UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"
CÂMPUS DE BOTUCATU
INSTITUTO DE BIOCIÊNCIAS
Programa de Pós-graduação em Ciências Biológicas (Zoologia)

NADAYCA THAYANE BONANI MATEUSSI

FILOGENÔMICA DE SERRASALMIDAE E ESTUDOS MORFOLÓGICOS E MOLECULARES EM *CATOPRION* E *PYGOCENTRUS* (TELEOSTEI: CHARACIFORMES)

Botucatu/SP
2019

NADAYCA THAYANE BONANI MATEUSSI

FILOGENÔMICA DE SERRASALMIDAE E ESTUDOS MORFOLÓGICOS E MOLECULARES EM *CATOPRION* E *PYGOCENTRUS* (TELEOSTEI: CHARACIFORMES)

Tese apresentada ao Programa de Pós-Graduação em Ciências Biológicas (Zoologia) do Instituto de Biociências da Universidade Estadual Paulista "Júlio Mesquita Filho" como requisito parcial para obtenção do título de Doutora em Ciências Biológicas (Zoologia)**.**
Área de concentração: Biodiversidade

Orientador: Dr. Claudio de Oliveira

Co-orientador: Dr. Bruno Francelino de Melo

Botucatu/SP

2019

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Mateussi, Nadayca Thayane Bonani.
Filogenômica de Serrasalmidae e estudos morfológicos e moleculares em *Catoprion* e *Pygocentrus* (Teleostei: Characiformes) / Nadayca Thayane Bonani Mateussi. - Botucatu, 2019

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu
Orientador: Claudio de Oliveira
Coorientador: Bruno Francelino de Melo
Capes: 20405006

1. Código de Barras de DNA taxonômico. 2. Zoologia - Classificação. 3. Peixes - Identificação. 4. Filogenia.

Palavras-chave: DNA barcode; Elementos Ultraconservados; Sistemática; Taxonomia.

A todos os professores que tive ao longo desse extenso caminho de apredizagem, da alfabetização à pós-graduação. A dedicação de cada um me trouxe até aqui. É tudo culpa deles.

“Everything in evolution makes better sense in the light of phylogenetics.”

(to paraphrase Dobzhansky 1973;

Sytsma & Pires 2001)

AGRADECIMENTOS

São muitas as pessoas e instituições que contribuíram, direta ou indiretamente, para a realização deste trabalho. Sou imensamente grata a cada uma delas, em especial:

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de Doutorado Sanduíche que me permitiu fazer parte do doutorado no exterior.

À Universidade Estadual Paulista, ao Instituto de Biociências de Botucatu, ao Programa de Pós-graduação em Ciências Biológicas (Zoologia) e ao Departamento de Morfologia pela oportunidade e apoio logístico.

Ao Dr. Claudio de Oliveira, pela orientação em cada fase do projeto e apoio desde o mestrado, pela confiança e oportunidade.

Ao Dr. Bruno F. Melo pela coorientação, conselhos e amizade, essenciais para a minha formação.

Ao Dr. Patrice Pruvost por me receber no Museum National d'Histoire Naturelle durante o Doutorado Sanduíche e aos técnicos Zoha e Hugo por toda ajuda prestada. *Je vous remercie*.

Ao Dr. Fausto Foresti, pelos ensinamentos e exemplo.

Aos colaborados e amigos Fábio Roxo e Luz Eneida Ochoa, pela disposição em ensinar, discutir e contribuir para o desenvolvimento da tese.

À Dra. Rafaela Ota, por ter me apresentado aos Serrasalmidae ainda no mestrado, pela amizade e pela disposição em colaborar comigo desde meu primeiro dia de iniciação científica (ainda ontem) até a finalização da presente tese.

Pelas visitas às instituições, empréstimo/doação de material, fotografias e/ou radiografias de exemplares, auxílio nas identificações, agradeço aos curadores, técnicos, alunos e demais pesquisadores e instituições: Mark Sabaj (ANSP), Francisco Villa (CZUT-IC), Juan Albornoz e Carlos DoNascimiento (IAvH), Lúcia Rapp Py-Daniel e Shizuka Hashimoto (INPA), Flávio Bockmann (LIRP), Carlos Alberto Lucena (MCP-PUCRS), Patrice Pruvost (MNHN), Paulo Buckup (MNRJ), Alessio Datovo e Michel Gianeti (MZUSP), Anja Palandacick (NMW), Carla S. Pavanelli (NUP), Fernando Carvalho (UFMS), Marcelo Andrade (UFPA), Carolina Dória e Dayana Brito (UNIR).

A todos do Laboratório de Biologia e Genética de Peixes, pelo companheirismo ao longo do doutorado, por toda ajuda, pelas conversas (produtivas ou não) e pelos momentos de descontração. Em especial a Aisni Mayumi, Beatriz Boza, Bruno Campos, Bruno Melo, Camila Souza, Cristhian Conde, Cristiane Shimabukuru Dias, Duílio Zerbinato, Fabilene Gomes, Fábio Roxo, Fernando Yuldi, Gabriel Costa e Silva, Giovana Ribeiro, Guilherme Costa Silva, Luz E. Ochoa, Jason J. Fernandez, Maria Lígia de Oliveira, Silvana Melo e Vanessa Paes.

A Camila Souza e Luz Ochoa pela paciência nos meus momentos de estresse com a tese, por toda a ajuda em solucionar os problemas de laboratório e pelo companheirismo. Se eu não perdi meu cabelo durante o doutorado foi graças a elas.

Aos meus amigos-família PDSE, Ana Paula Chorobura, Bruno C. Ribeiro, Diego Reis, Larissa Rézio, Luiz Arthur Faria, Juliana Mourão, Paula Bianchi, Rafael e Ruben Lemke, por transformarem uma grande experiência profissional em uma inesquecível experiência de vida.

A Alexandre Arvigo, Alexandre Ribeiro, Aline Acosta, Bárbara Martins, Bruno Tayar, Eduardo dal Farra, Jairon Moura, Isabel Soares, Juliana De Biasi, Marciano Venâncio, Maria Isabel Muller e Mariana Antunes, por serem a minha família em Botucatu. Em especial, à Isabela Moraes por ser a melhor *housemate* que alguém poderia esperar; do hospital ao carnaval.

A todos os meus amigos, por todo carinho, companheirismo, apoio e sorrisos que sempre garantiram que eu estivesse bem amparada nesse caminho.

E, finalmente, à toda minha família, pelo carinho e suporte. Principalmente à minha base, meus pais Emerson Mateussi e Alessandra Bonani Mateussi e irmão Victor Hugo Bonani Mateussi, pelo apoio e amor incondicionais. E Isabela Cracco, por ser minha irmã de coração e estar sempre ao meu lado.

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 literature. 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.

Filogenômica de Serrasalmidae e estudos morfológicos e moleculares em *Catoprion* e *Pygocentrus* (Teleostei: Characiformes)

Resumo

A família Serrasalmidae é endêmica da região Neotropical, amplamente distribuída na América do Sul, e possui 97 espécies válidas distribuídas em 16 gêneros recentes e um fóssil. Embora a monofilia de Serrasalmidae seja bem estabelecida, assim como dos três clados que a compõe, várias relações entre os subgrupos são ainda conflitantes, há confusão sobre as posições dos gêneros dentro destes clados e restam dúvidas sobre a monofilia, composição e validade dos gêneros e espécies. Ainda, a imprecisão da descrição de muitas das espécies e seus atuais status taxonômicos dificultam o entendimento e o estudo para diversas áreas. Sendo assim, o presente trabalho visou elucidar as relações filogenéticas da família através de análises filogenômicas utilizando elementos ultraconservados (UCEs) e delimitar espécies de *Catoprion* e *Pygocentrus* através de análises integrativas, morfológicas e moleculares. Os resultados corroboram a monofilia da família, bem como dos três clados principais: (1) *Colossoma, Mylossoma* e *Piaractus*, (2) o "clado *Myleus*" formado por *Acnodon, Myleus, Mylesinus, Myloplus, Ossubtus, Tometes* e *Utiaritichthys* e (3) o "clado piranha", composto por *Catoprion, Metynnis*, *Pristobrycon, Pygocentrus*, *Pygopristis* e *Serrasalmus.* As relações dentro do clado "*Colossoma+Mylossoma+Piaractus*" são bem suportadas, com *Piaractus* irmão do clado *Colossoma* e *Mylossoma*, sendo este clado grupo irmão de todos os outros serrasalmídeos. O "clado *Myleus*" apresenta gêneros não-monofiléticos, como *Myleus, Mylesinus, Myloplus, Tometes* e *Utiaritichthys*. E o "clado piranha" é recuperado como um grupo monofilético com *Metynnis* irmão das demais piranhas, os gêneros *Catoprion, Pygopristis* e *Pygocentrus* são monofiléticos enquanto *Serrasalmus* e *Pristobrycon* são parafiléticos. Análises moleculares de delimitação de espécies reconheceram três espécies previamente identificadas morfologicamente em *Pygocentrus*; além de corroborar a existência de ao menos quatro populações estruturadas em *P. nattereri* ao longo de sua distribuição continental. E por fim, *Catoprion mento* foi redescrita e teve sua distribuição restrita às bacias dos rios Orinoco, alto Paraguai e tributários da margem direita do Amazonas; enquanto uma nova espécie foi descrita para as demais sub-bacias Amazônicas e o rio Essequibo. As espécies de *Catoprion* podem ser diagnosticadas por contagens de escamas e apresentam 7.3% ± 0.02 de divergência genética (K2P).

Phylogenomics of Serrasalmidae and morphological and molecular studies in *Catoprion* and *Pygocentrus* (Teleostei: Characiformes)

Abstract

Serrasalmidae is an endemic Neotropical fish family, broadly distributed in South America, with 97 valid species allocated in 16 extant genera and one fossil. Even though the monophyly of the family is well stablished, as the formation of three major clades composing it, several aspects of relationships between subgroups are discordant among authors and are observed conflicts in the interrelationships and monophyly of some genera, and in the validity of some species. Furthermore, the imprecision of the species description and its current taxonomic status make it difficult the understanding and study for several areas. Thus, the present study aimed to elucidate the phylogenetic relationships within the family through a phylogenomic analysis using ultraconserved elements (UCEs) and to delimit species of *Catoprion* and *Pygocentrus* through morphological and molecular integrative analysis. The results corroborate the monophyly of family as well as the three main clades: (1) *Colossoma*, *Mylossoma* and *Piaractus*, (2) the "*Myleus* clade" composed by *Acnodon, Myleus, Mylesinus, Myloplus, Ossubtus, Tometes* and *Utiaritichthys*, and (3) the "piranha clade", composed by *Catoprion, Metynnis, Pristobrycon, Pygocentrus, Pygopristis* and *Serrasalmus.* The relationships within the "*Colossoma+Mylossoma+Piaractus*" clade are well supported, with *Colossoma* and *Mylossoma* as sister group of *Piaractus*, and this clade sister to all other Serrasalmids. The "*Myleus* clade" contains several non-monophyletic genera as *Myleus, Mylesinus, Myloplus, Tometes* and *Utiaritichthys.* And the "piranha clade" is recovered as a monophyletic group with *Metynnis* as sister to the other genera, with *Catoprion*, *Pygopristis* and *Pygocentrus* monophyletic and *Serrasalmus* and *Pristobrycon* as paraphyletic. Molecular analyzes of species delimitation recognized three species previously identified morphologically in *Pygocentrus*; besides corroborating the existence of at least four populations structured in *P. nattereri* throughout its continental distribution. And finally, *Catoprion mento* was redescribed and had its distribution restricted to the Orinoco, Upper Paraguay, and right bank tributary of the Amazon basin; while a new species was described for the other Amazon sub-basins and the Essequibo River. *Catoprion* species can be diagnosed by counting of scales and present 7.3% ± 0.02 genetic divergence (K2P).

SUMÁRIO

Introdução geral 11
A família Serrasalmidae 12
Filogenômica usando elementos ultraconservados 17
DNA barcoding 18
Justificativa 20
Objetivos 21
Bibliografia 22

Capítulo 1 – Phylogenomics of the Neotropical fish family Serrasalmidae (Ostariophysi: Characiformes) based on ultraconserved elements 28
1.1. Introduction 29
1.2. Material and methods 34
1.3. Results 40
1.4. Discussion 45
1.5. Bibliography 49

Capítulo 2 – Molecular species delimitation in the Piranha genus *Pygocentrus* (Characiformes: Serrasalmidae) 65
2.1. Introduction 67
2.2. Material and methods 70
2.3. Results 75
2.4. Discussion 78
2.5. Bibliography 81

Capítulo 3 – Species delimitation and taxonomic revision of *Catoprion* Müller & Troschel, 1844 (Characiformes: Serrasalmidae) with the description of a new species .. 85
3.1. Introduction 87
3.2. Material and methods 88
3.3. Results 92
3.4. Discussion 111
3.5. Bibliography 113

Introdução Geral

Peixes constituem o grupo mais diverso e numeroso entre os vertebrados, com cerca de 35.000 espécies válidas e muitas ainda a serem descritas, com 417 novas espécies descritas somente em 2018 (Fricke *et al.* 2019). Mais de 40% desse grupo habita ambientes de água doce (Lowe-McConnell 1987), sendo que, nesse ambiente, a maior diversidade de espécies se encontra na região Neotropical (Vari & Malabarba 1998; Lundberg *et al.* 2000; Lévêque *et al.* 2008), com estimativas superiores a 8.000 espécies (Schaefer 1998; Reis *et al.* 2003, 2016). Nesse cenário, a maioria dos peixes de água doce Neotropicais estão distribuídos principalmente em cinco grupos dominantes: Characiformes (lambaris, piabas, piranhas, matrinchãs, curimbas, piaus, traíras, pacus, dourados, etc), Siluriformes (bagres e cascudos), Cichliformes (acarás, tucunarés, jacundás), Cyprinodontiformes (guarus e peixes anuais) e Gymnotiformes (tuviras, ituís, poraquês) (Rosa & Lima 2008).

Destes, Characiformes é a ordem de maior riqueza de espécies e inclui cerca de 2.200 espécies válidas distribuídas em 24 famílias (Fricke *et al.* 2019), compreende espécies herbívoras, iliófagas, lepidófagas, onívoras e carnívoras, variando desde alguns milímetros (e.g., *Priocharax*) até mais de um metro de comprimento (e.g., *Colossoma*). Os Characiformes podem ser diferenciados externamente dos demais grupos de peixes de água-doce Neotropicais por possuírem o corpo coberto de escamas, apresentarem nadadeiras pélvicas em posição ventral, situadas, geralmente, bem atrás da inserção das nadadeiras peitorais, raios das nadadeiras moles (não transformados em espinhos pungentes) e, geralmente, a presença de uma nadadeira adiposa (Britski *et al.* 2007). Wiley & Johnson (2010) destacam sete sinapomorfias para o grupo, incluindo presença de foramen auditório proótico (*sensu* Weitzman 1962), presença de abertura dorsomedial na fossa pós-temporal, dentes de reposição para série externa do dentário ou alguns dentes do pré-maxilar formados em trincheiras ou criptas e dentes multicuspidados nas mandíbulas.

A família Serrasalmidae

Serrasalmidae, composta pelos popularmente conhecidos como pacus, piranhas e tambaqui, é uma das famílias de Characiformes, endêmica da região Neotropical, amplamente distribuída nos principais sistemas de rios da América do Sul, com ocorrência em todos os biótopos de água doce, exceto riachos muito estreitos e área bentônica de rios profundos (Jégu

2003), com registros de introdução intencional ou acidental em outros continentes (Hensel 2004; Caleta *et al.* 2011). Os serrasalmídeos podem ser identificados por um conjunto de caracteres, como corpo alto e comprimido, presença de uma serra ventral com espinhos oriundos da modificação de escamas e um espinho pré-dorsal, exceto em *Colossoma*, *Mylossoma* e *Piaractus*. Serrasalmidae agrega 97 espécies válidas distribuídas em 16 gêneros (Fricke *et al.* 2019), mais um gênero monotípico fóssil do Mioceno Superior, a †*Megapiranha paranensis* (Cione *et al.* 2009) (Tabela 1).

Tabela 1. Gêneros válidos e números de espécies válidas em Serrasalmidae (*sensu* Fricke *et al.* 2019).

Gênero	Espécies válidas
Acnodon Eigenmann, 1903	3
Catoprion Müller & Troschel, 1844	1
Colossoma Eigenmann & Kennedy, 1903	1
†*Megapiranha* Cione, Dahdul, Lundberg & Machado-Allison, 2009	1
Metynnis Cope, 1878	15
Mylesinus Valenciennes, 1850	3
Myleus Müller & Troschel, 1844	7
Myloplus Gill, 1896	12
Mylossoma Eigenmann & Kennedy, 1903	5
Ossubtus Jégu, 1992	1
Piaractus Eigenmann, 1903	2
Pristobrycon Eigenmann, 1915	5
Pygocentrus Müller & Troschel, 1844	3
Pygopristis Müller & Troschel, 1844	1
Serrasalmus Lacepède, 1803	27
Tometes Valenciennes, 1850	8
Utiaritichthys Miranda Ribeiro, 1937	3

Os Serrasalmidae apresentam grande importância econômica na América do Sul, tanto para a pesca de subsistência de populações ribeirinhas, principalmente na Amazônia, quanto para a pesca comercial (Ruffino 2004; Freeman *et al.* 2007; Ota *et al.* 2013) e aquicultura (Castagnolli 2000). Ademais, também apresentam valor na ornamentação, sendo apreciados

no aquarismo (e.g. mundopiranha.com), além de serem vendidas taxidermizadas como *souvenirs* (Freeman *et al.* 2007).

Embora o notório interesse popular despertado pela família devido a histórias e lendas propagadas em livros, filmes e músicas, ataques registrados a banhistas são ocasionais sendo raros aqueles causando danos mais sérios (Mol 2006). Assim, os reais danos causados por piranhas são maiores de fato à pesca, uma vez que atacam os peixes presos às redes, eliminando ou reduzindo seu valor comercial além dos danos causados ao próprio material de pesca (Agostinho *et al.* 1997). Ainda assim, apesar da importância e curiosidade voltados à família, o entendimento da sistemática do grupo permanece incipiente, restando muitas dúvidas sobre as atribuições de espécies e relações entre elas.

O *status* do grupo como Serrasalminae, subfamília de Characidae, ou Serrasalmidae, uma família separada, foi motivo de controvérsia entre pesquisadores ao longo dos anos; ainda que a monofilia seja consensual e bem estabelecida entre todos os autores (Machado-Allison 1983; Jégu 2004; Ortí *et al.* 1996, 2008; Thompson *et al.* 2014). Mais recentemente, estudos sugerem que os serrasalmídeos sejam mais proximamente relacionados a Anostomoidea do que a outras subfamílias de Characidae, de forma que atualmente o reconhecimento de Serrasalmidae como família é consensual e suportada por ambas hipóteses morfológicas e moleculares (Calcagnotto *et al.* 2005; Mirande 2010; Oliveira *et al.* 2011; Arcila *et al.* 2017).

A primeira espécie de serrasalmídeo a ter uma descrição formal foi *Salmo rhombeus* Linnaeus, 1766 (válida como *Serrasalmus rhombeus*), sendo a maioria das demais espécies descritas no século XIX e XX (Machado-Allison 2002; Jégu 2004). Pela falta de entendimento acerca da complexidade da ictiofauna Neotropical na época, as descrições frequentemente eram baseadas em poucos exemplares, sem atribuição precisa de localidade, sem exemplares tipo ou apenas com o tipo, entre outros problemas que contribuem enormemente para a confusão taxonômica da família (Machado-Allison 2002).

A primeira tentativa de classificação deste grupo foi realizada por Eigenmann (1915), que a dividiu em duas subfamílias dentro de Characidae: a subfamília Serrasalminae, composta por seis gêneros apresentando uma única série de dentes em ambas as maxilas; e Myleinae, composta por nove gêneros apresentando duas séries de dentes no pré-maxilar e frequentemente um par de dentes posterior à série labial da mandíbula. Subsequentemente, Norman (1929) uniu os dois grupos em uma única subfamília (Serrasalmoninae), baseado em caracteres morfológicos como abdômen quilhado e serrilhado, corpo alto e comprimido lateralmente, nadadeira dorsal longa e escamas pequenas. Gosline (1951) também considera

uma uma única subfamília, Serrasalminae, dentro de Characidae, reconhecendo *Mylossoma* e *Colossoma* (com *Piaractus*) formando um grupo separado dos demais pacus. Posteriormente, Géry (1972), com base nas séries de dentes, escamas, comprimento das nadadeira dorsal, espinho pré-dorsal e espinhos ventrais, elevou o grupo a nível de família e a subdividiu em três subfamílias: Catoprioninae, Myleinae e Serrasalminae, diferindo na classificação de Eigenmann apenas por separar *Catoprion* em uma nova sub-família baseada na dentição singular do gênero. Géry (1972) também agrupou todas as piranhas em um único gênero, *Serrasalmus*, com cinco subgêneros.

Os agrupamentos dentro de Serrasalmidae foram propostos principalmente baseados na dentição (Eigenmann 1915). Partindo dessa observação, Gosline (1951) hipotetizou que os dentes de Myleinae (*i.e.* duas séries de dentes molariformes, sendo a externa com cinco e a interna com dois dentes) seriam derivados dos dentes de Tetragonopterinae; enquanto os dentes de Serrasalminae seriam derivados da condição de Myleinae, com a junção das duas fileiras e perda de um dente, formando assim a única fileira de seis dentes do pré-maxilar das piranhas, ao passo que a condição de *Catoprion* seria "aberrante". Apesar de considerar a própria hipótese como lógica, o autor não argumenta contra ou a favor da mesma devido à inexistência de um caráter intermediário, posteriormente encontrado na descrição de †*Megapiranha paranensis* (Cione *et al.* 2009). O fóssil do Mioceno superior apresenta um arranjo truncado na disposição dos dentes do pré-maxilar ("zig-zag"), sendo considerado um padrão intermediário entre a condição de duas fileiras de dentes (pacus) e uma fileira de dentes encontrada (piranhas), sugerindo que o rearranjo dos dentes de um ancestral com duas fileiras de dentes poderia ter originado a linhagem com uma fileira. Porém, diferente da hipótese de perda de um dente durante a junção de Gosline (1951), Cione *et al.* (2009) argumentam a favor de uma fusão entre o quarto e quinto dentes da série externa dos Myleinae para a formação do sexto dente observado nas piranhas modernas.

Machado-Allison (1982) realizou o primeiro estudo cladístico de Serrasalmidae baseado em caracteres osteológicos, miológicos e de padrão de coloração. O autor considerou novamente o grupo como subfamília de Characidae e obteve uma filogenia que dividia o grupo em duas linhagens, a das piranhas, consistindo em *Catoprion*, *Metynnis*, *Pristobrycon*, *Pygocentrus*, *Pygopristis* e *Serrasalmus* e a dos pacus com *Acnodon*, *Colossoma*, *Mylesinus*, *Myleus*, *Mylossoma*, *Piaractus* e *Utiaritichthys*.

No primeiro estudo baseado em caracteres moleculares, Ortí *et al.* (1996) utilizaram sequências de genes ribossomais mitocondriais 12S e 16S para obter uma árvore com três clados, o primeiro contendo *Colossoma*, *Mylossoma* e *Piaractus*, o segundo sendo o das

piranhas, com indicação de que *Serrasalmus* e *Pristobrycon* são polifiléticos, e um terceiro clado, chamado clado "*Myleus*" para incluir *Mylesinus, Myleus* e *Myloplus*. Nessa hipótese, *Acnodon* é recuperado como irmão dos clados *Myleus* e das piranhas, e o clado *Colossoma+Mylossoma+Piaractus* é o clado menos derivados irmão dos demais. Os três clados propostos por Ortí *et al.* (1996) foram recuperados em vários trabalhos posteriores, tanto morfológicos quanto moleculares, porém com diferenças significativas nas relações internas e com diferentes suportes estatísticos para a monofilia dos gêneros (Jégu 2004; Dahdul 2007; Freeman *et al.* 2007; Ortí *et al.* 2008; Thompson *et al.* 2014). Entre esses trabalhos, as maiores dúvidas são em relação à monofilia de alguns gêneros, como *Myleus, Pristobrycon* e *Serrasalmus*; enquanto os maiores problemas metodológicos se referem à baixa representação taxonômica e baixo suporte para alguns grupos, de forma que muitas das designações genéricas dos clados carecem de maior suporte ou são claramente contraditórias com os dados obtidos.

A filogenia mais recente da família (Thompson *et al.* 2014), realizada com abordagem multilocus (10 genes nucleares e a região controle do DNA mitocondrial), também recupera a relação do clado *Colossoma/Mylossoma/Piaractus* irmão do grande grupo com clado *Myleus*+clado piranha. No entanto, a baixa representatividade de espécies utilizadas (38 de um total de 97) impede uma maior compreensão das relações entre espécies de determinados gêneros. Além de uma reduzida matriz de dados moleculares comparada com as filogenias obtidas com dados filogenômicos, este trabalho ainda deixa várias lacunas sobre as relações entre gêneros e entre espécies, principalmente nos gêneros *Pristobrycon*, *Serrasalmus*, *Mylesinus, Myleus* e *Tometes*. Além disso, diferentes análises recuperam diferentes topologias, inviabilizando uma maior compreensão da posição de diversos gêneros, principalmente no clado *Myleus*.

Thompson *et al.* (2014) ainda sugerem que o clado das piranhas é muito mais recente (Mioceno, 15-20 Ma) que o restante da família, tendo começado a se diversificar durante o Cretáceo (65-75 Ma). Embora haja inconsistência quanto ao tempo estimado para a diversificação da família, todos os trabalhos com relógio molecular recuperam o tempo de origem da diversificação do clado piranhas como relativamente recente: 9 Ma (Hubert *et al.* 2007), 13,5 Ma (Cione *et al.* 2009) e 10 Ma (Melo *et al.* in prep.). Para Thompson *et al.* (2014), essa diversificação recente pode ser a causa da grande confusão entre os táxons, uma vez que os mesmos podem apresentar uma maior proximidade filogenética. Um exemplo dessa diversificação pode ser observado na única espécie fóssil de Serrasalmidae, do Mioceno

superior, †*Megapiranha paranensis*, que apresenta caracteres intermediários entre pacus e piranhas (Cione *et al.* 2009).

Baseado em diversos estudos prévios sobre Serrasalmidae, Freeman *et al.* (2007) atribuíram a inconsistência nos estudos taxonômicos de Serrasalmidae à (1) quantidade limitada de material comparativo, (2) ausência de caracteres externos confiáveis para a distinção das espécies, (3) ampla variação intraespecífica na morfologia e padrão de colorido, geralmente relacionados a mudanças ontogenéticas, estágio reprodutivo e influências ambientais, (4) sobreposição nos padrões de colorido e caracteres morfológicos entre diferentes espécies, (5) distribuição geográfica pouco conhecida, (6) espécies pobremente descritas, com material tipo em más condições, perdido ou inexistente, (7) provável existência de complexos de espécies, (8) incerteza sobre a alocação de algumas espécies em *Serrasalmus* e *Pristobrycon* e (10) existência de muitos sinônimos nunca avaliados detalhadamente. Como citado acima, há várias inconsistências taxonômicas na família, mais evidente em alguns gêneros como *Myleus* e *Serrasalmus*, em que a identificação das espécies é bastante incerta e confusa, dificultando qualquer trabalho filogenético, fisiológico ou ecológico que se almeje realizar.

Filogenômica usando elementos ultraconservados (UCEs)

Os elementos ultraconservados (*ultraconserved elements*, UCEs), como seu nome diz, são regiões do genoma extremamente conservadas, e assim compartilhadas entre grupos pertencentes a linhagens muito distintas filogeneticamente. Os UCEs foram primeiramente descritos por Bejerano *et al.* (2004), que encontraram cerca de 480 segmentos maiores de 200 pb que eram absolutamente (100%) conservados em regiões ortólogas de humanos e camundongos e altamente conservados nos genomas de galinhas e cachorros (95 e 99%, respectivamente). Os UCEs foram mais frequentemente localizados em regiões próximas a éxons envolvidos no processamento de RNA, íntrons, ou ainda próximos de genes envolvidos na regulação da transcrição e desenvolvimento (Bejerano *et al.* 2004). Ainda no mesmo estudo, Bejerano *et al.* (2004) relatam que além desses elementos, mais de 5000 sequências de cerca de 100 pb são também absolutamente conservadas entre humanos, ratos e camundongos, portanto, muito mais conservadas que as codificantes de proteínas. Estudos posteriores mostraram que esses UCEs também se encontram em outros vertebrados, insetos,

vermes e fungos (Siepel *et al.* 2005; Faircloth *et al.* 2012).

Embora seu papel no genoma ainda não seja totalmente esclarecido (Dermitzakis *et al.* 2005), os UCEs têm sido associados à regulação gênica (Pennacchio *et al.* 2006) ou no desenvolvimento (Sandelin *et al.* 2004; Woolfe *et al.* 2004) e tem-se assumido que são importantes pela natureza extremamente conservada ao longo da árvore da vida. Entretanto, os silenciamentos gênicos de locus UCEs em ratos resultaram em uma prole viável e fértil (Ahituv *et al.* 2007), sugerindo que seu papel no desenvolvimento dos organismos ainda necessita ser melhor estudado.

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 nucleares, utilizando sequenciamento de nova geração. Como as sequências UCEs são altamente conservadas, elas são utilizadas para o anelamento de sondas (*primers* ou *probes*), a partir das quais as sequências flanqueadoras dos UCEs são lidas. Desde então, os UCEs têm sido utilizados em muitos níveis de comparação entre organismos, de populações até grandes grupos como ordens. Assim, por exemplo, McCormack *et al.* (2013) analisaram mais de 1500 locus de UCEs de 32 espécies de Neoaves e encontraram novas relações entre alguns membros, como a relação próxima entre passarinhos e papagaios. McCormack *et al.* (2012) estudaram 183–917 loci em mamíferos placentários e encontraram uma boa resolução nas árvores obtidas, com resultados importantes para uma melhor compreensão da história evolutiva do grupo. Crawford *et al.* (2012) analisaram a posição das tartarugas com o uso de 1145 locus de UCEs e os resultados suportaram a hipótese de que as tartarugas evoluíram de um ancestral comum compartilhado por Archosauria (crocodilianos e aves), rejeitando assim a hipótese de relação entre as tartarugas e os Lepidosauria (lagartos, serpentes e tuataras).

O primeiro estudo utilizando UCEs para o entendimento de peixes, realizado por Faircloth *et al.* (2013), contou com o sequenciamento de 500 locus de cerca de 30 espécies de Actinopterygii, obtendo filogenias com nós totalmente 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 dentro de Teleostei. Burress *et al.* (2018) utilizaram 5000 loci de UCEs para avaliar a diversificação de ciclídeos Neotropicais na bacia do rio Uruguai, encontrando um suporte robusto para a monofilia do grupo e aliando os resultados filogenômicos a análises de diversificação ecológica. Alfaro *et al.* (2018) utilizaram mais de 1000 loci no estudo de Acanthomorpha elucidando a relação de mais de 120 linhagens dentro do grupo, além de

apresentar novas hipóteses acerca do seu tempo de diversificação. Dessa maneira, é evidente que o uso de marcadores UCEs no estudo de grupos grandes e complexos, como os peixes de água doce neotropicais, tem se mostrado muito promissor.

DNA Barcoding

Técnicas moleculares têm se provado eficientes em estudos de biodiversidade, principalmente naqueles em que as ferramentas tradicionais são insuficientes ou incapazes de identificar espécies (Pereira *et al.* 2011). Assim, as sequências de DNA podem ser usadas como ferramenta auxiliar à taxonomia e delimitação de espécies, tanto corroborando evidências sobre hipóteses existentes ou como ponto de partida para novos testes (DeSalle *et al.* 2005).

Uma metodologia que tem se mostrado bastante eficiente na delimitação de grupos taxonômicos é o DNA *barcoding* utilizando sequências parciais do gene codificante mitocondrial *citocromo oxidase c subunidade I* (*COI*), proposto por Hebert *et al.* (2003) como um método de identificação de espécies animais. Esta ferramenta fornece evidências sobre as Unidades Taxonômicas Operacionais (Operational Taxonomy Units - OTUs) através do reconhecimento de linhagens genéticas evidenciando novos grupos taxonômicos potenciais (Baldwin & Weigt 2012), suportando os estudos taxonômicos tradicionais e descrição de novas espécies, incluindo peixes de água-doce neotropicais (e.g., Benine *et al.* 2009; Melo *et al.* 2011; Silva *et al.* 2013; Costa-Silva *et al.* 2015).

A grande variação intraespecífica nas espécies de Serrasalmidae, principalmente no que diz respeito à alometria e padrões de coloração durante seu desenvolvimento ou estágio reprodutivo, resulta em uma grande dificuldade no reconhecimento dos limites entre as espécies. Assim, nos últimos anos, trabalhos envolvendo abordagens integrativas têm contribuído de forma significativa para a compreensão da diversidade de espécies na família. Mateussi *et al.* (2016) utilizaram o DNA *barcoding* na identificação de espécies cis-andinas de *Mylossoma*, sugerindo a subdivisão de *M. duriventre* (Cuvier, 1818) em quatro espécies e embasando a subsequente revisão taxonômica do gênero (Mateussi *et al.* 2018) e o reconhecimento de mais duas espécies. Andrade *et al.* (2017) utilizaram uma abordagem integrativa de dados morfológicos e moleculares para a descrição de uma nova espécie de *Tometes* contribuindo para a delimitação das espécies que compõem o gênero. Finalmente, Machado *et al.* (2018) analisaram mais de mil exemplares representando 69 espécies de

Serrasalmidae de grandes bacias sulamericanas e reconheceram uma diversidade altamente subestimada para a família, identificando diversas linhagens não reconhecidas anteriormente. Assim, estes trabalhos confirmam a eficiência do método para o entendimento da família ao reafirmar que a utilização de abordagens baseadas no gene COI auxilia na identificação das espécies e permite uma maior discussão sobre os limites das espécies.

Nesse contexto, dois gêneros pertencentes ao clado das piranhas, *Catoprion* e *Pygocentrus*, carecem de estudos integrativos que suportem o reconhecimento e delimitação de suas espécies. *Catoprion* tem sido considerada monotípica desde sua descrição (Fricke *et al.* 2019) e, embora tenha sido hipotetizada anteriormente como um complexo de espécies (Taphorn 1992, 2003), carece de estudos conduzidos com exemplares representantes de toda sua área de ocorrência. *Pygocentrus* consiste em um gênero de piranha de grande interesse no aquarismo e interesse relativo na pesca regional, de importância principalmente nas comunidades ribeirinhas (Santos *et al.* 2006), com três espécies válidas atualmente (Fricke *et al.* 2019). Estudos para o gênero, apesar de representativos, focam apenas em populações de determinadas bacias e principalmente em uma única espécie, *P. nattereri* Kner, 1858 (Luz *et al.* 2015; Santos *et al.* 2016), de forma que a lacuna no conhecimento acerca do gênero é bastante evidente. Sendo assim, o DNA *barcoding* será utilizado no estudo de ambos os gêneros a fim de se elucidar as unidades taxonômicas que os compõem.

Justificativa

Várias análises filogenéticas morfológicas e moleculares foram realizadas para estabelecer relações intergenéricas e interespecíficas em Serrasalmidae (Machado-Allison 1982; Ortí *et al.* 1996, 2008; Jégu 2004; Freeman *et al.* 2007; Thompson *et al.* 2014). Apesar da maioria destes estudos indicarem a monofilia dos clados que compõem a família, há controvérsias sobre a monofilia dos gêneros bem como sobre as relações entre eles. Além disso, observa-se, nos estudos já desenvolvidos, uma baixa representatividade de táxons além de um reduzido número de caracteres moleculares.

Sendo assim, embora muito se tenha avançado no entendimento de Serrasalmidae no último século, a incongruência entre os resultados apresentados evidencia a necessidade de estudos adicionais, utilizando maior representatividade de táxons e de marcadores moleculares. Além disso, parte da inconsistência nos estudos taxonômicos de Serrasalmidae se deve à ausência de caracteres morfológicos para a distinção das espécies, ampla variação

e/ou sobreposição nos padrões de colorido e caracteres morfológicos e existência de sinônimos e complexos de espécies nunca avaliados detalhadamente (Freeman *et al.* 2007). Ainda, é recorrente a perpetuação do reconhecimento de gêneros popularmente conhecidos e com poucos representantes, mas com ampla distribuição geográfica, como sendo monotípicas sem uma avaliação sobre sua variação ou assumindo-se que esta variação é muito maior do que de fato é. Portanto, a abordagem de gêneros, individualmente, almejando a resolução de problemas taxonômicos pontuais dentro de Serrasalmidae, é importante para melhor delimitar as espécies e auxiliar em estudos filogenéticos e taxonômicos posteriores.

Objetivos

Objetivo geral:

O principal objetivo deste trabalho foi inferir hipóteses de relacionamento entre os táxons constituintes de Serrasalmidae usando caracteres moleculares.

Objetivos específicos:

1) Testar a monofilia e relações entre os gêneros de Serrasalmidae com ampla representatividade de táxons e caracteres genômicos;
2) Delimitar espécies de *Pygocentrus* e detectar a existência de populações estruturadas e linhagens genéticas adicionais;
3) Realizar o estudo integrativo de dados morfológicos e moleculares e realizar uma revisão taxonômica de *Catoprion*.

Bibliografia:

Agostinho C.S., Agostinho A.A., Marques E.E. & Bini L.M. (1997) Abiotic factors influencing Piranha attacks on netted fish in the upper Paraná river, Brazil. *North American Journal of Fisheries Management* 17, 712–718.

Ahituv N., Zhu Y., Visel A., Holt A., Afzal V., Pennacchio L.A., *et al.* (2007) Deletion of Ultraconserved Elements Yields Viable Mice. *PLoS Biology* 5, e234.

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

Andrade M.C., Machado V.N., Jégu M., Farias I.P. & Giarrizzo T. (2017) A new species of *Tometes* Valenciennes 1850 (Characiformes: Serrasalmidae) from Tocantins-Araguaia River Basin based on integrative analysis of molecular and morphological data. *PLoS ONE* 12, 1–21.

Arcila D., Ortí G., Vari R., Armbruster J.W., Stiassny M.L.J., Ko K.D., *et al.* (2017) Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nature Ecology and Evolution* 1, 1–10.

Baldwin C.C. & Weigt L.A. (2012) A New Species of Soapfish (Teleostei: Serranidae: Rypticus), with Redescription of R. subbifrenatus and Comments on the Use of DNA Barcoding in Systematic Studies. *Copeia* 2012, 23–36.

Bejerano G., Pheasant M., Makunin I., Stephen S., Kent W.J., Mattick J.S., *et al.* (2004) Ultraconserved elements in the human genome. *Science* 304, 1321–1325.

Benine R.C., Mariguela T.C. & Oliveira C. (2009) New species of *Moenkhausia* Eigenmann, 1903 (Characiformes: Characidae) with comments on the *Moenkhausia oligolepis* species complex. *Neotropical Ichthyology* 7, 161–168.

Britski H.A., Silimon K.Z. de S. & Lopes B.S. (2007) *Peixes do Pantanal: manual de identificação*, 2nd ed. Embrapa Informação Tecnológica, Brasília. 184 p.

Burress E.D., Alda F., Duarte A., Loureiro M., Armbruster J.W. & Chakrabarty P. (2018) Phylogenomics of pike cichlids (Cichlidae: *Crenicichla*): the rapid ecological speciation of an incipient species flock. *Journal of Evolutionary Biology* 31, 14–30.

Calcagnotto D., Schaefer S.A. & DeSalle R. (2005) Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. *Molecular Phylogenetics and Evolution* 36, 135–153.

Caleta M., Tutman P., Buj I., Zanella D., Mustafic P., Marcic Z., *et al.* (2011) How was a

Pirapitinga, *Piaractus brachypomus* (Serrasalmidae) introduced in Croatian freshwaters? *Cybium* 35, 259–261.

Castagnolli N. (2000) Brazil fish culture. In: *Encyclopedia of Aquaculture*. (Ed. R.R. Stickney), pp. 123–128. John Wiley & Sons, Inc, New York.

Cione A.L., Dahdul W.M., Lundberg J.G. & Machado-Allison A. (2009) *Megapiranha paranensis* , a new genus and species of Serrasalmidae (Characiformes, Teleostei) from the upper Miocene of Argentina. *Journal of Vertebrate Paleontology* 29, 350–358.

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, 1–12.

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, 783–786.

Dahdul W.M. (2007) *Phylogenetics and diversification of the Neotropical Serrasalminae (Ostariophysi: Characiformes)*. Faculties of the University of Pennsylvania. 107 p.

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.

DeSalle R., Egan M.G. & Siddall M. (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360, 1905–1916.

Eigenmann C.H.. (1915) The Serrasalminae and Myleinae. *Annals of the Carnegie Museum* 9, 226–272.

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.

Freeman B., Nico L.G., Osentoski M., Jelks H.L. & Collins T.M. (2007) Molecular systematics of Serrasalmidae: Deciphering the identities of piranha species and unraveling their evolutionary histories. *Zootaxa* 38, 1–38.

Fricke R., Eschmeyer W.N. & Fong J.D. (2019) Species by family/subfamily. (http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version accessed 09 Jan 2019.

Fricke R., Eschmeyer W.N. & van der Laan R. (2019) Catalog of fishes: genera, species, references. (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 09 Jan 2019.

Géry J. (1972) Poissons Characoïdes des Guyanes. I. Généralités. II. Famille des Serrasalmidae. *Zoologische Verhandelingen* 122, 1–250.

Gosline W.A. (1951) Notes on the characid fishes of the subfamily Serrasalminae. *Proceedings of the California Academy of Sciences* 27, 17–64.

Hebert P.D.N., Cywinska A., Ball S.L. & deWaard J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* 270, 313–321.

Hensel K. (2004) First record of the pirapitinga *Piaractus brachypomus* (Actinopterygii : Serrasalmidae) in Slovakia. *Biologia* 59, 205–210.

Hubert N., Duponchelle F., Nuñez J., Garcia-Davila C., 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, 2115–2136.

Jégu M. (2003) Subfamily Serrasalminae. In: *Check List of the Freshwater Fishes of South and Central America*. (Eds R.E. Reis, S.O. Kullander & C.J. Ferraris), pp. 185–199. EDIPUCRS, Porto Alegre.

Jégu M. (2004) *Taxonomie des Serrasalminae phytophages et Phylogenie des Serrasalminae (Teleostei: Characiformes: Characidae).* Museum National d'Histoire Naturelle.

Lévêque C., Oberdorff T., Paugy D., Stiassny M.L.J. & Tedesco P.A. (2008) Global diversity of fish (Pisces) in freshwater. *Hydrobiologia* 595, 545–567.

Lowe-McConnell R.H. (1987) *Ecological studies in tropical fish communities*. Cambridge University Press, Cambridge. 382 pp.

Lundberg J.G., Kottelat M., Smith G.R., Stiassny M.L.J. & Gill A.C. (2000) So Many Fishes, So Little Time: An Overview of Recent Ichthyological Discovery in Continental Waters. *Annals of the Missouri Botanical Garden* 87, 26–62.

Luz L.A., Reis L.L., Sampaio I., Barros M.C. & Fraga E. (2015) Genetic differentiation in the populations of red piranha, *Pygocentrus nattereri* Kner (1860) (Characiformes: Serrasalminae), from the river basins of northeastern Brazil. *Braz. J. Biol.* 75, 838–845.

Machado-Allison, A. (1982). *Studies on the Systematics of the Subfamily Serrasalminae (Pisces-Characidae)*. Section II. National Museum of Natural History, Smithsonian, pp. 96–267.

Machado-Allison A. (2002) Los Peces Caribes De Venezuela : Una Aproximación a Su Estudio Taxonómico. *Academia de Ciencias Fisicas, Matematicas y Naturales* LXII, 35–88.

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, 1–12.

Mateussi N.T.B., Oliveira C. & Pavanelli C.S. (2018) Taxonomic revision of the Cis-Andean species of *Mylossoma* Eigenmann & Kennedy, 1903 (Teleostei: Characiformes: Serrasalmidae). *Zootaxa* 4387, 275–309.

Mateussi N.T.B., Pavanelli C.S. & Oliveira C. (2016) Molecular identification of cryptic diversity in species of cis-Andean *Mylossoma* (Characiformes: Serrasalmidae). *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis* 28, 778–780.

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, 746–754.

McCormack J.E., Harvey M.G., Faircloth B.C., Crawford N.G., Glenn T.C. & Brumfield R.T. (2013) A Phylogeny of Birds Based on Over 1,500 Loci Collected by Target Enrichment and High-Throughput Sequencing. *PLoS ONE* 8, e54848.

Melo B.F., Benine R.C., Mariguela T.C. & Oliveira C. (2011) A new species of *Tetragonopterus* Cuvier, 1816 (Characiformes: Characidae: Tetragonopterinae) from the rio Jari, Amapá, northern Brazil. *Neotropical Ichthyology* 9, 49–56.

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

Mol J.H. (2006) Attacks on humans by the piranha *Serrasalmus rhombeus* in Suriname. *Studies on Neotropical Fauna and Environment* 41, 189–195.

Norman J.R. (1929) The South American Characid fishes of the subfamily Serrasalmoninae with a revision of the genus *Serrasalmus* Lacepede. *Proceedings of the Zoological Society of London* 52, 781–829.

Oliveira C., Avelino G.S., Abe K.T., Mariguela T.C., Benine R.C., Ortí G., *et al.* (2011) Phylogenetic relationships within the speciose family Characidae (Teleostei: Ostariophysi: Characiformes) based on multilocus analysis and extensive ingroup sampling. *BMC Evolutionary Biology* 11, 275.

Ortí G., Petry P., Porto J.I.R., JEGU M. & Meyer A. (1996) Patterns of nucleotide change in Mitochondrial Ribosomal RNA genes and the Phylogeny of Piranhas. *Journal of*

Molecular Evolution 42, 169–182.

Ortí G., Sivasundar A., Dietz K. & Jégu M. (2008) Phylogeny of the Serrasalmidae (Characiformes) based on mitochondrial DNA sequences. *Genetics and Molecular Biology* 31, 343–351.

Ota R.P., Röpke C.P., Zuanon J. & Jégu M. (2013) Serrasalmidae. In: *Peixes do rio Madeira v.II*. (Eds L.J. Queiroz, G. Torrente-Vilara, W.M. Ohara, T.H. da S. Pires, J. Zuanon & C.R. da C. Doria), pp. 14–48. Santo Antônio Energia, São Paulo - Brasil.

Pennacchio L.A., Ahituv N., Moses A.M., Prabhakar S., Nobrega M.A., Shoukry M., *et al.* (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444, 499–502.

Pereira L.H.G., 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 Paran Basin of Brazil. *Mitochondrial DNA* 22, 87–96.

Reis R.E., Albert J.S., Di Dario F., Mincarone M.M., Petry P. & Rocha L.A. (2016) Fish biodiversity and conservation in South America. *Journal of fish biology* 89, 12–47.

Reis R.E., Kullander S.O. & Ferraris-Jr. C.J. (2003) *Check list of the freshwater fishes of South and Central America*. 729 p.

Rosa R.S. & Lima F.C.T. (2008) Peixes. In: *Livro Vermelho da Fauna Brasileira Ameaçada de Extinção*, 1st edn. (Eds A.B.M. Machado, G.M. Drummond & A.P. Paglia), pp. 8–285. MMA, Brasília, DF.

Ruffino M.L. (2004) A pesca e os recursos pesqueiros na Amazônia brasileira. Manaus: Ibama/ProVárzea, 272 p.

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.

Santos C.H. dos A. dos, Leitão C.S. de S., Paula-Silva M. de N. & Almeida-Val V.M.F. (2016) Genetic differentiation in red-bellied piranha populations (*Pygocentrus nattereri* , Kner , 1858) from the Solimões-Amazonas River. *Ecology and Evolution* 6, 4203–4213.

Santos G.M., Ferreira E.J.G. & Zuanon J.A.S. (2006) *Peixes comerciais de Manaus*. ProVárzea/Ibama, Manaus, 144 p.

Schaefer S.A. (1998) Conflict and resolution: impact of new taxa on Phylogenetic studies of the Neotropical cascudinhos (Siluroidei: Loricariidae). In: *Phylogeny and Classification of Neotropical Fishes*. (Eds L.R. Malabarba, R.E. Reis, R.P. Vari, Z.M.S. Lucena & Luce), pp. 375–400. EDIPUCRS, Porto Alegre.

Siepel A., Bejerano G., Pedersen J.S., Hinrichs A.S., Hou M., Rosenbloom K., *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Research* 15, 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, 1613–1631.

Taphorn D.C. (2003) *Manual de Identificación y Biología de los Peces Characiformes de la Cuenca del Río Apure en Venezuela*. BioCentro, Universidad Nacional Experimental de los Llanos Occidentales "Ezequiel Zamora," Guanare, Venezuela, 393 p.

Thompson A.W., Betancur R., López-Fernández H. & Ortí G. (2014) A time-calibrated, multi-locus phylogeny of piranhas and pacus (Characiformes: Serrasalmidae) and a comparison of species tree methods. *Molecular Phylogenetics and Evolution* 81, 242–257.

Vari R.P. & Malabarba L.R. (1998) Neotropical Ichthyology: An Overview. In: *Phylogeny and Classification of Neotropical Fishes*. (Eds L.R. Malabarba, R.E. Reis, R.P. Vari, Z.M.S. Lucena & Luce), pp. 1–11. EDIPUCRS, Porto Alegre.

Weitzman S.H. (1962) The osteology of *Brycon meeki*, a generalized characoid fish, with a definition of the family. *Stanford Ichthyology Bulletin.* 8: 1–77.

Wiley E.O. & Johnson D. (2010) A teleost classification based on monophyletic groups. In: *Origin and Phylogenetic Interrrelationships of Teleosts.* (Eds. Nelson, J.S., Schultz, H.P. & Wilson, M.V.H.), pp. 123–182. Verlag Dr. Friedrich Pfeil, München, Germany.

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

Phylogenomics of the Neotropical fish family Serrasalmidae (Ostariophysi: Characiformes) based on ultraconserved elements

Phylogenomics of the Neotropical fish family Serrasalmidae (Ostariophysi: Characiformes) based on ultraconserved elements

Abstract

The Neotropical fish family Serrasalmidae comprises 16 extant genera widespread through major Neotropical rivers with relevant importance for regional fisheries and aquaculture. The monophyly of Serrasalmidae is consensual between morphological and molecular studies, as well as the recognition of three main clades, the "pacu clade" composed by *Colossoma*, *Mylossoma* and *Piaractus*; the "*Myleus* clade", composed by all other pacus except *Metynnis*; and the "piranha clade", composed by *Metynnis* plus five piranha genera. However, both intergeneric and interspecific relationships within each clade remain uncertain. We used a species-dense matrix of ultraconserved elements, multiple nuclear orthologous loci widely applied to phylogenetic studies, to construct a new phylogeny for the family. A total of 69 species (72%) and all genera plus 16 related taxa as outgroup were used to perform maximum likelihood (ML), Bayesian inference (BI) and species tree analyses. We obtained highly-supported phylogenies corroborating the three major-clade hypotheses (i.e. "pacu" clade, *Myleus* clade and "piranha" clade) herein recognized at the subfamily and tribe levels. Our results highlight the complexity for resolving the relationships mainly within the "*Myleus* clade" with *Myleus, Mylesinus, Myloplus, Tometes* and *Utiaritichthys* as non-monophyletic genera as well as *Serrasalmus* and *Pristobrycon* in the "piranha clade". Some taxonomic changes are proposed here to better accommodate species in monophyletic clades.

Key words: biodiversity; phylogeny; pacu, piranha, taxonomy; systematics, UCEs.

Introduction

Serrasalmidae is a family of Neotropical freshwater fishes comprising 17 genera, including †*Megapiranha* from the Upper Miocene, and 97 valid species (Fricke *et al.* 2019b), represented by the popularly known "pacus", "piranhas" and the large "tambaqui". Serrasalmids are broadly distributed through major South American basins and possess a wide diversity of dietary specializations such as carnivory, herbivory and lepidophagy, resulting in extremely polymorphic morphological traits, largely used in taxonomy and systematics (Goulding 1980; Sazima 1983; Sazima & Machado 1990; Nico & Taphorn 1998;).

During the 20th century, the taxonomic status of serrasalmids was discordant, being ranked as a family by some authors (Géry 1972, 1976; Géry *et al.* 1987; Jégu & dos Santos 1988) or a subfamily within Characidae by others (Machado-Allison 1983, 1985; Merckx *et al.* 2000; Jégu 2003, 2004). More recently, morphological (Mirande 2009, 2010), molecular studies (Calcagnotto *et al.* 2005; Oliveira *et al.* 2011; Arcila *et al.* 2017) and total evidence (Mirande 2018) suggested Serrasalmidae to be more closely related to the superfamily Anostomoidea and relatives rather than to any other Characidae subfamily.

Although the majority of the nominal species of Serrasalmidae were described in the 19th century (Jégu *et al.* 2004), the first classification of serrasalmids was made by Eigenmann (1915), who proposed two subfamilies in Characidae: the Serrasalminae, composed by species having a single teeth row on each jaw, and Myleinae, with species presenting two teeth rows in the premaxilla. Subsequent authors recognized basically the same groups, differing mainly on the relative rank of some taxa (Norman 1929; Nelson 1961; Géry 1972). Despite the taxonomic advance, these classifications recognized non-monophyletic groups as valid taxonomic units.

Machado-Allison (1982) conducted the first cladistic analysis for the family based on an extensive study of morphological characters, obtaining very similar results to those proposed by Eigenmann (1915). Later, Machado-Allison (1983) suggested Serrasalminae as a monophyletic subfamily of Characidae based on 27 synapomorphies, divided in two main clades. The "lineage A", composed by the so-called "pacus" (*Acnodon*, *Colossoma*, *Mylesinus*, *Myleus*, *Mylossoma*, *Piaractus* and *Utiaritichthys*), which comprised most of Myleinae proposed by Eigenmann (1915), and the "lineage B", constituted by all "piranhas" (*Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Serrasalmus*), corresponding to the Serrasalminae of Eigenmann (1915) plus *Metynnis* and *Catoprion* (Fig. 1A). The "lineage B" was slightly changed after the revision of some piranha genera (Machado-Allison *et al.* 1989)

(Fig. 1B). In conclusion, Machado-Allison (1983) suggested a further revision of some paraphyletic genera.

The first molecular analysis of Serrasalmidae was performed by Ortí *et al.* (1996) based on sequences of 12S and 16S mitochondrial ribosomal genes, with 37 species including all genera except *Ossubtus, Pygopristis* and *Utiaritichthys*, resulting three major clades rather than two as previously proposed by the morphological hypothesis. The first clade, composed by *Colossoma*, *Mylossoma* and *Piaractus*, appeared sister to the remaining serrasalmids corroborating previous hypothesis (Gosline 1951; Nelson 1961). *Acnodon* formed the sister group of two other groups, the "*Myleus* clade", that includes *Myleus*, *Mylesinus* and a new genus (*Tometes* according to Jégu *et al.* (2002)), and the "true piranhas" group, which is formed by the remaining serrasalmids reinforcing the hypothesis of non-monophyly of *Serrasalmus* and *Pristobrycon* (Fig. 1C). However, relatively low levels of sequence divergence among rRNA genes resulted in a poor resolution within these groups. Moreover, despite the absence of some genera (*Ossubtus, Pygopristis* and *Utiaritichthys*), the results highlighted the necessity of revisions of some genera such as *Myleus* and *Pristobrycon*.

Performing a new phylogeny for Serrasalmidae, Ortí *et al.* (2008) used 44 sequences of the mtDNA control region (D-loop) along with 12S and 16S sequences of 74 taxa, in order to represent all serrasalmid genera except *Utiaritichthys*. The authors corroborated the arrangement of Serrasalmidae composed by the same three main clades, denominated "pacu" clade as sister group of all other serrasalmids; the "*Myleus*" clade, composed by *Mylesinus*, *Myleus*, *Ossubtus* and *Tometes*; and the "piranha" clade, with *Metynnis* as sister group of *Catoprion*, *Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Serrasalmus* (Fig. 1C). Not surprisingly, the monophyly of some genera of the "*Myleus"* clade was not supported as well as the uncertain position of *Acnodon* (Ortí *et al.* 2008). For the "piranhas" clade, the authors obtained *Serrasalmus gouldingi* Fink & Machado-Allison, 1992 and *Pristobrycon* (not *P. striolatus*) as a monophyletic group, sister to *Serrasalmus* and *Pygocentrus*; this was sister to the clade with *Catoprion*, *Pygopristis* and *Pristobrycon striolatus*. According to the authors, many of generic designations need higher support or are clearly contradicted by the obtained data; besides concluding that the weak resolution obtained is consequence of a poor taxonomic sampling.

Similar results were obtained by Jégu (2004) based on 278 morphological characters of 38 serrasalmid taxa. Differences involved the resolution of *Colossoma+Piaractus* as sister group of *Mylossoma* and all other serrasalmids, *Acnodon* as sister group of the "*Myleus* clade"

and *Pristobrycon* as sister group of *Pygocentrus* and *Serrasalmus*. Nevertheless, Jégu (2004) also did not include *Utiaritichthys* in his analysis.

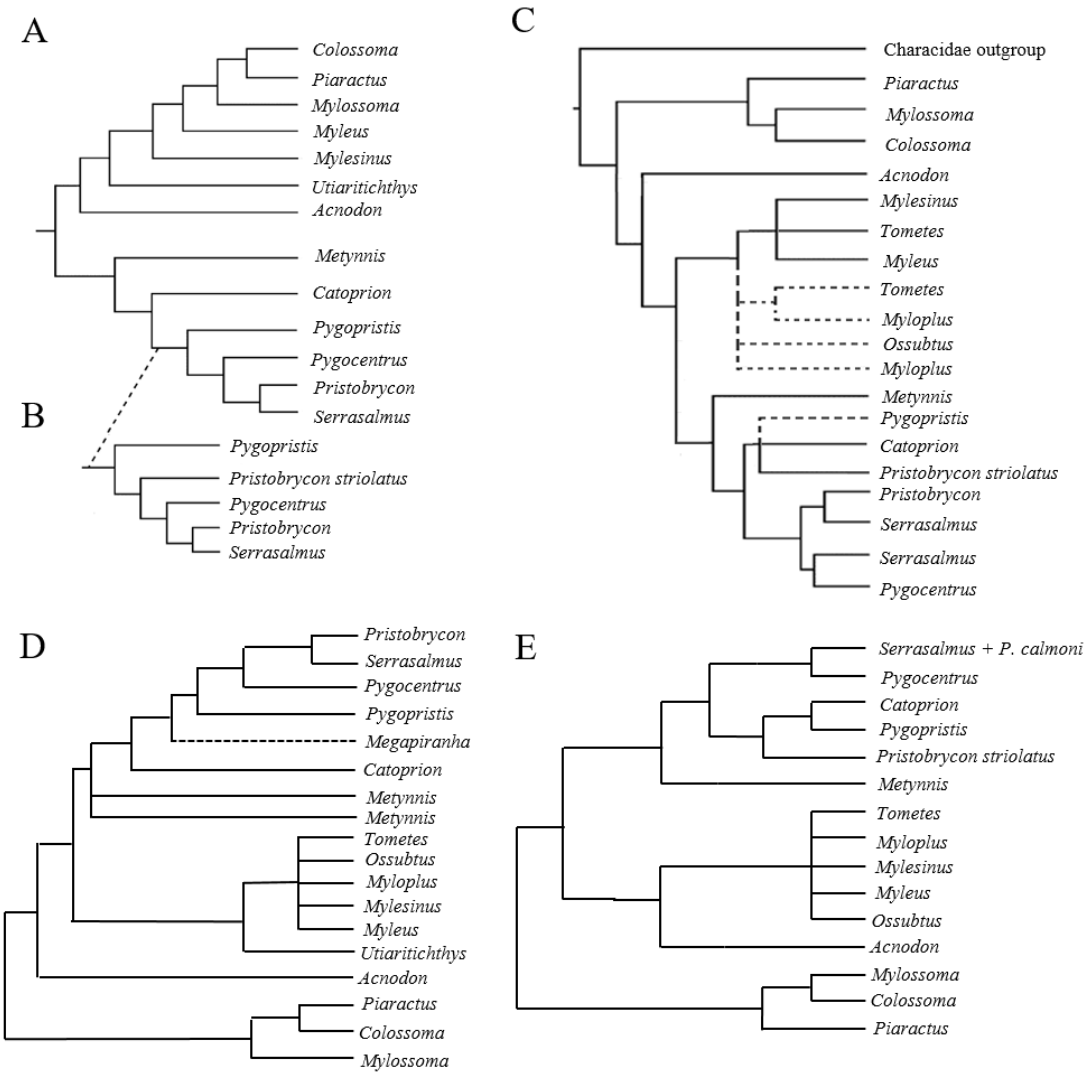

Figure 1. Previous intergeneric hypotheses within Serrasalmidae. (A) Morphological phylogeny of Machado-Allison (1982). (B) Morphological phylogeny of Machado-Allison *et al.* (1989), slightly different from A. (C) Molecular phylogeny of Ortí *et al.* (1996; 2008) based on mtDNA; dotted lines indicate the changes recovered by Ortí *et al.* (2008). Ortí *et al.* (1996) placed a new genus with *Myleus* and *Mylesinus*, identified latter as *Tometes* by Jégu (2002). (D) Morphological phylogeny of Cione *et al.* (2009) after Dahdul (2007); dotted line indicates the fossil genus. (E) Molecular phylogeny of Thompson *et al.* (2014) based on 10 nuclear and mitochondrial loci.

Dahdul (2007) performed both morphological and molecular phylogenies in an effort to understand relationships of Serrasalmidae. For the morphological analysis, all genera were represented from a wide geographical range and the results recovered basically the same three groups of Ortí *et al.* (1996, 2008). However, she recovered *Metynnis* as paraphyletic and sister

group of all other piranhas (Fig. 1D). The molecular analysis, using nuclear markers *rag 1* and *rag 2* and mitochondrial sequences of 12S and 16S, recovered the same three groups. However, in the molecular analysis, *Utiaritichthys* was also absent.

More recently, Thompson *et al.* (2014) performed a multilocus calibrated phylogenetic analysis of Serrasalmidae including 38 nominal species (40% of the family diversity) and all genera except *Utiaritichthys*. The authors corroborated the monophyly of the three main clades and the non-monophyly of several taxa (e.g. *Pristobrycon*, *Tometes*, *Myloplus* and *Serrasalmus rhombeus*) (Fig. 1E). They also suggested a much more recent "piranha" clade (Miocene, 15-20 Ma) than those others of the family, which instead, began to diversify during the Cretaceous (65-75 Ma). However, the clade was much older than proposed by Hubert *et al.* (2007) and Cione *et al.* (2009). The recent phylogenomic approach of Characiformes returned a topology with *Piaractus* sister to all serrasalmids, and then three main clades: *Mylossoma/Colossoma*, the *Myleus* clade, and the piranha clade (Betancur *et al.* 2018).

Despite the existence of several studies dedicated to elucidate the phylogenetic relationships within Serrasalmidae, their focus are on either morphological characters (Machado-Allison *et al.* 1989) or on a few molecular markers (Ortí *et al.* 1996, 2008; Thompson *et al.* 2014) and investigations applying genome approaches specifically to Serrasalmidae were never conducted. One of these genomic markers used in molecular systematics is the ultraconserved elements (UCEs), which consist in genomic regions extremely conserved and shared among distinct groups along the tree of life.

Faircloth *et al.* (2012) introduced the UCEs as a new class of molecular markers on phylogenetic studies through the enrichment of genomic libraries containing hundreds or thousands of nuclear loci using next-generation sequencing. Since UCE sequences are highly conserved, they are used as target to probes so that the polymorphic flanking regions can be used to reconstruct phylogenies. Therefore, UCEs has been used at many comparison levels among organisms, from population to large groups, for all sort of animals, as for example invertebrates (Faircloth *et al.* 2015; Zhang *et al.* 2018b), amphibians (Streicher *et al.* 2018), fishes (Alfaro *et al.* 2018; Faircloth *et al.* 2013), non-avian reptiles (Crawford *et al.* 2012, 2015), birds (Barker 2017; Bruxaux *et al.* 2018; Smith *et al.* 2014), and mammals (McCormack *et al.* 2012), proving to be highly resolutive with strongly supported phylogenies. Studies with fishes had shown high resolution of nodes at all levels (ancient and recent clades), as seen for actinopterygians (Faircloth *et al.* 2013), Neotropical cichlids (Burress *et al.* 2018) and acantomorphs (Alfaro *et al.* 2018). As such, the use of UCEs on

studies for complex groups, as Neotropical freshwater fishes, is very promising and highly effective.

In view of the foregoing, although the monophyly of Serrasalmidae is well stablished, several aspects of relationships between groups or genera of the family are still conflicting. Despite the agreement on the presence of three main clades, politomies and doubts about the monophyly and taxonomic attribution of several genera and subgroups are still observed. Therefore, the present study aims to elucidate the relationships of serrasalmids using, for the first time, all serrasalmids genera and a large number of species in a phylogenomic analysis with thousands of ultraconserved elements.

Material and Methods

Taxon sampling and library preparation. Tissues of 69 species of serrasalmids were obtained (Table 1), including all genera. Outgroup taxa include related characiform families Alestidae, Anostomidae, Chilodontidae, Curimatidae, Cynodontidae, Erythrinidae, Hemiodontidae, Parodontidae and Prochilodontidae (Table 1) based on previous molecular phylogenetic hypotheses (Oliveira et al. 2011; Arcila et al. 2017). Tissue samples and voucher specimens are deposited in the Academy of Natural Sciences, Philadelphia (ANSP), Coleção de Peixes do Museu de Ciências e Tecnologia da PUCRS, Porto Alegre (MCP), Coleção Zoológica de Referência da Universidade Federal de Mato Grosso do Sul, Campo Grande (ZUFMS-PIS), Laboratório de Biologia e Genética de Peixes, Universidade Estadual Paulista, Botucatu (LBP), Oregon State University, Corvallis (OS), Tropical Research Institute, Smithsonian Institution, Panama (STRI) and Universidade Federal de Rondônia, Porto Velho (UFRO-ICT).

Whole genomic DNA was extracted from ethanol-preserved muscle samples with the DNeasy Tissue Kit (Qiagen) and quantified using the Qubit dsDNA broad range Assay Kit (Invitrogen, Life Technologies) following manufacturer's protocol. The probeset was developed for ostariophysan fishes to generate sequence data for about 2,500 UCE loci (Faircloth *et al*. in prep.). Library preparation, sequencing, and raw data pipelining were performed by Arbor Biosciences® (Ann Arbor, Michigan, USA; http://www.arborbiosci.com) using the following protocol: DNA library preparation by modifying the Nextera (Epicentre Biotechnologies) library preparation protocol for solution-based target enrichment following Faircloth *et al.* (2012) and increasing the number of PCR cycles following the tagmentation

reaction to 20 as recommended by Faircloth *et al.* (2012). Nextera library was used to preparation protocol of *in vitro* transposition followed by PCR to prune the DNA and attach sequencing adapters (Adey *et al.* 2010), then the Epicentre Nextera kit was used to prepare transposase-mediated libraries with insert sizes averaging 100 bp (95% CI: 45 bp) (Adey *et al.* 2010).

To prepare libraries, whole genomic DNA (concentration of 40 ng/µl) was first sheared with a QSonica Q800R instrument and selected to modal lengths of approximately 500nt using a dual-step SPRI bead cleanup. 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 connection 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. After purification with SPRI beads, libraries were quantified with the Quant-iT(TM) Picogreen(R) dsDNA Assay kit (ThermoFisher). Pools were enriched comprising 100 ng each of 8 libraries (800 ng total) using the MYbaits(R) Target Enrichment system (MYcroarray) following manual version 3.0. After capture cleanup, the bead-bound library was resuspended in the recommended solution and amplified for 10 cycles using a universal P5/P7 primer pair and KAPA HiFi reagents. After purification, each captured library pool was quantified with PicoGreen and combined with all other pools in projected equimolar ratios prior to sequencing. Sequencing was performed across two Illumina HiSeq paired-end 100 bp lanes using v4 chemistry.

Table 1. Taxon, vouchers and locality information of the analyzed specimens of Serrasalmidae and related taxa.

Taxon	Voucher	Specimen number	Locality, basin	Coordinates	Country
Acnodon normani	LBP 19140	77187	Rio Paranã, Tocantins basin	12°37'31"S 47°52'59"W	Brazil
Catoprion mento	LBP 7556	35626	Rio Cuiabá, Paraguay basin	16°11'39"S 55°48'25"W	Brazil
Catoprion sp. N.	LBP 21615	63665	Igarapé Au-au, Amazon basin	02°56'19"N 61°03'06"W	Brazil
Colossoma macropomum	LBP 12837	54043	Rio Tapajós, Amazon basin	04°16'49"S 59°59'26"W	Brazil
Colossoma macropomum	LBP 12838	54052	Rio Tapajós, Amazon basin	04°16'49"S 59°59'26"W	Brazil
Metynnis altidorsalis	LBP 9997	43151	Rio Pelehojo, Orinoco basin	07°32'22"N 66°08'31"W	Venezuela
Metynnis cuiaba	LBP 8580	43371	Rio Paraguai, Paraguay basin	15°04'33"S 57°11'04"W	Brazil
Metynnis fasciatus	LBP 3978	22824	Lagoa Fazenda Taboca, Araguaia basin	20°05'07"S 50°58'59"W	Brazil

Metynnis guaporensis	UFRO-ICT 5396	6018	Igarapé Jatuarana, Madeira basin	08°49'55"S 64°02'55"W	Brazil
Metynnis hypsauchen	LBP 18390	43152	Rio Pelehojo, Orinoco basin	07°32'22"N 66°08'31"W	Venezuela
Metynnis lippincottianus	LBP 5429	27164	Rio Jari, Amazon basin	00°56'00"S 52°32'30"W	Brazil
Metynnis longipinnis	LBP 15530	63920	Rio Takutu, Amazon basin	03°17'57"N 59°55'38"W	Brazil
Metynnis luna 1	LBP 18398	42589	Rio Guamá, Amazon basin	01°34'00"S 47°09'51"W	Brazil
Metynnis luna 2	LBP 9394	43863	Rio Guamá, Amazon basin	01°34'00"S 47°09'51"W	Brazil
Metynnis maculatus	UFRO-ICT 5395	3986	Rio Jaciparaná, Madeira basin	09°17'03"S 64°23'57"W	Brazil
Metynnis mola	LBP 8447	42354	Rio Paraguai, Paraguay basin	16°03'13"S 57°48'31"W	Brazil
Metynnis polystictus	LBP 8015	37703	Rio Arinos, Amazon basin	14°08'21"S 56°04'19"W	Brazil
Mylesinus paraschomburgkii	LBP 20484	80923	Rio Jari, Amazon basin	00°32'01"S 52°35'05"W	Brazil
Mylesinus paucisquamatus	LBP 12839	53447	Rio Tapajós, Amazon basin	04°33'09"S 56°17'59"W	Brazil
Mylesinus sp.	LBP 6853	32707	Rio Negro, Amazon basin	00°49'S 62°49'W	Brazil
Myleus micans	LBP 24011	91323	Córrego Catolé, São Francisco basin	15°12'40"S 44°49'54"W	Brazil
Myleus setiger	LBP 16523	61481	Rio Xingu, Amazon basin	03°15'24"S 52°05'47"W	Brazil
Myloplus arnoldi 1	LBP 20266	79748	Rio Apiacás, Amazon basin	10°19'45"S 56°59'17"W	Brazil
Myloplus arnoldi 2	LBP 13827	57247	Rio Tapajós, Amazon basin	04°33'09"S 56°17'59"W	Brazil
Myloplus asterias	LBP 9072	42585	Rio Guamá, Amazon basin	01°34'17"S 47°10'10"W	Brazil
Myloplus levis	ZUFMS-PIS 5567	01	Rio Indaiá Grande, upper Paraná basin	19º43'32"S 51º55'48"W	Brazil
Myloplus planquettei	ANSP 179808	230	Yukanopito falls, Essequibo basin	01º54'53"N 58º31'14"W	Guyana
Myloplus rhomboidalis	ANSP 185981	10924	Rio Iriri, Amazon basin	03º50'32"S 52º44'03"W	Brazil
Myloplus rubripinnis	UFRO-ICT Uncatalogued	5782	Madeira basin	-	Brazil
Myloplus schomburgkii	OS 18990	PE10-044	Rio Nanay, Amazon basin	-	Peru
Myloplus ternetzi	ANSP 188689	1626	Lawa river, Maroni basin	03º19'31"N 54º03'48"W	Suriname
Myloplus tiete	INPA 53243	53243	Córrego Caracu, upper Paraná basin	22º40'00"S 53º15'00"W	Brazil
Mylossoma acanthogaster	LBP 24311	91508	Rio Sardinata, Catatumbo basin	08°39'25"N 72°37'46"W	Colombia
Mylossoma albiscopum	LBP 12089	51722	Rio Madeira, Amazon basin	08°51'42"S 64°03'49"W	Brazil
Mylossoma aureum	UFRO-ICT 19346	10260	Igarapé Jatuarana, Madeira basin	08°45'54"S 64°02'39"W	Brazil
Mylossoma duriventre	LBP 3741	21919	Rio Negro, Paraguay basin	19°34'33"S 56°14'49"W	Brazil
Mylossoma unimaculatum	LBP 12745	41182	Rio Araguaia, Tocantins basin	13°18'37"S 50°36'47"W	Brazil
Mylossoma sp.N 1	LBP 18873	75056	Laguna del Castilleros,	07º30'50"N	Venezuela

			río Orinoco	66º09'19"W	
Mylossoma sp.N 2	LBP 2190	15518	Laguna des Castilleros, Orinoco basin	07º30'50"N 66º09'19"W	Venezuela
Ossubtus xinguense	LBP 16563	64065	Rio Xingu, Amazon basin	03°15'24"S 52°05'47"W	Brazil
Piaractus mesopotamicus	LBP 4255	23804	Rio Paraná, upper Paraná basin	-	Brazil
Pristobrycon calmoni	LBP 2191	15554	Laguna des Castilleros, Orinoco basin	07º30'50"N 66º09'19"W	Venezuela
Pristobrycon striolatus 1	ANSP 188672	1631	Litanie rivier, Maroni basin	03°17'24"N 54°04'38"W	Suriname
Pristobrycon striolatus 2	ANSP 197515	10494	Rio Xingu, Amazon basin	02°51'55"S 51°59'21"W	Brazil
Pristobrycon striolatus 3	LBP 15050	61685	Rio Tapajós, Amazon basin	04°34'07"S 56°18'49"W	Brazil
Pristobrycon sp.	UFRO-ICT 7235	6010	Lago Cuniã, Madeira basin	08º19'56"S 63º29'00"W	Brazil
Pygocentrus cariba	LBP 10225	43107	Rio Apure, Orinoco basin	07°37'24"N 66°24'48"W	Venezuela
Pygocentrus nattereri	LBP 12693	43552	Rio Araguaia, Amazon basin	13°19'S 50°37'W	Brazil
Pygocentrus piraya 1	LBP 11336	45523	Lagoa da Tiririca, São Francisco basin	17°13'33"S 44°48'27"W	Brazil
Pygocentrus piraya 2	LBP 11286	48749	Rio São Francisco, São Francisco basin	09°56'46"S 37°06'15"W	Brazil
Pygocentrus piraya 3	LBP 11286	48750	Rio São Francisco, São Francisco basin	09°56'46"S 37°06'15"W	Brazil
Pygopristis denticulata 1	LBP 21609	62390	Rio Branco, Amazon basin	03°08'16"N 60°16'33"W	Brazil
Pygopristis denticulata 2	LBP 15529	63916	Rio Takutu, Amazon basin	03°17'57"N 59°55'38"W	Brazil
Serrasalmus brandtii	LBP 11269	48707	Rio São Francisco, São Francisco basin	09°51'23"S 37°06'30"W	Brazil
Serrasalmus compressus	UFRO-ICT 24705	2546	Rio Jaciparaná, Madeira basin	09°16'58"S 64°23'53"W	Brazil
Serrasalmus eigenmanni	LBP 16210	66983	Rio Tracuá, Amazon basin	04°28'11"S 56°17'01"W	Brazil
Serrasalmus elongatus 1	UFRO-ICT 1807	125	Lago do Cuniã, Madeira basin	08°20'36"S 63°30'48"W	Brazil
Serrasalmus elongatus 2	LBP 18025	72569	Igarapé Alencar, Amazon basin	03°05'57"S 58°27'18"W	Brazil
Serrasalmus gibbus	LBP 12811	41289	Rio Araguaia, Tocantins basin	13°18'37"S 50°36'47"W	Brazil
Serrasalmus gouldingi	LBP 20976	81155	Igarapé Carucaru, rio Jari, Amazon basin	00°56'00"S 52°32'30"W	Brazil
Serrasalmus hollandi	UFRO-ICT 5631	6805	Igarapé Jatuarana, Madeira basin	08°46'25"S 64°02'49"W	Brazil
Serrasalmus humeralis	LBP 4221	22719	Rio Juruá, Amazon basin	07°09'49"S 73°43'29"W	Brazil
Serrasalmus irritans	LBP 10229	43109	Rio Apure, Orinoco basin	07°37'24"N 66°24'48"W	Venezuela
Serrasalmus maculatus 1	LBP 3821	22107	Rio Negro, Paraguay basin	19°34'33"S 56°14'49"W	Brazil
Serrasalmus maculatus 2	LBP 3821	22108	Rio Negro, Paraguay basin	19°34'33"S 56°14'49"W	Brazil
Serrasalmus manueli	LBP 15032	61627	Rio Tapajós, Amazon basin	04°33'09"S 56°17'59"W	Brazil
Serrasalmus marginatus	LBP 19855	79019	Lagoa do Guaraná, rio Baiá, upper Paraná	22°43'16"S 53°18'09"W	Brazil

			basin		
Serrasalmus medinai	LBP 22502	86574	Lago Yahuarcaca, Amazon basin	04°11'45"S 69°57'20"W	Colombia
Serrasalmus odyssei	UFRO-ICT 20212	10000	Rio Pacaás-novos, Madeira basin	10°51'47"S 65°16'21"W	Brazil
Serrasalmus rhombeus 1	LBP 15365	62022	Rio Culuene, Amazon basin	13°31'02"S 53°04'41"W	Brazil
Serrasalmus rhombeus 2	LBP 15314	63320	Ribeirão Taquarussu, Tocantins basin	10°17'20"S 48°20'00"W	Brazil
Serrasalmus rhombeus 3	LBP 14239	59423	Igarapé Montanha, Amazon basin	04°55'58"S 56°51'51"W	Brazil
Serrasalmus cf. *serrulatus*	LBP 21606	61612	Rio Tapajós, Amazon basin	04°34'07"S 56°18'49"W	Brazil
Serrasalmus sp. "lauzannei"	UFRO-ICT 17954	8091	Rio Karipuna, Madeira basin	09°11'30"S 64°37'21"W	Brazil
Tometes ancylorhynchus	LBP 17403	69141	Rio Francisco Dumont, São Francisco basin	17°19'10"S 44°17'13"W	Brazil
Tometes camunani	LBP 20460	80790	Rio Iratapuru, Amazon basin	00°34'05"S 52°34'41"W	Brazil
Tometes kranponhah	ANSP 200868	13586	Rio Xingu, Amazon basin	03°35'00"S 51°49'24"W	Brazil
Tometes lebaili	ANSP 188681	1634	Litanie rivier, Maroni basin	03°17'24"N 54°04'38"W	Suriname
Tometes trilobatus 1	LBP 5196	26788	Rio Iratapuru, Amazon basin	00°34'03"S 52°34'41"W	Brazil
Tometes trilobatus 2	LBP 21012	81965	Rio Calçoene, Amazon basin	02°31'08"N 51°00'52"W	Brazil
Utiaritichthys longidorsalis	UFRO-ICT 19099	9141	Rio Roosevelt, Amazon basin	07°50'59"S 60°57'51"W	Brazil
Utiaritichthys sennaebragai	MCP 44061	1	Rio Juruena, Amazon basin	13°41'00S 59°00'00W	Brazil
Alestes inferus	AMNH 242137	333238	Mpozo river, Congo basin	05°50'05.45S 13°29'42.13E	DR Congo
Anodus elongatus	LBP 4244	22720	Rio Juruá, Amazon basin	07°09'49"S 73°43'29"W	Brazil
Apareiodon ibitiensis	LBP 2890	18635	Rio Grande, Paraná basin	21°19'37"S 47°14'19"W	Brazil
Curimata vittata	LBP 13846	57302	Rio Tapajós, Amazon basin	04°16'49"S 59°59'26"	Brazil
Cynodon sp.	LBP 10227	43105	Rio Apure, Orinoco basin	07°37'24"N 66°24'48"W	Venezuela
Erythrinus erythrinus	LBP 6625	31955	Rio Paraná, Paraná basin	22°39'01"S 53°04'43"W	Brazil
Hemiodus quadrimaculatus	LBP 21151	82973	Rio Amazonas, Amazon basin	03°04'49"N 51°28'50"W	Brazil
Hepsetus cuvieri		353404			
Hoplias aimara	LBP 20412	80691	Igarapé Pacanari, Jari basin	00°41'10"S 52°36'11"W	Brazil
Laemolyta garmani	OS 18777	PE10-089	Rio Nanay, Amazon basin		Peru
Leporellus vittatus	ANSP 182609	P6322	Rio Nanay, Amazon basin		Peru
Parodon hilarii	LBP 10408	48929	Córrego Joaninha, São Francisco basin	17°19'32"S 44°46'01"W	Brazil
Prochilodus lineatus	Uncatalogued	1	Rio Mogi-Guaçu, Paraná basin		Brazil
Rhaphiodon vulpinus	LBP 12660	43557	Rio Araguaia, Amazon basin	13°19'S 50°37'W	Brazil

Semaprochilodus knerii	ANSP 187277	P4298	Río Apure, Orinoco basin	Venezuela
Steindachnerina argentea	STRI-4270	BFD01625	Turure River	Trinidad & Tobago

Raw data analysis. After sequencing, adapter contamination, low quality bases and the sequences containing ambiguous base calls were trimmed using the Illumiprocessor (Faircloth 2016; https://github.com/faircloth-lab/illumiprocessor). After trimming, Illumina reads were assembled into contigs on a species-by-species basis using ABySS (Simpson *et al.* 2009) (phyluce_assembly_assemblo_abyss). After sequence assembly, a custom Python program (match_contigs_to_probes.py) available at the Phyluce was used, integrating LASTZ (Harris 2007) to align species-specific contigs to our probe-UCE set. This last program creates a relational database of matches to UCE loci by taxon. Then, the get_match_counts.py command (also included in Phyluce) was used to query the database and generate fasta files for UCE loci that are identified across all taxa. A custom Python program (seqcap_align_2.py) was used to align contigs using the Muscle algorithm (Edgar 2004) and to perform edge trimming (phyluce_align_seqcap_align). To ensure that the sequence data aligned together are from the same loci, locus names from sequence lines were removed (phyluce_align_remove_locus_name_from_nexus_lines). Phylogenetic analysis were performed with varying amounts of data (phyluce_align_get_only_loci_with_min_taxa) (50%, 75% and 90% of UCEs that are presented in the completed alignment matrices) to explore the potentially strong effect of missing data on phylogenetic reconstruction (Hosner *et al.* 2016; Streicher *et al.* 2016). And, finally, in order to prepare alignment data for analysis, a concatenated phylip file was created (phyluce_align_format_nexus_files_for_raxml).

Phylogenetic analysis. A Maximum likelihood analysis (ML), a Bayesian inference (BI), and a species tree were performed for the 50%, 75% and 90% complementary matrices. All threes were rooted in *Alestes inferus* + *Hepsetus cuvieri*. For ML analyses and BI, the data was partitioned to account for variation in rates and patterns of molecular evolution among sites using PartitionUCE (Tagliacollo & Lanfear 2018).

The ML analysis of the concatenated alignment was performed using the autoMRE function for the extended majority-rule consensus tree criterion, available in RAxML v8 (Stamatakis, 2014) to access bootstrap support for individual nodes. This option allows the bootstrap convergence to determine if pseudoreplicates reach sufficient stable support values (Pattengale *et al.* 2009). The best tree search was performed under the parameter –N=20,

which specifies the number of alternative runs on distinct parsimony starting trees.

The BI of the concatenated alignment was performed using ExaBayes v1.4 (Aberer *et al.* 2014) with two independent runs (two chains each) of 2,000,000 generations each was performed for the concatenated matrix using the GTR+G model, for different complementary matrices and for trimmed-alignment matrices. Tree space was sampled every 200 generations to yield a total of 10,001 trees. Parameter estimates and ESS values were visualized in Tracer v 1.6 (Rambaut & Drummond 2009). This allows to visualize the log of posterior probability within and between independent runs and to ensure that the average standard deviation of split frequencies was <1%, effective sample sizes (ESS) score >200, and the potential scale reduction factor for estimated parameters approximately 1.0. The 50% most credible set of trees was generated from the posterior distribution of possible topologies using TreeAnnotator (burn-in: 25%).

The species tree was inferred from individual gene trees using ASTRAL-III (Zhang *et al.* 2018a). The individual gene trees used as input to ASTRAL-III were estimated using a ML analysis using a RAxML bootstrapped using the parameter –N=5 and GTRGAMMA model.

Results

Following sequencing and cleaning data, we obtained an average of 4,247,931 cleaned reads per sample from 98 specimens of 85 taxa with a total of 2.29E+09 bp (Table S.1). The matrices with 50%, 75% and 90% complete alignments produced nearly identical topologies and strong node support for the three methods used (*i.e.* ML, BI and species tree). The phylogeny in Figures 2, 3 and 4 corresponds to the results of ML, BI and species tree for the 75% complete matrix, respectively. Remaining trees are available as appendices (Figs. S.1 to S.4). Most nodes of the ML tree presented 100% bootstrap support (Fig. 2), while all nodes in the Bayesian inference received posterior probabilities above 0.9 (Fig. 3) and the species tree showed high posterior probabilities values above 0.9 on most nodes, despite being the analysis with lower values of statistical support (Fig. 4). Values lower than 95% of bootstrap for ML and values lower than 0.9 for posterior probabilities in the BI and species tree analyses are indicated near nodes in the trees.

The UCE phylogeny partially corroborates previous morphological and molecular hypotheses, with differences in the relationships among genera and the monophyly of some

taxa. The monophyly of Serrasalmidae and the three main clades first proposed by Ortí *et al.* (1996) received 100% node support in all analyses performed.

The "pacu" clade (*Colossoma*, *Mylossoma* and *Piaractus*) was also recovered as the first lineage to diversify within Serrasalmidae corroborating previous molecular studies (Ortí *et al.* 1996, 2008; Thompson *et al.* 2014). This three-genus clade appeared sister to all remaining genera with *Piaractus* sister to *Colossoma+Mylossoma*. Internal relationships received maximum statistical support. This is the first phylogeny to include all *Mylossoma* species. The trans-Andean *M. acanthogaster* was recovered as sister to all cis-Andean *Mylossoma* species in addition with a new species from the Orinoco basin (Figs. 2, 3, 4).

The second lineage, the *Myleus* clade, is composed by seven genera: *Acnodon*, *Mylesinus*, *Myleus*, *Myloplus*, *Ossubtus*, *Tometes* and *Utiaritichthys*, all together represented herein as the first time in a molecular phylogeny, with maximum values of statistical supports for almost all relationships. *Mylesinus*, *Myleus*, *Myloplus*, *Tometes* and *Utiaritichthys* appeared as non-monophyletic genera. The monotypic *Ossubtus* was recovered as sister to *Myloplus schomburgkii*, and *Acnodon normani* was recovered as sister either to *Myloplus arnoldi* (species tree and 90%BI) or *Myloplus rhomboidalis* (ML and 75%BI) (Figs. 2, 3, 4).

Third lineage, the "piranha" clade, is composed by *Catoprion*, *Metynnis*, *Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Serrasalmus*. Within the clade, *Metynnis* was recovered as monophyletic, and sister to all other genera. For the first time, a molecular phylogeny of *Metynnis* is presented, and the species were divided into two major clades, with *M. altidorsalis*, *M. cuiaba*, *M. maculatus*, and *M. mola* as sister group to remaining species analyzed herein (Figs. 2, 3, 4).

Then, *Catoprion* is sister to *Pygopristis* and this two-genera clade is sister to the "true piranhas". *Pygocentrus*, hypothesized by several authors (Dahdul 2007; Freeman *et al.* 2007; Ortí *et al.* 2008) as non-monophyletic, had its monophyly supported with maximum statistic values for the three methods of analyses. The monophyly of *Serrasalmus* is not supported here because *Pristobrycon calmoni* and *Pristobrycon* sp. "Madeira" (*sensu* Ota *et al.* 2013), was recovered nested within *Serrasalmus*, also corroborating previous molecular hypothesis (Dahdul 2007; Freeman *et al.* 2007; Hubert *et al.* 2007; Thompson *et al.* 2014) (Figs. 2, 3, 4).

Figure 2. Phylogenetic relationships of Serrasalmidae based on a maximum likelihood analysis and inferred from a 75% complete supermatrix of ultraconserved elements. All nodes have bootstrap values above 95% except where indicated. Green branches correspond to "pacu" clade, blue branches to *Myleus* clade and red branches to "piranha" clade. Asterisks indicate type species.

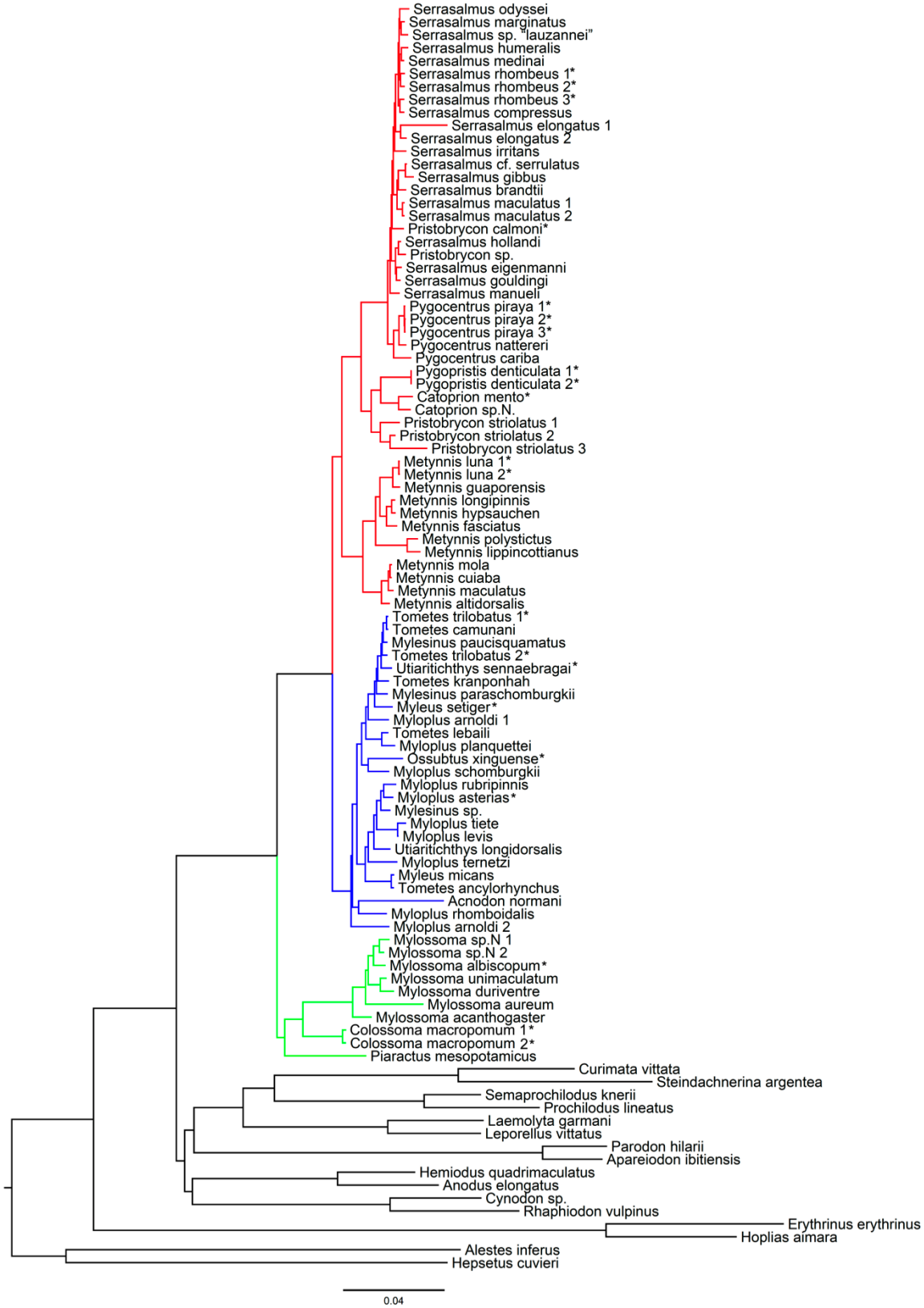

Figure 3. Bayesian phylogeny of Serrasalmidae based on a 75% complete supermatrix of ultraconserved elements. Posterior probabilities are equal or above 0.99 in all nodes. Green branches correspond to the "pacu" clade), blue branches to the *Myleus* clade and red branches to "piranha" clade. Asterisks indicate type species.

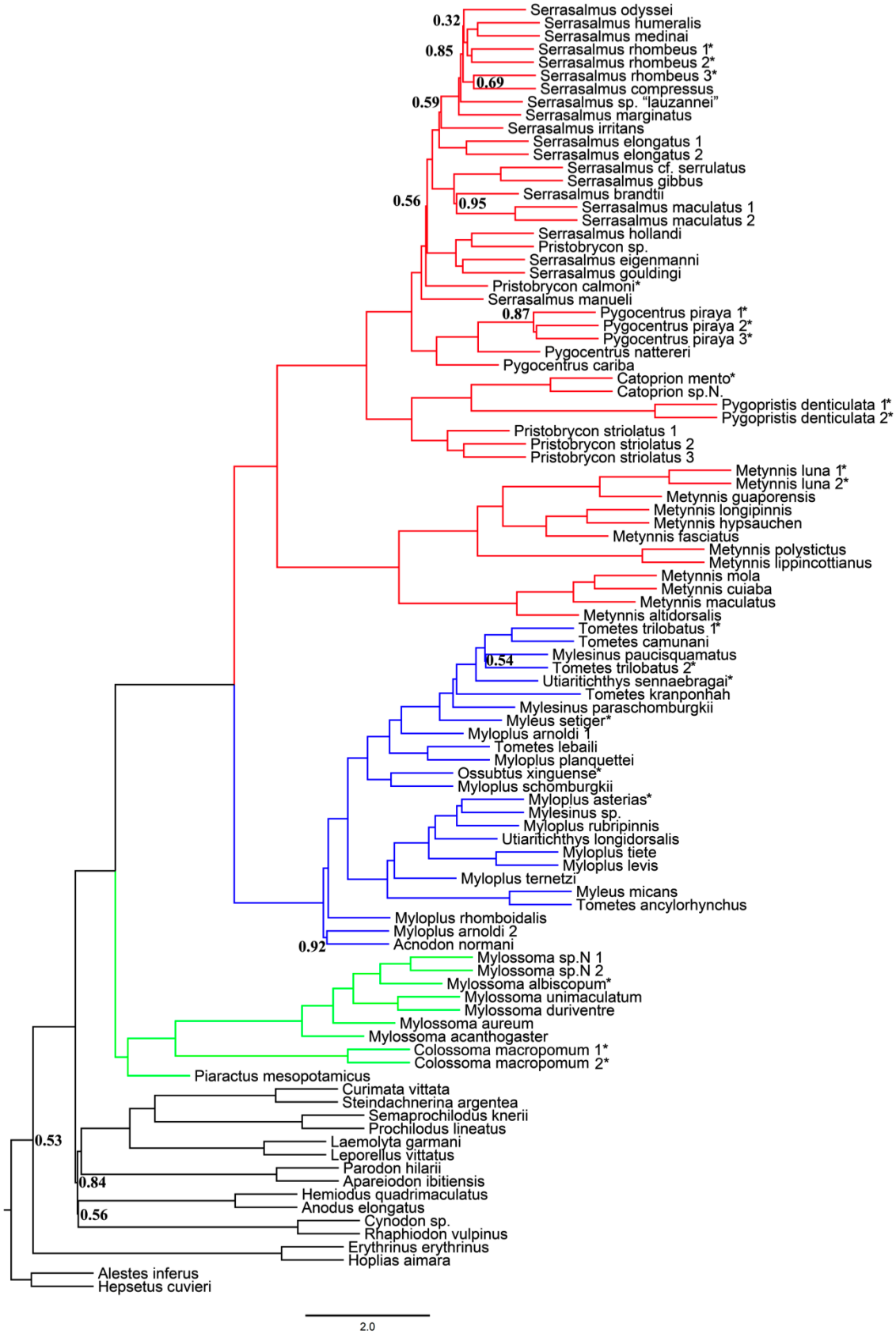

Figure 4. ASTRAL Species tree of Serrasalmidae based on a 75% complete supermatrix of ultraconserved elements. All nodes have posterior probabilities above 0.99 except where indicated. Green branches correspond to the "pacu" clade, blue branches to the *Myleus* clade and red branches to "piranha" clade. Asterisks indicate type species.

Discussion

The present phylogenomic analysis corroborate the well-stablished monophyly of Serrasalmidae (Ortí *et al.* 1996, 2008; Jégu 2004; Oliveira *et al.* 2011; Thompson *et al.* 2014; Arcila *et al.* 2017), using three different methods of phylogenetic inference (ML, BI and species three). The topology recovered herein is mostly congruent with previous morphological and molecular studies (Ortí *et al.* 1996, 2008; Cione *et al.* 2009; Thompson *et al.* 2014). The monophyly of the three main clades is a consensus among Serrasalmidae phylogenies, although intergeneric and interspecific relationships within each of those clades contain differences. The coincident evidence of monophyly of each of those three clades either in previous studies and here allow the recognition of two subfamilies and two tribes with the following arrangement:

- **Mylossominae** (nom. nov.): *Colossoma*, *Mylossoma*, and *Piaractus.*
- **Serrasalminae**:
    - **Myleini** (nom. nov.): *Acnodon*, *Mylesinus*, *Myleus*, *Myloplus*, *Ossubtus*, *Tometes*, and *Utiaritichthys.*
    - **Serrasalmini** (nom. nov.): *Catoprion*, †*Megapiranha*, *Metynnis*, *Pygocentrus*, *Pygopristis*, and *Serrasalmus* (including *Pristobrycon*).

Mylossominae (former "*pacu*" clade) is represented by three genera lacking a pre-dorsal spine. Uncertainties about topologies for the Mylossominae and its genera are recurrent. Jégu (2004) did not recover the "pacu" clade, but *Colossoma* and *Piaractus* as sister to all other Serrasalmidae based on eleven sinapomorphies (*e.g.* presence of dark maculas on flanks, five branchiostegal rays), while Machado-Allison (1982) and Cione *et al.* (2009) recovered *Mylossoma* sister to *Colossoma* and *Piaractus*. Herein, *Piaractus* appeared sister to *Colossoma* and *Mylossoma*, corroborating molecular data of Dahdul (2007), Ortí *et al*. (1996, 2008) and Thompson *et al.* (2014). This topology seems to be more congruent with the molecular hypothesis, since it corroborates all previous molecular studies (Dahdul 2007; Ortí *et al.* 1996, 2008; Thompson *et al.* 2014) while all preceding morphological hypothesis (Machado-Allison 1982; Jégu 2004; Dahdul 2007; Cione *et al.* 2009) recovers *Colossoma* as sister to *Piaractus* regardless their position in the phylogeny.

A taxonomic revision of the cis-Andean species of *Mylossoma* recognized four species (Mateussi *et al.* 2018), presented herein for the first time in a phylogeny. The trans-Andean species *M. acanthogaster*, endemic from the Lago Maracaibo, appeared sister to all cis-

Andean *Mylossoma* species in addition to an undescribed species (under description; our unpublished data) from the Orinoco basin. *Mylossoma aureum* appears as sister to the group *M. albiscopum* + *Mylossoma* sp.N and *M. duriventre* + *M. unimaculatum*, in partial agreement with a recent barcoding study of *Mylossoma* (Mateussi *et al.* 2017), in which *M. aureum* was recovered as sister to *Mylossoma* sp.N and this two-species group as sister to the remaining cis-Andean species.

The Myleini (former *Myleus* clade) differs from the first proposal as subfamily (Eigenmann 1915) – based on presence of two series of teeth in premaxillary – by the exclusion of *Metynnis* and *Catoprion*, that were recovered here (as previous hyphothesis) as more closely related to remaining *piranhas*. Myleini represents a complex group, perhaps with the most controversial and taxonomically problematic taxa in Serrasalmidae. Molecular studies have evidenced paraphyly for the genera *Mylesinus*, *Myloplus* and *Tometes* (Ortí *et al.* 1996; Thompson *et al.* 2014). Herein, none of them, in addition to *Myleus* and *Utiaritichthys*, appeared as monophyletic.

Among the genera of Myleini, *Acnodon* has been one with the most controversial positions in the phylogenies, either recovered alone and sister to Serrasalminae (Ortí *et al.* 1996, 2008; Cione *et al.* 2009), or sister to all other Myleini genera (Jégu 2004; molecular data of Dahdul 2007; Freeman *et al.* 2007; Thompson *et al.* 2014). Herein, *Acnodon* was sister to *Myloplus arnoldi* (90% BI; species tree) or *M. rhomboidalis* (75% BI; ML), both as basal taxa in Myleini. Only one species of the genus could be used for this study. Increasing the taxon sampling may improve the topology for this clade and clarify their interspecific relationships. Overall, it is evident that this tribe needs several taxonomic efforts to better understand each taxon before a deep analysis of their relationships.

All previous phylogenetic hypotheses for Serrasalmidae, both morphological and molecular, recovered *Metynnis* as the first clade to diversify within Serrasalmini (the "piranha" clade) sister to all piranha genera (*i.e. Catoprion*, *Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Serrasalmus*) (Cione *et al.* 2009; Dahdul 2007; Freeman *et al.* 2007; Jégu 2004; Machado-Allison 1983; Ortí *et al.* 1996, 2008; Thompson *et al.* 2014). Morphological sinapomorphies for this clade include reduction on number of infraorbital bones and reduction of a section the adductor mandibularis (Machado-Allison 1983; Jégu 2004). However, Cione *et al*. (2009) obtained a paraphyletic *Metynnis*, while our analyses support its monophyly, including herein, for the first time a higher number of representatives of the genus (11 of the 15 valid species).

The relationship between *Catoprion* and *Pygopristis* recovered here contradicts the hypotheses of *Catoprion* as sister to *Pygopristis*, and the later as sister to *Pygocentrus* and *Serrasalmus* (Ortí *et al.* 2008). However, it corroborates several previous studies (Dahdul 2007; Freeman *et al.* 2007; Hubert *et al.* 2007; Thompson *et al.* 2014), in which *Catoprion* is sister to *Pygopristis* and, almost always, they also appear as sister to *Pristobrycon striolatus* (Freeman *et al.* 2007; Hubert *et al.* 2007; Thompson *et al.* 2014).

Previous hypotheses proposed using mtDNA (Freeman *et al.* 2007; Hubert *et al.* 2007; Ortí *et al.* 1996, 2008) recuperated *Serrasalmus* and *Pygocentrus* as non-monophyletic genera. While Thompson *et al.* (2014), including nuclear genes, obtained the monophyly of *Pygocentrus* and of *Serrasalmus* + *Pristobrycon calmoni* with robust statistical support. This arrangement is corroborated herein with maximum values of bootstrap and posterior probabilities for all analyzes performed. Therefore, we suggest *Serrasalmus* is the senior synonym of *Pristobrycon* as betters discussed on a specific subtopic below.

On the last decade, the diversity of Serrasalmidae increased significantly, with the description of 12 species in this period (*e.g.* Andrade *et al.* 2016, 2018; Ota *et al.* 2016), some of them revealed by molecular techniques (*e.g.* Andrade *et al.* 2017). This number in only inferior to those observed for Characidae, Crenuchidae and Curimatidae within the 24 families of Characiformes (Fricke *et al.* 2019b). Machado *et al.* (2018) in a broad coverage DNA barcoding of Serrasalmidae comprising more than one thousand specimens recovered an extraordinary underestimated diversity within the family.

This hidden richness complicates the studies of phylogenetic relationships within the family, which are even more aggravated by morphological diagnosis for genera that are possibly not applicable (*e.g.* Jégu *et al.* 2004). These investigations are still impaired by the high level of variation in body shape and color pattern presented by most of the genera within Serrasalmidae (*e.g.* Myleini and *Metynnis*) during development and breeding period (Jégu 2003; Ota *et al.* 2016). Thus, studies including molecular tools are essential to recognize and delimit species within the family.

Besides the several taxonomic inconsistencies in Serrasalmidae shown by previous hypothesis, this study represents the most comprehensive molecular phylogeny of the family and uses a robust method to provide resolution of shallow nodes. However, the placement of some taxa remains uncertain. Therefore, the results presented herein provide a strong foundation for the urgent taxonomic revision of some groups, and highlights the urgent necessity of integrative studies using morphological and molecular tools to resolve generic problems within Myleini.

Taxonomy of Pristobrycon

Pristobrycon was described by Eigenmann (1915) to include five species with characters that he called "intermediate" between *Serrasalmus* Lacèpede, 1903 and *Rooseveltiella* Eigenmann, 1915 (*Pygocentrus nattereri*) (*e.g.* ectopterygoid toothless or with few poorly developed teeth, snout short and flattened anteriorly, cheecks covered by infraorbitals), designating *Serrasalmo* (*Pygocentrus*) *calmoni* Steindachner (1908) as type-species. The currently recognized species of *Pristobrycon* are *P. aureus*, *P. calmoni*, *P. careospinus*, *P. maculipinnis* and *P. striolatus*, distributed throughout Amazon, Orinoco and Guiana Shield rivers basins (Fricke *et al.* 2019a).

Nico *et al.* (2017) called *Pristobrycon* an "artificial genus" and divided in two groups based on presence or absence of an anal spine, discussing the disarray among different authors who treated some species of *Pristobrycon* as *Serrasalmus* and vice versa. All previous molecular hypotheses recovered *Pristobrycon calmoni* included within *Serrasalmus* while *P. striolatus* was recovered as sister to *Pygocentrus* and *Serrasalmus* (Dahdul 2007) or to *Catoprion* and *Pygopristis* (Freeman *et al.* 2007; Hubert *et al.* 2007; Ortí *et al.* 1996, 2008; Thompson *et al.* 2014) as herein.

We could perform a preliminary morphological analysis of *Pristobrycon* species and, in accordance to molecular and morphological data, both from previous authors and presented herein, we suggest that *Pristobrycon* must be allocated as junior synonym of *Serrasalmus* since its type species (*P. calmoni*) was found among *Serrasalmus* species. Our analysis also suggests that while species presenting pre-anal spine (*i.e. P. aureus* and *P. calmoni*) really belong to *Serrasalmus*, species lacking pre-anal spine (*i.e. P. careospinus*, *P. maculipinnis* and *P. striolatus*) clearly represents a different group that must be described as a new genus.

Bibliography

Aberer, A.J., Kobert, K. & Stamatakis, A. (2014) ExaBayes: Massively Parallel Bayesian Tree Inference for the Whole-Genome Era. *Molecular Biology and Evolution* 31, 2553–2556.

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 Biology* 11: R119.

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. *Nature Ecology & Evolution*, 2(4): 688–696.

Andrade, M.C., Jégu, M. & Gama, C.S. (2018) A new species of Myloplus Gill (Characiformes, Serrasalmidae) from the Tumucumaque mountain range, Brazil and French Guiana, with comments on M. rubripinnis. *Zootaxa* 4403, 111–122.

Andrade, M.C., Jégu, M. & Giarrizzo, T. (2016) Tometes kranponhah and Tometes ancylorhynchus (Characiformes: Serrasalmidae), two new phytophagous serrasalmids, and the first Tometes species described from the Brazilian Shield. *Journal of fish biology* 89, 467–494.

Andrade, M.C., Machado, V.N., Jégu, M., Farias, I.P. & Giarrizzo, T. (2017) A new species of Tometes Valenciennes 1850 (Characiformes: Serrasalmidae) from Tocantins-Araguaia River Basin based on integrative analysis of molecular and morphological data. *PLoS ONE* 12, 1–21.

Arcila, D., Ortí, G., Vari, R., Armbruster, J.W., Stiassny, M.L.J., Ko, K.D., Sabaj, M.H., Lundberg, J., Revell, L.J. & Betancur, R.R. (2017) Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nature Ecology and Evolution* 1, 1–10.

Barker, F.K. (2017) Molecular phylogenetics of the wrens and allies (Passeriformes: Certhioidea), with comments on the relationships of Ferminia. *Proceedings of the National Academy of Sciences*, 1–28.

Betancur, R.R., Arcila, D., Vari, R.P., Hughes, L.C., Oliveira, C., Sabaj, M.H. & Ortí, G. (2018) Phylogenomic incongruence, hypothesis testing, and taxonomic sampling: The mnophyly of characiform fishes. *Evolution*, 1–17.

Bruxaux, J., Gabrielli, M., Ashari, H., Prŷs-Jones, R., Joseph, L., Milá, B., Besnard, G. & Thébaud, C. (2018) Recovering the evolutionary history of crowned pigeons (Columbidae: Goura): Implications for the biogeography and conservation of New Guinean lowland birds. *Molecular Phylogenetics and Evolution* 120, 248–258.

Burress, E.D., Alda, F., Duarte, A., Loureiro, M., Armbruster, J.W. & Chakrabarty, P. (2018) Phylogenomics of pike cichlids (Cichlidae: *Crenicichla*): the rapid ecological speciation of an incipient species flock. *Journal of Evolutionary Biology* 31, 14–30.

Calcagnotto, D., Schaefer, S.A. & DeSalle, R. (2005) Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. *Molecular Phylogenetics and Evolution* 36, 135–153.

Cione, A.L., Dahdul, W.M., Lundberg, J.G. & Machado-Allison, A. (2009) *Megapiranha paranensis* , a new genus and species of Serrasalmidae (Characiformes, Teleostei) from the upper Miocene of Argentina. *Journal of Vertebrate Paleontology* 29, 350–358.

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, 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.

Dahdul, W.M. (2007) Ph.D. Phylogenetics and diversification of the Neotropical Serrasalminae (Ostariophysi: Characiformes). Faculties of the University of Pennsylvania. 107 pp.

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.

Eigenmann, C.H. (1915) The Serrasalminae and Myleinae. *Annals of the Carnegie Museum* 9,

226–272.

Faircloth, B.C. (2016) PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics* 32, 786–788.

Faircloth, B.C., Branstetter, M.G., White, N.D. & Brady, S.G. (2015) Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among hymenoptera. *Molecular Ecology Resources* 15, 489–501.

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(6): e65923.

Freeman, B., Nico, L.G., Osentoski, M., Jelks, H.L. & Collins, T.M. (2007) Molecular systematics of Serrasalmidae: Deciphering the identities of piranha species and unraveling their evolutionary histories. *Zootaxa* 38, 1–38.

Fricke, R., Eschmeyer, W. & van der Laan, R. (2019a) Catalog of fishes: genera, species, references. Available from: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (January 21, 2019).

Fricke, R., Eschmeyer, W.N. & Fong, J.D. (2019b) Species by Family/Subfamily. Available from: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp (January 15, 2019).

Géry, J. (1972) Poissons Characoïdes des Guyanes. I. Généralités. II. Famille des Serrasalmidae. *Zoologische Verhandelingen* 122, 1–250.

Géry, J. (1976) Les genres de Serrasalmidae (Pisces, Characoidei). *Bulletin Zoölogisch Museum, Universiteit Van Amsterdam* 5, 47–54.

Géry, J., Mahnert, V. & Dlouhy, C. (1987) Poissons Characoïdes non Characidae du Paraguay

(Pisces, Ostariophysi). *Revue Suisse Zoologie* 94, 357–464.

Gosline, W.A. (1951) Notes on the characid fishes of the subfamily Serrasalminae. *Proceedings of the California Academy of Sciences* 27, 17–64.

Goulding, M. (1980) *The fishes and the forest : explorations in Amazonian natural history*. University of California Press, Berkeley and Los Angeles, California.

Harris, H. (2007) *Improved pairwise alignment of genomic DNA*. The Pennsylvania State University.

Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L. & Kimball, R.T. (2016) Avoiding missing data biases in phylogenomic inference: An empirical study in the landfowl (Aves: Galliformes). *Molecular Biology and Evolution* 33, 1110–1125.

Hubert, N., Duponchelle, F., Nuñez, J., Garcia-Davila, C., 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, 2115–2136.

Jégu, M. (2003) Subfamily Serrasalminae. In: R. E. Reis, S. O. Kullander, and C. J. Ferraris (Eds), *Check List of the Freshwater Fishes of South and Central America*. EDIPUCRS, Porto Alegre, pp. 185–199.

Jégu, M. (2004) Taxonomie des Serrasalminae phytophages et Phylogenie des Serrasalminae (Teleostei: Characiformes: Characidae). Museum National d'Histoire Naturelle. 421 pp.

Jégu, M., Hubert, N. & Belmont-Jegu, E. (2004) Réhabilitation de Myloplus asterias (Müller & Troschel, 1844), espèce-type de Myloplus Gill, 1896 et validation du genre Myloplus Gill (Characidae : Serrasalminae). *Cybium* 28, 119–157.

Jégu, M. & dos Santos, G.M. (1988) Une nouvelle espèce du genre Mylesinus (Pisces, Serrasalmidae), M. paucisquamatus, décrite du bassin du Rio Tocantins (Amazonie, Brésil). *Cybium* 12, 331–341.

Jégu, M., Santos, G.M., Keith, P. & Le Bail, P.-Y. (2002) Description complèmentaire et rèhabilitation de *Tometes trilobatus* Valenciennes, 1850, espèce-type de *Tometes* Valenciennes (Characidae: Serrasalminae). *Cybium* 26, 99–122.

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, 1–12.

Mateussi, N.T.B., Oliveira, C. & Pavanelli, C.S. (2018) Taxonomic revision of the Cis-Andean species of Mylossoma Eigenmann & Kennedy, 1903 (Teleostei: Characiformes: Serrasalmidae). *Zootaxa* 4387, 275–309.

Mateussi, N.T.B., Pavanelli, C.S. & Oliveira, C. (2017) Molecular identification of cryptic diversity in species of cis-Andean Mylossoma (Characiformes: Serrasalmidae). *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis* 28, 778–780.

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, 746–754.

Merckx, A., Jégu, M. & Santos, G.M. (2000) Une nouvelle espece de Serrasalmus (Teleostei: Characidae: Serrasalminae), S. altispinis n. sp., decrite du rio Uatumã (Amazonas, Bresil) avec une description complementaire de S. rhombeus (Linnaeus, 1766) du Plateau Guyanais. *Cybium* 24, 181–201.

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 taxonymy. *Neotropical Ichthyology* 8(3), 385–568.

Mirande, J.M. (2018) Morphology, molecules and the phylogeny of Characidae (Teleostei, Characiformes). *Cladistics,* 1–19.

Nelson, E.M. (1961) The swin bladder in the Serrasalminae: with notes on additional morphological features. *Fieldiana: Zoology* 39, 603–624.

Nico, L.G., Jégu, M. & Andrade, M.C. (2017) Family Serrasalmidae - Piranhas and Pacus. In: P. Van der Sleen and J. Albert (Eds), *Field guide to the fishes of the Amazon, Orinoco, and Guianas*. D&N Publishing, Wiltshire, UK, pp. 172–196.

Nico, L.G. & Taphorn, D.C. (1998) Food Habits of Piranhas in the Low Llanos of Venezuela. *Biotropica* 20, 311–321.

Norman, J.R. (1929) The South American Characid fishes of the subfamily Serrasalmoninae with a revision of the genus Serrasalmus Lacepede. *Proceedings of the Zoological Society of London* 52, 781–829.

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 Evolutionary Biology* 11, 275.

Ortí, G., Petry, P., Porto, J.I.R., JEGU, M. & Meyer, A. (1996) Patterns of nucleotide change in Mitochondrial Ribosomal RNA genes and the Phylogeny of Piranhas. *Journal of Molecular Evolution* 42, 169–182.

Ortí, G., Sivasundar, A., Dietz, K. & Jégu, M. (2008) Phylogeny of the Serrasalmidae (Characiformes) based on mitochondrial DNA sequences. *Genetics and Molecular Biology* 31, 343–351.

Ota, R.P., Py-Daniel, L.H.R. & Jégu, M. (2016) A new silver dollar species of *Metynnis* Cope, 1878 (Characiformes: Serrasalmidae) from Northwestern Brazil and Southern Venezuela. *Neotropical Ichthyology* 14(4): e160023.

Ota, R.P., Röpke, C.P., Zuanon, J. & Jégu, M. (2013) Serrasalmidae. In: L. J. Queiroz, G. Torrente-Vilara, W. M. Ohara, T. H. da S. Pires, J. Zuanon, and C. R. da C. Doria (Eds), *Peixes do rio Madeira v.II*. Santo Antônio Energia, São Paulo - Brasil, pp. 14–48.

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R.P., Moret, B.M.E. & Stamatakis, A. (2009) How many bootstrap replicates are necessary? *Lecture Notes in Computer Science* 5541, 184–200.

Rambaut, A. & Drummond, A.J. (2009) Tracer V1.5. *Available from http://beast.bio.ed.ac.uk/Tracer*.

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

Sazima, I. & Machado, F.A. (1990) Underwater observations of piranhas in western Brazil. *Environmental Biology of Fishes* 28, 17–31.

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M. & Birol, I. (2009) ABySS: A parallel assembler for short read sequence data. *Genome Research* 19, 1117–1123.

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.

Streicher, J.W., Miller, E.C., Guerrero, P.C., Correa, C., Ortiz, J.C., Crawford, A.J., Pie, M.R. & Wiens, J.J. (2018) Evaluating methods for phylogenomic analyses, and a new phylogeny for a major frog clade (Hyloidea) based on 2214 loci. *Molecular Phylogenetics and Evolution* 119, 128–143.

Streicher, J.W., Schulte, J.A. & Wiens, J.J. (2016) How Should Genes and Taxa be Sampled for Phylogenomic Analyses with Missing Data? An Empirical Study in Iguanian Lizards. *Systematic Biology* 65, 128–145.

Tagliacollo, V.A. & Lanfear, R. (2018) Estimating Improved Partitioning Schemes for Ultraconserved Elements J. Townsend (Ed). *Molecular Biology and Evolution* 35, 1798–1811.

Thompson, A.W., Betancur, R., López-Fernández, H. & Ortí, G. (2014) A time-calibrated, multi-locus phylogeny of piranhas and pacus (Characiformes: Serrasalmidae) and a comparison of species tree methods. *Molecular Phylogenetics and Evolution* 81, 242–257.

Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. (2018a) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19, 153.

Zhang, S.-Q., Che, L.-H., Li, Y., Dan Liang, Pang, H., Ślipiński, A. & Zhang, P. (2018b) Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nature Communications* 9, 205.

Species	Voucher	Specimen number	Number of trimmed reads	Contigs assembled	Total bp contigs	UCE contigs	Total bp UCES	Mean length
Acnodon normani	LBP 19140	77187	10,568,593	208,675	43,663,765	1,128	1,016,226	900
Catoprion mento	LBP 7556	35626	2,966,961	153,117	21,475,685	1,827	1,268,628	694
Catoprion sp, N,	LBP 21615	63665	4,445,599	118,815	25,706,807	1,743	1,737,540	996
Colossoma macropomum 1	LBP 12837	54043	7,140,735	525,678	75,592,404	1,878	1,141,274	607
Colossoma macropomum 2	LBP 12838	54052	2,066,137	109,180	16,196,947	1,850	1,383,476	747
Metynnis altidorsalis	LBP 9997	43151	2,151,079	74,044	14,544,577	1,925	1,816,404	943
Metynnis cuiaba	LBP 8580	43371	4,829,794	88,228	18,388,805	1,963	1,644,883	837
Metynnis fasciatus	LBP 3978	22824	1,881,346	58,158	10,928,239	1,925	1,472,655	765
Metynnis guaporensis	UFRO-ICT 5396	6018	2,360,357	72,953	13,990,506	1,919	1,629,922	849
Metynnis hypsauchen	LBP 18390	43152	2,961,284	64,836	12,695,349	1,942	1,381,497	711
Metynnis lippincottianus	LBP 5429	27164	2,267,401	64,828	12,012,224	1,942	1,429,820	736
Metynnis longipinnis	LBP 15530	63920	921,074	31,431	6,277,012	1,846	1,354,087	733
Metynnis luna 1	LBP 18398	42589	2,101,237	52,078	10,792,072	1,940	1,543,426	795
Metynnis luna 2	LBP 9394	43863	3,136,520	158,597	23,365,290	1,908	1,474,444	772
Metynnis maculatus	UFRO-ICT 5395	3986	4,109,535	88,304	17,164,808	1,609	1,177,152	731
Metynnis mola	LBP 8447	42354	2,098,273	74,350	14,322,530	1,830	1,640,403	896
Metynnis polystictus	LBP 8015	37703	1,224,456	35,706	7,586,034	1,917	1,500,122	782
Mylesinus paraschomburgkii	LBP 20484	80923	6,034,639	486,635	80,228,481	1,799	1,638,325	910
Mylesinus paucisquamatus	LBP 12839	53447	2,656,252	67,828	13,320,651	1,839	1,429,460	777
Mylesinus sp,	LBP 6853	32707	2,608,470	86,424	15,942,404	1,765	1,424,273	806
Myleus micans	LBP 24011	91323	2,710,054	69,639	14,194,321	1,813	1,672,127	922

Myleus setiger	LBP 16523	61481	4,252,597	96,010	17,787,400	1,748	1,225,596	701
Myloplus arnoldi 1	LBP 20266	79748	1,635,379	48,552	9,180,872	1,876	1,388,188	739
Myloplus arnoldi 2	LBP 13827	57247	2,477,940	129,855	18,202,527	1,775	1,091,571	614
Myloplus asterias	LBP 9072	42585	1,541,117	61,796	10,641,499	1,579	1,060,607	671
Myloplus levis	ZUFMS-PIS 5567	1	10,538,217	39,099	9,090,093	1,879	1,029,721	548
Myloplus planquettei	ANSP 179808	230	15,017,639	295,155	58,678,238	977	702,684	719
Myloplus rhomboidalis	ANSP 185981	10924	10,791,612	210,958	43,227,525	1,428	1,254,501	878
Myloplus rubripinnis	UFRO-ICT	5782	7,896,208	35,747	7,867,357	1,796	889,779	495
Myloplus schomburgkii	OS 18990	PE10-044	6,093,498	193,904	27,975,924	1,753	956,245	545
Myloplus ternetzi	ANSP 188689	1626	10,596,643	175,906	37,027,139	1,218	836,962	687
Myloplus tiete	INPA 53243	53243	12,093,705	474,046	85,008,802	1,084	730,798	674
Mylossoma acanthogaster	LBP 24311	91508	1,606,554	52,022	10,591,912	1,914	1,853,319	968
Mylossoma albiscopum	LBP 12089	51722	1,546,997	91,408	13,300,718	1,734	1,173,732	676
Mylossoma aureum	UFRO-ICT 19346	10260	5,945,081	169,307	33,260,735	1,268	960,925	757
Mylossoma duriventre	LBP 3741	21919	7,695,851	278,246	52,062,658	1,259	1,021,790	811
Mylossoma unimaculatum	LBP 12745	41182	4,734,531	417,797	67,804,936	1,779	1,653,286	929
Mylossoma sp.N 1	LBP 18873	75056	6,394,546	185,031	36,464,028	1,433	1,222,720	853
Mylossoma sp.N 2	LBP 2190	15518	1,597,330	88,037	13,287,269	1,809	1,325,913	732
Ossubtus xinguense	LBP 16563	64065	2,335,276	101,511	15,120,720	1,742	1,305,476	749
Piaractus mesopotamicus	LBP 4255	23804	9,515,921	347,544	63,880,056	1,173	967,064	824
Pristobrycon	LBP 2191	15554	1,792,409	88,765	13,570,795	1,812	1,425,219	786

calmoni								
Pristobrycon striolatus 1	ANSP 188672	1631	6,475,798	181,530	35,669,461	1,728	1,437,006	831
Pristobrycon striolatus 2	ANSP 197515	10494	6,120,921	472,425	78,247,439	1,872	1,881,689	100
Pristobrycon striolatus 3	LBP 15050	61685	2,971,355	110,925	17,609,682	1,269	471,048	371
Pristobrycon sp,	UFRO-ICT 7235	6010	7,795,913	154,773	32,487,545	1,429	1,168,164	817
Pygocentrus cariba	LBP 10225	43107	1,827,536	59,122	11,013,018	1,639	1,125,015	686
Pygocentrus nattereri	LBP 12693	43552	1,668,421	51,999	10,496,743	1,909	1,686,870	883
Pygocentrus piraya 1	LBP 11336	45523	938,112	45,999	7,319,295	1,664	1,145,682	688
Pygocentrus piraya 2	LBP 11286	48749	1,945,483	114,327	16,257,319	1,711	1,142,438	667
Pygocentrus piraya 3	LBP 11286	48750	1,034,372	36,722	7,558,820	1,826	1,561,485	855
Pygopristis denticulata 1	LBP 21609	62390	1,740,295	57,080	11,086,268	1,893	1,519,125	802
Pygopristis denticulata 2	LBP 15529	63916	4,086,705	113,971	21,684,830	1,886	1,639,505	869
Serrasalmus brandtii	LBP 11269	48707	2,681,474	66,414	14,006,853	1,942	1,735,532	893
Serrasalmus compressus	UFRO-ICT 24705	2546	4,711,494	124,056	21,389,721	1,840	1,209,670	657
Serrasalmus eigenmanni	LBP 16210	66983	2,582,392	59,902	11,902,219	1,966	1,423,631	724
Serrasalmus elongatus 1	UFRO-ICT 1807	125	1,823,441	61,963	10,882,828	1,028	467,165	454
Serrasalmus elongatus 2	LBP 18025	72569	1,822,352	46,806	9,608,719	1,920	1,455,661	758
Serrasalmus gibbus	LBP 12811	41289	10,370,766	397,865	72,489,685	1,201	904,323	752
Serrasalmus gouldingi	LBP 20976	81155	3,211,776	94,961	18,360,838	1,953	1,813,005	928
Serrasalmus hollandi	UFRO-ICT 5631	6805	1,396,935	48,840	8,540,334	1,644	990,460	602
Serrasalmus humeralis	LBP 4221	22719	3,433,294	93,165	17,148,280	1,528	955,315	625

Serrasalmus irritans	LBP 10229	43109	2,957,388	78,582	15,774,569	1,792	1,408,722	786
Serrasalmus maculatus 1	LBP 3821	22107	3,196,815	68,890	13,961,502	1,945	1,558,899	801
Serrasalmus maculatus 2	LBP 3821	22108	5,193,350	112,557	23,416,251	1,923	1,810,178	941
Serrasalmus manueli	LBP 15032	61627	1,576,615	49,137	10,307,206	1,760	1,528,394	868
Serrasalmus marginatus	LBP 19855	79019	5,166,986	27,037	5,983,543	1,889	902,379	477
Serrasalmus medinai	LBP 22502	86574	3,780,242	81,716	17,546,119	1,964	1,779,379	905
Serrasalmus odyssei	UFRO-ICT 20212	10000	3,583,710	277,059	45,526,081	1,330	1,042,832	784
Serrasalmus rhombeus 1	LBP 15365	62022	1,014,752	34,671	7,713,053	1,813	1,676,292	924
Serrasalmus rhombeus 2	LBP 15314	63320	1,493,239	56,258	11,229,459	1,871	1,757,137	939
Serrasalmus rhombeus 3	LBP 14239	59423	1,662,084	42,572	8,329,703	1,942	1,258,031	647
Serrasalmus cf, *serrulatus*	LBP 21606	61612	2,293,430	68,224	14,209,119	1,743	1,560,347	895
Serrasalmus sp, "lauzannei"	UFRO-ICT 17954	8091	11,246,981	241,943	49,698,841	1,394	1,300,492	932
Tometes ancylorhynchus	LBP 17403	69141	6,979,336	250,303	46,467,560	1,578	1,363,671	864
Tometes camunani	LBP 20460	80790	3,797,275	221,307	30,964,450	1,849	1,203,603	650
Tometes kranponhah	ANSP 200868	13586	4,833,475	373,477	61,170,123	1,922	1,925,178	100
Tometes lebaili	ANSP 188681	1634	13,029,055	201,345	40,526,570	1,325	932,247	703
Tometes trilobatus 1	LBP 5196	26788	4,738,987	106,285	20,073,131	1,955	1,612,689	824
Tometes trilobatus 2	LBP 21012	81965	2,583,873	80,247	16,001,827	1,878	1,727,653	919
Utiaritichthys longidorsalis	UFRO-ICT 19099	9141	1,375,159	46,172	8,780,215	1,861	1,429,038	767
Utiaritichthys sennaebragai	MCP 44061	-	15,941,850	356,618	68,136,307	1,124	947,878	843

Alestes inferus	AMNH	333238	1,847,106	54,586	10,841,727	1,770	1,301,771	735
Anodus elongatus	LBP 4244	22720	3,116,484	67,158	10,949,292	1,588	790,539	497
Apareiodon ibitiensis	LBP 2890	18635	3,990,904	109,257	14,573,850	1,592	556,059	349
Curimata vittata	LBP 13846	57302	5,740,998	153,249	21,163,438	1,721	811,274	471
Cynodon sp,	LBP 10227	43105	1,365,513	44,264	9,879,399	1,909	2,074,001	108
Erythrinus erythrinus	LBP 6625	31955	3,639,480	120,233	22,493,182	1,879	1,556,540	828
Hemiodus quadrimaculatus	LBP 21151	82973	1,631,554	51,307	11,070,851	1,935	1,708,269	882
Hepsetus cuvieri		353404	6,075,880	16,592	3,343,078	1,907	647,872	339
Hoplias aimara	LBP 20412	80691	1,853,591	64,451	12,795,021	1,805	1,668,022	924
Laemolyta garmani	OS 18777	PE10-089	4,261,677	98,031	14,314,375	1,613	781,340	484
Leporellus vittatus	ANSP 182609	P6322	3,928,683	89,881	13,114,650	1,622	703,794	433
Parodon hilarii	LBP 10408	48929	2,148,993	58,036	12,226,843	1,926	1,836,328	953
Prochilodus lineatus	Uncatalog ued	1	2,012,868	94,573	15,125,729	1,783	1,035,757	580
Rhaphiodon vulpinus	LBP 12660	43557	2,299,180	67,201	14,070,460	1,864	2,049,069	109
Semaprochilodus knerii	ANSP 187277	P4298	5,396,483	151,165	21,357,715	1,800	782,454	434
Steindachnerina argentea	STRI-4270	BFD01625	4,565,598	121,995	17,021,405	1,834	977,664	533
		Total	416,297,276	13,441,549	2,405,321,817	171,397	129,707,619	71,663
		Average	4,247,931	133,085	23,815,067	1,697	1,284,234	710

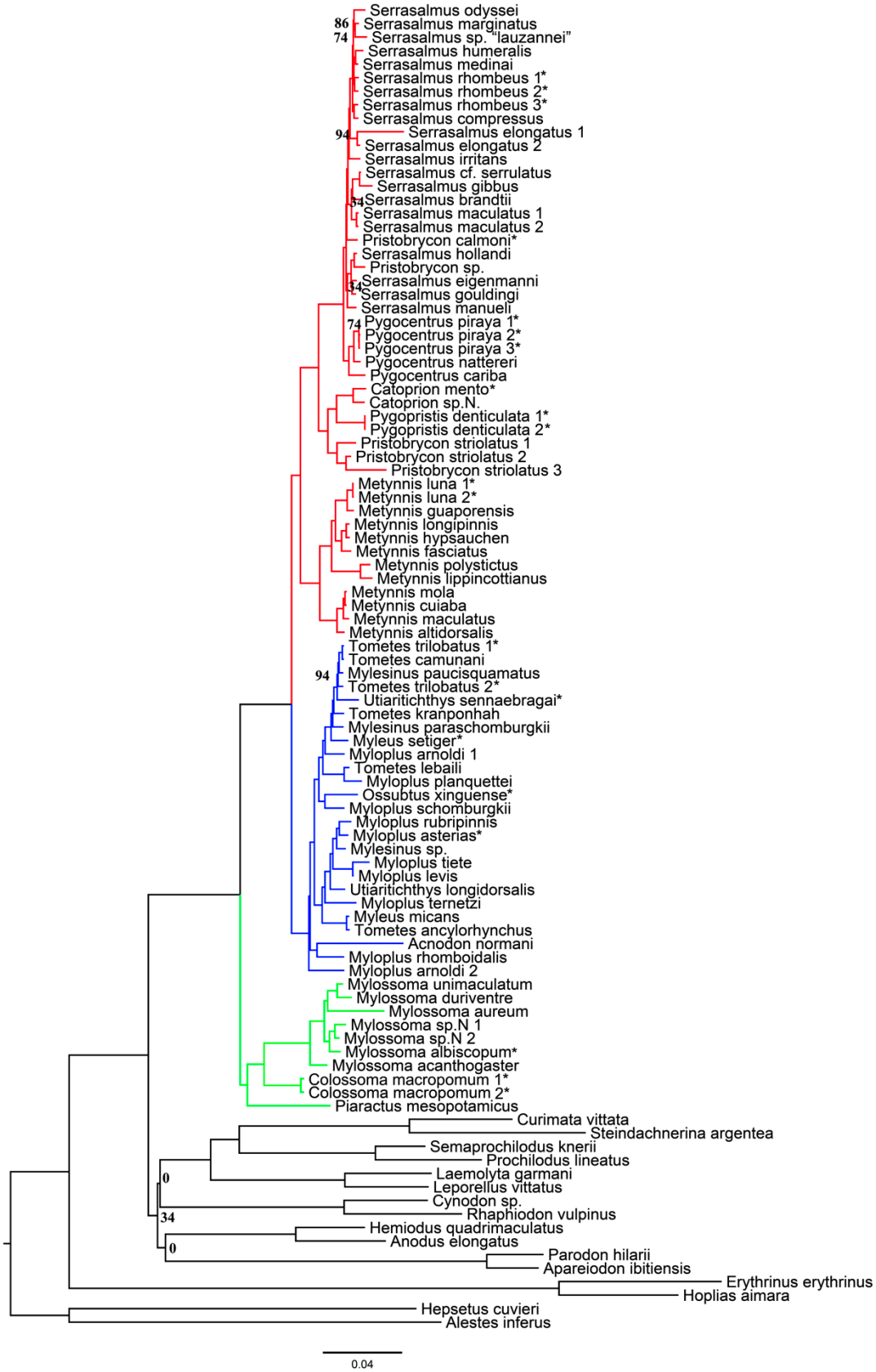

Figure S1. Phylogenetic relationships of Serrasalmidae based on a maximum likelihood analysis and inferred from a 50% complete super matrix of ultraconserved elements. All nodes have bootstrap values above 95% except where indicated. Green branches corresponds to the Mylossominae, blue branches to the Myleinae and red branches to Serrasalminae. Asterisks indicate type species.

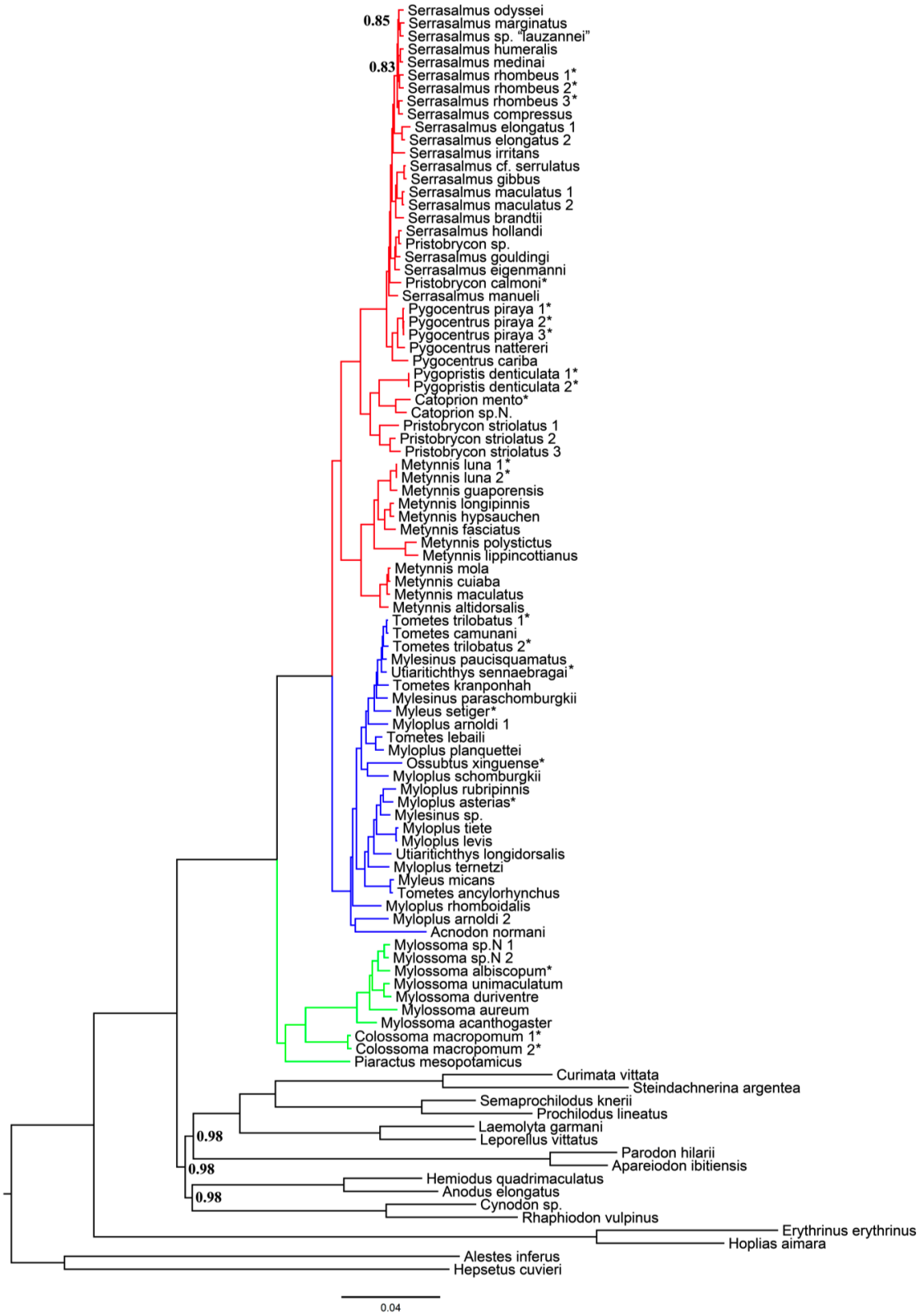

Figure S.2. Bayesian phylogeny of Serrasalmidae based on a 90% complete supermatrix of ultraconserved elements. Posterior probabilities lower than 0.99 are indicated in nodes. Green branches corresponds to the Mylossominae, blue branches to the Myleinae and red branches to Serrasalminae. Asterisks indicate type species.

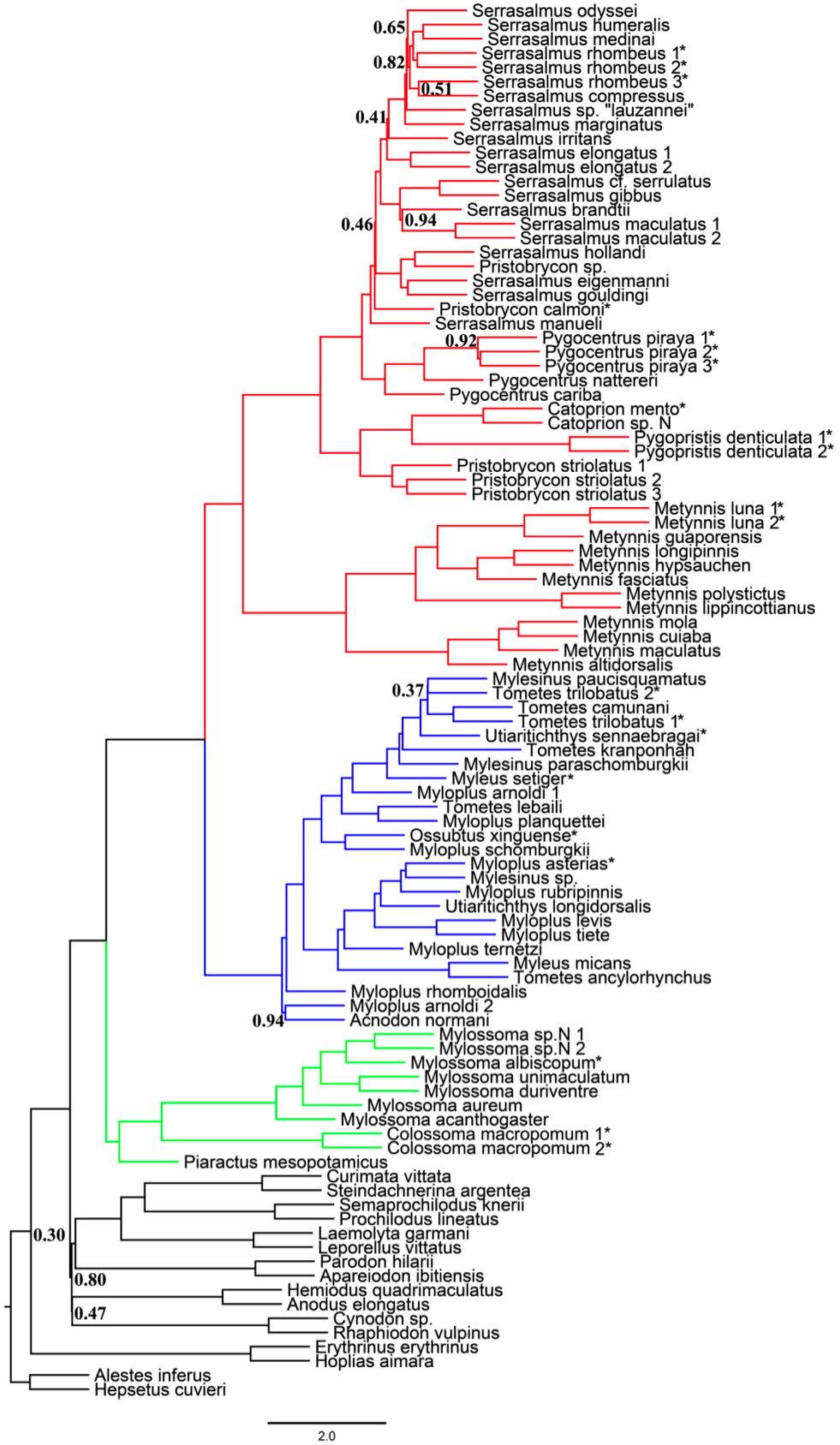

Figure S.3. Species tree of Serrasalmidae based on a 50% complete supermatrix of Ultraconserved elementes. Posterior probabilities lower than 0.99 are indicated in nodes. Green branches corresponds to the Mylossominae, blue branches to the Myleinae and red branches to Serrasalminae. Asterisks indicate type species.

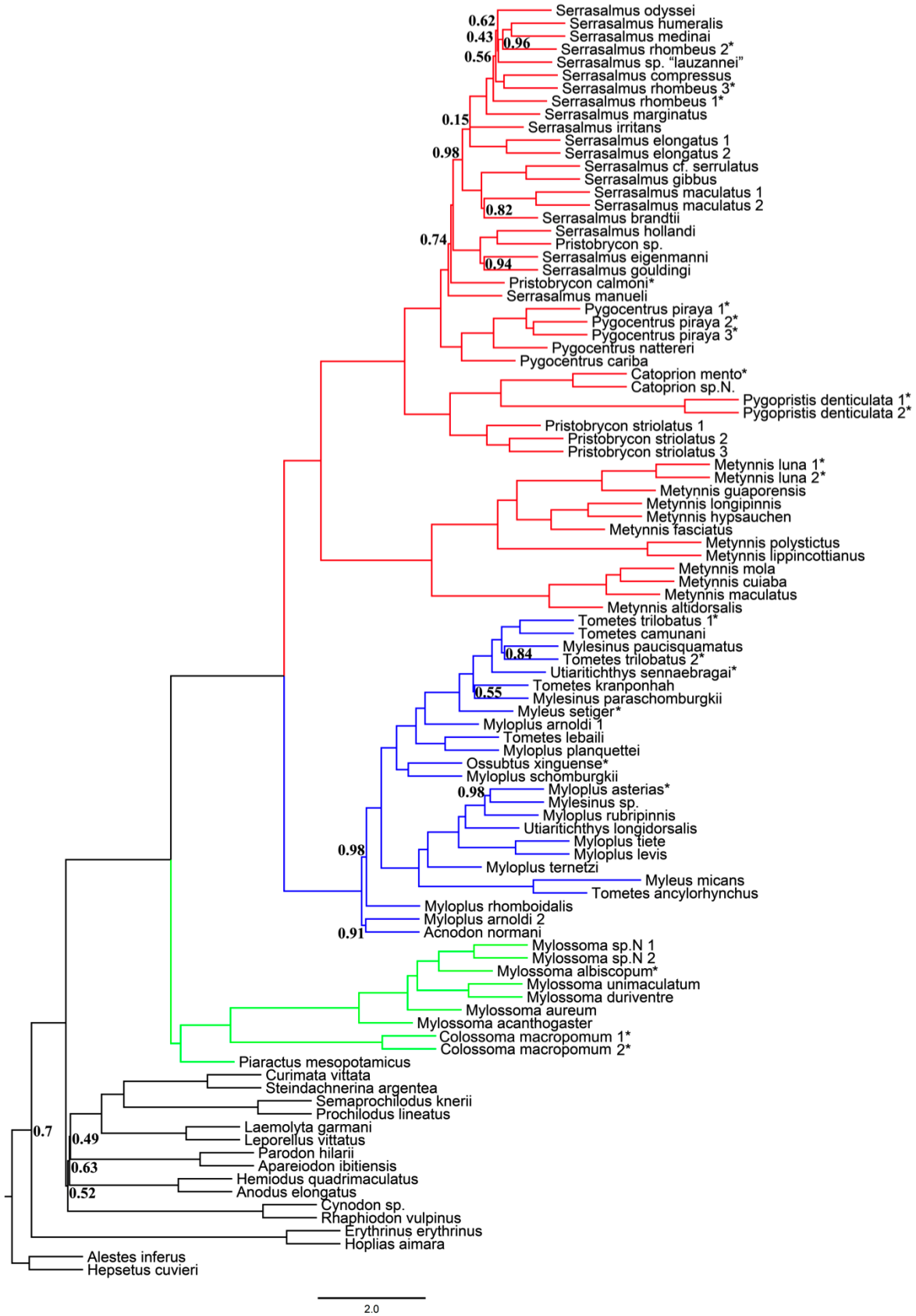

Figure S.4. Species tree of Serrasalmidae based on a 90% complete supermatrix of Ultraconserved elementes. Posterior probabilities lower than 0.99 are indicated in nodes. Green branches corresponds to the Mylossominae, blue branches to the Myleinae and red branches to Serrasalminae. Asterisks indicate type species.

Chapter 2

Species delimitation in piranhas of the genus *Pygocentrus* and intraspecific genetic variation in *P. nattereri* (Teleostei, Characiformes)

Species delimitation in piranhas of the genus *Pygocentrus* and intraspecific genetic variation in *P. nattereri* (Teleostei, Characiformes)

Abstract

Pygocentrus is a "piranha" genus of Serrasalmidae, commonly found in major basins of South America with three valid species: *P. cariba* from the Orinoco basin, *P. piraya* from the São Francisco basin, and *P. nattereri*, widely distributed in the Amazon, Essequibo, Paraná-Paraguay, and coastal rivers of northeastern Brazil. Although much appreciated in the aquarium trade and in regional fisheries, few studies have focused on the genetic diversity and systematics of the genus and doubts concerning the validity and species boundaries remain to be explored. DNA barcoding and species delimitation methods were applied to delimit species of *Pygocentrus*, test the phylogeographic structure of *P. nattereri*, and to access the origin of introduced specimens of *P. nattereri* in Asia. The final matrix included 108 terminals of the three species from a large geographic range in South America. Species delimitation analyses involved a maximum-likelihood tree reconstruction, genetic distances, PTP, BIN, ABGD, and GMYC. The maximum likelihood, genetic distance analyses, PTP, and BIN support a previous morphological hypothesis in recognizing the three species of *Pygocentrus*, whereas ABGD and GMYC recognized multiple species in *P. piraya* and *P. nattereri*. Overall, the morphology-defined species of *Pygocentrus* are corroborated with additional evidence that *P. nattereri* contain at least four structured populations along the continental distribution. Results also indicate that introduced specimens in Asia descend from the lineage of *P. nattereri* from the main rio Amazonas.

Key words: Biodiversity, DNA barcode; Neotropical region; systematics

Introduction

The Neotropical fish family Serrasalmidae contains 16 genera and 97 species (Fricke *et al.* 2019) of ecomorphologically diverse freshwater fishes commonly known as "pacus" and "piranhas". The the four genera of carnivorous "piranhas" are *Pristobrycon* Eigenmann, 1915, *Pygocentrus* Müller & Troschel, 1844, *Pygopristis* Müller & Troschel, 1844, and *Serrasalmus* Lacepède, 1803, which are closely related to the lepidophagous genus *Catoprion* Müller & Troschel, 1844 (Jégu 2003). The monophyly of this five-genera group is supported on the basis of both morphological (Machado-Allison 1982; Cione *et al.* 2009) and multilocus molecular data (Thompson *et al.* 2014).

The genus *Pygocentrus* includes the largest species of "piranhas" that can reach up 50 cm SL (Britski *et al.* 2007) that are highly appreciated in the ornamental trade and with a reduced but still representative economic importance in regional fisheries and aquaculture (Santos *et al.* 2006; Lin *et al.* 2015). *Pygocentrus* can be morphologically distinguished from other Serrasalmidae genera by a substantially wider head, dorsal profile moderately to strongly convex, presence of a preanal spine (sometimes undetectable externally), tricuspid teeth and lack of ectopterygoid teeth, except in small juveniles (Fink 1993; Nico *et al.* 2017). The main synapomorphy of the genus is the presence of crests around the lateral-sensorial system of the frontal, parietal and pterotic bones (Machado-Allison & Fink 1991). The genus is monophyletic and hypothesized to be the sister of *Serrasalmus* plus *Pristobrycon calmoni* (Hubert *et al.* 2007; Thompson *et al.* 2014; see Chapter 1).

The taxonomic revision of *Pygocentrus* (Fink 1993) validated three species: *P. cariba* (Humboldt, 1821), endemic from the Río Orinoco basin; *P. nattereri* Kner, 1858, widely-distributed in the Amazon, Essequibo, Paraná-Paraguay, and coastal rivers of northeastern Brazil; and *P. piraya* (Cuvier, 1819), the type-species of the genus, endemic from the rio São Francisco basin (Fig. 1). This three-species hypothesis was only recently tested in a broader barcoding study of the entire family Serrasalmidae that recognized both *P. cariba* and *P. piraya* but presented variable numbers of entities within *P. nattereri* depending on the delimitation analyses (Machado *et al.* 2018). These analytical inconsistencies and the limited taxon sampling from relatively few Amazonian regions imply the necessity of an intrageneric analysis to refine the species delimitation analyses including samples from multiple South American regions.

The variable number of species delimited for the red-bellied piranha *Pygocentrus nattereri* (Machado *et al.* 2018) derives probably from the genetic structure of lineages from distinct river systems. For example, the phylogeographic study of *Pygocentrus* using the mtDNA control region found *P. nattereri* with structured lineages, in which the Paraná, Ucayali, and Madeira lineages appeared genetically closer to each other than to the mainstream Amazonas lineage (Hubert *et al.* 2007). Intraspecific population genetic studies in *P. nattereri* from the northeastern Brazil (Luz *et al.* 2015) and along the Solimões-Amazonas river (Santos *et al.* 2016) have also shown high levels of genetic diversity and significant genetic differentiation among populations. In addition, *P. nattereri* has been introduced in Asia, more specifically in Bangladesh (Rahman & Ahmed 2007), China (Lin *et al.* 2015) and Philippines (Guerrero 2014; Sarmiento *et al.* 2016), but the reports did not determine the precise geographic origin of those lineages in South America.

Here, we used partial sequences of the mitochondrial gene cytochrome oxidase c subunit I (COI) and modern species delimitation methods in order to (1) test the morphological hypothesis supporting the presence of three species of *Pygocentrus* (Fink 1993), (2) test the population genetic hypothesis that *P. nattereri* contains multiple genetically-structured populations (Hubert *et al.* 2007; Luz *et al.* 2015; Santos *et al.* 2016), and (3) determine the original river system of recently introduced specimens in Asia (Sarmiento *et al.* 2016).

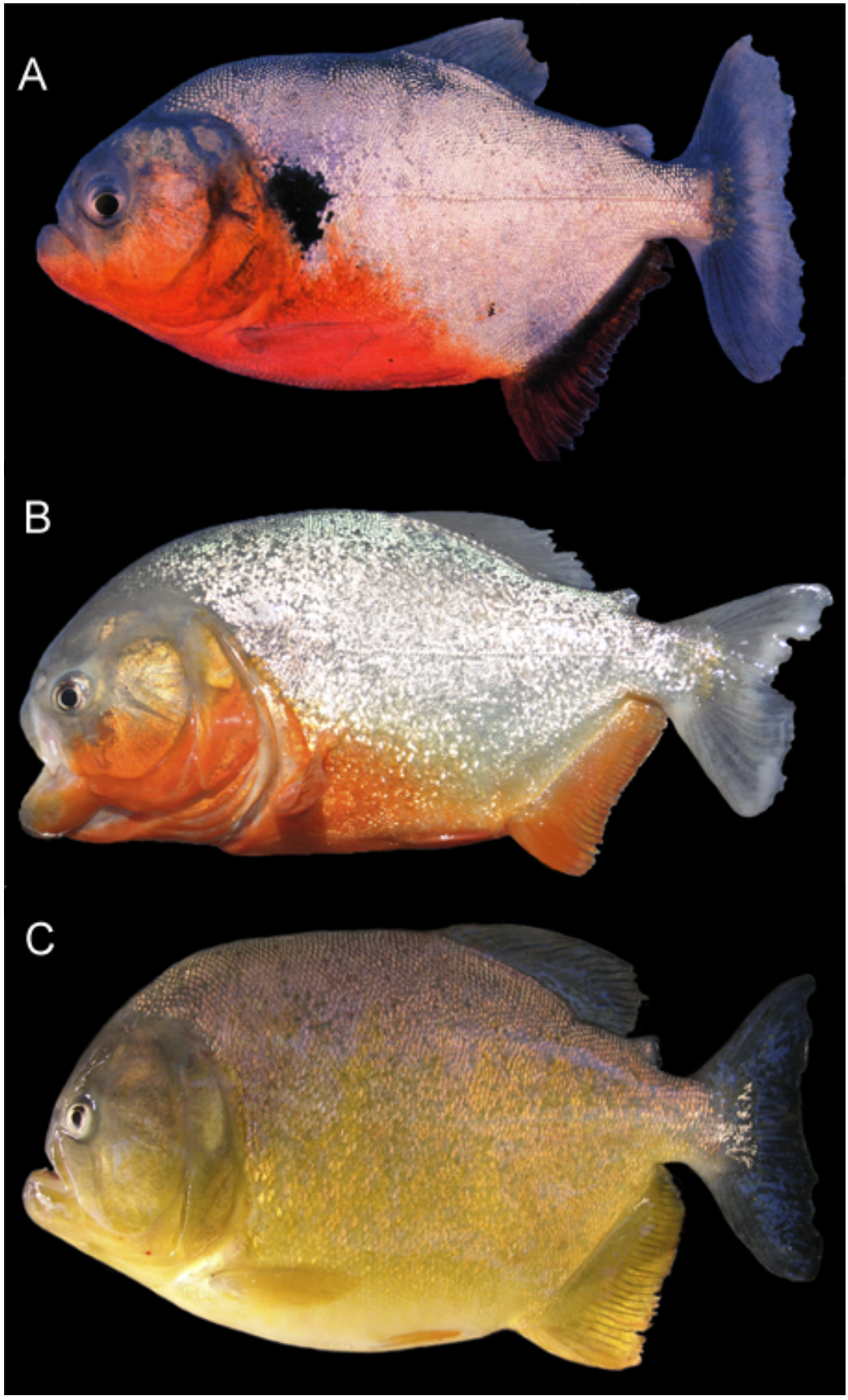

Figure 1. Representative specimens of *Pygocentrus*: A) *P. cariba*, Río Orinoco. B) *P. nattereri*, Rio das Mortes, Araguaia basin. C) *P. piraya*, Rio São Francisco. Photographs A by Mauro Nirchio, B and C by Alec Zeinad. Specimens not preserved.

Material and methods

Taxon sampling

Specimens were collected or obtained from fish collections, and morphologically identified consulting the taxonomic literature and identification keys (Fink 1993). Vouchers were fixed in 95% ethanol or 10% formalin and transferred to 70% ethanol for permanent storage and posteriorly deposited in the Laboratório de Biologia e Genética de Peixes, Universidade Estadual Paulista, Botucatu, Brazil (LBP), and Colección de Zoología, Universidad del Tolima, Ibagué, Colombia (CZUT-IC) (Table 1). Additional sequences were obtained from GenBank (ncbi.nlm.nih.gov/genbank) and BOLD (boldsystems.org) databases (Table 2). We attempted to obtain samples from all river basins in order to sample the total genetic diversity of each species. We also included available sequences of *Pygocentrus nattereri* introduced in the Philippines (Sarmiento *et al.* 2016) to identify the original region that served as sources of those introduced specimens. One sample of *Serrasalmus marginatus* was used to root the trees.

Table 1. List of used specimens of *Pygocentrus* and *Serrasalmus*.

Taxon	Voucher	Specimens number	Locality, basin	Coordinates	Country
P. cariba	LBP 10225	43107, 43108	Rio Apure, Orinoco basin	07°37'24"N 66°24'48"W	Venezuela
	LBP 2229	15663, 15664, 15666	Rio Punta Brava, Orinoco basin	07º37'20"N 66º06'28"W	Venezuela
	LBP 2290	15815	Rio Orinoco, Orinoco basin	07º39'06"N 66º10'34"W	Venezuela
	CZUT-IC 12810	836	Río Ariporo, Orinoco basin	05º54'09"N 71º29'15"W	Colombia
	CZUT-IC 12871	853	Caño el Oso, Orinoco basin	05º56'02"N 71º29'12"W	Colombia
	CZUT-IC 11395	951	Caño Materro, Orinoco basin	04º39'25"N 72º06'44"W	Colombia
P. nattereri	LBP 12738	41012, 41050, 41051, 41053	Rio Araguaia, Tocantins-Araguaia basin	13°18'37"S 50°36'47"W	Brazil
	LBP 4000	23086, 23087, 23091	Lago Morto, Tocantins-Araguaia basin	11°40'09"S 50°51'00"W	Brazil

	LBP 2978	19616, 19617, 19618, 19619, 19620	Lagoa da Égua, Tocantins-Araguaia basin	13°20'05"S 50°42'16"W	Brazil
	LBP 12693	43551, 43552, 43554	Rio Araguaia, Tocantins-Araguaia basin	13°19'00"S 50°37'00"W	Brazil
	LBP 12641	47072	Rio Cuiabá, Paraguay basin	17°30'49"S 37°14'26"W	Brazil
	LBP 19961	79280	Rio Paraguay, Paraguay basin	27°26'01"S 56°58'38"W	Paraguay
	LBP 21836	83905, 83906, 83908, 83961	Rio Negro, Amazon basin	03°12'00"S 59°57'39"W	Brazil
	LBP 22328	86477, 86478, 86479	Rio Solimões, Amazon basin	04°17'32"S 69°54'55"W	Brazil
	LBP 1697	12780, 12781	Lago do Vanico, Amazon basin	03°09'17"S 59°53'12"W	Brazil
	LBP 20651	81252, 81255	Lago Pracuúba, Atlântico basin	01°44'49"N 50°46'59"W	Brazil
	LBP 20977	81157, 81161	Rio Jari, Amazon basin	00°56'00"S 52°32'30"W	Brazil
	LBP 22815	87632, 87633, 87634	Rio Solimões, Amazon basin	03°17'35"S 60°18'26"W	Brazil
P. piraya	LBP 11286	48749, 48750, 48751	Rio São Francisco, São Francisco basin	09°56'46"S 37°06'15"W	Brazil
	LBP 21613	47336, 47337, 47338, 47339	Rio São Francisco, São Francisco basin	17°13'33"S 44°48'27"W	Brazil
	LBP 21612	59752, 59753	Rio São Francisco, São Francisco basin	09°55'28"S 37°07'22"W	Brazil
	LBP 11300	42931	Rio São Francisco, São Francisco basin	20°20'53"S 46°04'10"W	Brazil
	LBP 11336	45522	Lagoa da Tiririca, São Francisco basin	17°13'33"S 44°48'27"W	Brazil
	LBP 11337	45546, 45547, 45548	Lago Maria Joana, São Francisco basin	17°19'29"S 44°45'57"W	Brazil
S. marginatus	LBP 19865	79045	Lagoa das Garças, Paraná basin	22°43'27"S 53°13'04"W	Brazil

Table 2. Specimens obtained from GenBank and BoldSystem.

Taxon	GenBank/ BOLD number	Obtained from	Locality
P. piraya	DSFRE115-08 to 118-08	Bold System	Unknown
	DSFRE351-08 to 352-08	Bold System	Unknown
	DSFRE387-08 to 388-08	Bold System	Unknown
	HQ600843 to 849	GenBank	São Francisco basin, Minas Gerais, Brazil
	KP256432 to 433	GenBank	São Francisco river basin, Brazil
	KP256435	GenBank	São Francisco river basin, Brazil
	HM405211	GenBank	São Francisco basin, Minas Gerais, Brazil
P. nattereri	DSFRE372-08 to 74-08	Bold System	Unknown
	FCOD001-15 to 03-15	Bold System	Philippines, Metro Manila
	FCOD006-15 to 09-15	Bold System	Philippines, Metro Manila
	GBMTG1931-16	Bold System	Unknown
	ITAPE137-15 to 150-15	Bold System	Itapecuru river, Maranhão, Brazil
	KP256424 to 427	GenBank	Upper Paraguay basin, Brazil
	KP256428 to 431	GenBank	Tocantins, Brazil

DNA extraction, amplification and sequencing

Tissue samples were taken from livers, gills, fins or muscles. The total DNA was isolated using the Qiagen "DNeasy Blood & Tissue" kit according to manufacturer's instructions. Partial segments of the *COI* gene were amplified by PCR using the primers Fish F1 (5'-TCAACCAACCACAAAGACATTGGCAC-3') and Fish R1 (5'–TAGACTTCTGGGTGGCCAAAGAATCA–3') described by Ward *et al.* (2005). The PCR was performed on a thermocycler with a final volume of 12 µl containing of 8.175 µl distilled water, 0.5 µl dNTP (8 mM), 1.25 µl 10X Taq buffer (500 mM KCl; 200 mM Tris-HCl), 0.375 µl of $MgCl_2$, 0.25 µl of each primer (10 µM) and 0.2 µl of PHT Taq polymerase. PCR conditions consisted of an initial denaturation at 95ºC for 5 minutes, followed by 35 cycles including denaturation at 95ºC for 45s, annealing at 52ºC for 45s and extension at 68ºC for 120s, and a final extension at 68ºC for 5 minutes. Amplified products were checked on 1% agarose gel.

The PCR products were purified with ExoSAP-IT (USB Corporation, Cleveland, OH, USA) following the manufacturer's protocol. The purified product was used as template to sequence both DNA strands. The cycle sequencing reaction was carried out using a BigDye Terminator v3.1 Cycle Sequencing Ready Reaction kit (Applied Biosystems) in a final volume of 7 µl containing 0.35 µl primer (10 mM), 1.05 µl buffer 5X, 0.7 µl BigDye mix, and 3.9 µl distilled water. The cycle sequencing conditions were initial denaturation at 96ºC for 2 min followed by 30 cycles of denaturation at 96ºC for 45s, annealing at 50ºC for 60s, and extension at 60ºC for 4 min. The sequencing products were purified following the protocol suggested in the BigDye Terminator v3.1 Cycle Sequencing kit's manual (Applied Biosystems). All samples were sequenced on an ABI 3130 Genetic Analyzer (Applied Biosystems) following the manufacturer's instructions.

Species delimitation analyses

Sequences were assembled and edited in Geneious 4.8 (Kearse *et al.* 2012) to obtain a single consensus sequence for each specimen and also check for deletions, insertions and stop-codons. Then, sequences were aligned with Muscle algorithm (Edgar 2004) implemented in Geneious 4.8. Specimens were binned into groups according to a neighbor-joining tree using the Kimura-2-parameters model (K2P; Kimura 1980) in MEGA X (Kumar *et al.* 2018); subgroups of *Pygocentrus nattereri* were separated in order to test prior hypothesis of multiple isolated populations. Three approaches of genetic distances were obtained using the K2P model in MEGA X: the overall mean distance, intraspecific distances, and interspecific distances. The neighbor-joining tree was then generated in MEGA and tested by 1,000 bootstrap pseudoreplicates.

We used four distinct species delimitation methods (Poisson Tree Process, Barcode Index Number, Automatic Barcode Gap Discovery, and General Mixed Yule Coalescent Model) for our dataset using either sequence-based estimations or topological analyses resulted from the maximum likelihood (ML) or Bayesian analyses. The maximum likelihood (ML) analysis was performed through RaxML v8.2 (Stamatakis 2014) with the GTRGAMMA model, and *a posteriori* analysis of bootstrap was conducted with the autoMRE function at RAxML and the best-fitting ML tree was reconciled with the bootstrap replicates. The resulting ML tree was used as input tree for the Poisson Tree Process (PTP) model that delimits species using non-ultrametric trees since the speciation rate is modeled

directly by the number of nucleotide substitutions (Zhang *et al.* 2013). The analysis was performed at the PTP webserver (http://species.h-its.org/ server) using 100,000 MCMC generations and a 0.1 burn-in rate.

Then, we used the Barcode Index Number (BIN), at the BOLD webserver (http://www.boldsystems.org), which clusters *COI* sequences into Operational Taxonomic Units (OTUs; referred as BINs) independent of prior taxonomic assignment. For BINs designation, the webserver performs a Refined Single Linkage (RESL) Analysis, a staged clustering process that employs single linkage clustering as a tool for the preliminary assignment of records to an OTU and a subsequent finishing step that employs the graph analytical approach of Markov Clustering (Ratnasingham & Hebert 2013).

In addition, we performed the Automatic Barcode Gap Discovery (ABGD) analysis. The ABGD is an automatic procedure that sort sequences into hypothetical species based on the barcode gap. The method infers from the data a model-based confidence limit for intraspecific divergence and detects the barcode gap as the first significant gap beyond this limit and uses it to partition the data. Inference of the limit and gap detection are then recursively applied to previously obtained groups to get finer partitions until there is no further partitioning (Puillandre *et al.* 2012). The analysis was performed at the webserver (http://www.abi.snv.jussieu.fr/public/abgd/abgdweb.html) with Kimura (K80; 2.0) distance model with Pmin = 0.001 and Pmax=0.1.

Finally, we ran the General Mixed Yule Coalescent Model (GMYC), a likelihood method that delimits species by fitting within- and between species branching models to reconstructed gene trees (Fujisawa & Barraclough 2013). Before the analysis, the online tool ElimDupes (http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html) was used to remove the duplicated haplotypes; since GMYC requires no politomies, removing the repeated haplotypes improves de algorithm and maximizes the computational time analysis. Then, a Bayesian inference of phylogeny was estimated with a relaxed lognormal clock with a speciation birth/death model, on an arbitrary timescale, using BEAUTi and BEAST v.1.8.4 (Drummond *et al.* 2012). The nucleotide evolution model used to estimate the ultrametric tree was the HKY+I model, as estimated by PartitionFinder (Lanfear *et al.* 2012). A random tree was used as a starting tree for the MCMC searches with two independent runs of 250,000,000 generations, and a tree was sampled every 25,000th generation. The distribution of log-likelihood scores was examined to determine the stationary phase for each search and to decide whether extra runs were required to achieve convergence using the program Tracer v.1.7.1 (Rambaut *et al.* 2018). All sampled topologies beneath the asymptote were discarded

as part of a burn-in procedure (burn-in: 25%), and the remaining trees were used to construct a 50% majority-rule consensus tree in TreeAnnotator v1.8.4. The resulting tree was visualized in FigTree v.1.4.3, and the resultant topology was implemented in the GMYC analysis. The GMYC delimitation analysis was performed at the webserver (https://species.h-its.org/gmyc/R) with a single threshold method and other parameters set as default.

Results

Barcode sequences with more than 500 bp were obtained from 57 specimens (uploaded to Bold System under the IDs PYGO001-18 – 048-18; PYGO049-19 – 057-19) in addition to 51 sequences obtained from BOLD and GenBank, resulting in a final matrix with 108 sequences. Stop codons, deletions or insertions were not identified in any sequence. Following alignment and editing, the final matrix had 525 characters of which 475 positions were conserved and 50 were variable, with 23.0% adenine, 31.6% cytosine, 27.7% thymine and 17.7% guanine. The genetic distance analysis clearly splits the three species of *Pygocentrus* with 0.062±0.011 of distance between *P. cariba* and *P. piraya*, 0.055±0.011 between *P. cariba* and *P. nattereri*, and 0.027±0.006 between *P. piraya* and *P. nattereri* (Table 3). Subgroups of *P. nattereri* presented genetic distances ranging from 0.010 (Amazonas-Paraguay and Amazonas-Itapecuru) to 0.019 (Itapecuru-Paraguay) (Table 4). Results also reveal low intraspecific genetic variation within each lineage (0.000–0.002) (Table 4).

Table 3. Genetic distances among *Pygocentrus* species. Bold numbers represent intraspecific values.

	P. cariba	*P. piraya*	*P. nattereri*
P. cariba	**0.000±0.000**		
P. piraya	0.062±0.011	**0.003±0.001**	
P. nattereri	0.055±0.011	0.027±0.006	**0.009±0.003**

Table 4. Genetic distances among populations of *Pygocentrus nattereri*. Bold numbers represent intrapopulation values. Lineages are represented by their main river basins.

	Araguaia	Paraguay	Amazonas	Itapecuru
Araguaia	**0.000±0.000**			
Paraguay	0.012±0.005	**0.000±0.000**		
Amazonas	0.012±0.005	0.010±0.004	**0.002±0.002**	
Itapecuru	0.018±0.006	0.019±0.007	0.010±0.004	**0.000±0.000**

The ML gene tree shows long branches structuring each of the three species of *Pygocentrus* and the clear separation of sublineages in *P. nattereri* (Fig. 2). Species delimitation analyses returned distinct numbers of species. Both PTP and BIN resulted in three species, *P. cariba*, *P. nattereri* and *P. piraya*, despite low posterior probabilities supporting nodes in the PTP. On the other hand, ABGD resulted in six partitions that ranged from 10 (P = 0.001) to one lineage (P = 0.012), with three partitions supporting the presence of six lineages of *Pygocentrus* (P = 0.002 - 0.007), with *P. nattereri* divided in four distinct groups: Amazon, Araguaia, Itapecuru and Paraguay. The GMYC revealed a disproportionate number of 11 species, one lineage for *P. cariba*, four lineages for *P. piraya* and six lineages for *P. nattereri*. The threshold time obtained in the GMYC analysis was -6.47E-04T (where T is the time from present to the time of the root). Additionally, we included seven sequences of introduced specimens of *Pygocentrus nattereri* in the Phillippines (Sarmiento *et al.* 2016), and the topology shows that they appear nested within group Amazon where probably was the source of the introduction.

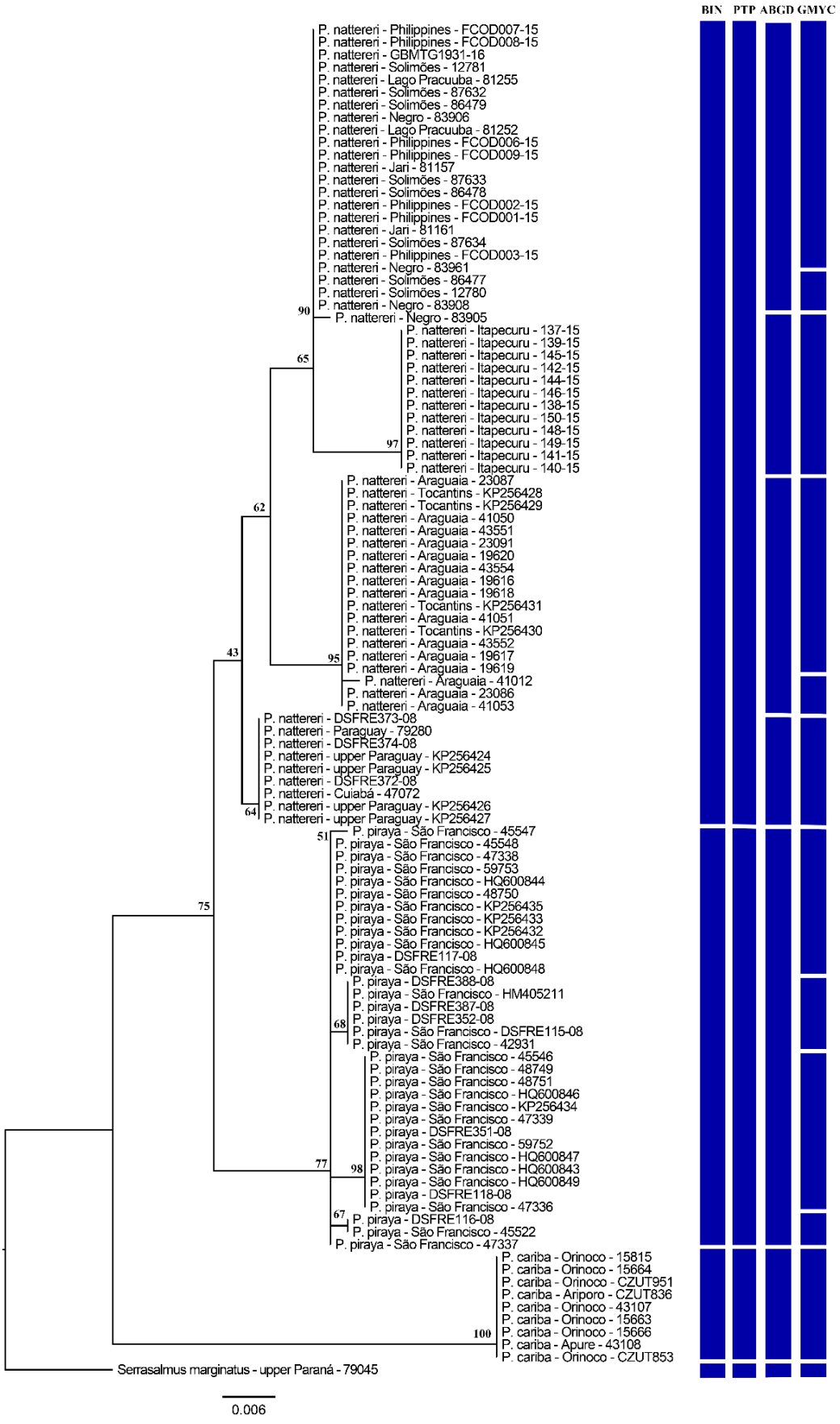

Figure 2. Maximum likelihood tree based on the *cytochrome oxidase c subunit I* gene for *Pygocentrus* species. Vertical columns after taxa names represent the limits among species according to BIN, PTP, ABGD and GMYC methods.

Discussion

Species delimitation in Pygocentrus

Morphological characters have been used traditionally to discriminate Serrasalmidae species, even though the strong influence of allometry during ontogeny and high variability of color patterns generally complicates precise species identifications (Fink 1993). The integrative use of molecular approaches, more specifically the DNA barcoding method, has helped to identify, discriminate and describe the species diversity in several serrasalmid clades (Mateussi *et al.* 2017; Machado *et al.* 2018). Our species delimitation analyses contribute to the advancement of knowledge in this field.

Species delimitation results support the recognition of the two endemic species *P. cariba* (río Orinoco) and *P. piraya* (rio São Francisco), but differ in respect to the widely distributed *P. nattereri*. While genetic distances, PTP, and BIN recognize three species of *Pygocentrus*, corroborating Fink (1993)'s morphological hypothesis, the ABGD and GMYC evidenced multiple species within the present concept of *P. nattereri*, with splits for each subdrainage (main Amazonas, Araguaia, Itapecuru and Paraguay). We conservatively recognize only three species of *Pygocentrus* (with *P. nattereri* being a single species) until the conduction of a broader study with focus on the morphological variation of *P. nattereri*.

The results achieved herein are partially congruent with the most recent barcoding study of the family Serrasalmidae that included all species of *Pygocentrus* (Machado *et al.* 2018). Both *P. cariba* and *P. piraya* appear as valid species, while *P. nattereri* presents multiple sublineages depending on the delimitation approach. Machado *et al.* (2018) found two well-defined lineages of *P. nattereri* (Tocantins/Araguaia and Branco/Madeira/Tapajós) with GMYC recognizing a third lineage from the Rio Guaporé (Madeira basin). Our results corroborate those same lineages and added two more lineages: the Paraguay and Itapecuru lineages (Fig. 2).

Previous studies have demonstrated that GMYC overestimate the number of species. Monaghan *et al.* (2009), for example, applied the GMYC to 12 families of insects in Madagascar and delineated 370 genetic GMYC species to 348 morphospecies. Although it may indeed represent cryptic or yet unrecognized species, Monaghan *et al.* (2009) argued that the number of species diagnosable within GMYC groups could have been reduced if spatially intermediate populations in widespread species contained a combination of what were diagnostic haplotypes. In other words, the addition of haplotypes of intermediate divergence

in widespread species could weaken the distinction between discrete clusters resulting in fewer GMYC groups.

Phylogeographic structure of Pygocentrus nattereri and origin of introductions

Pygocentrus nattereri is the most abundant and widely distributed species of *Pygocentrus* and, accordingly, has controversial species boundaries and carries a historic of doubts about its diagnostic features, validity and taxonomic status. Fink (1993) performed a revision of *Pygocentrus* and, although analyzing *P. nattereri* from all drainages, the author could not find any exclusive character supporting species status, but a combination of characters such as absence of umeral blotch and rays at the adipose fin in adults. The identification of a lineage is proved to be dependent on the type of used method (Costa-Silva *et al.* 2015). Our present data corroborates the hypothesis that *P. nattereri* represents a single species with structured populations along the large continental distribution (Fink 1993; Hubert *et al.* 2007).

Subgroups within *Pygocentrus nattereri* exhibited values below 2% of K2P genetic distance, a cutoff commonly used for Neotropical freshwater fishes (Carvalho *et al.* 2011; Pereira *et al.* 2013), which confirm the interpretation of a single species. Although intraspecific values found for *P. nattereri* are below the threshold, they may still be considered high (1.0–1.9%) and should be interpreted with caution. High levels of intraspecific divergence may be related to phylogeographic history or geographic structure, as expected for freshwater fishes (Ward *et al.* 1994). The effect of highly restricted gene flow attributable to the fragmented nature of freshwater ecosystems may lead each lineage to have independent evolutionary histories not reflecting patterns of genetic variation suggestive of a single large population (April *et al.* 2011). The intraspecific diversification among lineages from different geographic regions demonstrates that individuals from some taxa can be identified not only according to species but linked to a particular watershed (April *et al.* 2011). This appears to be the case of the four distinct lineages of *P. nattereri*, corroborating previous phylogeography and population genetic hypotheses (Hubert *et al.* 2007; Luz *et al.* 2015; Santos *et al.* 2016). However, samples from the Essequibo and Parnaíba should be added in future studies to complete the total geographic range of *P. nattereri*.

Hubert *et al.* (2007) performed a phylogeographic study of the piranha genera *Serrasalmus* and *Pygocentrus* and, indeed, obtained a recent time of diversification for

species of *Pygocentrus*. For example, the cladogenetic events leading to *P. nattereri* and *P. piraya* dated at around 2.63±0.2 Ma, the split of *P. nattereri* from the Amazon and that from the upper Paraguay at around 1.8 Ma and that from the Paraná at about 1.77 ± 0.3 Ma, and the differentiation of the lineages from the upper Amazon (Ucayali and Madeira) at around 0.79±0.1 Ma, which suggest a rapid differentiation of piranha lineages.

Pygocentrus species are widely introduced outside their native ranges, impacts in environments where they are exotic are reported mainly with respect to predation of native species and damage to fishing nets and other fish species (Latini & Petrere Jr. 2004; Trindade & Jucá-Chagas 2008). Herein, sequences of *Pygocentrus* introduced in the Phillippines (Sarmiento *et al.* 2016) were included in our analyses and the results indicate that those introduced lineages represent descendent of the Amazonian lineage. Since introduction of fish species may lead to ecological damage (*e.g.* competition for food, space and spawning sites) and the introduction of diseases, parasites and pathogens, the accurate information about origin of introduced specimens of *P. nattereri* might contribute for future local management purposes.

Bibliography

April J., Mayden R.L., Hanner R.H. & Bernatchez L. (2011) Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences of the United States of America* 108, 10602–10607.

Britski H.A., Silimon K.Z.S. & Lopes B.S. (2007) *Peixes do Pantanal: manual de identificação*, 2nd ed. Embrapa Informação Tecnológica, Brasília. 227 pp.

Cagauan A.G. (2007) Red-bellied Pacu in the Philippines. *Journal of Environmental Science and Management* 10, 42–47.

Carvalho D.C., Oliveira D.A.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, 80–86.

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, 1–12.

Drummond A.J., Suchard M.A., Dong Xie & Rambaut A. (2012) Bayesian phylogenetics with BEAUTi and the BEAST 1.7. *Molecular Biology and Evolution* 29, 1969–1973.

Edgar R.C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32, 1792–1797.

Fink W.L. (1993) Revision of the piranha genus *Pygocentrus* (Teleostei, Characiformes). *Copeia* 1993, 665–687.

Fricke R., Eschmeyer W.N. & Fong J.D. (2019) *Species by Family/Subfamily*. (http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version accessed 09 Jan 2019.

Fujisawa T. & Barraclough T.G. (2013) Delimiting Species Using Single-Locus Data and the Generalized Mixed Yule Coalescent GMYC) Approach : A Revised Method and Evaluation on Simulated Data Sets. *Systematic Biology* 62, 707–724.

Guerrero III R.D. (2014) Impacts of Introduced Freshwater Fishes in the Philippines (1905-2013): A Review and Recommendations. *Philippine Journal of Science* 143, 49–59.

Hubert N., Duponchelle F., Nuñez J., Garcia-Davila C., 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, 2115–2136.

Jégu M. (2003) Subfamily Serrasalminae: pacus and piranhas. In: Reis R, Kullander S.O. & Ferraris Jr C.J., editors. *Check list of the freshwater fishes of South and Central America*. Porto Alegre: Edipucrs, p. 182–196.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P. & Drummond A. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 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, 111–20.

Kumar S., Stecher G., Li M., Knyaz C. & Tamura K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* 35, 1547–1549.

Lanfear R., Calcott B., Ho S.Y.W. & Guindon S. (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular biology and evolution* 29, 1695–701.

Latini A.O. & Petrere Jr. M. (2004) Reduction of a native fish fauna by alien species : An example from Brazilian freshwater tropical lakes Reduction of a native fish fauna by alien species : an example from Brazilian freshwater tropical lakes. *Fisheries Management and Ecology* 11, 71–79.

Lin Y., Gao Z. & Zhan A. (2015) Introduction and use of non-native species for aquaculture in China: status, risks and management solutions. *Reviews in Aquaculture* 7, 28–58.

Luz L.A., Reis L.L., Sampaio I., Barros M.C. & Fraga E. (2015) Genetic differentiation in the populations of red piranha, *Pygocentrus nattereri* Kner (1860) (Characiformes: Serrasalminae), from the river basins of northeastern Brazil. *Brazilian Journal of Biology* 75, 838–845.

Machado-Allison, A. & Fink, W.L. (1991) Notas sobre la sistemática del género *Serrasalmus* y géneros relacionados. Parte II: el género *Pygocentrus* en Venezuela. *Acta Biologica Venezuelica* 13(1-2), 109–135.

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, 1–12.

Mateussi N.T.B., Oliveira C. & Pavanelli C.S. (2018) Taxonomic revision of the Cis-Andean species of Mylossoma Eigenmann & Kennedy, 1903 (Teleostei: Characiformes: Serrasalmidae). *Zootaxa* 4387, 275–309.

Mateussi N.T.B., Pavanelli C.S. & Oliveira C. (2017) Molecular identification of cryptic diversity in species of cis-Andean Mylossoma (Characiformes: Serrasalmidae). *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis* 28, 778–780.

Monaghan M.T., Wild R., Elliot M., Fujisawa T., Balke M., Inward D.J.G., Lees D.C., Ranaivosolo R., Eggleton P., Barraclough T.G. & Vogler A.P. (2009) Accelerated species Inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology* 58, 298–311.

Nico L.G., Jégu M. & Andrade M.C. (2017) Family Serrasalmidae - Piranhas and Pacus. In: Van der Sleen P. & Albert J., editors. *Field guide to the fishes of the Amazon, Orinoco, and Guianas*. D&N Publishing, Wiltshire, UK. p. 172–196.

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

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

Rahman M.M. & Ahmed A.T.A. (2007) Studies on breeding and larval development of red bellied piranha, *Pygocentrus nattereri* Kner, 1858 in Bangladesh. *Bangladesh J. Zool.* 35, 193–203.

Rambaut A., Drummond A.J., Xie D., Baele G. & Suchard M.A. (2018) Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* 0, 1–5.

Ratnasingham S. & Hebert P.D.N. (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE* 8, e:66213.

Santos C.H.A., Leitão C.S.S., Paula-Silva M.N. & Almeida-Val V.M.F. (2016) Genetic differentiation in red-bellied piranha populations (*Pygocentrus nattereri* Kner, 1858) from the Solimões-Amazonas River. *Ecology and Evolution* 6, 4203–4213.

Santos G.M., Ferreira E.J.G. & Zuanon J.A.S. (2006) *Peixes comerciais de Manaus*. ProVárzea/Ibama, Manaus. 144 pp.

Sarmiento K.P., Ventolero M.F.H. & Santos M.D. (2016) Identification of prohibited red piranha, *Pygocentrus nattereri* (Kner 1858) from confiscated juvenile fish in Manila, Philippines. *International Journal of Philippine Science and Technology* 9, 23–29.

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

Suchard M.A., Lemey P., Baele G., Ayres D.L., Drummond A.J. & Rambaut A. (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* 4, vey016.

Thompson A.W., Betancur R., López-Fernández H. & Ortí G. (2014) A time-calibrated, multi-locus phylogeny of piranhas and pacus (Characiformes: Serrasalmidae) and a comparison of species tree methods. *Molecular Phylogenetics and Evolution* 81, 242–257.

Trindade M.E. de J. & Jucá-Chagas R. (2008) Diet of two serrasalmin species, *Pygocentrus piraya* and *Serrasalmus brandtii* (Teleostei: Characidae), along a stretch of the rio de Contas, Bahia, Brazil. *Neotropical Ichthyology* 6, 645–650.

Ward R.D., Woodwark M. & Skibinski D.O.F. (1994) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *Journal of Fish Biology* 44, 213–232.

Ward R.D., Zemlak T.S., Innes B.H., Last P.R. & Hebert P.D.N. (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360, 1847–1857.

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

Chapter 3

Species delimitation and taxonomic revision of *Catoprion* Müller & Troschel, 1844 (Characiformes: Serrasalmidae) with the description of a new species

Species delimitation and taxonomic revision of *Catoprion* Müller & Troschel, 1844 (Characiformes: Serrasalmidae) with the description of a new species

Abstract

Catoprion is a lepidophagous genus in Serrasalmidae that is easily recognized by possessing the lower jaw much more prominent than the upper jaw and by the presence of mamiliform teeth arranged in single irregular series on both premaxilla and dentary. Although presenting a well-supported phylogenetic position within the Serrasalmidae, *Catoprion* has not been the subject of studies dealing with species delimitation that are widespread throughout South American freshwaters. Molecular analyses through DNA barcoding with 49 sequences of *Catoprion* and one additional sequences of the related genus *Pygopristis*, were conducted in order to delimit species of *Catoprion* through genetic distances and conventional likelihood, PTP, ABGD and GMYC methods. All species delimitation methods recovered *Catoprion* as two distinct lineages with the first occurring in the Paraguay, Purus and Madeira rivers, and the other in the Branco, Jatapu, Negro, Uatumã, Nhamundá, Trombetas, Tapajós and Xingu rivers of the Amazon basin. Morphological analyses were performed in 178 specimens of *Catoprion* as well as in other comparative serrasalmid genera, and corroborate the split of these two lineages as two distinct species. Considering that analysis of the type species, *C. mento*, indicate the species as belonging to Paraguay or Tocantins rivers, the other lineage was considered an undescribed species. *Catoprion* sp. n. differs from *C. mento* by presenting 86 to 94 perforated scales in the lateral line (*vs.* 65–86 scales) and 35 to 40 circumpeduncular scales (*vs.* 29–34 scales).

Keywords: DNA barcode, Biodiversity, Neotropical region, Systematics

Introduction

Serrasalmidae is a Neotropical freshwater fish family comprising 97 species within 16 extant genera (Fricke *et al.* 2019), morphologically identified by a combination of characters that include a deep, laterally compressed body, a ventral keel composed by spines derived from modified abdominal scales and presence of a predorsal spine in almost all genera, with exception of *Colossoma* Eigenmann & Kennedy, 1903, *Mylossoma* Eigenmann & Kennedy, 1903 and *Piaractus* Eigenmann, 1903 (Jégu 2003). The diet specializations in serrasalmids, which include carnivory, frugivory, herbivory and lepidophagy, result in disparate morphological traits related to teeth shape and buccal apparatus (Goulding 1980; Sazima 1983; Sazima & Machado 1990; Nico & Taphorn 1998). Among these, the genus *Catoprion* Müller & Troschel, 1844 stands out for its singular lepidophagous habit (Roberts 1970).

Catoprion was described to allocate *Serrasalmus mento* Cuvier, 1819 and has been considered monotypic, with the single species *Catoprion mento*, since its description (Müller & Troschel 1844). The genus is easily recognized from the other serrasalmid genera by presenting a lower jaw much more prominent than the upper jaw and a distinctive single row of mamiliform teeth in both maxilla and dentary (Géry 1972), considered as an "aberrant condition" in the family (Gosline 1951). Based on this condition, Géry (1972) proposed the subfamily Catoprioninae to include *Catoprion*, an arrangement followed by subsequent Géry's studies (Géry 1976, 1977) until the first cladistic studies in Serrasalmidae that, based on morphological characters, proposed *Catoprion* as more related to the *piranhas* (*Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Serrasalmus*) (Machado-Allison 1982). This hypothesis was recovered by subsequent morphology- and molecular-based phylogenies (Ortí *et al.* 1996, 2008; Jégu 2004; Freeman *et al.* 2007; Cione *et al.* 2009; Thompson *et al.* 2014; Mateussi *et al.* Chapter 1) that place *Catoprion* within the "piranha clade" (*sensu* Ortí *et al.* 1996).

Although the phylogenetic position of the genus is well supported and corroborated by several studies, no efforts aiming to explore the putative presence of multiple widespread species throughout South America, on Amazon, Orinoco, Essequibo, and upper Paraguay river basin (Jégu 2003) has been performed yet. Here, we use the DNA barcoding method that allows the identification and recognition of species and flagging new candidate species (Costa-Silva *et al.* 2015; Mateussi *et al.* 2016) to highlight putative multiple taxonomic units in *Catoprion*. Then, the taxonomic revision of the genus was performed based on the

examination of 178 preserved specimens of *Catoprion*, with the redescription of *Catoprion mento* and the description of a new species from the Amazon basin.

Material and Methods

Taxon sampling

Specimens were collected or obtained from previous collections. Institutional abbreviations follow Sabaj (2016). Specimens used in molecular analyses are listed in Table 1. Additional sequences were obtained from GenBank (accession numbers MG751919–MG751937; Table 2).

TABLE 1. List of specimens of *Catoprion* and related taxa used in the molecular analyses.

Taxon	Voucher	Specimens	Locality, basin	Coordinates	Country
Catoprion mento	LBP 7556	35625, 35626	Rio Cuiabá, Paraguay basin	16°11'39" S 55°48'25" W	Brazil
	LBP 26133	93487–93490	Rio Paraguay, Paraguay basin	16°27'02" S 55°19'12" W	Brazil
	LBP 11127	50884–50888	Rio Purus, Amazon basin	07°29'34" S 63°35'03" W	Brazil
	UFRO 7501	5909, 5912	Rio Jaci-Paraná, Rio Madeira, Amazon basin	09°17'03" S 64°23'43" W	Brazil
Catoprion sp.n.	ANSP 198245	P 26431	Rio Xingu, Amazon bazin	01°46'30"S 52°12'57"W	Brazil
	INPA 37207	P 17948	Rio Jatapu, rio Uatumã, Amazon basin	02°10'31"S 58°10'26"W	Brazil
	INPA 37258	P 19799	Rio Jatapu, rio Uatumã, Amazon basin	02°01'03"S 58°10'26"W	Brazil
	INPA 47302	P 26437	Rio Xingu, Amazon bazin	01°46'30"S 52°12'57"W	Brazil
	INPA 47347	P 26450	Rio Acaraí, rio Xingu, Amazon basin	02°20'13"S 52°31'13"W	Brazil
	INPA 37862	P 27287	Rio Aiuanã, rio Negro, Amazon basin	00°35'24"S 64°55'10"W	Brazil
	LBP 4342	24074	Rio Uraricoera, Rio Branco, Amazon basin	03°11'00" N 60°33'20" W	Brazil
	LBP 15148	62389, 62391, 62392	Rio Branco, Amazon basin	03°08'16" N 60°16'33" W	Brazil
	LBP 21615	63663–63666	Rio Cauamé, Rio Negro, Amazon basin	02°56'19" N 61°03'06" W	Brazil
	LBP 15534	63933–63935	Rio Tacutu, Rio Branco, Amazon basin	03°22'55" N 59°51'28" W	Brazil
	MCP 51844	261	Rio Tapajós, Amazon basin	02°53'29"S 55°10'32"W	Brazil
Pygopristis denticulata	LBP 15609	64256	Rio Tacutu, Rio Branco, Amazon basin	03°17'24" N 59°53'48" W	Brazil

TABLE 2. List of specimens of *Catoprion* added from GenBank database used in the molecular analyses.

Taxon	GenBank number	Locality, all from Amazon basin	Country
Catoprion sp.n.	MG751919, MG751923-25	Rio Branco	Brazil
	MG751920-22, MG751937	Rio Nhamundá	Brazil
	MG751926-27	Rio Negro	Brazil
	MG751928	Rio Aripuanã	Brazil
	MG751929-32	Rio Jatapu	Brazil
	MG751933-35	Rio Trombetas	Brazil
	MG751936	Rio Uatumã	Brazil

DNA extraction, PCR and sequencing

Samples were taken from muscle tissues. Total DNA was isolated using the Qiagen "DNeasy Blood and Tissue kit" according to manufacturer's instructions. Then, partial segments of the *Cytochrome oxidase c subunit I* (COI) were amplified by PCR using the primers Fish F1 (5'–TCAACCAACCACAAAGACATTGGCAC–3') and Fish R1 (5'–TAGACTTCTGGGTGGCCAAAGAATCA–3') described by Ward *et al.* (2005). The PCR was performed on a thermocycler with a final volume of 12 µl containing 8.175 µl distilled water, 0.5 µl dNTP (8 mM), 1.25 µl 10X Taq buffer (500 mM KCl; 200 mM Tris–HCl), 0.375 µl $MgCl_2$, 0.25 µl each primer (10 µM) and 0.2 µl Taq PHT DNA polymerase. PCR conditions consisted on an initial denaturation at 94ºC for 5 minutes, followed by 35 cycles including denaturation at 95ºC for 60s, annealing at 52ºC for 60s and extension at 68ºC for 2 minutes, with a final extension at 68ºC for 10 minutes. Amplified products were checked on 1% agarose gel.

The PCR products were then purified with ExoSAP–IT (USB Corporation, Cleveland, OH, USA) following the manufacturer's protocol. The purified product was used as template to sequence both DNA strands. The cycle sequencing reaction was carried out using a BigDye Terminator v3.1 Cycle Sequencing Ready Reaction kit (Applied Biosystems) in a final volume of 7 µl containing 0.35 µl primer (10 mM), 1.05µl buffer 5X, 0.7 µl BigDye mix, and 3.9 µl distilled water. The cycle sequencing conditions were initial denaturation at 96ºC for 2

minutes followed by 30 cycles of denaturation at 96ºC for 45s, annealing at 50ºC for 60s, and extension at 60ºC for 4 minutes. The sequencing products were purified following the protocol suggested in the BigDye Terminator v3.1 Cycle Sequencing kit's manual (Applied Biosystems). All samples were sequenced on an ABI 3130 Genetic Analyzer (Applied Biosystems) following the manufacturer's instructions.

Analyses

All sequences were analyzed on Geneious v4.8 (Kearse *et al.* 2012) to obtain the consensus sequence and check for deletions, insertions and stop codons. After downloading sequences of *Catoprion* available at the GenBank, all sequences were aligned using Muscle algorithm (Edgar 2004) implemented on Geneious. To determine genetic distance between species, we group sequences based on the resulted topology from the distance analyses. Genetic distances were obtained using the Kimura-2-Parameter (K2P) model (Kimura 1980) at MEGA X (Kumar *et al.* 2018). A neighbor-joining tree was generated using the K2P model with 1,000 bootstrap replicates.

A maximum likelihood (ML) analysis was performed in RaxML v8.2 (Stamatakis 2014) with the GTRGAMMA model. The best tree was accessed through five searches and a *posteriori* analysis of bootstrap replicates was conducted with RAxML with the autoMRE tool. The resulting ML tree was used as input tree for the Poisson Tree Process (PTP) analysis (Zhang *et al.* 2013). The PTP analysis was performed at the webserver (http://species.h–its.org/ server) using 100,000 MCMC generations and a 0.1 burn-in rate. In addition, we performed the Automatic Barcode Gap Discovery (ABGD) analysis (Puillandre *et al.* 2012) at the ABGD webserver (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html) with Kimura (K80; 2.0) distance model and other parameters at default (Pmin = 0.001; Pmax = 0.1).

A Bayesian inference was estimated with a lognormal relaxed clock (Drummond *et al.* 2006) and a speciation birth/death model, on an arbitrary timescale, using BEAUTi and BEAST v.1.8.4 (Drummond *et al.* 2012). The nucleotide evolutionary model used to estimate the ultrametric tree was the GTR+G model, as estimated by Partition Finder (Lanfear *et al.* 2012). A random tree was used as a starting tree for the MCMC searches with two independent runs for 300,000,000 generations, and a tree was sampled every 30,000 generations. Tracer v.1.7.1 (Rambaut *et al.* 2018) was used to examine the distribution of log-likelihood scores and determine the stationary phase for each search and to decide whether

extra runs were required to achieve convergence. All sampled topologies beneath the asymptote were discarded as part of a burn-in procedure (10%), and the remaining trees were used to construct a 50% majority-rule consensus tree in TreeAnnotator v1.8.4. The resulting tree was checked in FigTree v.1.4.3 and used as input file for the General Mixed Yule Coalescent Model (GMYC; Fujisawa & Barraclough 2013; Pons *et al.* 2006) analysis performed at the webserver (https://species.h–its.org/gmyc/R) with a single threshold method.

Morphological analyses

Morphometric and meristic data were taken whenever possible on the left side of specimens using digital calipers (precision of 0.1 mm). Measurements partially followed Géry (1972) and are described in detail at the Appendix S1. All measurements were presented as percentages of standard length (SL), except for subunits of head, which were presented as percentages of head length (HL). Specimens cleared and stained (cs) were used for obtaining osteological data and were prepared according to Taylor & VanDyke (1985). Radiographs (rd) were taken with a Faxitron MX-60 digital X-ray system. Vertebrae incorporated into the Weberian apparatus were counted as four elements. Ventral-keel spines are divided in prepelvic spines (extending to origin of pelvic fin and not including the spine lying over the pelvic-fin origin), post-pelvic spines (including the spine over the pelvic-fin origin, plus those from the pelvic-fin origin to the double pair of spines), and anal spines (double pairs of spines around anus). In the description, each count is followed by its frequency in parentheses. The type counts are marked with asterisks. In the list of examined specimens, the total number of specimens is first reported, followed by the number of examined specimens in parentheses (if different of the total number), and the number of specimens prepared as dry skeletons, cleared and stained or from which radiographs were taken. The map provided was developed through QGIS 2.18 program.

Results

Species delimitation through DNA barcode

Barcode sequences with more than 500 bp were obtained from 31 specimens (uploaded to Bold System under the IDs CATOP001-19 to 025-19) and nineteen additional sequences were included from GenBank, totalizing 50 sequences in the final matrix. Stop codons, deletions or insertions were not observed in any of the sequences. Following alignment and editing, the final matrix had 553 characters, of which 469 positions were conserved and 81 were variable, with 23.2% adenine, 31.5% cytosine, 27.2% thymine and 18.1% guanine. The genetic distance analysis showed that the two species of *Catoprion* differs between each other from 7.3% ± 0.02 K2P distance.

The ML tree inferred through RAxML recovered the two groups consistent with the genetic distance analysis and the previous morphological identification of species and exhibited strong node support (Fig. 1). The species delimitation analyses PTP, ABGD and GMYC confirm the presence of two species of *Catoprion*, the first, *C. mento*, occurring in the Paraguay, Purus and Madeira rivers and the second, an undescribed species occurring in the Branco, Japurá, Negro and Trombetas rivers. The ABGD analysis resulted in nine partitions that ranged from two (P = 0.059) to 11 (P = 0.001) lineages, with five partitions (P = 0.007 – 0.059) suggesting two lineages of *Catoprion*. The threshold time obtained in the GMYC analysis was -6.67E-03T (where T is the time from present to the time of the root).

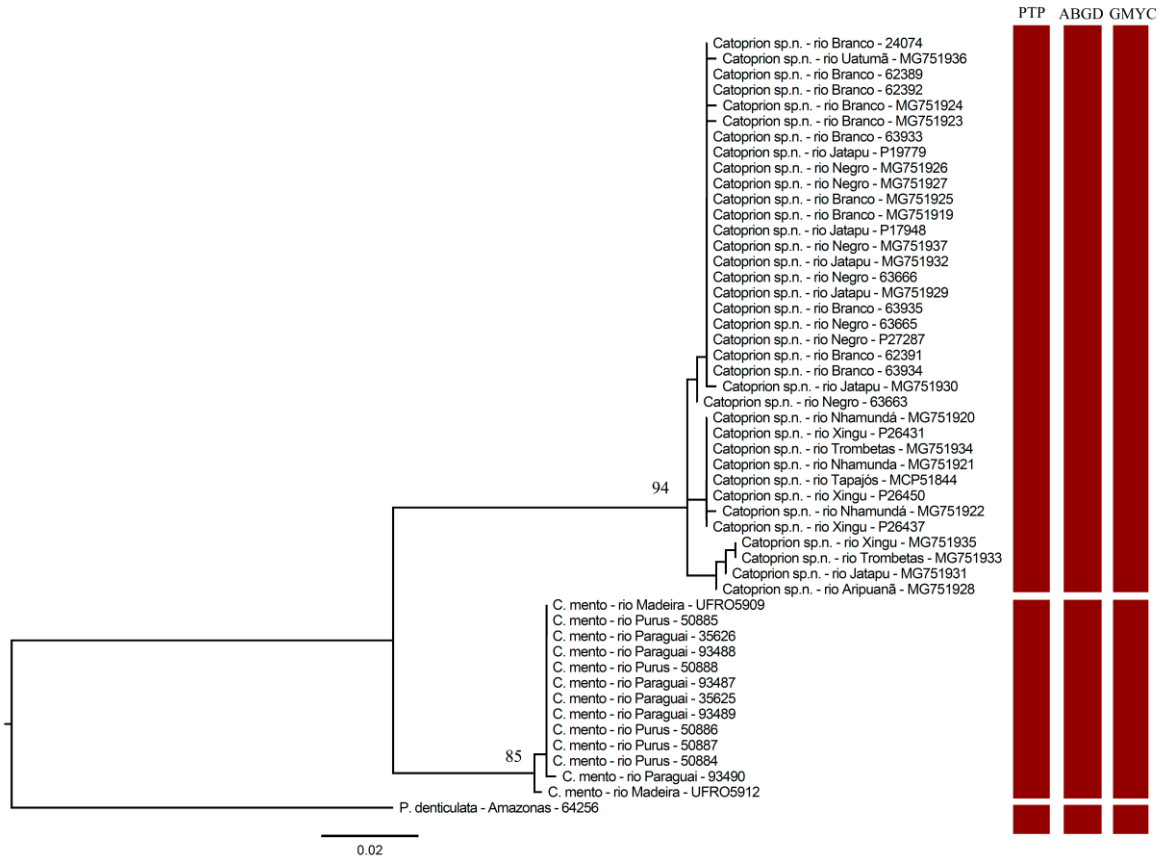

Figure 1. Maximum likelihood tree based on the *cytochrome oxidase c subunit I* gene for *Catoprion* species. Columns represent the limits among species according to PTP, ABGD and GMYC models.

Taxonomic revision

***Catoprion* Müller & Troschel, 1844**

Catoprion Müller & Troschel, 1844: 96 [original description, type by monotypy: *Serrasalmus mento* Cuvier, 1819]. —Géry, 1972: 207 [Guyanas; Catoprioninae]. —Géry, 1976: 54 [checklist]. —Géry, 1977: 294 [Serrasalmidae; Catoprioninae]. Taphorn, 1992: 149 [Apure]. —Jégu, 2003: 183 [checklist]. —Jégu & Ingenito, 2007: 40 [checklist]. —Nico *et al.*, 2017: 181 [Amazon, Orinoco and Guyanas].

Diagnosis. *Catoprion* can be distinguished from the remaining Serrasalmidae genera by presenting mouth superior, with a strongly prognathic jaw, and a single irregular row of five

mamiliform (rounded base and narrower tip) teeth in the premaxilla that can be projected out of the mouth. It can also be discriminated by the combination of three non-exclusive characters: 1) presence of a spine preceding the dorsal fin; 2) first rays of the dorsal and anal fins quite prolonged, sometimes reaching the caudal fin; 3) well-developed spines in the ventral keel.

Distribution. *Catoprion* is widely distributed in lowlands of the Orinoco, Essequibo, Amazon, Tocantins-Araguaia and Paraguay river basins (Fig. 2).

Etymology. *Cato*, from the Greek *kata*, meaning down; *prion*, from the Greek *prion*, meaning saw.

Remarks. In the original description, Müller & Troschel (1844) described two series of teeth in premaxilla: “Dentes ossis intermaxillaris biseriales”. The morphology of *Catoprion*’s dentition is an intermediate state between the condition “two rows of molariform/incisiviforms teeth” observed in “pacus” and the condition “single row of sharp multi cuspidated teeth” observed in all other “piranhas” (*Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Serrasalmus*).

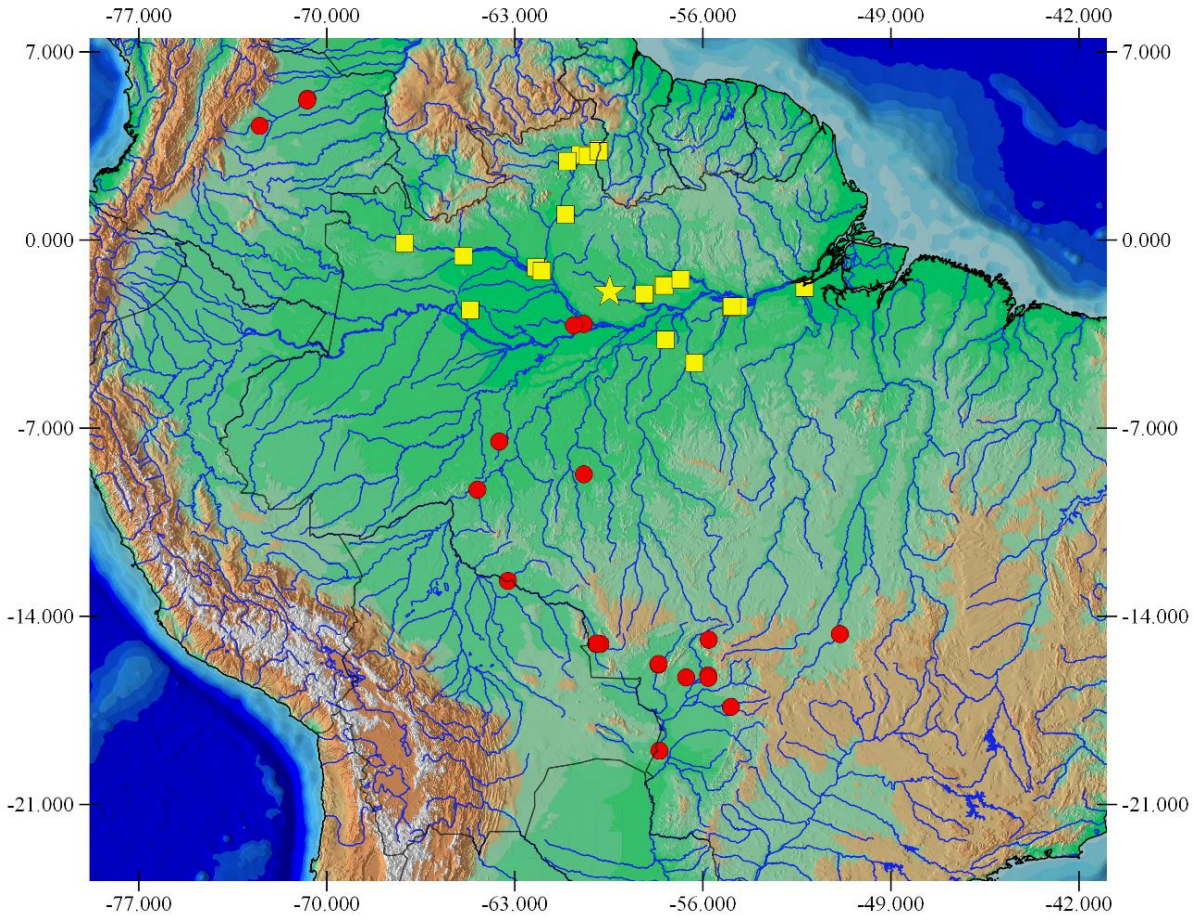

Figure 2. South American map showing the geographic distribution of analyzed specimens of *Catoprion mento* (red circles) and *Catoprion* sp. n. (yellow squares). Stars represent type localities. One symbol may represent more than one locality. *Catoprion* sp. N is also recorded for the Essequibo basin (see text).

***Catoprion mento* (Cuvier, 1819)**

Figs. 3–6; Table 3

Serrasalmus mento Cuvier, 1819: 369, pl.28 [original description; type-locality "Brésil"].

Catoprion mento.—Kner, 1859: 26 [description]. —Eigenmann, 1910: 442 [*partim*; listed]. —Géry, 1972: 207 [brief description]. —Géry, 1976: 54 [checklist]. —Britski *et al.* 2007: 78 [Pantanal; brief description]. —Ota *et al.* 2013: 17 [Madeira basin, Brazil; brief description, photo].

Mylesinus macropterus Ulrey, 1894: 612 [original description; type-locality: "Brazil"]. —Ulrey, 1895: 296 [description].

Figure 3. Holotype of *Serrasalmus mento* (=*Catoprion mento*), MNHN A.9869, 79.8 mm SL, Cabinet d'Ajuda (see Remarks for type locality). Photo by MNHN: L. Randrihasipara. Scale bar = 1 cm.

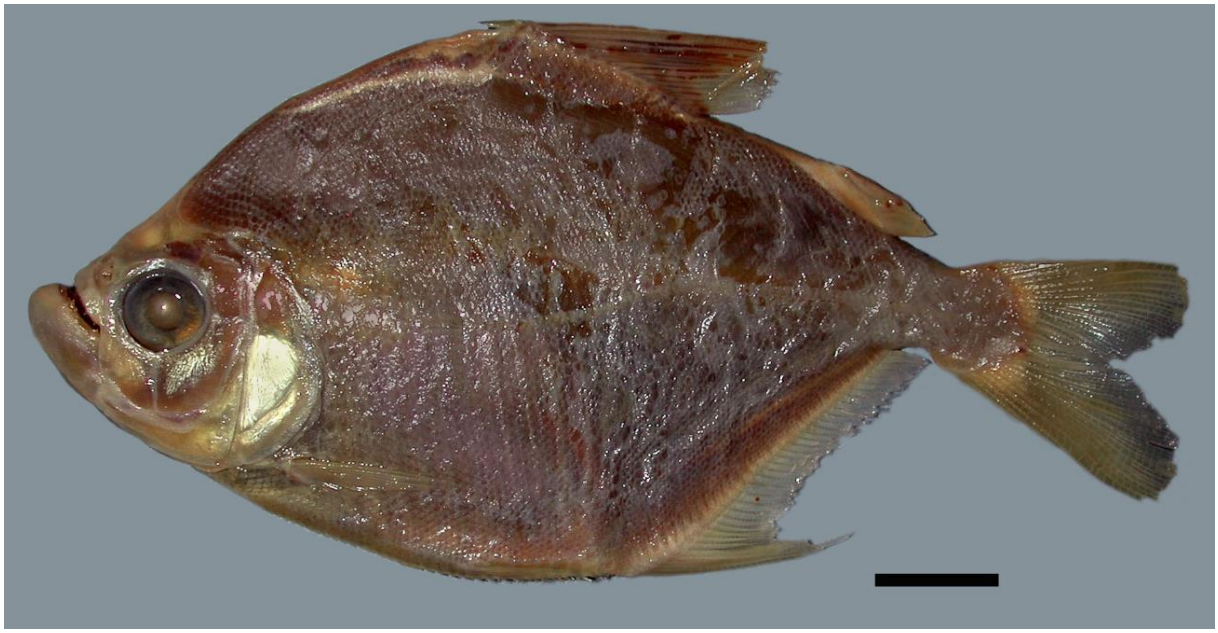

Figure 4. Holotype of *Mylesinus macropterus*, CUMV 3267, Brazil. Photo by CUMV. Scale bar = 1 cm.

Figure 5. *Catoprion mento*. LBP 7556, Brazil, Mato Grosso, Barão de Melgaço, lake at margin of rio Cuiabá, rio Paraguay basin: **a.** male, 74.0 mm SL; **b.** female, 65.0 mm SL.

Diagnosis. *Catoprion mento* differs from *Catoprion* sp. n. by presenting 65–86 perforated scales in the lateral line (*vs.* 86–94 scales) and 29–34 circumpeduncular scales (*vs.* 35–40 scales).

Description. Morphometric data presented in Table 3. Body deep, compressed laterally; general shape rhomboid. Highest body depth at vertical through dorsal-fin origin. Dorsal

profile slightly concave along head and convex between head and dorsal-fin origin. Dorsal profile of body fairly sloping at base of dorsal and adipose fins, and straight at interval between these two fins. Ventral profile convex from isthmus to end of anal fin. Ventral keel with a serra of spines. Pre-pelvic spines 17 (2), 18 (1), 19 (6), 20 (8), 21 (17), 22 (20), 23* (12), 24 (7) or 25 (2). Post-pelvic spines 7 (1), 8 (2), 9 (32), 10* (33) or 11 (4). Anal spines 1 (3), 2 (58), 3* (12) or 4 (1). Caudal peduncle deeper than long.

Eyes lateral at middle of head; upper margin of eyes above longitudinal line at lateral line origin. Frontal and parietal fontanels expanded laterally. Snout very short and slightly pointed in lateral view. Nostrils dorsolateraly positioned, at longitudinal axis through upper margin of eye. Mouth upturned. Jaw strongly prognathous. Premaxillary and dentary teeth mamiliform, with a rounded base and narrower tip. Premaxilla with five teeth; two large forward directed teeth and three smaller teeth, in an irregular single row. Dentary with a single row of six teeth. Maxilla edentulous.

Body completely covered by small cycloid scales. Lateral line complete, with 65 (1), 66 (1), 68 (1), 69 (1), 71 (2), 72 (1), 73 (2), 74 (1), 75 (2), 76 (4), 77 (5), 78 (6), 79 (5), 80 (5), 81 (5), 82 (3), 83 (3), 84 (3), 85 (2) or 86 (2) perforated scales; first to tenth scale larger than remaining. Scales series above lateral line 24 (1), 27 (1), 29 (1), 30 (2), 31 (1), 32 (7), 33 (7), 36 (1) or 37 (1). Scales below lateral line 21 (1), 24 (3), 25 (1), 26 (5), 27 (3), 28 (2), 29 (2), 30 (4) or 31 (1). Circumpeduncular scales 29 (1), 30 (6), 31 (5), 32 (13), 33 (4) or 34 (5).

Dorsal fin preceded by spine, its origin approximately equidistant from tip of snout and end of hypural plate. First dorsal-fin rays prolonged; in some specimens surpassing caudal-fin base. Branched dorsal-fin rays gradually decreasing in size, branched dorsal-fin rays 13 (1), 14 (30), 15* (34) or 16 (9). Adipose fin short, longer than deep. Pectoral fin falcate, with i + 11 (7), 12* (25), 13 (19), 14 (11) or 15 (1) rays. Pelvic fin small, with i + 6* (58) branched rays. First anal-fin rays prolonged; in some specimens surpassing caudal-fin base. Branched anal-fin rays gradually decreasing in size; branched anal-fin rays 30 (1), 32 (5), 33* (14), 34 (26), 35 (24), 36 (6) or 37 (2). Caudal fin bifurcated, lobes of similar size, with 17 (52) branched rays.

First gill arch with elongated gill rakers, almost as long as filaments, decreasing in size towards the extremities; 8 (12), 9 (16), 10 (7) or 11 (1) gill rakers on upper limb, 9 (1), 10 (3), 11 (13), 12 (19) or 13 (4) gill rakers on lower limb. Lower gill rakers longer than upper. Vertebrae 36 (7), 37* (18) or 38 (2). Supraneurals 4* (29).

TABLE 3. Morphometric data of *Catoprion mento*. H = holotype; N = number of specimens; SD = standard deviation.

	H	N	Range	Mean	SD
Standard length (mm)	79.8	65	24.2–119.5	62.9	–
Percentages of standard length					
Adipose-fin base	10.9	65	7.2–12.5	10.4	0.01
Anal-fin base	39.3	65	32.9–46.9	41.6	0.10
Anal-fin length	–	59	18.1–60.7	32.6	0.10
Body depth	64.6	65	42.2–71.2	62.4	0.06
Caudal-peduncle depth	10.5	65	8.0–11.3	10.0	0.01
Caudal-peduncle length	10.1	63	5.7–08.9	07.5	0.01
Dorsal-fin base	22.0	64	17.7–26.1	22.1	0.02
Dorsal-fin length	32.6	61	23.9–74.2	45.6	0.13
Head length	29.6	65	27.1–37.0	30.5	0.02
Interdorsal length	16.6	63	12.6–19.3	15.3	0.01
Pectoral-fin length	–	65	14.7–23.6	20.3	0.02
Pelvic-fin length	12.6	64	12.1–17.2	14.9	0.01
Postdorsal length	60.9	65	48.8–61.8	57.1	0.02
Predorsal length	55.4	65	53.8–60.9	57.0	0.01
Percentages of head length					
Interorbital width	34.47	65	20.4–40.6	32.8	0.05
Maxilla length	39.7	64	25.1–43.9	36.5	0.04
Orbital diameter	31.6	52	25.4–37.9	32.5	0.02
Snout length	20.4	65	14.1–28.7	22.5	0.03

Color in alcohol. General body color brownish-yellow, darker in dorsal portion. Black or brown humeral blotch vertically elongated, sometimes inconspicuous or absent. Fins hyaline to yellowish. Caudal-fin base dark brown; in V–shape with the apex directed towards the head, advancing through the caudal peduncle. First rays of dorsal fin often dark. First rays and distal margin of anal fin often dark. Vertical dark band through the eye. (Fig. 5)

Color in life. General body color silvery, darker dorsally. Conspicuous orange blotch on opercle. Dorsal, pectoral, pelvic and adipose fins hyalines. First ray of dorsal and pectoral fins often dark grey. Anal fin yellow to orange, mainly on first rays, gradually turning to hyaline on last rays. Caudal-fin base black, in V–shape with the apex directed towards the head. Vertical dark band through the eye (Fig. 6).

Fig. 6. *Catoprion mento*, LBP 3820, 55 mm SL, rio Negro, Aquidauana, Mato Grosso do Sul, Brazil. Photo by C. Oliveira.

Geographic distribution. *Catoprion mento* occurs in the río Orinoco, upper río Paraguay, and right tributaries of the Amazon basin, including the upper rio Araguaia, rio Madeira, rio Purus and the lower rio Negro (Fig. 2).

Sexual dimorphism. *Catoprion mento* presents an anterior lobe in the anal fin developed from the elongation of the distal tips of the seventh to thirteenth anal-fin rays in males (Figs. 5 and 6), whereas females present an elongation restricted at the anterior rays of the anal fin. The prolongation of the first dorsal and anal-fin rays were observed in both males and females and, thus, not correlated to sexual dimorphism.

Ecological notes. *Catoprion mento* feeds primarily on fish scales, whereas aquatic insect larvae, fish flesh and fins and plants are reported as rare items (Sazima 1983; Nico & Taphorn 1998). As behavioral approach, it may stalks, use clumps or watersheds as covers to ambush its preys or yet linger around the prey and attack from close quarters (Sazima 1983). In the Orinoco basin, it is considered unusual, with preferential habitat reported as dark water with abundant vegetation and clear water, and spawning at the beginning of the rainy season, dispersing eggs on the aquatic vegetation (Taphorn 2003).

Conservation status. *Catoprion mento* is relatively common throughout the area of occurrence and widespread along large drainage basins of central and northwestern portions of South America. No specific threats were detected. Accordingly, we suggest that the species remain classified as Least Concern (LC) according to IUCN criteria (IUCN 2018) as in the last evaluation performed by the ICMBio.

Remarks. In the original description of *Serrasalmus mento,* the type locality was assigned as Brazil although brought from Lisbon ("venu de Lisbonne, paroît aussi originaire du Brèsil" - Cuvier 1819). Both the MNHN database and the label of the type assign the specimens to "Cabinet d'Ajuda", a collection from Museu da Ajuda (Lisbon, Portugal) known for been ransacked during the Napoleonic Era (Myer 1950; Vanzolini 2004). The collection, made by Alexandre Rodrigues Ferreira during his traveling to Brazil at the behest of the Portuguese Crown, presents severe doubts about the exact source of specimens, which did not have its origins indicated at the drawings or records at the Museu d'Ajuda (Vanzolini 2004). Finding the locality of the type has being a hard work since no other record was found. However, meristic data taken directly to the type indicate about 80 perforated scales in the lateral line (counting is approximate due the really poor conditions of the type), which match the counts of specimens from the upper Paraguay, Araguaia and right tributaries of the Amazon basin. Thus, we conclude that *Serrasalmus mento* was most likely to be collected in one of these abovementioned localities and thus, the other genetic lineage and morphologically distinct species should be considered an undescribed species.

In the "Poissons Characoides des Guyanes", Géry (1972) did not use any specimen from the Guianas, so the counts and measures were taken from specimens from the Bolivian Amazon. His counting confirms the identification of *Catoprion mento* (80 perforated scales in the lateral line) and its occurrence cannot be confirmed for the Guianas. Although we could

not examine specimens from the Bolivian Amazon, we consider the species occurring there as *Catoprion mento* based on Géry's (1972) counting and based on the confirmed occurrence of this species in the middle portions of the Madeira and Purus rivers.

Mylesinus macropterus was described based on a single specimen by Ulrey (1894), and then based on two specimens by Ulrey (1895), with both descriptions assigning it to "Brazil". The Ulrey (1895)'s paper was based on fishes collected by Charles Frederick Hartt in Brazil, during the "Thayer Expedition". Since the expedition achieved the area of occurrence of both *C. mento* and *Catoprion* sp. n. (Agassiz 1868; Hartt 1870) and the counts present in the original description indicates about 83 perforated scales in the lateral line (type whereabouts unknown), *Mylesinus macropterus* Ulrey, 1894 must remain a junior synonym of *Catoprion mento* (Cuvier, 1819).

Material examined. Types. MNHN A.9869, 1 (rd), 79.8 mm SL, holotype of *Serrasalmus mento*, Brazil. CUMV 3267, 1 (rd), holotype of *Mylesinus macropterus*, whereabouts unknown, Brazil, Jan 1860, C. Hartt. **Non-types. Brazil. Amazon basin. Amazonas.** MZUSP 9570, 2, 90.0–102.0 mm SL, fish market, Manaus, 17 Sep 1968, Expedição Permanente da Amazônia. **Rio Araguaia basin. Goiás.** MZUSP 54525, 1, 76.7 mm SL, rio Araguaia, 1997–1998, Coleção rio Araguaia. MZUSP 89136, 1 (rd), 50.3 mm SL, Aruanã, lake at margin of Araguaia river, 14°39'23"S 50°54'03"W, 25 Jul 2005, CBE team. **Pará.** MZUSP 20574, 2, 57.0–70.9 mm SL, Jurunundéua lake (trib. rio Capim), 19 Aug 1970, Expedição Permanente da Amazônia. **Rio Madeira basin. Amazonas.** INPA 33717, 1, 92.3 mm SL, Apuí, rio Guariba (trib. rio Aripuanã), 08°42'42"S 60°25'53"W, Pedroza WS *et al.* **Mato Grosso.** MZUSP 37511, 2, 46.4–104.2 mm SL, Vila Bela da Santíssima Trindade, rio Guaporé (trib. Rio Mamoré), 28 Sep 1984, Garavelo JC & Polonoroeste team. MZUSP 37664, 2, 27.7–28.6 mm SL, Vila Bela da Santíssima Trindade, rio Guaporé (trib. rio Mamoré), 10 Oct 1984, Garavelo JC & Polonoroeste team. MZUSP 64951, 1, 97.5 mm SL, Vila Bela da Santíssima Trindade, rio Guaporé (trib. Rio Mamoré), 15°01'17"S 59°58'30"W, Oct 1997, Machado FA *et al*. MZUSP 77203, 2, 84.7–106.5 mm SL, Panelas, rio Roosevelt, 17 Jul 1997, Machado FA *et al.* MZUSP 95321, 1, 83.9 mm SL, Vila Bela da Santíssima Trindade, rio Guaporé (trib. Rio Mamoré), 15°01'37"S 59°49'09"W, 13 Oct 2006, Machado FA *et al*. MZUSP 115655, 1, 40.9 mm SL, Vila Bela da Santíssima Trindade, rio Guaporé (trib. rio Mamoré), 15°00'18"S 59°57'19"W, 28 Aug 2013, Oyakawa O *et al*. **Rondônia.** UFRO–ICT 7501, 2, 87.3–119.5 mm SL, Jaciparaná, Madalena lake (trib. rio Madeira), 09°17'03"S 64°23'43"W, 25 Nov 2011, Matsuzaki A. **Rio Negro basin. Amazonas.** MZUSP 6717, 5, 64.8–68.6 mm

SL, Manaus, rio Negro, 23 Nov 1967, Expedição Permanente da Amazônia. MZUSP 43350, 1, 90.9 mm SL, Paricatuba, lake at margin of rio Negro, 11 Nov 1972, Expedição Permanente da Amazônia. MZUSP 54517, 2, 43.4–57.2 mm SL, Paricatuba, lake at margin of rio Negro, 11 Nov 1972, Expedição Permanente da Amazônia. **Rio Paraguay basin. Mato Grosso.** LBP 7556, 4, 57.4–74.0 mm SL, Barão de Melgaço, lake at margin of rio Cuiabá, 16°11'39"S 55º48'25"W, 29 Jan 2009 Oliveira C *et al.* LBP 26133, 4 (rd), 30.1–30.8 mm SL, Pantanal, rio Paraguay, 16°27'02.6" S 55°19'12.6" W, Flausino Jr N *et al*. MZUSP 19941, 1, 64.9 mm SL, Itiquira, baía Grande, rio Itiquira (trib. rio Piquiri), 29 Oct 1978, Oliveira JC. MZUSP 27274, 6 (1 cs), 50.9–64.9 mm SL, Poconé, rio Paraguay, 11 May 1983, Sazima I *et al.* MZUSP 35874, 1, 42.1 mm SL, Itiquira, rio Itiquira, 30 Sep 1979, Medeiros JHB & Oliveira, JC. MZUSP 90217, 1, 39.0 mm SL, Cáceres, rio Sepotuba, 15°47'33"S 57°39'20"W, 2 Mar 2002, Britski HA *et al.* MZUSP 75237, 1, 48.0 mm SL, Itiquira, lake between rivers Piquira and Itiquira, Oct 1984, Medeiros JHB & Oliveira, JC. MZUSP 90282, 4, 26.8–38.4 mm SL, Cáceres, rio Sepotuba, 15°46'07"S 57°38'54"W, 4 Mar 2002, Britski HA *et al.* MZUSP 95079, 1 (rd), 60.6 mm SL, Barão de Melgaço, rio Mutum, 16º19'30"S 55º49'59"W, 30 Sep 2006, Machado FA & Lima FCT. MZUSP 96679, 1, 54.5 mm SL, Barão de Melgaço, Pantanal de Paiaguás, 16°17'00"S 55º48'00"W, 30 Sep 2006, Machado FA & Lima FCT. NUP 1044, 5, 37.6–108.2 mm SL, Chapada dos Guimarães, Reservatório Manso, 14º52'S 55º47'W, 2000–2004, Nupélia. **Mato Grosso do Sul.** LBP 3820, 1 (rd), 55 mm SL, Aquidauana, rio Negro, 19º34'33"S 56º14'49"W, 01 Aug 2006, Oliveira C. *et al.* MZUSP 36412, 2, 60.6–74.9 mm SL, Nhecolândia, Corumbá, baía dos Búfalos, Mar 1985, Mourão GM & Bastos EK. MZUSP 36417, 1, 42.4 mm SL, Nhecolândia, Corumbá, faz. Nhumirim, 24 May 1985, Bastos EK. MZUSP 36418, 4, 29.6–38.1 mm SL, Nhecolândia, Corumbá, farm Nhumirim, May 1985, Mourão GM & Bastos EK. MZUSP 48306, 2, 39.7–51.8 mm SL, Pantanal de Paiaguás, Farm Santo Antônio, Liparelli T. **Rio Purus basin. Amazonas.** LBP 11127, 13 (rd), 24.2–41.0 mm SL, Lábrea, rio Purus, 07º29'34"S 63º35'03"W, 24 Aug 2010, Oliveira C *et al*. **Rio Solimões basin. Amazonas.** MZUSP 5878, 1, 97.1 mm SL, Manacapuru, Manacapuru lake (trib. rio Solimões), 27 Mar 1967, Expedição Permanente da Amazônia. MZUSP 6884, 2, 67.1–106.3 mm SL, Manaus, Januari lake (trib. rio Solimões), 20 Nov 1967, Expedição Permanente da Amazônia. **Colombia. Río Orinoco basin. Casanare.** CZUT–IC 9675, 1 (rd), 36.5 mm SL, río Meta, 05°15'31"N 70°43'27"W. **Meta.** MNHN 2007 0229, 1, 78.6 mm SL, Puerto López, Caño estero (trib. río Meta), 04º15'00"N 72º30'00"W, 04 Apr 1995, Pedreros SP. **Vichada.** CZUT–IC 8982, 3 (2) (rd), 30.3–30.4 mm SL, Santa Rosalia, Laguna La Portuguesa (trib. río

Vichada), 05º11'49"N 70º44'09"W. **Venezuela.** MNHN 87 792, 1, 45.7 mm SL, río Orinoco, Chaffanjon. MNHN 87 793, 1, 50.9 mm SL, río Orinoco, Chaffanjon.

***Catoprion* sp. n.**

Figs. 7–8; Table 4

Catoprion mento (non Cuvier, 1819). —Müller & Troschel, 1844: 96 [description]. —Müller & Troschel, 1845: 22 [description]. —Eigenmann, 1910: 442 [*partim*; listed]. —Eigenmann, 1912: 387 [British Guiana; brief description]. —Fowler, 1914: 251 [Rupununi river, British Guiana; listed]. —Lowe–McConnell, 1964: 142 [Rupununi river, British Guiana; listed]. —Machado *et al*. 2018: 7 [Barcoding of Serrasalmidae].

Figure 7. *Catoprion* sp. n., MZUSP 50015, holotype, male, 93.7 mm SL, Brazil, Amazonas, Presidente Figueiredo, rio Uatumã, about 500m downstream UHE Balbina.

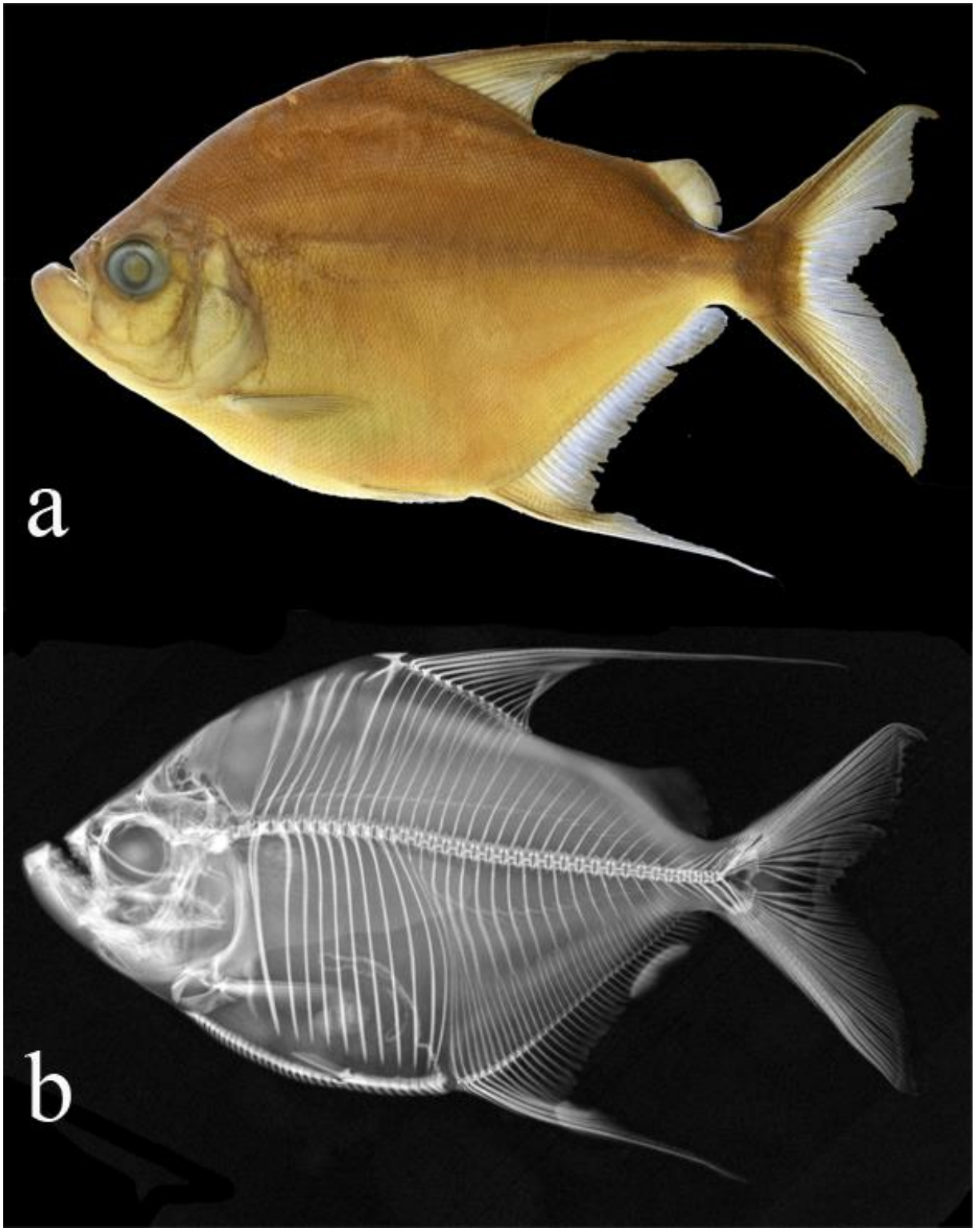

Figure 8. *Catoprion* sp. n., MZUSP 113015, paratype, 78.3 mm SL, Brazil, Roraima, Caracaraí, igarapé do Campo, tributary of rio Jufari, rio Branco, Amazon basin. **a.** preserved specimen; **b.** radiography.

Holotype. MZUSP 50015, 93.7 mm SL, male, Brazil, Amazonas, Presidente Figueiredo, rio Uatumã, about 500m downstream the UHE Balbina, 01º55'12"S 59º28'22"W, 21 Jul 1995, Lima FCT & UHE Balbina team.

Paratypes. Brazil. Rio Branco basin. Roraima. LBP 4342, 1, 75,0 mm SL, Boa Vista, igarapé do Cajual, rio Uraricoera, 03º11'00"N 60º33'20"W, 11 Nov 2006, Devidé R *et al.* LBP 15582, 1, 109.1 mm SL, Bonfim, igarapé Chidaua, rio Tacutu, Rio Branco, 03º18'28"N 59º53'19"W, 23 Apr 2012, Britzke R & Melo BF. **Rio Negro basin. Amazonas.** INPA 37862, 1, 100.3 mm SL, Santa Isabel do rio Negro, rio Aiuanã, 00º35'24"S 64º55'10"W, 04 Apr 2012, Silva R *et al.* **Roraima.** MZUSP 113015, 2 (rd), 70.3–78.3 mm SL, Caracaraí, igarapé do Campo, rio Jufari, rio Branco, 01º03'54"S 62º07'46"W, 4 Sep 2011, Oyakawa OT *et al*. **Rio Nhamundá basin. Amazonas.** INPA 46313, 1, 131.2 mm SL, Nhamundá, rio Nhamundá, 01º40'30"S 57º28'36"W, 13 Nov 2013, Machado V. *et al.* INPA 46314, 2, 124.3–125.6 mm SL, Nhamundá, rio Nhamundá, 01º41'27"S 57º25'20"W, 11 Nov 2013, Machado V. *et al*. **Rio Uatumã basin. Amazonas.** INPA 37246, 5 (1), 123.5 mm SL, São Sebastião do Uatumã, rio Jatapu, 02º01'03"S 58º10'26"W, 01 Oct 2011, Rapp Py–Daniel L. *et al.* **Rio Xingu basin. Pará.** INPA 47302 (rd), 1, 113.1 mm SL, Porto de Moz, rio Xingu, 01º46'30"S 52º12'57"W, Sabaj M *et al.*

Non-types. Brazil. Rio Amazonas basin. Amazonas. MZUSP 7328, 1, 75.2 mm SL, Maués, Igarapé at rio Marau, 03 Dec 1967, Expedição Permanente da Amazônia. **Pará.** MZUSP 9178, 3, 82.5–94.5 mm SL, Santarém, rio Maicá, 19 Oct 1971, Expedição Permanente da Amazônia. MZUSP 9507, 2, 99.1–100.1 mm SL, Monte Alegre, rio Amazonas, 14 Aug 1968, Expedição Permanente da Amazônia. **Rio Branco basin. Roraima.** INPA 35670, 2 (rd), 102.6–102.7 mm SL, Caracaraí, rio Anaua, rio Jauari, 00º57'12"N 61º06'49"W, 16 Apr 2007. LBP 15148, 3 (rd), 30.2–37.1 mm SL, Boa Vista, rio Branco, 03º08'16"N 60º16'33"W, 20 Apr 2012, Britzke R & Melo BF. LBP 15534, 22 (21 rd, 1 cs), 31.6–54.3 mm SL, Bonfim, Lago Fazenda Romer, Rio Tacutu, 03º22'55"N 59º51'28"W, 22 Apr 2012, Britzke R & Melo BF. LBP 21608, 1 (rd), 47 mm SL, Bonfim, Fazenda Romer lake, rio Tacutu, 03º22'55"N 59º51'28"W, 22 Apr 2012, Britzke R. & Melo BF. LBP 21615, 5, 19.9–47.0 mm SL, Boa Vista, Igarapé Au-Au, 02º56'19"N 61º03'06"W, 26 Apr 2012, Britzke R & Melo BF. MZUSP 20246, 11, 24.4–79.0 mm SL, Boa Vista, lake at Tepequem road, 10 Feb 1969, Roberts T. **Rio Japurá basin. Amazonas.** MZUSP 36162, 1, 110.7 mm SL, Igarapé Cacar, lago Amanã near mouth of rio Japurá, 06 Sep 1979, Barthem R. **Rio Negro basin. Amazonas.** INPA

52533, 1, 96.2 mm SL, São Gabriel da Cachoeira, 00º07'13"S 67º06'07"W, 02 Feb 2014, Machado V. *et al.* **Roraima.** MZUSP 112544, 1, 126.6 mm SL, Caracaraí, igarapé do Campo (trib. Rio Jufari), 01º04'01"S 62º07'40"W, 28 Aug 2011, Oyakawa O *et al.* MZUSP 113085, 1, 112.1 mm SL, Caracaraí, igarapé Santa Fé (trib. rio Jufari), 01º00'34"S 62º13'04"W, 01 Sep 2011, Oyakawa O *et al.* MZUSP 113599, 1, 34.9 mm SL, Caracaraí, praia do Paricá, rio Jufari, 01º08'41"S 61º59'57"W, 10 Sep 2011, Oyakawa O *et al.* **Rio Tapajós basin. Pará.** LBP 15054, 1, 97.7 mm SL, Itaituba, rio Tapajós, 04º34'07"S 56º18'49"W, 10 Jun 2012, Britzke R. *et al.* MZUSP 8451, 8 (1 cs), 44.3–102.5 mm SL, Alter do Chão, igarapé Jacundá, 02º30'00"S 54º57'00"W, Expedição Permanente da Amazônia. MZUSP 57576, 1, 111.9 mm SL, Santarém, lake near Alter do Chão, 02º28'05"S 54º55'34"W, Westneat M. *et al.* **Rio Trombetas basin. Pará.** INPA 46254, 2 (1), 119.6 mm SL, Porto Trombetas, igarapé Água Fria, 01º27'56"S 56º49'42"W, 16 Nov 2013 Soares I. *et al.* MZUSP 5439, 2, 95.7–98.4 mm SL, Oriximiná, rio Trombetas, Feb 1967, Expedição Permanente da Amazônia. MZUSP 5512, 1, 99.2 mm SL, Oriximiná, lago Jacupã, Feb 1967, Expedição Permanente da Amazônia. MZUSP 8210, 1, 100.0 mm SL, Oriximiná, lake Jacupã, 17 Dec 1967, Expedição Permanente da Amazônia. **Rio Uatumã basin. Amazonas.** INPA 37419, 1, 144.7 mm SL, São Sebastião do Uatumã, igarapé Três, rio Jatapu, 02º00'07"S 58º11'35"W, 27 Sep 2011, Rapp Py-Daniel L. *et al.* MZUSP 9546, 1, 70.5 mm SL, São Sebastião do Uatumã, 08 Sep 1968, Expedição Permanente da Amazônia.

Diagnosis: *Catoprion* sp. n. differs from *Catoprion mento* by presenting 86–94 perforated scales in the lateral line (*vs.* 65–86 scales) and 35–40 circumpeduncular scales (*vs.* 29–34 scales).

Description. Morphometric data presented in Table 4. Body deep, compressed laterally; general shape rhomboid. Highest body depth at vertical line through dorsal–fin origin. Dorsal profile slightly concave along head and convex between head and dorsal-fin origin. Dorsal profile of the body fairly sloping at the base of the dorsal and adipose fins, and straight at interval between these two fins. Ventral profile convex from isthmus to end of anal fin. Ventral keel with a serra of spines. Pre-pelvic spines 19 (3), 20 (3), 21 (9), 22 (10), 23* (21), 24 (15), 25 (6), 26 (2) or 27 (1). Post-pelvic spines 8 (1), 9* (22), 10 (30), 11 (15) or 12 (3). Anal spines 1 (9), 2* (53) or 3 (8). Caudal peduncle deeper than long.

Eyes lateral at middle of head; upper margin of eyes above longitudinal line at lateral-line origin. Frontal and parietal fontanels expanded laterally. Snout very short and slightly

pointed in lateral view. Nostrils dorsolateraly positioned, at longitudinal axis through upper margin of eye. Mouth upturned. Jaw strongly prognathous. Premaxillary and dentary teeth mamiliform, with a rounded base and narrower tip. Premaxilla with five teeth, two large forward-directed teeth and three smaller teeth, in an irregular single row. Dentary with a single row of six teeth. Maxilla edentulous.

Body completely covered by small cycloid scales. Lateral line complete, with 86 (4), 87 (6), 88 (8), 89* (10), 90 (7), 91 (9), 92 (6), 93 (7) or 94 (2) perforated scales; first to tenth scale larger than remaining. Scales series above lateral line 32 (1), 34 (1), 35 (1), 36 (5), 37 (6), 38 (7), 39 (5) or 40 (4). Scales below lateral line 26 (1), 27 (1), 28 (2), 29 (3), 30 (2), 31 (7), 32 (7) or 33 (3). Circumpeduncular scales 35 (3), 36 (9), 37 (10), 38* (18), 39 (6) or 40 (3).

Dorsal fin preceded by spine, with origin approximately equidistant from tip of snout and end of hypural plate. First dorsal-fin rays prolonged; in some specimens surpassing caudal-fin base. Branched dorsal-fin rays gradually decreasing in size, branched dorsal-fin rays 13 (4), 14 (20), 15 (34) or 16* (13). Adipose fin short, longer than deep. Pectoral fin falcate, with i + 11 (4), 12 (17), 13 (26), 14* (20) or 15 (1) rays. Pelvic fin small, with i + 6* (62) branched rays. First anal-fin rays prolonged; in some specimens surpassing caudal-fin base. Branched anal-fin rays gradually decreasing in size; branched anal-fin rays 32 (4), 33 (19), 34 (26), 35* (12) or 36 (3). Caudal fin bifurcated, lobes of similar size, with 17* (60) branched rays.

First gill arch with elongated gill rakers, almost as long as filaments, decreasing in size towards the extremities; 6 (2), 8 (20), 9 (18) or 10 (3) gill rakers on upper limb, 11 (6), 12 (19), 13 (19) or 14 (2) gill rakers on lower limb. Lower gill rakers longer than upper. Vertebrae 36 (5), 37 (23) or 38 (2). Supraneurals 4 (29) or 5 (1).

Table 4. Morphometric data of *Catoprion* sp. n. H = holotype; N = number of specimens; SD = standard deviation. Range includes the holotype.

	H	N	Range	Mean	SD
Standard length (mm)	93.7	72	30.2–144.7	72.5	–
Percentages of standard length					
Adipose-fin base	10.3	72	6.1–12.8	10.2	0.01
Anal-fin base	42.5	72	33.7–45.8	40.1	0.03
Anal-fin length	32.6	60	17.9–56.6	32.4	0.09
Body depth	68.4	72	47.3–75.4	60.2	0.08
Caudal-peduncle depth	9.9	72	8.2–11.2	09.0	0.01
Caudal-peduncle length	8.6	72	6.4–09.7	07.9	0.01
Dorsal-fin base	23.1	72	17.2–24.9	21.6	0.02
Dorsal-fin length	55.9	71	24.6–71.0	42.0	0.13
Head length	30.0	72	27.4–33.1	29.8	0.01
Interdorsal length	18.1	72	10.8–18.9	15.4	0.01
Pectoral-fin length	21.1	70	15.0–24.8	19.6	0.02
Pelvic-fin length	14.2	68	11.6–17.6	14.2	0.01
Postdorsal length	60.8	72	50.2–63.6	56.5	0.03
Predorsal length	55.2	72	53.1–60.9	56.9	0.02
Percentages of head length					
Interorbital width	37.7	56	27.1–39.4	33.3	0.03
Maxilla length	34.2	72	27.8–47.5	37.8	0.04
Orbital diameter	39.2	72	24.5–41.0	33.9	0.05
Snout length	25.6	72	18.0–27.7	22.9	0.02

Color in alcohol. General body color brownish–yellow, darker in dorsal portion. Conspicuous brown or black band on the end of caudal peduncle. Black or brown humeral blotch vertically elongated, sometimes inconspicuous or absent. Fins hyaline to yellowish. Caudal-fin base dark brown; in V–shape with the apex directed towards the head. First rays of dorsal fin black or brown. First rays and distal margin of anal fin often dark. Vertical, dark band through the eye (Figs. 7 and 8).

Color in life. General body color silvery, darker dorsally. Conspicuous orange blotch on opercle. Dorsal, pectoral, pelvic and adipose fins hyaline. First ray of dorsal fin often dark grey. Anal fin yellow to orange, mainly in the anterior rays, gradually turning to hyaline on posteriormost rays. Caudal fin base black, in V–shape with the apex directed towards the head. Vertical dark band through the eye.

Geographic distribution. *Catoprion* sp. n. occurs in the Amazon basin, including the Branco, Negro, Uatumã, Nhamundá, Trombetas, Tapajós, and Xingu drainages (Fig. 2). *Catoprion* sp. n. probably occurs in the Essequibo basin, Guyana, based on the fact that former studies have identified *C. mento* for this basin (Eigenmann 1912; Fowler 1914; Lowe-McConnell 1964; Vari *et al.* 2009).

Sexual Dimorphism. *Catoprion* sp. n. presents anterior lobe developed from the elongation of distal border of the anal fin in males, between about the 7 and 13th branched rays (Fig. 9), whereas females present only the firsts rays of anal fin elongated. Vieira & Géry (1979) consider the extension of the dorsal fin as exclusive to males. Here, we analyzed both females and males with elongated dorsal and anal fins that extend to the caudal fin, which led us to conclude that the elongation of fins are not related to sexual dimorphism.

Ecological notes. *Catoprion* sp. n. has a lepidophagous behavior, with adults feeding preferentially on scales and juveniles also feeding on insects (Vieira & Géry 1979).

Conservation status. *Catoprion* sp. n. is widespread along major drainages of the Amazon basin, where it is relatively common where it occurs. No specific threats were detected on those immense regions, which suggest that the species can be categorized as Least Concern (LC) according to IUCN criteria (IUCN 2018).

Remarks. We were unable to examine specimens of *Catoprion* from the Essequibo basin, but data presented by Eigenmann (1912) indicate 89–94 scales in the lateral line in four examined specimens from Rockstone and Konawaruk rivers, which fit the herein determined counts of *Catoprion* sp. n. The new species, *Catoprion* sp. n. is very similar to *C. mento* in body shape, color pattern, sexual dimorphism and ecological features, being the reason of why the two species were treated as conspecific for a long time.

Discussion

An integrative approach was performed herein to investigate species limits of *Catoprion*. This integrative analysis of morphological and molecular features has been proven as highly informative and valuable to study the lineages of Serrasalmidae flagging new candidate species and improving description and discussion of species boundaries (Mateussi *et al.* 2016; 2018; Andrade *et al.* 2017).

Previous authors had already hypothesized the existence of more than one species within *Catoprion mento* (Taphorn 2003) but without examining specimens from a wide range of localities or a complete analysis of morphological features that would allow them to diagnose or describe other candidate species. Although previous molecular studies had used specimens of *Catoprion*, the lack of extensive taxon sampling from several South American regions did not allow them to discriminate distinct genetic lineages/species (Machado *et al.* 2018). Herein, the combined analysis of about 50 morphometric and meristic characters of 178 specimens of *Catoprion* from the entire distribution of the genus except the Essequibo basin, in addition to a molecular analysis of mitochondrial data allow us to unambiguously discriminate two distinct species, *C. mento* from Paraguay, Araguaia, Madeira, Purus and Orinoco, and *Catoprion* sp. n from northern portions of the Amazon basin, plus the Tapajós and Xingu rivers.

Catoprion mento and *Catoprion* sp. n. are very similar in body shape, coloration, and ecologically behavior (lepidophagy) but they differ in lateral line and circumpeduncular scale counts, which are slightly larger (thus fewer) in *C. mento* than those in *Catoprion* sp. n. Moreover, molecular data strongly supports the existence of two species based on three modern species delimitation analyses: PTP, ABGD and GMYC, in addition to 7.3% of pairwise K2P genetic distance. Traditionally, a 2% cutoff of interspecific distance has been used, at least initially, in species delimitation based on the K2P genetic distance (Carvalho *et al.* 2011; Pereira *et al.* 2013). This mean value is concordant with mean values found in other serrasalmids, *e.g.* 1.4–9.0% in *Mylossoma* (Mateussi *et al.* 2016), 2.7–6.2% in *Pygocentrus* (see Chapter 2) and 1.6–9.3% in *Tometes* (Andrade *et al.* 2017).

Our results reinforce the integrative approach of molecular and morphological methods in taxonomy of Neotropical freshwater fishes as a powerful way to describe new species. This represent a continuation of combined-evidence studies with the diversity of

Serrasalmidae complemented with description of new taxa (Mateussi *et al*. 2016, 2018) and new efforts might reveal similar results with other serrasalmid genera.

Comparative material examined. *Acnodon senai*, MNHN 1989 312, paratypes, 4, 45–82 mm SL, rio Jari, Pará, Brazil. *Colossoma macropomum*, UFRO-ICT 11118, 1, 152 mm SL, rio Jaciparaná, Jaciparaná, Rondônia, Brazil. *Metynnis guaporensis*, UFRO-ICT 5396, 1, 111 mm SL, igarapé Jatuarana, Porto Velho, Rondônia, Brazil. *Metynnis maculatus*, 1, UFRO-ICT 7528, 1, 122 mm SL, lake Cuniã, Porto Velho, Rondônia, Brazil. *Mylesinus paraschomburgkii*, MNHN 1987 1403, paratypes, 2, 270–295 mm SL, rio Cachorro, rio Trombetas, Pará, Brazil. *Myleus setiger*, LBP 24509, 4, 171–174 mm SL, rio Jamanxim, Tapajós, Novo Progresso, Pará, Brazil. *Myloplus arnoldi*, MNHN 2017 0251, 2, 127–151 mm SL, rio Xingu, Altamira market place, Pará, Brazil. *Myloplus planquettei*, MNHN 2001 1226, paratype, 400 mm SL, riv. Litany, Saut Loue, French Guyana. *Mylossoma aureum*, LBP 21832, 2, 103–109 mm SL, Catalão, rio Negro, Manaus, Amazonas, Brazil. *Mylossoma albiscopum*, LBP 18175, 3 (rd), 59–71 mm SL, rio Solimões, Manacapuru, Amazonas, Brazil. *Ossubtus xinguense*, MNHN 1992 0004, 2 paratypes, 149–164 mm SL, rio Xingu, Altamira, Pará, Brazil. *Pristobrycon calmoni*, MNHN 2007 0225, 1, 196 mm SL, Caño Laguna Catagena de Chaira, rio Caguán, Cartagena de Chaira, Caquetá, Colombia. *Pygocentrus cariba*, LBP 2229, 36–50 mm SL, Punta Brava, río Orinoco, Caicara del Orinoco, Bolívar, Venezuela. *Pygocentrus nattereri*, LBP 7785, 3 (rd), 101–108 mm SL, lake Boca Franca, rio Araguaia, Cocalinho, Mato Grosso, Brazil. *Pygopristis denticulata*, MNHN 2017 0242, 2, 171–187 mm SL, Paraná–Maxiparaná, rio Xingu, Brazil. *Serrasalmus altispinnis*. MNHN 1997 113, 1 paratype, 133 mm SL, rio Quarenta Ilhas, rio Uatumã, Amazonas, Brazil. *Serrasalmus compressus*, MNHN 1986 0615, holotype, 110 mm SL, Laguna Mocovi, Bolivia. *Serrasalmus rhombeus*, UFRO-ICT 5498, 4, 83–90 mm SL, igarapé Jatuarana, rio Madeira, Porto Velho, Rondônia, Brazil. *Tometes lebaili*, MNHN 2001 1212, 1 paratype, 139 mm SL, riv. Litany, village Antecume Pata, French Guyana. *Tometes mahkwe*, MNHN 2001 1231, 2 paratypes, 157–159 mm SL, fleuve Maroni, village Antecume Pata, French Guyana. *Utiaritichthys* sp. MNHN 1991 0704, 1, 193 mm SL, fleuve Sinnamony, French Guyana.

Bibliography

Agassiz L. (1868) *A journey in Brazil*. 4th ed. Boston: Ticknor and Fields. 540 pp.

Andrade M.C., Machado V.N., Jégu M., Farias I.P. & Giarrizzo T. (2017) A new species of *Tometes* Valenciennes 1850 (Characiformes: Serrasalmidae) from Tocantins-Araguaia River Basin based on integrative analysis of molecular and morphological data. *PLoS ONE* 12, 1–21.

Carvalho D.C., Oliveira D.A.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, 80–86.

Cione A.L., Dahdul W.M., Lundberg J.G. & Machado-Allison A. (2009) *Megapiranha paranensis*, a new genus and species of Serrasalmidae (Characiformes, Teleostei) from the upper Miocene of Argentina. *Journal of Vertebrate Paleontology* 29, 350–358.

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, 1–12.

Edgar R.C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32, 1792–1797.

Eigenmann C.H. (1910) Catalogue of the freshwater fishes of tropical and south temperate America. In: Reports of the Princeton University Expeditions to Patagonia 1896–1899. *Zoology* 3(4), 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. *Memoirs of the Carnegie Museum* 5, i-xxii + 1-578, Pls. 1-103.

Freeman B., Nico L.G., Osentoski M., Jelks H.L. & Collins T.M. (2007) Molecular systematics of Serrasalmidae: Deciphering the identities of piranha species and unraveling their evolutionary histories. *Zootaxa* 38, 1–38.

Fricke R., Eschmeyer W.N. & Fong J.D. (2019) *Species by Family/Subfamily*. (http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version accessed 09 Jan 2019.

Fowler H.W. (1914) Fishes from Rupununi river, British Guiana. *Proceedings of the Academy of Natural Sciences of Philadephia* 66: 229–284.

Géry J. (1977) *Characoids of the world*. T.F.H. Publications, Inc., Neptune city, N.J. 672 pp.

Géry J. (1976) Les genres de Serrasalmidae (Pisces, Characoidei). *Bulletin Zoölogisch Museum, Universiteit Van Amsterdam* 5, 47–54.

Géry J. (1972) Poissons Characoïdes des Guyanes. I. Généralités. II. Famille des Serrasalmidae. *Zoologische Verhandelingen* 122, 1–250.

Gosline W.A. (1951) Notes on the characid fishes of the subfamily Serrasalminae. *Proceedings of the California Academy of Sciences* 27, 17–64.

Goulding M. (1980) *The fishes and the forest: explorations in Amazonian natural history*. University of California Press, Berkeley and Los Angeles, California.

Hartt C.F. (1870) *Thayer Expedition*. Boston: Fields, Osgood & Co. 620 pp.

IUCN (2018). *The IUCN Red List of Threatened Species*. Version 2018-1. <http://www.iucnredlist.org>. Downloaded on 05 July 2018.

Jégu M. (2003) Subfamily Serrasalminae. In: *Check List of the Freshwater Fishes of South and Central America*. (Eds R.E. Reis, S.O. Kullander & C.J. Ferraris), pp. 185–199. EDIPUCRS, Porto Alegre.

Jégu M. (2004) *Taxonomie des Serrasalminae phytophages et Phylogenie des Serrasalminae (Teleostei: Characiformes: Characidae).* Museum National d'Histoire Naturelle. 421 pp.

Jégu M. & Ingenito L.F.S. (2007) Família Characidae: Serrasalminae. In: *Catálogo das espécies de peixes de água doce do Brasil*. (Eds Buckup P.A., Menezes N.A. & Ghazzi M.S.), pp. 40–43. Museu Nacional, Rio de Janeiro.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P. & Drummond A. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 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, 111–20.

Kner, R. (1859) Zur Familie der Characinen. III. Folge der Ichthyologischen Beiträge. *Denkschriften der Kaiserlichen Akademie der Wissenschaften in Wien, Mathematisch-Naturwinssenschaftkiche Classe* 17: 137–182.

Kumar S., Stecher G., Li M., Knyaz C. & Tamura K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* 35, 1547–1549.

Lanfear R., Calcott B., Ho S.Y.W. & Guindon S. (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular biology and evolution* 29, 1695–701.

Lowe-McConnell, R.H. (1964) The fishes of the Rupununi savanna district of British Guiana,

South America. Part 1. Ecological groupings of fish species and effects of the seasonal cycle on the fish. *Journal of Linnean Society (Zoology)* 45(304), 103–144.

Machado-Allison, A. (1982) *Studies on the systematics of the Subfamily Serrasalminae (Pisces – Characidae).* The George Washington University, Washington DC, USA.

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, 1–12.

Mateussi N.T.B., Oliveira C. & Pavanelli C.S. (2018) Taxonomic revision of the Cis-Andean species of Mylossoma Eigenmann & Kennedy, 1903 (Teleostei: Characiformes: Serrasalmidae). *Zootaxa* 4387, 275–309.

Mateussi N.T.B., Pavanelli C.S. & Oliveira C. (2016) Molecular identification of cryptic diversity in species of cis-Andean Mylossoma (Characiformes: Serrasalmidae). *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis* 28, 778–780.

Müller J. & Troschel F.H. (1844) Synopsis generum et specierum familiae Characinorum (Prodomus descriptionis novorum generum et specierum). *Archiv für Naturgeschichte* 10, 81–99.

Müller J. & Troschel F.H. (1845) Horae Ichthyologicae. *Beschreibung und Abbildung neuer Fische Die Familie der Characinen.* Erstes und Zweites Heft. Viet & Comp., Berlin. 1, 1–40.

Nico L.G. & Taphorn D.C. (1998) Food Habits of Piranhas in the Low Llanos of Venezuela. *Biotropica* 20, 311–321.

Nico L.G., Jégu M. & Andrade, M.C. (2017) Family Serrasalmidae – Piranhas and Pacus. In: *Field guide to the fishes of the Amazon, Orinoco & Guianas*. (Eds Van der Sleen, P. & Albert, J.), pp. 172–195. Princeton University Press, Princeton and Oxford.

Ortí G., Petry P., Porto J.I.R., JEGU M. & Meyer A. (1996) Patterns of nucleotide change in Mitochondrial Ribosomal RNA genes and the Phylogeny of Piranhas. *Journal of Molecular Evolution* 42, 169–182.

Ortí G., Sivasundar A., Dietz K. & Jégu M. (2008) Phylogeny of the Serrasalmidae (Characiformes) based on mitochondrial DNA sequences. *Genetics and Molecular Biology* 31, 343–351.

Ota R.P., Röpke C.P., Zuanon J. & Jégu M. (2013) Serrasalmidae. In: *Peixes do rio Madeira v.II*. (Eds L.J. Queiroz, G. Torrente-Vilara, W.M. Ohara, T.H. da S. Pires, J. Zuanon & C.R. da C. Doria), pp. 14–48. Santo Antônio Energia, São Paulo - Brasil.

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

QGIS Development Team (2019) QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org.

Rambaut A., Drummond A.J., Xie D., Baele G. & Suchard M.A. (2018) Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* 1–5.

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

Sazima I. & Machado F.A. (1990) Underwater observations of piranhas in western Brazil. *Environmental Biology of Fishes* 28, 17–31.

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

Suchard M.A., Lemey P., Baele G., Ayres D.L., Drummond A.J. & Rambaut A. (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* 4, vey016.

Taphorn D.C. (2003) *Manual de Identificación y Biología de los Peces Characiformes de la Cuenca del Río Apure en Venezuela*. BioCentro, Universidad Nacional Experimental de los Llanos Occidentales "Ezequiel Zamora," Guanare, Venezuela. 393 pp.

Taylor W. & VanDyke G. (1985) Revised procedures for staining and clearing small fishes and other vertebrates for bone nd cartilage study. *Cybium* 9, 107–119.

Thompson A.W., Betancur R., López-Fernández H. & Ortí G. (2014) A time-calibrated, multi-locus phylogeny of piranhas and pacus (Characiformes: Serrasalmidae) and a comparison of species tree methods. *Molecular Phylogenetics and Evolution* 81, 242–257.

Ulrey A.B. (1894) Preliminary descriptions of some new South American Characinidae. *American Naturalist* 28, 608–615.

Ulrey A.B. (1895) The South American Characinidae collected by Charles Frederick Hartt. *Annals of the New York Academy of Sciences* 8, 257–300.

Vari R.P., Ferraris Jr. C.J., Radosavljevic A. & Funk V.A. (2009) *Checklist of freshwater fishes of the Guiana Shield*. Bulletin of the Biological Society of Washington, no. 17, 93 pp.

Vieira I. & Géry J. (1979) Crescimento diferencial e nutrição em *Catoprion mento* (Characoidei). Peixe lepidófago da Amazônia. *Acta Amazonica* 9, 143–146.

Ward R.D., Zemlak T.S., Innes B.H., Last P.R. & Hebert P.D.N. (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360, 1847–1857.

Supplementary Material

Appendix 1. Measurements:

1. Anal-fin length – distance between the origin of the first ray of the anal fin and the distal end of its longest ray;
2. Body depth – distance measured vertically between the origin of the dorsal fin and the ventral midline;
3. Caudal-peduncle depth – distance measured vertically in the narrower portion of caudal peduncle;
4. Caudal-peduncle length – distance between the origin of the last ray of the anal fin to the end of the hypural complex;
5. Dorsal-fin length – distance between the origin of the first ray of the dorsal fin and the distal end of its longest ray;
6. Head length – distance between the anterior end of the snout and the posterior margin of the bony operculum;
7. Interdorsal length – distance between the end of the dorsal fin base, measured at the origin of the last dorsal-fin ray, and the beginning of the adipose fin;
8. Interorbital width – smaller distance between orbits;
9. Length of adipose-fin base – length of adipose-fin base;
10. Length of anal-fin base – distance between the origins of the first and last rays of the anal fin;
11. Length of dorsal-fin base – distance between the origins of the first and last rays of the dorsal fin;
12. Maxilla length – distance between the pre-maxillary symphysis and the ventral end of the maxillary;
13. Orbital diameter – distance measured horizontally between the anterior and posterior limits of the orbit;
14. Pectoral-fin length – distance between the origin of the first ray of the pectoral fin and the distal end of its longest ray,
15. Pelvic-fin length – distance between the origin of the first ray of the pelvic fin and the distal end of its longest ray,
16. Postdorsal length – distance between the origin of the first ray of the dorsal fin to the end of the hypural complex;

17. Predorsal length – distance between the origin of the first ray of the dorsal fin to the anterior end of the snout;
18. Snout length – distance between the anterior end of the snout and the anterior limit of the orbit;
19. Standard length – distance between the anterior end of the snout, including the upper lip, and the end of the hypural complex.