil_LBP12701 R_descalvadensis_Paraguay_Brazil_LBP26152 R_descalvadensis_Paraguay_Brazil_LBP26153 R_descalvadensis_Paraguay_Brazil_LBP26154 R_descalvadensis_Paraguay_Brazil_LBP45493 R_descalvadensis_Parana_Argentina_LBP79211 R_descalvadensis_Parana_Argentina_LBP79300 R_descalvadensis_Parana_Brazil_LBP26391 R_descalvadensis_Parana_Brazil_LBP26390 R_descalvadensis_Parana_Brazil_LBP26389 R_descalvadensis_Argentina_Brazil_LBP79252 R_descalvadensis_Paraguay_Brazil_LBP45492 R_descalvadensis_Parana_Brazil_LBP45656 R_descalvadensis_Parana_Brazil_LBP26387 R_descalvadensis_Paraguay_Brazil_LBP45363 R_descalvadensis_Parana_Brazil_LBP26388 R_descalvadensis_Paraguay_Brazil_LBP19482 R_descalvadensis_Paraguay_Brazil_LBP58429 R_descalvadensis_Paraguay_Brazil_LBP58431 R_descalvadensis_Paraguay_Brazil_LBP22784 R_descalvadensis_Paraguay_Brazil_LBP22254 R_descalvadensis_Parana_Brazil_LBP17316 R_descalvadensis_Parana_Brazil_LBP17314
Group[12] n: 3 ;id: R_sazimai_Mearim_Brazil_CICCAA021113 R_sazimai_Mearim_Brazil_CICCAA021111 R_sazimai_Mearim_Brazil_CICCAA021112
Group[13] n: 11 ;id: R_affinis_Araguaia_Brazil_LBP28119 R_affinis_Araguaia_Brazil_LBP41030 R_affinis_Araguaia_Brazil_LBP12908 R_affinis_Araguaia_Brazil_LBP36826 R_affinis_Araguaia_Brazil_LBP44276 R_affinis_Araguaia_Brazil_LBP41031 R_affinis_Tocantins_Brazil_LBP93882 R_affinis_Araguaia_Brazil_LBP43610 R_affinis_Araguaia_Brazil_LBP68750 R_affinis_Araguaia_Brazil_LBP68752 R_affinis_Araguaia_Brazil_LBP68753
Group[14] n: 23 ;id: R_affinis_Madeira_Brazil_LBP95100 R_affinis_Madeira_Brazil_LBP95188 R_affinis_Madeira_Brazil_LBP95207 R_affinis_Madeira_Brazil_LBP95051 R_affinis_Madeira_Brazil_LBP95208 R_affinis_Madeira_Brazil_LBP95102 R_affinis_Madeira_Brazil_LBP95187 R_affinis_Jurua_Brazil_LBP23568 R_affinis_Jurua_Brazil_LBP23657 R_affinisi_Solimoes_Brazil_LBP87153 R_affinis_Amazonas_Brazil_LBP57588 R_affinis_Amazonas_Brazil_LBP87949 R_affinis_Solimoes_Brazil_LBP87170 R_affinis_Solimoes_Brazil_LBP87173 R_affinis_Amazon_Brazil_LBP87883 R_affinis_Amazon_Brazil_LBP87882 R_affinis_Solimoes_Brazil_LBP87172 R_affinis_Amazon_Brazil_LBP87171 R_affinis_Amazon_Brazil_LBP57585 R_affinis_Amazon_Brazil_LBP87948 R_affinis_Jurua_Brazil_LBP23521 R_affinis_Amazon_Brazil_LBP87947 R_affinis_Jurua_Brazil_LBP22929

https://bioinfo.mnhn.fr/abi/public/abgd/temp/22336.2147471855/groupe6.init.html 1/1

Supplementary Fig. 2. ABGD dataset reporting the number of groups delimitated to *Roeboides*.

Supplementary Fig. 3. Bayesian Poisson Tree Processes (bPTP) delimitation of species of *Charax*.



```
# Max likilhood partition
Species 1 (support = 1.000)
      T_carvalhoi_Jari_Brazil_LBP27022

Species 2 (support = 0.889)
      R_dispar_Acre_Brazil_LBP49266,R_dispar_Acre_Brazil_LBP49270

Species 3 (support = 0.356)

R_myersii_Amazon_Brazil_MCP52389,R_myersii_Amazon_Colombia_LBP87569,R_myersii_Amazon_Colombia_LBP87570,R_microlepis_Paraguay_Brazil_LBP44482,R_microlepis_Paraguay_Brazil_LBP44481,R_microlepis_Paraguay_Brazil_LBP26124,R_microlepis_Paraguay_Brazil_LBP26125,R_margareteae_Mearim_Brazil_CICCAA02092,R_margareteae_Mearim_Brazil_CICCAA020751

Species 4 (support = 1.000)
      R_numerosus_Orinoco_Venezuela_LBP47538

Species 5 (support = 0.000)

R_biserialis_Amazon_Brazil_LBP72600,R_biserialis_Amazon_Brazil_LBP72597,R_biserialis_Amazon_Brazil_LBP72495,R_biserialis_Amazon_Brazil_LBP72497,R_biserialis_Amazon_Brazil_LBP72599,R_biserialis_Amazon_Brazil_LBP72496,R_descalvadensis_Paraguay_Brazil_LBP22784,R_descalvadensis_Parana_Argentina_LBP79300,R_descalvadensis_Paraguay_Brazil_LBP58430,R_descalvadensis_Paraguay_Brazil_LBP26152,R_descalvadensis_Paraguay_Brazil_LBP58431,R_descalvadensis_Argentina_Brazil_LBP79252,R_descalvadensis_Parana_Brazil_LBP26391,R_descalvadensis_Parana_Brazil_LBP26387,R_descalvadensis_Paraguay_Brazil_LBP26154,R_descalvadensis_Paraguay_Brazil_LBP22254,R_descalvadensis_Parana_Argentina_LBP79213,R_descalvadensis_Parana_Brazil_LBP45656,R_descalvadensis_Paraguay_Brazil_LBP11701,R_descalvadensis_Paraguay_Brazil_LBP22253,R_descalvadensis_Paraguay_Brazil_LBP58429,R_descalvadensis_Parana_Brazil_LBP26390,R_descalvadensis_Paraguay_Brazil_LBP19482,R_descalvadensis_Paraguay_Brazil_LBP45363,R_descalvadensis_Paraguay_Brazil_LBP45492,R_descalvadensis_Parana_Brazil_LBP26389,R_descalvadensis_Paraguay_Brazil_LBP22653,R_descalvadensis_Paraguay_Brazil_LBP58428,R_descalvadensis_Parana_Brazil_LBP26388,R_descalvadensis_Paraguay_Brazil_LBP45493,R_descalvadensis_Paraguay_Brazil_LBP58432,R_descalvadensis_Parana_Brazil_LBP43168,R_descalvadensis_Parana_Brazil_LBP17314,R_descalvadensis_Parana_Brazil_LBP17316,R_descalvadensis_Parana_Brazil_LBP43169,R_descalvadensis_Arica_Mirim_Brazil_LBP28244

Species 6 (support = 0.002)

R_affinis_Amazon_Brazil_LBP87948,R_affinis_Solimoes_Brazil_LBP87153,R_affinis_Solimoes_Brazil_LBP87172,R_affinis_Amazon_Brazil_LBP87171,R_affinis_Solimoes_Brazil_LBP87173,R_affinis_Amazon_Brazil_LBP57585,R_affinis_Amazon_Brazil_LBP87883,R_affinis_Jurua_Brazil_LBP23657,R_affinis_Amazon_Brazil_LBP87947,R_affinis_Amazonas_Brazil_LBP57588,R_affinis_Amazonas_Brazil_LBP87949,R_affinis_Amazon_Brazil_LBP87882,R_affinis_Solimoes_Brazil_LBP87170,R_affinis_Jurua_Brazil_LBP23568,R_affinis_Jurua_Brazil_LBP23521,R_affinis_Jurua_Brazil_LBP22929,R_affinis_Madeira_Brazil_LBP95102,R_affinis_Madeira_Brazil_LBP95208,R_affinis_Madeira_Brazil_LBP95207,R_affinis_Madeira_Brazil_LBP95106,R_affinis_Madeira_Brazil_LBP95188,R_affinis_Madeira_Brazil_LBP95051,R_affinis_Madeira_Brazil_LBP95187

Species 7 (support = 0.325)
      R_sazimai_Mearim_Brazil_CICCAA021112,R_sazimai_Mearim_Brazil_CICCAA021111,R_sazimai_Mearim_Brazil_CICCAA021113

Species 8 (support = 0.110)

R_affinis_Araguaia_Brazil_LBP44276,R_affinis_Araguaia_Brazil_LBP41031,R_affinis_Araguaia_Brazil_LBP28119,R_affinis_Araguaia_Brazil_LBP41030,R_affinis_Araguaia_Brazil_LBP12908,R_affinis_Araguaia_Brazil_LBP36826,R_affinis_Tocantins_Brazil_LBP93882,R_affinis_Araguaia_Brazil_LBP43610,R_affinis_Araguaia_Brazil_LBP08750,R_affinis_Araguaia_Brazil_LBP08753,R_affinis_Araguaia_Brazil_LBP68752

Species 9 (support = 0.103)

R_xenodon_Sao_Francisco_Brazil_LBP59740,R_xenodon_Sao_Francisco_Brazil_LBP38312,R_xenodon_Sao_Francisco_Brazil_LBP59738,R_xenodon_Sao_Francisco_Brazil_LBP59739,R_xenodon_Sao_Francisco_Brazil_LBP49054,R_xenodon_Sao_Francisco_Brazil_LBP47330,R_xenodon_Sao_Francisco_Brazil_LBP48879,R_xenodon_Sao_Francisco_Brazil_LBP47328,R_xenodon_Sao_Francisco_Brazil_LBP49056,R_xenodon_Sao_Francisco_Brazil_LBP49055,R_xenodon_Sao_Francisco_Brazil_LBP49058,R_xenodon_Sao_Francisco_Brazil_LBP49057,R_xenodon_Sao_Francisco_Brazil_LBP47331,R_xenodon_Sao_Francisco_Brazil_LBP38313,R_xenodon_Sao_Francisco_Brazil_LBP47332

Species 10 (support = 0.211)

R_dientonito_Maracaibo_Venezuela_LBP29549,R_dientonito_Maracaibo_Venezuela_LBP29548,R_dientonito_Maracaibo_Venezuela_LBP29551,R_dientonito_Maracaibo_Venezuela_LBP29550,R_dientonito_Maracaibo_Venezuela_LBP29577,R_dientonito_Maracaibo_Venezuela_LBP29546,R_dientonito_Maracaibo_Venezuela_LBP29578,R_dientonito_Maracaibo_Venezuela_LBP29547,R_dientonito_Orinoco_Venezuela_LBP15635,R_dientonito_Orinoco_Venezuela_LBP15632,R_dientonito_Orinoco_Venezuela_LBP15636,R_dientonito_Orinoco_Venezuela_LBP15633

Species 11 (support = 1.000)
      P_franciscoensis_Sao_Francisco_Brazil_LBP49190

Species 12 (support = 1.000)
      P_calvertI_Parnaiba_Brazil_LBP27355

Species 13 (support = 0.877)
      R_carti_Mandinga_Panama_STRI321,R_carti_Playon_Chico_Panama_STRI4950

Species 14 (support = 0.518)
      R_occidentalis_Cocle_del_Sur_Panama_STRI17718,R_occidentalis_Santa_Maria_Panama_STRI3404

Species 15 (support = 0.334)
      R_lisaae_Salama_Nuevo_Costa_Rica_STRI2820,R_lisaae_Salama_Nuevo_Costa_Rica_STRI2818,R_bouchellei_Higueron_Costa_Rica_STRI2116,R_bouchellei_Canas_Costa_Rica_STRI2173,R_bouchellei_Grande_de_Matagalpa_Nicaragua_STRI14271,R_bouchellei_Pizote_Costa_Rica_STRI2150

Species 16 (support = 0.483)
      R_dayi_Magdalena_Colombia_LBP31161,R_dayi_Magdalena_Colombia_LBP30742

Species 17 (support = 0.199)
      R_guatemalensis_Chagres_Panama_STRI15205,R_guatemalensis_Llano_Sucio_Panama_LBP18532,R_guatemalensis_Cocoli_Panama_STRI9141,R_guatemalensis_Llano_Sucio_Panama_LBP18531

Species 18 (support = 0.498)
      R_loftini_Tambo_Panama_AM18,R_loftini_Cascajal_Panama_STRI15327
```


Supplementary Fig. 4. Bayesian Poisson Tree Processes (bPTP) delimitation of species of *Roeboides.*

So Many Fish, So Little Time... (*Lundberg, John G. et al.*, 2000)