UNIVERSIDADE ESTADUAL PAULISTA
INSTITUTO DE BIOCIÊNCIAS
UNESP – BOTUCATU

PROGRAMA DE PÓS-GRADUAÇÃO
CIÊNCIAS BIOLÓGICAS (GENÉTICA)

Estudo da diversidade e das relações filogenéticas do gênero *Astyanax* (Characiformes, Characidae) baseado em sequências de DNA

Bruno César Rossini

Botucatu-SP
2015

UNIVERSIDADE ESTADUAL PAULISTA
INSTITUTO DE BIOCIÊNCIAS
UNESP – BOTUCATU

PROGRAMA DE PÓS-GRADUAÇÃO
CIÊNCIAS BIOLÓGICAS (GENÉTICA)

Estudo da diversidade e das relações filogenéticas do gênero *Astyanax* (Characiformes, Characidae) baseado em sequências de DNA

Bruno César Rossini
Orientador: Prof. Dr. Claudio de Oliveira

Tese apresentada ao Instituto de Biociências, Câmpus de Botucatu, UNESP, para obtenção do título de Doutor no Programa de Pós-Graduação em Ciências Biológicas (Genética)

Botucatu-SP
2015

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

Rossini, Bruno César.
Estudo da diversidade e das relações filogenéticas do gênero Astyanax (Characiformes, Characidae) baseado em sequências de DNA / Bruno César Rossini. - Botucatu, 2015

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu
Orientador: Claudio de Oliveira
Capes: 20204000

1. Peixe - Distribuição geográfica. 2. Peixe - Identificação. 3. Astyanax (Peixe). 4. Filogenia. 5. Análise de DNA.

Palavras-chave: Astyanax; DNA barcode; complexos de espécies; filogenia; peixes neotropicais.

Aos meus pais, que sempre me encaminharam e ajudaram nas melhores escolhas, e à minha esposa, que sempre está ao meu lado

"O trabalho dá cansaço e suor de experiência,
Trabalhar por trabalhar é relaxar a competência..."

Viúva Rica

Tião Carreiro e Pardinho
Compositor: Tião Carreiro / Edward De Marchi

AGRADECIMENTOS

À Deus, que com sua compaixão, reduz todos os obstáculos empre dando força para superar os problemas.

Aos meus pais, Sônia e Antonio, por me ajudarem a chegar até aqui e sempre me apoiarem. em todas as estapas da minha vida.

À Leila, que com seu amor e compreensão fez com que esta tese e o trabalho fossem possíveis.

Ao Prof. Dr. Claudio Oliveira, sem o qual sem dúvida nada estaria escrito aqui. Ao me acolher em seu laboratório aceitando ser meu orientador mesmo não me conhecendo, mas mesmo assim depositando grande expectativas e disponibilidade para ajudar. À sua humildade e serenidade.

Ao meu supervisor, chefe e diretor do Instituto de Biotecnologia da UNESP, Prof. Dr. Celso Luis Marino, que pela compreensão e confiança que sempre me brindou, liberando-me, quando possível, de minhas atividades laborais junto ao Instituto para poder atuar na compsição deste trabalho. Minha gratidão e amizade.

Aos pesquisadores Dr. Vinícius Araújo Bertaco, Dr. Filipe A. G. Melo e ao MSc. Carlos A. M. Oliveira pelas identificações dos exemplares de *Astyanax*, um trabalho extremamente complicado e que sem dúvida estruturou as discussões deste trabalho.

Ao Dr. Juan Martín Días de Astarloa, da Universidad de Mar del Plata, pelas amostras de *Astyanax* da Argentina. À Dra. Ângela M. Zanata pelas amostras da Bahia. E a todos os muitos colegas que compartilharam amostras importantes para esse estudo.

A todos os meus amigos do laboratório, os quais proporcionaram bons momentos de descontração nas pesadas horas de trabalho. Em especial ao Guilherme (Varvito) e Fábio (Fio) pela inegável ajuda nas análises e discussões.

Ao programa de Pós-Graduação em Ciências Biológicas (Genética) que com a particpação de todos os docentes e funcionários contribuiram para o crescimento profissonal.

À CAPES e a FAPESP pelo apoio financeiro (FAPESP processo 2011/00264-2), sem o qual seria impossível a permanência dos estudos.

À equipe dos servidores e docentes do Instituto de Biociências de Botucatu, a qual agora faço parte.

RESUMO

A família Characidae é a mais especiosa entre os peixes da ordem Characiformes. Um desses especiosos gêneros, com muitos conflitos taxonômicos é *Astyanax*, o qual é composto por, atualmente, 142 espécies, das quais 52 foram descritas nos últimos dez anos. Nesse estudo apresentamos um extenso trabalho de amostragem geográfica do gênero *Astyanax*, com exemplares provenientes da América do Sul e Central, cobrindo em sua maioria a distribuição do gênero. Neste contexto foram analisados espécimes do gênero em duas abordagens, uma acerca da diversidade de espécies e outra com foco nas relações filogenéticas do grupo. O estudo de diversidade foi baseado em sequencias *barcode*, do qual a porção 5' do gene Citocromo Oxidase subunidade I, proposta como capaz de separar espécies em nível molecular, auxiliando assim os estudos taxonômicos e a descoberta de novas espécies. O uso de diferentes abordagens para a clusterização de sequências (análises ABGD, GMYC e BIN) mostram uma consistência dos resultados obtidos com o valor inicial de corte de 2%, mas GMYC tende a identificar um número maior de grupos do que as demais análises. Os resultados apontam para a existência de quatro grandes grupos no gênero, totalizando 122 grupos de espécies, mas, em muitos casos, diversas espécies estão agrupadas em único cluster, tornando a identificação por DNA barcode praticamente impossível. Para a filogenia do gênero a análise baseada em dados moleculares multilocus apontam que *Astyanax* é polifilético e com pelo menos quatro linhagens, as quais são formadas pelos complexos de espécies A. scabripinnis e A. fasciatus, outro pelo complexo A. bimaculatus, uma linhagem com os exemplares da América Central e outra com espécies de regiões costeiras do leste do Brasil. Além disso, algumas espécies de *Astyanax* estão relacionadas com gêneros como *Moenkhausia* e *Jupiaba*, conhecidamente similares ao gênero em estudo. Todos os *Astyanax* estão presentes no Clado C, com única exceção de *A. festae*, que ficou no Clado A, e assim propomos sua revisão. Os dados finais deste trabalho apontam para um cenário muito complexo da taxonomia e filogenia de *Astyanax*.

ABSTRACT

The Characidae family is the most species rich between the order Characiformes. One such species rich genera, with many taxonomic conflicts is *Astyanax*, which comprises currently 142 species, of which 52 were described in the last ten years. In this study we present an extensive geographic sampling of the genus *Astyanax* with specimens from South and Central America, covering almost the entire distribution of the genus. In this context, the genus was examined by two approaches, one about the diversity of species and another focusing on the phylogenetic relationships of the group. The diversity study was based on barcode sequences, of which the 5' portion of Cytochrome Oxidase subunit I gene, proposed as standard tool to separate species at the molecular level, thus helping the taxonomic studies and the discovery of new species. The use of different approaches to clustering sequences (ABGD, GMYC and BIN analysis) show a consistency of results obtained with the initial cutoff value of 2%, but GMYC tends to identify a larger number of groups than other analyzes. The results point to the existence of four major groups in the genus, comprising 122 species groups, but in many cases, several species are grouped into a single cluster, making identification by DNA barcode virtually impossible. For the phylogeny analysis of the genus based on multilocus molecular, data shows that *Astyanax* is polyphyletic and at least four lineages, which are formed by the species complexes A. scabripinnis and A. fasciatus, other by complex A. bimaculatus, a lineage with specimens from Central America and other with species from coastal regions of South America. In addition, some species of *Astyanax* are related with genera such as *Moenkhausia* and *Jupiaba*, known to be morpholgical similar. All *Astyanax* are present in Clade C, with only exception of *A. festae*, who was in Clade A, and then we propose its review. Final data from this study point to a very complex scenario of the taxonomy and phylogeny of *Astyanax*.

LISTA DE FIGURAS

Introdução geral

Capítulo 1

Capítulo 2

LISTA DE TABELAS

Introdução geral

Capítulo 1

Capítulo 2

SUMÁRIO

Introdução geral

O número de espécies de vertebrados classificados superava 45 mil há mais de uma década (Pough, 1999) passando a aproximadamente 62 mil espécies nos últimos anos (The World Conservation Union, 2010), dentre os quais o grupo dos peixes até hoje é predominante. Estes possuem tamanhos, formas e habitats variados, ocupando ambientes dulcícolas, estuarinos e marinhos. Dados recentes mostram que estão reconhecidas mais de 32 mil espécies de peixes válidas, dos quais os peixes de água doce correspondem a 47,3%, com uma taxa superior a 150 espécies descritas/ano (Eschmeyer et al., 2010).

Entre os peixes de água doce, ou continentais, as espécies neotropicais são bastante diversas e ocorrem desde a América Central até o sul da América do Sul distribuídas em 71 famílias (Reis et al., 2003), superando 7 mil espécies descritas atualmente (Albert & Reis, 2011). Destes, destacam-se dois grupos muito especiosos, os Characiformes e Siluriformes (Eschmeyer & Fong et al.,2014). Os peixes da ordem Characiformes exibem grande variedade de formas e tamanhos corporais, habitando lagos e rios do continente africano e região neotropical (Géry, 1977). Registros apontam desde espécies com mais de 100 cm de comprimento (*Salminus brasiliensis*; Godoy, 1975) até espécies miniaturas, não ultrapassando 25 mm (Weitzman e Fink, 1983; Weitzman e Vari, 1988). Dentro da ordem Characiformes o grupo mais diversificado é a família Characidae (Malabarba et al., 1998; Reis et al., 2003; Oliveira et al., 2011). Mais de 600 espécies estão distribuídas em 88 gêneros que na maioria dos casos apresenta colocação ainda incerta, o que reflete a baixa

compreensão das relações filogenéticas entre as espécies conhecidas. Muitas espécies são utilizadas para alimentação humana, aquariofilia e atividades de lazer como, por exemplo, a pesca esportiva (Lima et al., 2003). Recentemente, o número de espécies válidas de Characidae subiu para quase 1100, sendo que 203 desta foram descritas nos últimos dez anos (Eschmeyer e Fong, 2014).

A grande diversidade dos peixes neotropicais está associada às mudanças geomorfológicas ocorridas no continente sul-americano. Desde a separação da placa sul-americana da africana há cerca de 118 Ma, destacam-se os eventos de separação de bacias hidrográficas o soerguimento dos Andes (início a ~90 Ma) com a elevação da Cordilheira Leste (separação regiões Cis e Transandinas ~12 Ma), a separação do Orinoco e Maracaibo (~8 Ma) e Orinoco e Amazonas (~10 e 8 Ma). Além disso, incursões e regressões marinhas nas terras baixas do continente, bem como as oscilações no nível do mar e períodos de glaciações causaram diversos eventos vicariantes de espécies que culminaram na grande biodiversidade de peixes (Albert et al., 2006; Albert & Reis, 2011; Lundberg et al., 1998).

O gênero Astyanax

A família Characidae está entre as que apresentam o maior número de problemas taxonômicos e sistemáticos entre os peixes neotropicais. As relações dentro da família permanecem incertas e muitos gêneros têm sua monofilia questionada, dentre os quais está *Astyanax*. O gênero *Astyanax* é um dos mais especiosos da família Characidae e atualmente encontra-se como

Incertae sedis em Characidae (Lima et al., 2003). Este gênero foi descrito por Baird e Girard (1854), a partir da espécie *Astyanax argentatus* (hoje *A. mexicanus*) (Lima et al., 2003; Garavello e Sampaio, 2010). Eigenmann (1921, 1927) define *Astyanax* com uma combinação de caracteres, dos quais duas fileiras de dentes pré-maxilares, cinco dentes na série interna do pré-maxilar, linha lateral completa, nadadeira adiposa presente e nadadeira caudal nua.

Eigenmann (1921, 1927) fez uma revisão do gênero incluindo 74 espécies e subespécies, muitas das quais hoje estão incluídas em outros gêneros, como *Jupiaba*. Somente Géry (1977) estudou extensivamente o gênero depois desta revisão de Eigenmann, embora o mesmo tenha mantido basicamente o proposto por Eigenmann. Até 2003 o gênero possuía 86 espécies válidas (Lima et al., 2003), no entanto, este número já alcança 142 espécies reconhecidas como válidas (Tabela 1), sendo 34 destas descritas nos últimos 10 anos (Froese e Pauly, 2014; Eschmeyer e Fong, 2014).

Tabela 1. Lista das espécies válidas do gênero *Astyanax* segundo (Froese e Pauly, 2014).

No.	Nome válido	Autores
1.	***Astyanax abramis***	**(Jenyns, 1842)**
2.	*Astyanax aeneus*	(Günther, 1860)
3.	*Astyanax ajuricaba*	Marinho & Lima, 2009
4.	*Astyanax alburnus*	(Hensel, 1870)
5.	*Astyanax altior*	Hubbs, 1936
6.	*Astyanax altiparanae*	Garutti & Britski, 2000
7.	*Astyanax angustifrons*	(Regan, 1908)
8.	*Astyanax anterior*	Eigenmann, 1908
9.	*Astyanax aramburui*	Protogino, Miquelarena & López, 2006
10.	*Astyanax argyrimarginatus*	Garutti, 1999
11.	*Astyanax asuncionensis*	Géry, 1972
12.	*Astyanax atratoensis*	Eigenmann, 1907
13.	*Astyanax aurocaudatus*	Eigenmann, 1913
14.	*Astyanax bifasciatus*	Garavello & Sampaio, 2010
15.	*Astyanax bimaculatus*	(Linnaeus, 1758)
16.	*Astyanax biotae*	Castro & Vari, 2004
17.	*Astyanax bockmanni*	Vari & Castro, 2007
18.	*Astyanax bourgeti*	Eigenmann, 1908
19.	*Astyanax brachypterygium*	Bertaco & Malabarba, 2001
20.	*Astyanax brevirhinus*	Eigenmann, 1908
21.	*Astyanax burgerai*	Zanata & Camelier, 2009
22.	*Astyanax caucanus*	(Steindachner, 1879)
23.	*Astyanax chaparae*	Fowler, 1943
24.	*Astyanax chico*	Casciotta & Almirón, 2004
25.	*Astyanax clavitaeniatus*	Garutti, 2003
26.	*Astyanax cocibolca*	Bussing, 2008
27.	*Astyanax cordovae*	(Günther, 1880)
28.	*Astyanax correntinus*	(Holmberg, 1891)
29.	*Astyanax courensis*	Bertaco, Carvalho & Jerep, 2010
30.	*Astyanax cremnobates*	Bertaco & Malabarba, 2001

31.	*Astyanax daguae*	Eigenmann, 1913
32.	*Astyanax depressirostris*	Miranda Ribeiro, 1908
33.	*Astyanax dissensus*	Lucena & Thofehrn, 2013
34.	*Astyanax dissimilis*	Garavello & Sampaio, 2010
35.	*Astyanax dnophos*	Lima & Zuanon, 2004
36.	*Astyanax eigenmanniorum*	(Cope, 1894)
37.	*Astyanax elachylepis*	Bertaco & Lucinda, 2005
38.	*Astyanax endy*	Mirande, Aguilera & Azpelicueta, 2006
39.	*Astyanax epiagos*	Zanata & Camelier, 2008
40.	*Astyanax erythropterus*	(Holmberg, 1891)
41.	*Astyanax fasciatus*	(Cuvier, 1819)
42.	*Astyanax fasslii*	(Steindachner, 1915)
43.	*Astyanax festae*	(Boulenger, 1898)
44.	*Astyanax filiferus*	(Eigenmann, 1913)
45.	*Astyanax gisleni*	Dahl, 1943
46.	*Astyanax giton*	Eigenmann, 1908
47.	*Astyanax goyacensis*	Eigenmann, 1908
48.	*Astyanax goyanensis*	(Miranda Ribeiro, 1944)
49.	*Astyanax gracilior*	Eigenmann, 1908
50.	*Astyanax guaporensis*	Eigenmann, 1911
51.	*Astyanax guaricana*	Oliveira, Abilhoa & Pavanelli, 2013
52.	*Astyanax guianensis*	Eigenmann, 1909
53.	*Astyanax gymnodontus*	(Eigenmann, 1911)
54.	*Astyanax gymnogenys*	Eigenmann, 1911
55.	*Astyanax hastatus*	Myers, 1928
56.	*Astyanax henseli*	de Melo & Buckup, 2006
57.	*Astyanax hermosus*	Miquelarena, Protogino & López, 2005
58.	*Astyanax integer*	Myers, 1930
59.	*Astyanax intermedius*	Eigenmann, 1908
60.	*Astyanax ita*	Almirón, Azpelicueta & Casciotta, 2002
61.	*Astyanax jacobinae*	Zanata & Camelier, 2008
62.	*Astyanax jacuhiensis*	(Cope, 1894)
63.	*Astyanax janeiroensis*	Eigenmann, 1908
64.	*Astyanax jenynsii*	(Steindachner, 1877)
65.	*Astyanax jordanensis*	Vera Alcaraz, Pavanelli & Bertaco, 2009
66.	*Astyanax jordani*	(Hubbs & Innes, 1936)
67.	*Astyanax kennedyi*	Géry, 1964
68.	*Astyanax kompi*	Hildebrand, 1938
69.	*Astyanax kullanderi*	Costa, 1995
70.	*Astyanax lacustris*	(Lütken, 1875)
71.	*Astyanax latens*	Mirande, Aguilera & Azpelicueta, 2004
72.	*Astyanax laticeps*	(Cope, 1894)
73.	*Astyanax leonidas*	Azpelicueta, Casciotta & Almirón, 2002
74.	*Astyanax leopoldi*	Géry, Planquette & Le Bail, 1988
75.	*Astyanax lineatus*	(Perugia, 1891)
76.	*Astyanax longior*	(Cope, 1878)
77.	*Astyanax longirhinus*	Garavello & Sampaio, 2010
78.	*Astyanax maculisquamis*	Garutti & Britski, 1997
79.	*Astyanax magdalenae*	Eigenmann & Henn, 1916
80.	*Astyanax marionae*	Eigenmann, 1911
81.	*Astyanax maximus*	(Steindachner, 1876)
82.	*Astyanax megaspilura*	Fowler, 1944
83.	*Astyanax mexicanus*	(De Filippi, 1853)
84.	*Astyanax microlepis*	Eigenmann, 1913
85.	*Astyanax microschemos*	Bertaco & Lucena, 2006
86.	*Astyanax minor*	Garavello & Sampaio, 2010
87.	*Astyanax multidens*	Eigenmann, 1908
88.	*Astyanax mutator*	Eigenmann, 1909
89.	*Astyanax myersi*	(Fernández-Yépez, 1950)
90.	*Astyanax nasutus*	Meek, 1907
91.	*Astyanax nicaraguensis*	Eigenmann & Ogle, 1907
92.	*Astyanax obscurus*	(Hensel, 1870)
93.	*Astyanax ojiara*	Azpelicueta & Garcia, 2000
94.	*Astyanax orbignyanus*	(Valenciennes, 1850)
95.	*Astyanax orthodus*	Eigenmann, 1907
96.	*Astyanax pampa*	Casciotta, Almirón & Azpelicueta, 2005
97.	*Astyanax paraguayensis*	(Fowler, 1918)
98.	*Astyanax parahybae*	Eigenmann, 1908
99.	*Astyanax paranae*	Eigenmann, 1914
100.	*Astyanax paranahybae*	Eigenmann, 1911
101.	*Astyanax paris*	Azpelicueta, Almirón & Casciotta, 2002
102.	*Astyanax pedri*	(Eigenmann, 1908)
103.	*Astyanax pelecus*	Bertaco & Lucena, 2006
104.	*Astyanax pellegrini*	Eigenmann, 1907
105.	*Astyanax pirabitira*	Lucena, Bertaco & Berbigier, 2013

106.	*Astyanax pirapuan*	Tagliacollo, Britzke, Silva & Benine, 2011
107.	*Astyanax poetzschkei*	Ahl, 1932
108.	*Astyanax procerus*	Lucena, Castro & Bertaco, 2013
109.	*Astyanax puka*	Mirande, Aguilera & Azpelicueta, 2007
110.	*Astyanax pynandi*	Casciotta, Almirón, Bechara, Roux & Ruíz Díaz, 2003
111.	*Astyanax ribeirae*	Eigenmann, 1911
112.	*Astyanax rivularis*	(Lütken, 1875)
113.	*Astyanax robustus*	Meek, 1912
114.	*Astyanax ruberrimus*	Eigenmann, 1913
115.	*Astyanax rupununi*	Fowler, 1914
116.	*Astyanax saguazu*	Casciotta, Almirón & Azpelicueta, 2003
117.	*Astyanax saltor*	Travassos, 1960
118.	*Astyanax scabripinnis*	(Jenyns, 1842)
119.	*Astyanax schubarti*	Britski, 1964
120.	*Astyanax scintillans*	Myers, 1928
121.	*Astyanax serratus*	Garavello & Sampaio, 2010
122.	*Astyanax siapae*	Garutti, 2003
123.	*Astyanax stenohalinus*	Messner, 1962
124.	*Astyanax stilbe*	(Cope, 1870)
125.	*Astyanax superbus*	Myers, 1942
126.	*Astyanax symmetricus*	Eigenmann, 1908
127.	*Astyanax taeniatus*	(Jenyns, 1842)
128.	*Astyanax totae*	Ferreira Haluch & Abilhoa, 2005
129.	*Astyanax trierythropterus*	Godoy, 1970
130.	*Astyanax troya*	Azpelicueta, Casciotta & Almirón, 2002
131.	*Astyanax tumbayaensis*	Miquelarena & Menni, 2005
132.	*Astyanax tupi*	Azpelicueta, Mirande, Almirón & Casciotta, 2003
133.	*Astyanax turmalinensis*	Triques, Vono & Caiafa, 2003
134.	*Astyanax unitaeniatus*	Garutti, 1998
135.	*Astyanax utiariti*	Bertaco & Garutti, 2007
136.	*Astyanax validus*	Géry, Planquette & Le Bail, 1991
137.	*Astyanax varzeae*	Abilhoa & Duboc, 2007
138.	*Astyanax venezuelae*	Schultz, 1944
139.	*Astyanax vermilion*	Zanata & Camelier, 2009
140.	*Astyanax villwocki*	Zarske & Géry, 1999
141.	*Astyanax xavante*	Garutti & Venere, 2009
142.	*Astyanax xiru*	Lucena, Castro & Bertaco, 2013

A grande variedade anatômica e o alto número de espécies de *Astyanax* são possivelmente a principal causa da ausência de estudos mais abrangentes no gênero. O gênero *Astyanax* compreende peixes de corpo achatado e alongado, não atingindo mais que 200 mm e distribuídos da porção sul dos Estados Unidos até a Argentina central (Bertaco e Garutti, 2007).

O monofiletismo do gênero é duvidoso, sendo que vários autores (Rosen, 1972; Weitzman e Fink, 1983; Weitzman e Malabarba, 1998; Oliveira et al., 2011; entre outros) sugerem que possivelmente existem várias linhagens independentes dentro de *Astyanax*. Além disso, do ponto de vista citogenético, o gênero é muito variável, com 2n=36 a 2n=50 cromossomos, com a maioria

das espécies estudadas com número cromossômico conservado 2n=50 (Pazza e Kavalco, 2007).

Atualmente verifica-se a existência de inúmeras espécies do gênero *Astyanax* sendo tratadas somente como *Astyanax spp.* (Pie et al., 2009). Além disso, diversos autores indicam a existência de complexos de espécies, como por exemplo, para *A. bimaculatus*, *A. fasciatus*, *A. scabripinnis* e para as espécies endêmicas do rio Iguaçu e ainda para *A. hastatus* (Moreira-Filho e Bertollo, 1991; Garutti e Britski, 2000; Bertaco e Lucena, 2006; Melo e Buckup, 2006; Pie et al., 2009; Kavalco et al., 2009). Dados moleculares para as espécies do rio Iguaçu (mtDNA, D-loop) apontam divergências com dados morfológicos, sugerindo a presença de espécies crípticas (Pie et al., 2009). Para *A. scabripinnis*, análises citogenéticas e morfológicas sugerem que esta espécie corresponda a um complexo de espécies (Moreira-Filho e Bertollo, 1991) e que este complexo seria composto por pelo menos 15 espécies (Bertaco e Lucena, 2006). O complexo *A. fasciatus* seria composto por diversas espécies, das quais somente as provenientes do rio São Francisco seriam *A. fasciatus*, enquanto que as outras espécies do leste do Brasil, América Central e rio Paraná seriam espécies muito parecidas, nomeadas *Astyanax* sp. *aff. fasciatus* (Melo e Buckup, 2006; Pazza et al., 2008). O complexo *A. bimaculatus* seria representado por espécies com cromatóforos distribuídos em pelo menos dois níveis do integumento, um mais superficial, epidérmico e um mais profundo, dérmico, com uma mancha umeral horizontalmente ovalada. Este complexo é composto por aproximadamente 18 espécies e subespécies (Garutti, 1995; Garutti e Langeani, 2009).

Apesar destes estudos, ainda pouco se sabe sobre as relações dentro do grupo e entre os outros membros de Characidae. O primeiro estudo baseado em dados moleculares de sequências parciais de mtDNA (12S e 16S) na tentativa da elucidação das relações dos caraciformes foi realizado por Ortí e Meyer (1997). Após isso, um amplo estudo envolvendo análises dentro de Characidae e outros grupos, baseado em quatro genes nucleares e dois mitocondriais, com o alinhamento de 3660 pares de bases, revelou a presença de um clado bem suportado formado por *Astyanax*, *Astyanacinus*, *Moenkhausia*, *Inpaichthys*, *Hemigrammus* e *Hyphessobrycon* (Calcagnotto et al., 2005) (figura 1a).

Javonillo et al. (2010) baseado em dados moleculares de três genes mitocondriais (12S, 16S e COI) e um nuclear (RAG2) totalizando 2940 pares de bases, analisaram duas espécies do gênero *Astyanax* (*A. cremnobates* e *A. mexicanus*) e consideraram o grupo polifilético (figura 1b). Mais recentemente, em um estudo molecular mais extenso em Characidae, com dois genes mitocondriais (16S rRNA e citocromo b) e três nucleares (Myh6, RAG1 e RAG2) totalizando 4680pb, Oliveira et al. (2011) também analisaram a posição de três espécies de *Astyanax* (*A. aeneus*, *A. jordani* e *A. mexicanus*) entre os Characidae. Os resultados também incluem *Astyanax* no clado C (como proposto por Javonillo et al., 2010), juntamente com outros gêneros especiosos como *Hemigrammus*, *Hyphessobrycon*, *Moenkhausia*, *Knodus* e *Jupiaba* e duas subfamílias Stethaprioninae e Rhoadsiinae (figura 2). Ainda que este trabalho tenha envolvido um grande número de espécies da família Characidae, somente uma pequena parcela de espécies do gênero *Astyanax* foi analisada impossibilitando um estudo particular deste gênero.

Deste modo, com um pequeno número de espécies de *Astyanax* analisadas (2 espécies por Javonillo et al., 2010; 1 espécie por Ortí e Meyer, 1997 e 2 espécies por Calcagnotto et al., 2005 e 3 espécies por Oliveira et al., 2011), baseadas em sequências de DNA, estes resultados mostram a necessidade de maiores esforços para tentar resolver as relações deste enorme grupo de espécies de Characidae.

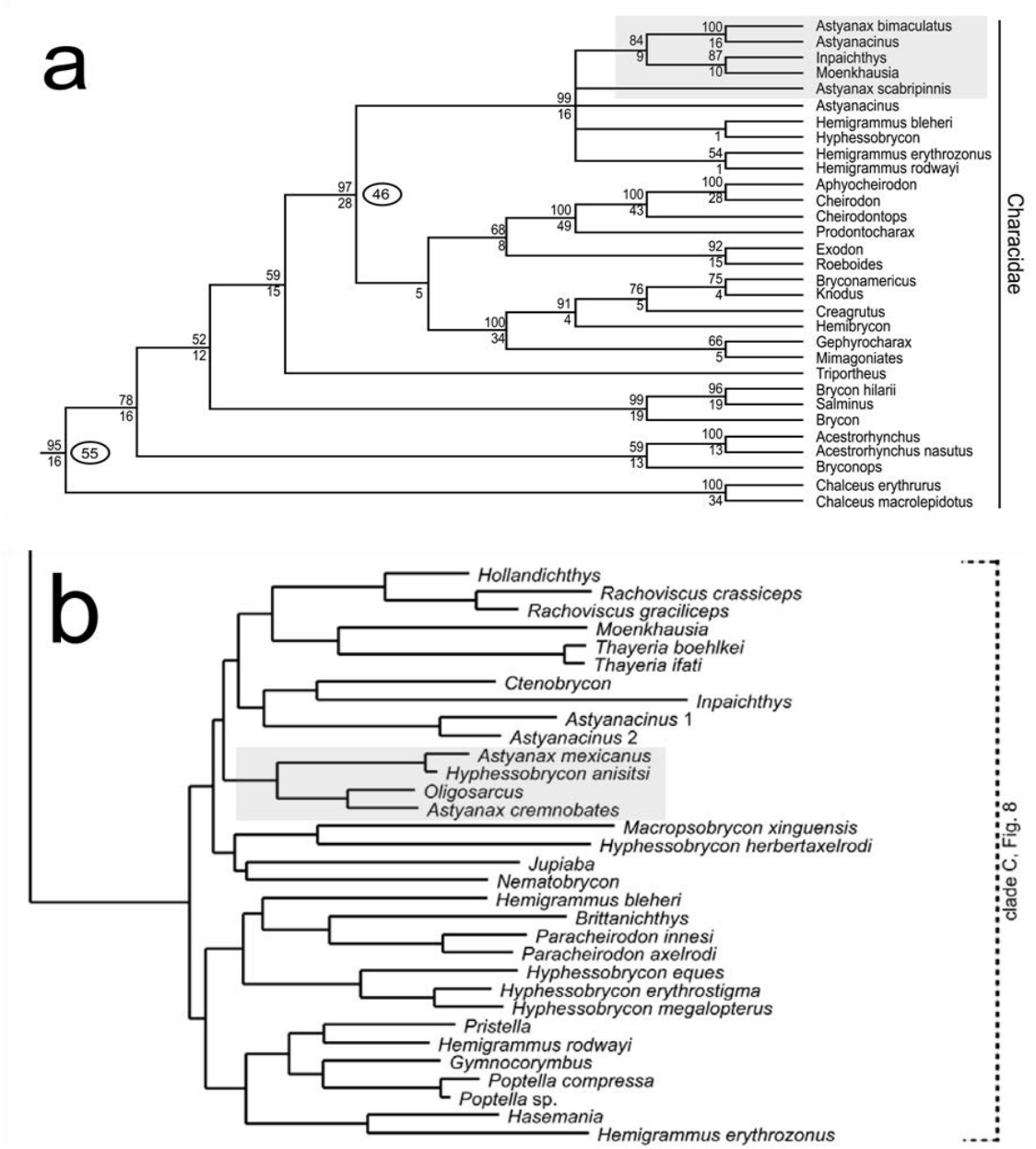

Figura 1. Hipóteses de relação dentro de Characidae mostrando a posição de *Astyanax*. (a) Calcagnotto et al. (2005); (b) Clado C de Javonillo et al. (2010). Em destaque, clados contendo *Astyanax*.

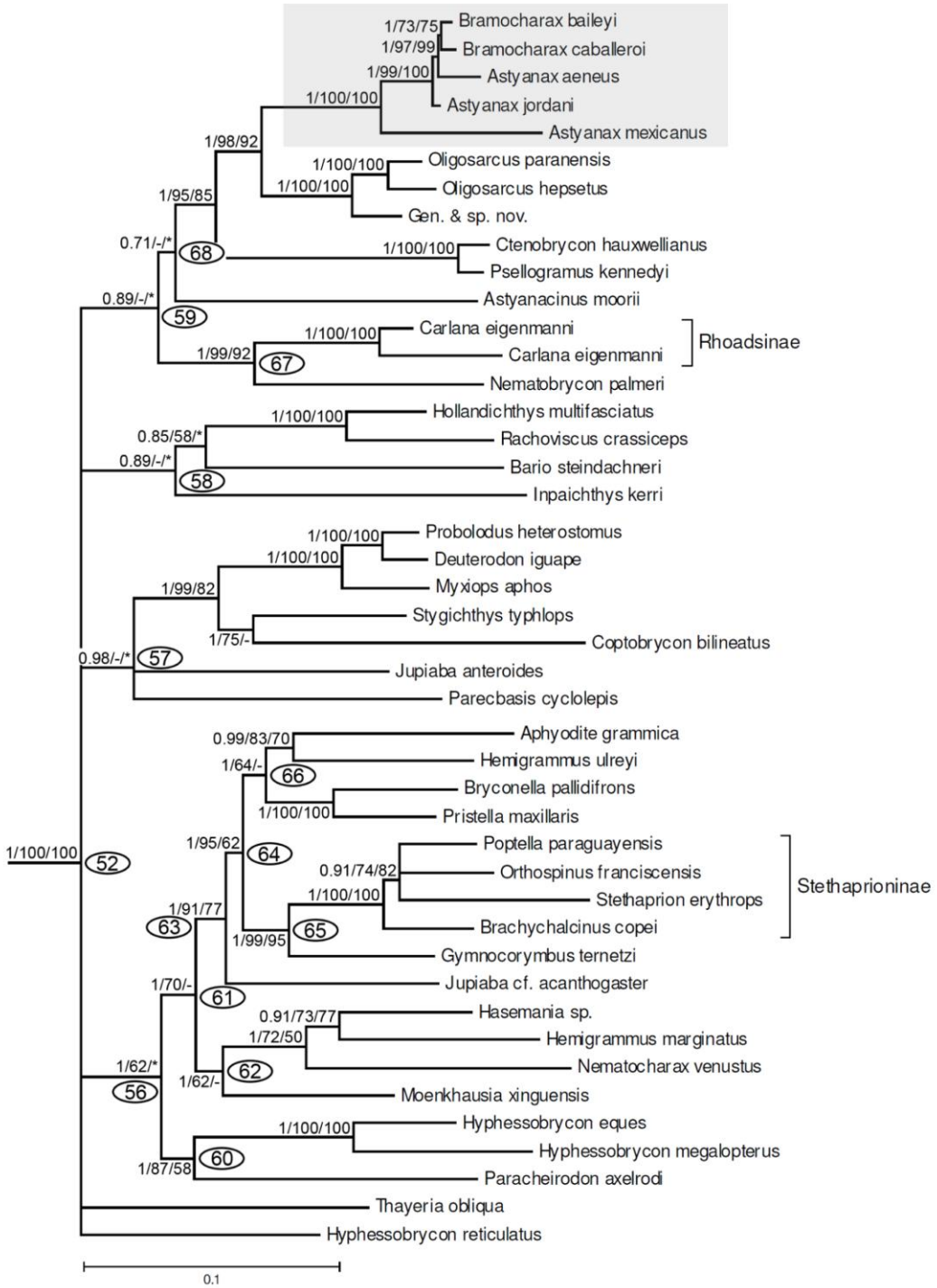

Figura 2. Relações filogenéticas dentro do clado incluindo *Astyanax* (em destaque) obtido por Máxima Verossimilhança. Retirado de Oliveira et al. (2011).

Do ponto de vista morfológico, em uma análise filogenética de Characidae envolvendo 160 espécies, sendo 17 do gênero *Astyanax*, e tendo como base o estudo de 360 caracteres morfológicos, Mirande (2010) mostrou que *Astyanax* é um gênero polifilético (figura 3). Segundo o autor citado as

espécies de *Astyanax* analisadas pertencem a três grupos, identificados como (1) "*Hyphessobrycon luetkenii* clade" com uma espécie de *Astyanax* mais relacionada a duas espécies de *Hyphessobrycon*, (2) "*Astyanax paris* clade" formado exclusivamente por *A. paris*, e (3) "*Astyanax* clade" formado por diversas espécies de *Astyanax* e espécies de *Hyphessobrycon*, *Psellogrammus* e *Markiana.* Ainda que poucas espécies de *Astyanax* tenham sido analisadas por Mirande (2010) este é o estudo filogenético mais amplo no gênero e reforça a idéia de que o mesmo é polifilético.

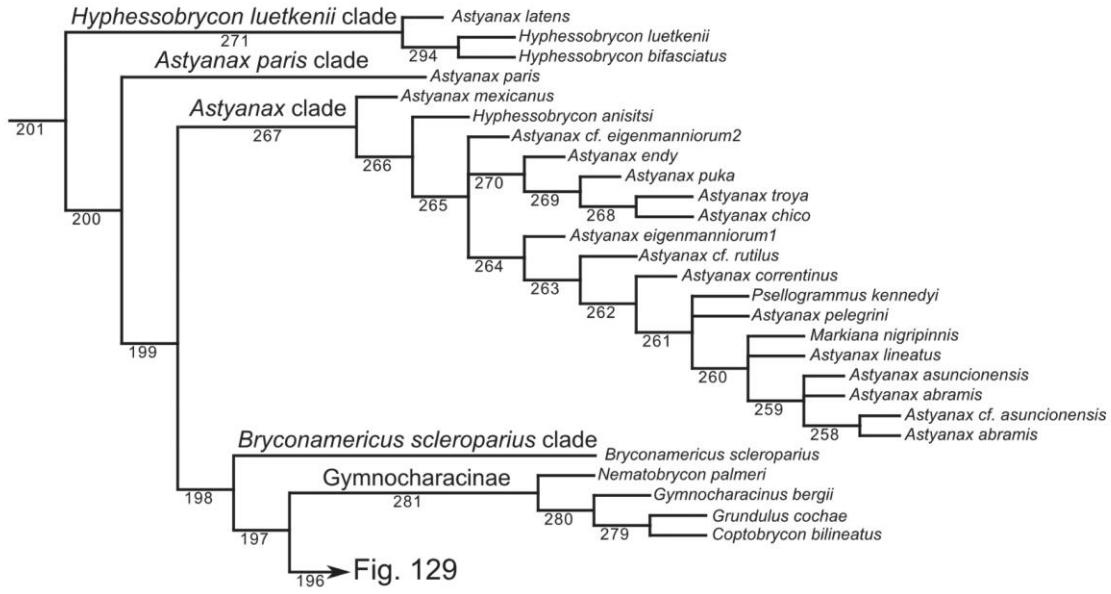

Figura 3. Hipótese parcial das relações dentro de Characidae mostrando a posição de algumas espécies de *Astyanax*, segundo Mirande (2010).

A resolução de grupos internos (gêneros e subfamílias) de Characidae é fundamental para o entendimento das relações evolutivas dentro da família. Os poucos estudos moleculares dentro do gênero *Astyanax*, os problemas de identificação de suas espécies, o número ainda incerto de espécies e o possível polifiletismo no gênero, fazem com que o estudo desse grupo seja necessário, para que possam ser estabelecidas as relações internas do

gênero, as relações deste com outros gêneros dentro da família e ainda, para o conhecimento da diversidade de espécies do grupo.

Esse grande número de espécies de *Astyanax* reflete uma alta complexidade em relação a identificação e descrição de novas espécies, uma vez que a plasticidade fenotípica e variabilidade genética dos caracteres utilizados na identificação morfológica de espécies podem levar a identificações errôneas. Deste modo, uma alternativa a catalogação e inferência de relações filogenéticas é a abordagem da genética molecular, como por exemplo a identificação molecular via DNA *barcode* (Hebert et al., 2003a).

DNA barcode

O número de espécies catalogadas no mundo é de aproximadamente 1.2 milhões (Bisby et al., 2010), porém as especulações sobre o real número de espécies existentes atinge dezenas de milhões (Odegaard, 2000; May e Harvey, 2009; Padial et al., 2010; Costello, et al., 2011; Mora et al., 2011). Mora et al. (2011) estimam que a biodiversidade atinja 8,75 milhões de espécies e, deste modo, seriam necessários mais de 1.200 anos para a descrição de todas as espécies da terra. Além disso, há questões sobre o conceito de espécie (de Queiroz, 2007) e em relação a identificação das espécies, dada a plasticidade fenotípica e possibilidade de existirem espécies crípticas e ainda a identificação em estágios específicos de vida (ex. larval, adulto) (Hebert et al., 2003a).

Apesar da existência de relatos de identificação molecular de espécies desde a década de 60 (Cowie, 1968), apenas no início deste século, com o proposto por Hebert et al. (2003a), é que um sistema de identificação biológica

baseado em sequências DNA tomou grandes proporções, tendo o intuito de ajudar nessa difícil tarefa de identificação através da biologia molecular. Dessa forma, foi sugerido que uma pequena região do genoma seria capaz de diferenciar todos, ou pelo menos, a vasta maioria do animais. Para tanto, nos animais, a região escolhida é o gene da citocromo *c* oxidase subunidade 1 (COX1 ou COI). Já para plantas, ainda há controvérsias e diversos genes e ainda, combinações de genes foram propostos para criar um mecanismo de identificação confiável, sendo os mais comuns o *rcbl* e *matk* (Hollingsworth et al., 2011). Para outro grupo de eucariotos, os fungos, a região padrão é a ITS (Internal Transcribed Spacers) (Schoch et al., 2012).

Comparada ao sistema de identificações de produtos comerciais, o código de barras, a porção 5' do gene COI com aproximadamente 650pb, produziria uma enorme quantidade de possibilidades de sequências dos nucleotídeos distribuídos ao longo de seu comprimento e deste modo, possibilitaria a identificação única de cada espécie, uma vez que foi apresentado que este gene possui uma alta variabilidade interespecífica, mas com reduzida variabilidade intraespecífica (Hebert et al., 2003a; Hebert et al., 2004). Este sistema tem por princípio o uso de uma região padrão do genoma e as premissas de escolha do marcador COI são devidas a (1) herança mitocondrial do gene, com baixa recombinação e ausência de íntrons (Hebert et al., 2003a), a (2) presença de *primers* bem estabelecidos para amplificação em grande parte dos animais (Folmer et al., 1994) e o (3) melhor potencial de sinal filogenético do que os outros genes mitocondriais, exibindo maiores taxas de evolução comparadas a outros genes como o 12S e 16S rDNA e citocromo *b* (Hebert et al., 2003a).

Com a padronização de um segmento do genoma para a identificação de espécies foi criado um banco de dados de referência, chamado de BOLD (*Barcode* of Life Data Systems, http://www.boldsystems.org), com o objetivo de prover ferramentas para análise e identificação das sequências *barcode* (Ratnasingham & Hebert, 2007). Este banco de dados é curado, com informações dos espécimes depositados, como dados de coleta, número de depósito em coleções reconhecidas, sequências consenso (mínimo de 500pb), *primers* utilizados na reação para obtenção das sequências e os eletroferogramas de sequenciamento. Atualmente, o banco de dados conta com aproximadamente 230 mil espécies e mais de 4,25 milhões de sequências depositadas (acesso Dezembro/2014). Em especial para os peixes, há o programa FISHBOL (the Fish Barcode of Life Campaign, http://www.fishbol.org/, Ward et al., 2009), uma campanha internacional com o intuito de sequenciar todos os peixes, que contém atualmente quase 11 mil espécies (mais de 30% do total de espécies de peixes reconhecidas como válidas) com sequências *barcode* e mais de 100 mil sequências depositadas (acesso Dezembro/2014).

No entanto, a abordagem de identificação possui algumas críticas, como em casos de hibridação e divergência muito recente (onde as substituições acumuladas não seriam suficientes para diferenciar as espécies) (Dawnay et al., 2007; Song et al., 2008; Ward et al., 2009b). Outro ponto é a presença de "NUMTs" (*Nuclear Mitochondrial Pseudogenes*), ou seja, a integração do genoma mitocondrial no genoma nuclear por uma duplicação, por exemplo, levando a possível amplificação de ambas as cópias do gene devido a conservação das regiões flanqueadoras. Neste caso, ambas as cópias do gene

COI poderiam ser amplificadas, ou ainda, somente a cópia nuclear, o que poderia levar a ambiguidade de resultados e alocação errônea da sequência, de forma que a cópia nuclear está sobre outras pressões seletivas, podendo acumular mais substituições do que a cópia original e grupos então poderiam ser formados somente com cópias de pseudogenes, levando a uma superestimação de número de espécies (Song et al., 2008). Porém, tais sequências são facilmente visualizadas devido a presença de *indels* (inserções e deleções), mutações pontuais e presença de *stop codons* (códons de parada de leitura) (Song et al., 2008; Ward et al., 2009b). Para Becker et al. (2011), estas críticas são pormenores da técnica, sendo muito mais preocupante a identificação errônea dos espécimes que levariam a complicações posteriores de identificação via DNA *barcoding*.

Além disso, após mais de uma década da proposta de Hebert et al. (2003), diversos valores de corte baseados em distâncias foram propostos, como por exemplo, de 1% (Ratnasingham & Hebert, 2007), 2% e 3% (mais comuns em estudos de peixes neotropicais; Lara et al., 2010; Carvalho et al., 2011; Mabragaña et al., 2011; Pereira et al., 2011a; Pereira et al., 2011b; Pereira et al., 2013), valores 10x da média intra-específica (Hebert et al., 2004) e ainda *barcoding gap* (Meyer & Paulay, 2005). Mais atualmente, outros trabalhos propõem outras análises que não são baseadas em distância como modelos para a delimitação de espécies (Pons et al., 2006; Fujisawa & Barraclough, 2013; Zhang et al., 2013).

Alguns estudos demonstram a presença de novas espécies e também de espécies crípticas. Em estudos com lepidópteros, Hebert et al. (2004) encontraram dez espécies crípticas através de DNA *barcoding* baseados em

uma única forma morfológica de adultos. Porém, estas diferem na aparência das larvas, habitats e preferências alimentares, demonstrando assim, a presença de um complexo de espécies. Resultados similares foram encontrados em larvas de dípteros (gênero *Chironomus)*, que vivem em simpatria e cuja morfologia das larvas é muito semelhante. A taxonomia é baseada em cromossomos politênicos, mas os dados de COI apontam a presença de um complexo de espécies envolvendo estes dípteros (Pfenninger et al., 2007). Outros estudos já possuem abordagens práticas, como identificações forenses de importantes espécies comerciais amazônicas com fins de conservação (Ardura et al., 2010) e controle de vendas de pescados em mercados sul-africanos (Cawthorn et al., 2012).

Em peixes, essa abordagem está sendo muito utilizada, principalmente após a criação do FISHBOL. Em estudos com peixes marinhos, 207 espécies foram estudadas e corretamente alocadas em espécies singulares, com poucas exceções que demonstraram haver incorreta identificação do espécime ou possível complexo de espécies (Ward et al., 2005). Resultados semelhantes foram encontrados na fauna ictiológica continental canadense, onde cerca de 95% das espécies (190 de 203 espécies no total, correspondendo a 85 gêneros e 28 famílias) foram sequenciadas, sendo a presença de possíveis espécies crípticas e hibridização em alguns casos foram apresentados (Hubert et al., 2008). Para *Astyanax*, esta ferramenta pode contribuir para estimar a diversidade de espécies e contribuir para a identificação e descrição de novas espécies.

Objetivos gerais

O presente trabalho teve como objetivos principais investigar a diversidade biológica presente no gênero *Astyanax* e testar a hipótese de monofiletismo do gênero, estudando as relações entre suas espécies e ampliando assim nosso conhecimento do grupo.

Objetivos específicos:

1 - Sequenciar genes mitocondriais e nucleares do maior número possível de espécies do gênero *Astyanax* e de gêneros relacionados, com o propósito de construir uma árvore filogenética robusta e representativa do grupo analisado.

2 - Para espécies com ampla distribuição pretende-se analisar o maior grupo de amostras possível, buscando avaliar o monofiletismo das mesmas, possibilitando, se for o caso, a descrição de novas espécies.

3 - Estabelecer as relações de *Astyanax* com gêneros relacionados de Characidae.

4 - Delimitar a distribuição geográfica das espécies e do gênero.

5 - Criar um banco de dados moleculares para permitir uma identificação molecular dos exemplares estudados.

Materiais e métodos

Amostras analisadas

Utilizou-se amostras de exemplares de espécies de *Astyanax* da coleção de tecidos do Laboratório de Biologia de Genética de Peixes – UNESP – Botucatu (LBP), credenciada no Ministério do Meio Ambiente como Fiel Depositária de Amostras do Patrimônio Genético. Os exemplares disponíveis somam 10410 exemplares, pertencentes a 996 lotes (sendo 738 lotes com identificação até o nível de espécie e o restante até o nível de gênero), coletados em diferentes bacias hidrográficas do Brasil, Colômbia, Guiana, México, Peru e Venezuela. Amostras de tecidos de exemplares da Argentina foram obtidas do museu da Universidad Nacional de Mar del Plata. Temos assim, 60 espécies nominais identificadas, sendo elas: *Astyanax abramis, Astyanax aeneus, Astyanax altiparanae, Astyanax anterior, Astyanax argyrimarginatus, Astyanax asuncionensis, Astyanax bifasciatus, Astyanax bimaculatus, Astyanax biotae, Astyanax bockmanni, Astyanax burgerai, Astyanax cf. giton, Astyanax cf. jequitinhonhae, Astyanax cf. lineatus, Astyanax cf. wappi, Astyanax dissimilis, Astyanax eigenmanniorum, Astyanax elachylepis, Astyanax erytropterus, Astyanax fasciatus, Astyanax fasciatus, Astyanax festae, Astyanax giton, Astyanax goyacensis, Astyanax goyanensis, Astyanax guaporensis, Astyanax gymnodontus, Astyanax gymnogenys, Astyanax hamatilis, Astyanax hastatus, Astyanax henseli, Astyanax intermedius, Astyanax jacuhiensis, Astyanax janeiroensis, Astyanax jordani, Astyanax lacustris, Astyanax laticeps, Astyanax magdalenae, Astyanax marionae, Astyanax metae, Astyanax mexicanus, Astyanax minor, Astyanax*

mutator, Astyanax pampa, Astyanax parahybae, Astyanax paranae, Astyanax pelecus, Astyanax pirapuan, Astyanax ribeirae, Astyanax rivularis, Astyanax saguazu, Astyanax scabripinnis, Astyanax schubarti, Astyanax serratus, Astyanax taeniatus, Astyanax trierytropterus, Astyanax varzeae, Astyanax venezuela, Astyanax vermilion, Astyanax xavante. Os espécimes sequenciados, mas que demonstraram ser de outros gêneros foram retirados das análises.

As amostras de tecido utilizadas para estudos moleculares, estão preservadas em etanol absoluto e mantidas sob refrigeração a -20°C. Os exemplares depositados foram fixados em solução de formol 10% e estão preservados em álcool 70%.

Extração de DNA genômico

O DNA total foi extraído de tecidos preservados em etanol. Foram seguidos diferentes protocolos, utilizando tanto o kit comercial Phire Animal Tissue Direct PCR (Finnzymes) quanto extração em placas (seguindo protocolo do Canadian Centre for DNA Barcoding-CCDB, disponível em http://www.ccdb.ca/docs/ CCDB_DNA_Extraction.pdf, última visualização em Dezembro 2013) e posteriormente, realizadas reações de PCR.

Amplificação das sequências barcode

As sequências parciais do gene Citocromo C Oxidase subunidade I (COI), utilizada como padrão para sequencias *barcode* em animais, foram

amplificadas com diversos conjuntos de *primers* foram utilizados para a obtenção dessas sequências (tabela 2).

As reações de PCR seguiram o seguinte perfil em um volume final de 12,5µl: 1µl de DNA (concentração de 50ng), 0.25µl de cada um dos primers Forward e Reverse (concentração 10mM), 1,25µl tampão de reação, 0.2µl de mix DNTPs 200mM, 0,37µl $MgCl_2$, 0,0625µl Taq DNA polimerase e água ultra pura q.s.p. para 12,5µl. As amplificações foram realizadas em termociclador Veriti 96 Well (Applied Biosystems) seguindo-se: desnaturação inicial de 5 minutos a 96ºC, 35 ciclos de 96ºC - 45 segundos, 54ºC - 45 segundos, 72ºC - 1 minuto, seguidos de uma etapa de extensão final a 72ºC por 1 minuto.

Tabela 2. Sequências dos primers utilizados para a obtenção das sequências *barcode* nos espécimes analisados.

Gene	Primer	Sequência (5'- 3')	Referência
COI	FishF1	TCAACCAACCACAAAGACATTGGCAC'	Ward et al. (2005)
	FishR1	TAGACTTCTGGGTGGCCAAAGAATCA	Ward et al. (2005)
	FishF2	TCGACTAATCATAAAGATATCGGCAC	Ward et al. (2005)
	FishR2	ACTTCAGGGTGACCGAAGAATCAGAA	Ward et al. (2005)
	L6252-Asn	AAGGCGGGGAAAGCCCCGGCAG	Melo et al., 2011
	H7271-COXI	TCCTATGTAGCCGAATGGTTCTTTT	Melo et al., 2011
	C_FishF1t1 – C_FishR1t1 (Fish cocktail)		
	VF2_t	TGTAAAACGACGGCCAGTCAACCAACCACAAAGACATTGGCAC	Ivanova et al. 2007
	FishF2_t	TGTAAAACGACGGCCAGTCGACTAATCATAAAGATATCGGCAC	Ivanova et al. 2007
	FishR2_t	CAGGAAACAGCTATGACACTTCAGGGTGACCGAAGAATCAGAA	Ivanova et al. 2007
	FR1d_t	CAGGAAACAGCTATGACACCTCAGGGTGTCCGAARAAYCARAA	Ivanova et al. 2007
	Sequencing primers for M13-tailed PCR product		
	*M13F (-21)	TGTAAAACGACGGCCAG	Messing, 1983
	*M13R (-27)	CAGGAAACAGCTATGA	Messing, 1983

* Primers utilizados apenas nas reações de sequenciamento dos produtos amplificados obtidos pelos conjuntos descritos por Ivanova et al. (2007).

Amplificação dos genes nucleares e mitocondriais utilizados na construção da filogenia

Sequencias parciais de genes nucleares e mitocondriais foram obtidas do 16S rRNA (16S), *cytochrome C oxidase subunit I* (COI), *ATP synthase 6 and 8* (ATPase 6/8) and *cytochrome b* (Cytb). Além disso, sequencias dos genes *myosin heavy chain gene 6* (Myh6), *recombination activating gene 1* (RAG1), *recombination activating gene 2* (RAG2) foram obtidas pelo protocolo de nested-PCR de acordo com Oliveira et al. (2011). Reações em cadeia da polimerase (PCR) seguiram o seguinte perfil com um volume final de 12,5µl: 1µl de DNA (concentração de 50ng), 0.25µl de cada um dos primers Forward e Reverse (concentração 10mM), 1,25µl tampão de reação, 0.2µl de mix DNTPs 200mM, 0,37µl $MgCl_2$, 0,0625µl Taq DNA polimerase e água ultra pura q.s.p. para 12,5µl. As amplificações foram realizadas em termociclador Veriti 96 Well (Applied Biosystems) seguindo-se: desnaturação inicial de 5 minutos a 96ºC, 35 ciclos de 96ºC por 45s, 45s a 48-58°C (de acorodo com o primer, tabela 3), 60-90s a 72ºC, seguidos de uma etapa de extensão final a 72ºC por 1 minuto.

Tabela 3. Sequencias dos primers de genes nucleares e mitocondriais utilizados para a construção da filogenia.

Gene	**Primer name**	**Primer sequence (5'-3')**	Source
ATPase subunits 6 e 8	ATP 8.2 - L8331	AAAGCRTYRGCCTTTTAAGC	Perdices et al., 2002
	CO3.2 - H9236	GTTAGTGGTCABGGCTTGGRTC	Perdices et al., 2002
16S	16Sa-L	ACGCCTGTTTATCAAAAACAT	Palumbi, 1996
	16Sb-H	CCGGTCTGAACTCAGATCACGT	Palumbi, 1996
Cytb	L14841	AAATCAAAGCATAACACTGAAGATG	Kocher et al., 1989
	H15915	CCAATTTGCATGGATGTCTTCTCGG	Irwing et al., 1991
	LNF	GACTTGAAAAACCAYCGTTGT	Oliveira et al., 2011
	H08R2	GCTTTGGGAGTTAGDGGTGGGAGTTAGAATC	Oliveira et al., 2011
Myh6	F329	CCGCMTGGATGATCTACAC	Li et al., 2007
1stPCR	A3R1	ATTCTCACCACCATCCAGTTGAA	Li et al., 2007
Myh6	A3F2	GGAGAATCARTCKGTGCTCATCA	Li et al., 2007
2ndPCR	A3R2	CTCACCACCATCCAGTTGAACAT	Li et al., 2007
	R1242	ACAGGATTGAGATGCTGTCCA	Li et al., 2007
	Myh6COF1	GACTGTTAACACCAAGAGAGT	Oliveira et al., 2011
	Myh6COF2	GTTATCCAGTATTTTGCAAGTATTGC	Oliveira et al., 2011
	Myh6COR1	TTGAACATCTTCTCATACAC	Oliveira et al., 2011
	Myh6COR2	TTCTCATACACTGACTTAGCCAGTGC	Oliveira et al., 2011
RAG1	2510F	TGGCCATCCGGGTMAACAC	Li & Ortí, 2007
1stPCR	4090R	CTGAGTCCTTGTGAGCTTCCATRAAYTT	Li & Ortí, 2007
RAG1	2535F	AGCCAGTACCATAAGATGTA	Li & Ortí, 2007
2ndPCR	4078R	TGAGCCTCCATGAACTTCTGAAGRTAYTT	Li & Ortí, 2007
	Rag1CF1	ACCCTCCGTACTGCTGAGAA	Oliveira et al., 2011
	Rag1CF2	TACCGCTGAGAAGGAGCTTC	Oliveira et al., 2011
	Rag1CF3	GAGAAGGAGCTTCTCCCAGG	Oliveira et al., 2011
	Rag1CF4	GCTTCCATCAGTTTGAGTGG	Oliveira et al., 2011
	Rag1CF5	CAGCTCTTGGAACATAGGCATCA	Oliveira et al., 2011
	Rag1CR1	CGTCGGAAGAGCTTGTTGCC	Oliveira et al., 2011
	Rag1CR2	TGTTGCCAGACTCATTGCCCTC	Oliveira et al., 2011
	Rag1CR3	CCCTCGCTGGCCCAGGCACC	Oliveira et al., 2011
	Rag1CR4	ATCTCGTTCCACAATCTCAGGC	Oliveira et al., 2011
	Rag1CR5	CATGGGCCAGTGTCTTGTGGAGGT	Oliveira et al., 2011
RAG2	164F	AGCTCAAGCTGCGYGCCAT	Oliveira et al., 2011
1stPCR	RAG2-R6	TGRTCCARGCAGAAGTACTTG	Lovejoy & Collette, 2001
RAG2	176R	GYGCCATCTCATTCTCCAACA	Oliveira et al., 2011
2ndPCR	Rag2Ri	AGAACAAAAGATCATTGCTGGTCGGG	Oliveira et al., 2011

Eletroforese dos produtos amplificados

Os produtos das reações de PCR foram submetidos à checagem de amplificação em gel de agarose 1% corado com GelRed (Biotium Inc.) na proporção de 1µl para cada 100ml de gel. O gel foi imerso em tampão TAE (Tris-Ácido Acético-EDTA) e submetido a 180 Volts por aproximadamente 20

minutos. Os produtos amplificados foram então submetidos às preparações para a reação de sequenciamento.

Limpeza das amostras amplificadas

Após a amplificação e corrida das amostras em gel de agarose, os produtos da PCR foram purificados com o kit *"ExoSap-IT*" (USB Corporation). Em um microtubo para PCR (0,2 ml) foram adicionados 5µl do produto amplificado juntamente com 2,0µl da solução de purificação (0,13 µl de ExoSap + 1,87 µl de água ultrapura). As amostras foram colocadas em termociclador por 1 hora a 37 ºC seguida de 15 minutos à 80ºC.

Reação de sequenciamento

A reação de sequenciamento foi realizada com o Kit *"Big Dye Terminator v.3.1 Cycle Sequencing Ready Reaction"* (Applied Biosystems). A reação consistiu de um volume final de 7µl: 1,0µl do produto amplificado (concentração de aproximadamente 40ng), 0,35µl de *primer* (Forward ou Reverse, 10 uM), 0,7µl de BigDye), 1,05µl de Tampão 5X para sequenciamento (kit BigDye) e 3,9µl de água ultrapura. A reação seguiu-se em um termociclador sob o seguinte perfil: desnaturação inicial por 2 minutos a 96°C, 35 ciclos a 96°C - 30 segundos, TM °C - 15 segundos e a 60°C por 4 minutos, etapa final a 12ºC até retirada do material do aparelho. Os primers utilizados na reação de sequenciamento foram os mesmos utilizados na PCR e presentes na tabelas 2 e 3. Ambas as direções (5' e 3') foram sequenciadas.

Preparação das amostras amplificadas por precipitação em EDTA/Acetato de sódio/etanol para a reação de sequenciamento

Após o término da reação de sequenciamento, seguiu-se uma etapa de limpeza dos amplicons para eliminação de excessos de reagentes utilizados na reação de sequenciamento, seguindo as etapas: adicionar 0,7µl de EDTA (125 mM) e depois adicionar 0,7µl de Acetato de Sódio (3 M), homogeneizar e centrifugar brevemente; então adicionar 17,5µl de etanol 100% e incubar por 15 minutos a temperatura ambiente; centrifugar por 15 minutos a 13.000 rpm à 25°C, descartar o etanol e secar em papel toalha; adicionar 24,5µl de etanol 70% gelado ecentrifugar por 10 minutos a 13.000 rpm à 20°C; descartar o etanol e secar em papel toalha; adicionar 24,5µl de etanol 100% gelado e centrifugar por 10 minutos a 13.000 rpm à 20°C; descartar o etanol e secar em papel toalha e secar em termociclador por 8 minutos a 96°C com os tubos e com a tampa do termociclador abertos (evaporação do etanol); guardar os tubos no freezer à 4°C envolto em papel alumínio, até o momento do sequenciamento.

Sequenciamento de DNA

As amostras foram analisadas em sequenciador de DNA automático, modelo ABI 3130-Genetic Analyzer (Applied Biosystems) presente em nosso laboratório. Para cada amostra foi adicionado 15ul de formamida Hi-Di (Applied Biosystems), seguindo de uma etapa de denaturação das amostras a 96ºC por dois minutos e resfriamento rápido em gelo por mais dois minutos. Logo em

seguida as amostras foram colocadas no aparelho para a realização da corrida de leitura da sequencia.

Análise dos dados

Obtenção das sequencias consenso, alinhamento e contrução da árvore de NJ (DNA barcode)

Ambas as sequências (forward e reverse) de cada espécime e de cada gene foram analisadas com o programa SeqScape versão 2.6 (Applied Biosystems) para a formação das sequências consenso e verificação de presença de *stop codons* (DNA *barcode*).

As sequências não demonstraram nenhum tipo de contaminação por DNA exógeno. Estas sequências foram editadas com o programa BioEdit 7.0.9.0 (Hall, 1999). Nos casos em que incertezas quanto ao correto nucleotídeo observado ocorrerem, as posições ambíguas foram identificadas utilizando a codificação IUPAC. As sequências barcode menores que 500pb foram retiradas das análises.

O alinhamento das sequências foi realizado com o programa MUSCLE (Edgar, 2004) implementado no MEGA 5 (Tamura et al., 2011). O alinhamento foi inspecionado visualmente para identificação de qualquer erro, que foi corrigido. Os valores de distâncias genéticas intra e interespecíficas foram calculados, utilizando-se o modelo de substituição Kimura-2-parameter (K2P) (Kimura, 1980). Os alinhamentos foram analisados pelo método de Neighbor-Joining (Saitou e Nei, 1987) usando o programa MEGA 5 (Tamura et al., 2011)

para identificar possíveis erros de sequenciamento. As possíveis espécies foram delimitadas baseado na distância molecular com corte de 2% seguindo os resultados apresentados por Ward et al. (2009).

Análises alternativas ABGD e GMYC para as sequencias barcode

As análises pelo programa ABGD foram realizadas via acesso remoto ao servidor no site http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html (acesso Maio/2014), seguindo o default do programa como sugerido por Puillandre et al. (2012a) e Kekkonenn e Hebert (2014) com o modelo Kimura-2P. Para maximizar a descoberta de possíveis espécies, o valor de X foi de 0.1 (X é o valor relativo do *gap*, que neste caso é de 0,1%), P mínimo de 0.005 e máximo de 0.1 (P refere-se à máxima distância interna de cada possível espécie). ABGD é um procedimento automático de identificação que forma clusters de sequências de possíveis espécies, baseado em distâncias par-a-par e nas diferenças entre as variações intra e inter-específicas (barcoding gap), detectando os limites da distribuição mesmo que esta tenha sobreposição (Puillandre et al., 2012a; Puillandre et al., 2012b).

Árvores ultramétricas foram geradas no programa BEAST v1.8.0 utilizando o modelo Yule (especiação), modelo de substituição nucleotídica GTR+G+I (como proposto pela análise do ModelTest no MEGA 5), partindo de uma árvore aleatória e com 50 milhões de gerações anotadas a cada 5 mil gerações. A convergência dos valores foram visualizadas no programa TRACER v1.6 (Rambaut et al., 2013). A análise GMYC (Pons et al., 2006) implementada no pacote SPLITS (Fujisawa & Barraclough, 2013) foi realizada

em ambiente R (R studio) com a opção "single threshold". Foram utilizados somente haplótipos únicos para as análises de GMYC, uma vez que problemas na análise com dados redundantes já foram relatados na literatura (Monaghan, et al., 2009; Fujisawa & Barraclough, 2013). Para isso, as sequências idênticas foram retiradas através da ferramenta ElimDupes (disponível em http://hcv.lanl.gov/ content/sequence/ELIMDUPES/elimdupes. html). Em todas as árvores um exemplar do gênero *Tetragonopterus* (*Tetragonopterus carvalhoi*) foi incluso como outgroup na análise bayesiana.

As sequências consenso foram depositadas no banco de dados do BOLD dentro do projeto "BAST- *Barcode Astyanax*", com números de acesso de BAST001-12 a BAST1593-15.

Análises filogenéticas

Para a construção da filogenia, foram utilizadas 109 espécies de *Astyanax*, dos quais foram separados pelos métodos de DNA *barcoding*. Destas, 15 espécies correspondem ao grupo de *Astyanax* da América Central, sendo que os dados foram obtidos do GenBank. Seguindo Oliveira et al. (2011), incluímos amostras dos Clados A, B e C, sendo *Chalceus erythrurus* como outgroup para o enraizamento das árvores.

As análises de Máxima Verossimilhança (ML) foram conduzidas no software RAxML 8.1.11 (Stamatakis, 2014) através do Web-Server CIPRES Science Gateway, usando o modelo de partição dos dados. O modelo de substituição nucleotídica adotado foi o default (GTR+GAMMA), partindo de uma árvore aleatória e com bootstrap de 1000 réplicas.

Para a análises Bayesiana, foi utilizado o programa MrBayes v. 3.2.2 com modelo de partição (Ronquist et al., 2012). Os dados foram particionados em 1ª, 2ª e 3ª posição dos códons para todos os genes, à exceção do gene 16S que foi mantido como partição única (total de 19 partições). O melhor modelo para cada gene e partição foi estimado através PartitionFinder v. 1.1.1 (Lanfear et al., 2012) sob o critério de BIC (Bayesian Information Criterion). Duas corridas independentes MCMC foram conduzidas sob 50 milhões de réplicas, amostrando uma árvore a cada 10000 passos. Os dados de convergência foram analisados no programa TRACER v1.6 (Rambaut et al., 2013). Como *burn-in* foram retiradas 30% das árvores e o restante das árvores foram utilizadas para a construção de uma árvore de consenso por maioria de 50% no programa PAUP* (Swofford, 2003).

REFERÊNCIAS BIBLIOGRÁFICAS

Albert, J., N. R. Lovejoy & W. G. R. Crampton. 2006. Miocene tectonism and the separation of cis- and trans-Andean river basins: Evidence from Neotropical shes. Journal of South American Earth Science, 21:14-27.

Albert JS, Reis RE. (2011). Historical biogeography of Neotropical Freshwater Fishes. Univ. of California Press, Berkeley, Los Angeles.

Ardura, A.; Linde, A.R.; Moreira, J.C.; Garcia-Vasquez, E. (2010). DNA barcoding for conservation and management of Amazonian commercial fish. Biological Conservation, 143: 1438-1443

Becker, S.; Hanner, R.; Steinke, D. (2011). Five years of FISH-BOL: Brief status report. Mitochondrial DNA, 22(S1): 3-9.

Bertaco, V.A. e Lucena, C.A.S. (2006). Two new species of Astyanax (Ostariophysi: Characiformes: Characidae) from eastern Brazil, with a synopsis of the Astyanaxscabripinnis species complex. Neotropical Ichthyology, 4(1):53-60.

Bertaco, V. A. & V. Garutti. 2007. New Astyanax from the upper rio Tapajós drainage, Central Brazil (Characiformes: Characidae). Neotropical Ichthyology, 5(1): 25-30.

Bisby FA, Roskov YR, Orrell TM, Nicolson D, Paglinawan LE, Bailly N, Kirk PM, Bourgoin T, Baillargeon G., eds (2010). Species 2000 & ITIS Catalogue of Life: 2010 Annual Checklist. DVD; Species 2000: Reading, UK.

Calcagnotto, D. et al. (2005). Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. Mol. Phylogenet. Evol. 36, 135-153.

Carvalho, D.C.; Oliveira, D.A.A.; Pompeu, S.A.; 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(S1): 80-86.

Cawthorn, D-M.; Steinman, H.A.; Witthuhn, R.C. (2012). DNA barcoding reveals a high incidence of fish species misrepresentation and substitution on the South African market. Food Research International, 46: 30-40.

Costello, M.J., Wilson, S., Houlding,B.(2011). Predicting total global species richness using rates of species description and estimates of taxonomic effort. Syst Biol doi:10.1093/sysbio/syr080

Cowie, W.P., 1968. Identification of fish species by thin-slab polyacrylamide gel electrophoresis of the muscle myogens. Journal of the Science of Food and Agriculture, Volume 19, (4):226-229.

Dawnay, N.; Ogden, R.; McEwing, R.; Carvalho, G.R.; Thorpe, R.S. (2007). Validation of the barcoding gene COI for use un forensic genetic species identification. Forensic Science International, 173:1-6.

De Queiroz, K. Species concepts and species delimitation. Systematic Biology, v. 56, n. 6, p. 879-886, 2007.

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

Eigenmann, CH. (1921). The American Characidae. Memoirs Museum of Comparative Zoology, vol. 43, no. 3, p. 209-310.

Eigenmann, CH. (1927). The American Characidae. Memoirs Museum of Comparative Zoology, vol. 43, no. 4, p. 311-428.

Eschmeyer, W. N. & Fong, J. D. (2014). SPECIES BY FAMILY/SUBFAMILY. (http://research.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version accessed 01 June 2014.

Eschmeyer, W.N., Fricke, R., Fong, J.D. e Polack, D.A. (2010); Marine fish diversity: history of knowledge and discovery. Zootaxa, 2525: 19-50.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294-299.

Frézal, L. E.; Leblois, R. (2008). Four years of DNA barcoding: current advances and prospects. Infection, Genetics and Evolution, v. 8, n. 5, p. 727-736.

Froese, R., e Pauly, D. (Eds.) (2014). FishBase, World Wide Web electronic publication:www.?shbase.org,ver. Accessed April 2014.

Fujisawa, T., Barraclough, T.G. 2013. Delimiting species using single locus data and the generalized mixed Yule coalescent approach. Systematic Biology, doi: 10.1093/sysbio/syt033.

Garavello, J.C. e Sampaio, F.A.A. (2010). Five new species of genus Astyanax Baird & Girard, 1854 from Rio Iguaçu, Paraná, Brazil (Ostariophysi, Characiformes, Characidae).Braz. J. Biol., 70(3): 847-865.

Garutti, V. & H. A. Britski. 2000. Descrição de uma espécie nova de Astyanax (Teleostei: Characidae) da bacia do Alto rio Paraná e considerações sobre as demais espécies do gênero na bacia. Comunicações do Museu de Ciências e Tecnologia da PUCRS, Série Zoologia, 13: 65-88.

Garutti, V. (1995). Revisão taxonômica dos Astyanax (Pisces, Characidae), com mancha umeral ovalada e mancha no pedúnculo caudal, estendendo-se à extremidade dos raios caudais medianos, das Bacias do Paraná, São Francisco e Amazônica. Tese (livre docência) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. São José do Rio Preto : [s.n.]. viii, 286 f. : il.

Garutti, V. e Langeani, F. (2009). Redescription of Astyanax goyacensis Eigenmann, 1908 (Ostariophysi: Characiformes: Characidae). Neotropical Ichthyology, 7(3):371-376.

Géry, J., 1977. Characoids of the World. Neptune City: T. F. H. Publications. 672 p.

Godoy, M.P. (1975). Peixes do Brasil. Subordem Characoidei. Bacia do Rio Mogi Guassu. Volume I. Editora Franciscana, Piracicaba, SP. p. 216.

Hall, T.A. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." Nucl. Acids Symp. Ser. 41: 95-98.

Hebert P. D. N., Cywinska A., Ball S. L., deWaard J. R. (2003a) Biological identification through DNA barcodes. Proc. R. Soc. Lond. B.;270:313-321.

Hebert, P. D. N.; Penton, E. H.; Burns, J. M.; Janzen, D. H.; Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astrapes fulgerator. Proc. Natl. Acad. Sci. USA, v. 101, p. 14812-14817.

Hollingsworth, P.M., Graham, S.W., Little, D.P. 2011. Choosing and Using a Plant DNA Barcode. PLoS ONE 6(5): e19254. doi:10.1371/journal.pone.0019254.

Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burridge M, Watkinson D, Dumont P, Curry A, Bentzen P, Zhang J, April J, Bernatchez L. 2008. Identifying Canadian freshwater fishes through DNA barcodes. PLoS One, 3:e2490.

Ivanova, N. V.; Zemlak, T. S.; Hanner, R. H.; Hebert, P. D. N. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes, v. 7, p. 544-548, 2007.

Javonillo, R. et al. (2010). Relationships among major lineages of characid fishes (Teleostei: Ostariophysi: Characiformes), based on molecular sequence data. Molecular Phylogenetics and Evolution 54: 498-511.

Kavalco, K.F.; Brandao, K.O.; Pazza, R.; Almeida-Toledo L.F. (2009). Astyanax hastatus Myers, 1928 (Teleostei, Characidae): A new species complex within the genus Astyanax?. Genetics and Molecular Biology, 32, 3, 477-483.

Kekkonenn, M., Hebert, P.D.N. (2014). DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. Mol Ecol Resour. 2014 Jan 30. doi: 10.1111/1755-0998.12233.

Kimura M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution, v. 16, p. 111-120.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29 (6): 1695-1701. http://dx.doi.org/10.1093/molbev/mss020

Lara, A., J. L. P. León, R. Rodríguez, D. Casane, G. Côté, L. Bernatchez & E. Machado. 2010. DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts. Molecular Ecology Resources, 10: 421-430.

Lima, F.C.T. et al. (2003). Genera Incertae Sedis in Characidae. Pp. 106-169. In: Reis, R. E., Kullander, S. O., & Ferraris, C. J. (Eds.). Check list of the freshwater fishes of South and Central America (CLOFFSCA). Porto Alegre, EDIPUCRS, 729 p.

Li C, Ortí G, Zhang G, Lu G: A practical approach to phylogenomics: The phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol Biol* 2007, 7:44.

Li C, Ortí G: Molecular phylogeny of Clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. *Mol Phylogenet Evol* 2007, 44:386-398.

Lovejoy NR, Collette BB: Phylogenetic relationships of new world needlefishes (Teleostei: Belonidae) and the biogeography of transitions between marine and freshwater habitats. *Copeia* 2001, 1: 324–338.

Lundberg JG. (1998). The temporal context for diversification of Neotropical fishes. In: Malabarba LR, Reis RE, Vari RP, Lucena CAS,. Lucena ZMS, eds. Phylogeny and Classification of Neotropical Fishes. Edipurus, Porto Alegre, Brazil. 67-91.

Mabragaña E, Astarloa JMD, Hanner R, Zhang J, Castro MG. 2011. DNA Barcoding Identifies Argentine Fishes from Marine and Brackish Waters. PLoS One, 6:e28655.

Malabarba, L.R.; Reis, R.E.; Vari, R.P.; Lucena, Z.M.S.; Lucena, C.A.S.(eds). (1998). Phylogeny and classification of neotropical fishes. Porto Alegre: EDIPUCRS. 603p.

May, R.M.; Harvey, P.H. (2009). Species Uncertainties. SCIENCE, 323, p:687. doi: 10.1126/science.1170937.

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(1):49-56.

Melo, F.A.G. e Buckup, P.A. (2006). Astyanax heseli, a new name for Tetragonopterus aeneus Hensel, 1870 from southern Brazil (Teleostei: Characiformes). Neotropical Ichthyology, 4(1):45-52.

Messing, J. (1983). New M13 vectors for cloning. Methods in Enzymology, v. 101, p. 20-78, 1983.

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biology, 3, 2229-2238.

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

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP. 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol 58:298-311.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., Worm, B., 2011. How Many Species Are There on Earth and in the Ocean?. PLoS Biology, volume 9. 8: e1001127.

Moreira-Filho, O. e Bertollo, L.A. (1991). Astyanax scabripinnis (Pisces, Characidae): a species complex. Revista Brasileira de Genética, 14(2): 331-357.

Odegaard, 2000. How many species of arthropods? Erwin's estimate revised. Biological Journal of the Linnean Society, 71:583-597.

Oliveira, C.; Avelino, G.S.; Abe, K.T.; Maariguela, T.C.; Beine, R.C.; Ortí, G.; Vari, R.P.; Correa 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 2011, 11:275..

Ortí, G., Meyer, A. (1997). The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. Syst. Biol., 46: 75-100.

Padial, J.m.; Miralles, A.; Riva, I.; Vences, M. (2010). The integrative future of taxonomy. Frontiers in Zoology, 7:16. http://www.frontiersinzoology.com/content/7/1/16

Pazza R and Kavalco KF (2007) Chromosomal evolution in the neotropical characin Astyanax (Teleostei, Characidae). Nucleus, 50:519-543.

Pazza, R.; Kavalco, S.A.F.; Penteado, P.R.; Kavalco, K.F.; Almeida-Toledo, L.F. (2008). The species complex Astyanax fasciatus Cuvier (Teleostei, Characiformes) - a multidisciplinary approach. Journal of Fish Biology, 72: 2002-2010. doi:10.1111/j.1095-8649.2008.01837.x

Perdices, A., Bermingham, E., Montilla, A., Doadrio, I. 2002. Evolutionary history of the genus Rhamdia(Teleostei: Pimelodidae) in Central America. Mol. Phylogenet. Evol. 25, 172-189.

Pereira LHG, Hanner R, Foresti F, Oliveira C (2013) Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics 14: 20 doi:10.1186/1471-2156-14-20.

Pereira LHG, Pazian MF, Hanner R, Foresti F, Oliveira C. (2011b) DNA barcoding reveals hidden diversity in the Neotropical freshwater fish Piabina argentea (Characiformes: Characidae) from the Upper Paraná Basin of Brazil. Mitochondrial DNA 2011, 22(Suppl 1):87-96.

Pereira, L.H.G.; Maia, G.M.G.; Hanner, R.; Foresti, F.; Oliveira, C. (2011a). DNA barcodes discriminate freshwater fishes from the Para?´ba do Sul River Basin, São Paulo, Brazil. Mitochondrial DNA, 22(S1): 71-79.

Pfenninger, M.; Nowak, C.; Kley, C.; Steinke, D.; Streit, B. (2007). Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic Chironomus(Diptera) species. Mol. Ecol., v. 16, p. 1957- 1968.

Pie, M.R. et al. (2009). Molecular data reveal a diverse Astyanax species complex in the upper Iguaçu River. Journal of Fish Biology, 75:2357-2362.

Pough, F.H.; Heiser, J.B.; Mc Farland, W.N. (1999). A vida dos vertebrados. 2.ed. São Paulo, Atheneu Editora São Paulo

Pons J., Barraclough T. G., Gomez-Zurita J., Cardoso A., Duran D. P., Hazell S., Kamoun S., Sumlin W. D., Vogler A. P. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 2006;55:595-609.

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012a. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol Ecol, 21(8):1864-1877.

Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier MC, Cruaud C, Holford M, Samadi S. 2012b. Large-scale species delimitation method for hyperdiverse groups. Mol Ecol, 21(11):2671-2691.

Rambaut, A., Suchard, M.A., Drummond, A.J. 2013. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/tracer

Ratnasingham, S.; Hebert, P.D.N. (2007). BOLD: The barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes, 7:355-364.

Reis, R.E., Kullander, S.O., Ferraris, C. (2003). Check List of the Freshwater Fishes of South and Central America (CLOFFSCA), EDIPUCRS, Porto Alegre, 729 pp.

Rosen, D.E. (1972). Origin of the Characid Fish Genus Bramocharax and a Description of a Second, More Primitive, Species in Guatemala. American Museum Novitates, No 2500.

Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. Systematic Biology. doi: 10.1093/sysbio/sys029

Saitou, M. e Nei, N. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. Mol. Biol. Evol. 4(4):406-425.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, et al. (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. PNAS 109: 6241-6246.

Song, H.; Buhay, J.E.; Whiting, M.F.; Crandall, K.A. (2008). Many species in one: DNA barcoding overestimates the numbers of species when nuclear mitochondrial pseudogenes are coamplified. PNAS, vl. 105, number 36: 13486-13491.

Stamatakis, A. (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics 10.1093/bioinformatics/btu033.

Swofford DL: PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Massachusetts: Sinauer Associates; 2003.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28: 2731-2739.

The World Conservation Union. 2010. IUCN Red List of Threatened Species. Summary Statistics for Globally Threatened Species. Table 1: Numbers of threatened species by major groups of organisms (1996-2010).

Ward, R. D. (2009b). DNA barcode divergence among species and genera of birds and fishes. Molecular Ecology Resources, v. 9, p. 1077-1085..

Ward, R. D. et al. (2005). DNA barcoding Australia´s fish species. Phil. Trans. R. Soc. B, v. 360, p. 1847-1857.

Ward, R. D., Hanner, R. and Hebert, P. D. N. (2009), The campaign to DNA barcode all fishes, FISH-BOL. Journal of Fish Biology, 74: 329-356. doi: 10.1111/j.1095-8649.2008.02080.x

Weitzman, S.H. e Fink, W.L. (1983). Relationships of the neon tetras, a group of south american freshwater fishes (Teleostei, Characidae), with comments on the phylogeny of new world characiformes. Bull. Mus. Comp. Zool., 150(6): 339-395.

Weitzman, S.H., Vari, R.P. (1988). Miniaturization in South American freshwater fishes; an overview and discussion. Proceedings of the Biological Society of Washington, 101 (2): 444-465.

Zhang, Jiajie, Kapli, P., Pavlidis, P., and Stamatakis, A. 2013. A General Species Delimitation Method with Applications to Phylogenetic Placements. Bioinformatics (Oxford, England), 29 (22): 2869-2876.

Capítulo 1

Astyanax species diversity through DNA barcoding: molecular data assessment and the species complexes

Running title: DNA barcoding of *Astyanax*

Keywords: freshwater fish, COI, DNA barcode, molecular identification

Abstract

The molecular identification through DNA barcoding has been extensively used in an attempt to solve taxonomic problems and help in the discovery of new species. The Neotropical fishes exhibits a wide variety of shapes and sizes, with many species yet to be cataloged and many species very difficult to identify. Characidae is the most species rich family of Characiformes, with many genera with taxonomic uncertainties. One of such species rich genus widely distributed and with many taxonomic conflicts is *Astyanax*. Here we present an extensive analysis of the genus in its almost entirely area of occurrence, based on DNA barcoding. The use of different approaches to clustering sequences (ABGD, GMYC and BIN) showed consistency in the results obtained by the initial cutoff value of 2%, but GMYC shows a tendency to identify a larger number of clusters than others. The results point to the existence of four large groups in the genus. However, whereas some groups, as that composed by the trans-Andean species, are composed of well-defined species in others there is a grouping of several species in a single cluster, making the species delimitation very difficult and in many cases impossible. The final results suggest a very complex scenario in the taxonomy of *Astyanax*.

Introduction

Neotropical freshwater fish species are quite diverse occurring from Central America to southern of South America, corresponding to 71 families (Reis et al., 2003) and exceeding 7000 species (Albert & Reis, 2011). Inside this richness two very species rich groups are Characiformes and Siluriformes (Eschmeyer & Fong, 2014). Characiformes exhibit a great variety of shapes and body sizes, inhabiting lakes and rivers of Neotropical Region and also the Africa (Géry, 1977). Within this order the most diverse family is Characidae that recently had the number of valid species reached nearly 1,100 (Eschmeyer and Fong, 2014).

Characidae is still one of the most problematic group among Neotropical fishes (Oliveira et al., 2011) with many taxonomic and systematic problems. Relationships within the family remain uncertain and many genera have questioned its monophyly, such as *Astyanax* (Mirande, 2010; Oliveira et al., 2011). *Astyanax* is one of the most species rich genus of this family and currently stands as *Incertae sedis* in Characidae (Lima et al., 2003). Eigenmann (1921, 1927) was the first one that made a revision of the genus including 74 species and subspecies, many of which are now re-described in other genus such as *Jupiaba*. Only Géry, in 1977, extensively studied the genus after the first review, and basically maintained the proposal of Eigenmann. By 2003 the genus had 86 valid species (Lima et al., 2003), but now this number has already increased to 142 recognized as valid, with 34 of those described in the past 10 years (Froese and Pauly, 2014; Eschmeyer and Fong, 2014).

Currently there are many species of *Astyanax* being treated only as *Astyanax sp.* (Pie et al., 2009). In addition, several authors indicate the existence of species complexes, such as *Astyanax bimaculatus* (A. bimaculatus complex)*, Astyanax fasciatus* (A. fasciatus complex)*, Astyanax scabripinnis* (A. scabripinnis complex), endemic species of the Iguaçu river and also to coastal species such *Astyanax hastatus* (Moreira-Filho and Bertollo, 1991; Garutti and Britski, 2000; Bertaco and Lucena, 2006; Melo and Buckup, 2006; Pie et al, 2009;. Kavalco et al, 2009). For the *A.* scabripinnis complex, cytogenetic and morphological analysis suggests that this group corresponds to a wide distributed complex (Moreira-Filho & Bertollo, 1991) and that this group could consist of at least 15 species (Bertaco and Lucena, 2006). A. fasciatus complex consist of several species of which only those from the São Francisco river may be considered real *A. fasciatus*, while other species from eastern Brazil, Central America and the Paraná River are morphologically similar species that should be named as *Astyanax sp. aff. fasciatus* (Melo and Buckup, 2006). A third well-known complex is that named A. bimaculatus, widely distributed in South America, in which chromatophores are distributed at least in two levels of the integument, one more superficial and a second deeper, with a horizontally oval humeral spot (Garutti, 1995). This complex consists of about 18 undescribed species and subspecies (Garutti, 1995; Garutti and Langeani, 2009).

The large number of species of *Astyanax* reflects a high complexity related to the identification and description of new Neotropical characiforms species, since the phenotypic plasticity and variability of the characters used in the morphological identification of species may lead to erroneous identification. Inside of this context, molecular identification could helps to address this

problem. DNA barcoding has been extensively used in the identification and resolution of many taxonomic problems in fishes. Many studies have demonstrated a impressive success in the use of that methodology, as in the first study with marine fishes, where the species identification success was 100%, with no overlapping distribution between the species analyzed (Ward et al., 2005). Other study of marine fishes of the Indian coast also correctly identified almost all the species (Lakra et al., 2011). But in freshwater fish species rates are lower, as in Mexico and Guatemala, the correct identification was 93% (Valdez-Moreno et al., 2009). In a study with freshwater species from Canada the authors also could identify 93% of specimens (Hubert et al., 2008). An exception is in the Upper Paraná basin (Brazil) where the value of correct discrimination reached 99.2% (Pereira et al., 2013).

After more than a decade of the molecular identification system proposed by Hebert et al. (2003), different cutoff values based on distances have been tested, for example, 1% (Ratnasingham & Hebert, 2007), 2% and 3% (most common in studies of Neotropical fishes: Lara et al., 2010; Carvalho et al., 2011; Mabragaña et al., 2011; Pereira et al., 2011a; Pereira et al., 2011b; Pereira et al., 2013), values 10x greater than the intraspecific average (Hebert et al. 2004) and barcoding gap (Meyer & Paulay, 2005). More recently new methods for automatic species identification have been proposed as the Automatic Barcode Gap Discovery (ABGD), the Barcode Index Number (BIN) and the Generalized Mixed Yule Coalescent model (GMYC). ABGD is an automatic identification procedure that forms clusters of sequences of possible species, based on distances and differences between intra and interspecific variations, detecting the boundaries even distribution has overlaps (Puillandre

et al, 2012a; Puillandre et al, 2012b.). BIN also works based on distance methods, clustering sequences with a 2% threshold followed by a Markov analysis (Ratnasingham and Hebert, 2013). But other studies suggests further analyzes that are not based only in distance as a model for species discrimination (Pons et al., 2006; Fujisawa & Barraclough, 2013; Zhang et al, 2013). The GMYC analysis use of an ultrametric tree for establishing species limits, with a mixture of Yule (1924) (pure-birth) and Kingman models (1982) (coalescence), where the algorithm computes the probability of splits of lineages based on rates of speciation, thus identifying a cutoff value for which it is possible to say where species and populations splits (Powell, 2012).

Considering the difficulties facing the morphological identification of *Astyanax* and the presence of several species complexes, the present work aims to investigate the biological diversity of this genus based on DNA barcode analysis, contributing to expand our knowledge of the Neotropical fish fauna and DNA barcode methodology.

Material and Methods

Sample Collections

Specimens of *Astyanax* were collected in different river basins from Argentina, Brazil, Colombia, Guyana, Peru and Venezuela (Figure 1 and Supplementary file 1). Additionally, data of 366 samples from GenBank were used. Argentine tissue samples were provided by the museum of the Universidad Nacional de Mar del Plata. The tissue samples used for molecular studies are preserved in absolute ethanol and kept under refrigeration at -20°C. The voucher specimens were fixed in 10% formalin solution and are preserved

in 70% ethanol. Morphological vouchers were deposited in the fish collection of the Laboratory of Biology and Genetic of Fish (LBP) of Universidade Estadual Paulista, Brazil or Universidad Nacional de Mar del Plata, Argentina. Consensus sequences were deposited in the database of the BOLD within the project "BAST- Barcoding *Astyanax*" with access numbers BAST001-12 the BAST1593-15.

DNA extraction, PCR amplification and sequencing

Total DNA was extracted from muscle fragments following the Canadian Center for DNA-Barcoding (CCDB) protocol (available http://www.ccdb.ca). The segment from the 5' region of the mitochondrial COI gene was amplified using different combination of primers: L5698-Asn (Miya & Nishida, 2000); FishF1, FishF2, FishR1 and FishR2 (Ward et al., 2005); C_FishF1t1–C_FishR1t1 cocktail (Ivanova et al., 2007); and H7271-COI (Melo et al., 2011). Polymerase chain reactions (PCR) were perfomed in 12.5 µl, containing: 1µl DNA (concentration 50 ng/µl), 0.25µl of each of the Forward and Reverse primers (concentration 10 mM), 1,25µl reaction buffer, 0.2µl of 200mM dNTPs mix, 0.37µl $MgCl_2$ and 0.0625µl (5 units/µl) Platinum Taq DNA polymerase (Invitrogen). The amplifications were performed in a thermocycler (Veriti® 96-Well Thermal Cycler, Applied Biosystems) following: initial denaturation of 5 minutes at 96°C, 35 cycles of 96°C for 45 seconds, 54°C for 45 seconds, 72ºC for 1 minute, and finally a extension step at 72°C for 1 minute. PCR amplified products were cleanup with ExoSAP-IT (USB Corporation) and sequenced in both directions using the BigDye Terminator v3.1 Cycle Sequencing Kit (Life

Technologies) in ABI3130 Genetic Analyzer automated sequencer (Applied Biosystems).

Data analysis

Sequences were edited with the BioEdit program 7.0.9.0 (Hall, 1999). The alignment of the sequences was performed using the program MUSCLE (Multiple Sequence Comparison by Log-Expectation; Edgar, 2004). MEGA 5 (Tamura et al., 2011) was used for the calculation of genetic distances values using the Kimura 2-parameter (K2P) substitution model (Kimura, 1980) and also to estimate the Neighbor-Joining (NJ) tree (Saitou and Nei, 1987) following the species cutoff value of 2%. Other analyzes were performed in ABGD program via remote access in the server http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html site, with the K2P model following the default program. To maximize the discovery of possible species, parameters have been changed (relative value gap X=0.1, Pmin=0.005 and Pmax=0.1), as proposed by Ratnasingham and Herbert (2013). BIN approach was carried in BOLD database only for the project BAST, excluding public records from GenBank and sequences of specimens from Argentina. For GMYC analysis ultrametric trees were generated using the program BEAST v1.8.0 with the Yule speciation model, GTR+G+I nucleotide substitution model (selected from the MEGA 5), starting from a random tree and 50 million generations recorded every 5000 generations. The convergence of the values was checked in TRACER v1.6 program (Rambaut et al., 2013). The GMYC analysis (Pons et al., 2006) implemented in the package *splits* (SPecies' LImits by Threshold Statistics; Fujisawa & Barraclough, 2013) was performed in R with the "single threshold"

option. Only unique haplotypes were used for this, once that problems in the analysis of redundant data have been reported in the literature (Monaghan et al., 2009). For that, the identical sequences were removed by ElimDupes tool (available http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html). One specimen of genus *Tetragonopterus* (*Tetragonopterus carvalhoi*) was included as outgroup in the bayesian analysis because it belongs to an different clade (Clade B) from *Astyanax* (Clade C) following Oliveira et al. (2011) Characidae phylogeny.

Results

A total of 1683 barcode sequences were obtained for *Astyanax* (including public records), which represent almost the entire area of occurrence of the genus, including 60 nominal species and about 65 species only identified at genus level (Supplementary table T1). The sequence average size is 633bp, with absence of stop codons, deletions or insertions.

The dendrogram obtained in the NJ analysis (cutoff 2%, Figures 2 to 5) revealed the presence of four major groups, named herein as Clade 1, 2, 3 and 4 respectively. In the first group species basically belong to the 'A. fasciatus' and 'A. scabripinnis' complexes (Clade 1); the second is represented by species of Central America (Clade 2); the third, mainly species from 'A. bimaculatus' complex (Clade 3) and finally other *Astyanax* species (Clade 4, more details on supplemental Figure F1). The Clade 4 besides the differences among the species was kept as unique as there is no clear distinction between complex of species or geographic pattern of distribution inside it. Based on these clades, 122 groups were identified (excluding an outgroup) corresponding to possible

species with intraspecific average of 0.047%, where 29 are singletons (only one specimen).

The overall mean distance between the four major clades is 13.5% (Table 1) and the average distance within each clade ranged from 3.36% (Clade 2) to 18.72% (Clade 4; Table 1). The overall distance between groups of species ranged from 2% to 30.9% (Supplementary Table T2). The average divergence inside each cluster is presented in the Supplementary Table T3.

The analysis of the dataset in ABGD using the default parameters revealed the presence of a maximum 39 groups of species. But when the intraspecific maximum distance was set to 0.5% (consistent with the calculated average intraspecific with cutoff 2%) and X = 0.1, the number of groups was 120. Thus, the second result is the most plausible explanation for the data in question. BIN results are partial since the GenBank sequences were not in BOLD systems, but for the clusters analyzed it showed similar results from ABGD and the NJ analysis. The GMYC analysis revealed the presence of 205 species (confidence interval: 201-228, threshold time: -0.01135164), a value higher than that observed by other methods (Table 2).

The groups of species with low values of genetic interspecific distance (<2%) are mainly present in Clade 1, but can also be found in other groups (at least 39 nominal species it is present in more than one cluster). Of these, several groups have the presence of at least two species, pointing to a geographical regionalization (i.e., cluster composed by *A. sp, A. fasciatus, A. cf. jequitinhonhae*) or even the existence of other groups with a wide distribution (i.e., cluster composed by *A. sp, A. abramis, A. altiparanae, A. asuncionensis, A. lacustris*). There is also the fact that species such as *A. anterior*, *A.*

bifasciatus, *A. bimaculatus*, *A. laticeps*, *A. scabripinnis* and *A. fasciatus* are present in more than one species group, even in more than one clade with divergence greater than 2%.

Discussion

The NJ analysis (Figure 2) and the Bayesian analysis (GMYC – data not showed) revealed the presence of four major groups. The genetic distance analyses (NJ – K2P model) show that the identified clades are quite distinct (Table 1). Three of them display a low internal genetic distance (4.08%, 3.36% and 5.82%- Clades 1, 2 and 3, respectively) and one display a very high internal genetic distance (18.72% – Clade 4). For Neotropical freshwater fishes, the mean values of divergence within genus are smaller than 10% (average 8.37% in Hubert et al., 2008; 6.8% in Pereira et al., 2013). However, values larger than 10% were also reported in *Tetragonopterus* where the distances, based on barcode sequences, reached values from 10.2 to 12.5% (Melo et al., 2011), but the higher values found in the present study are beyond the limits of intra-generic values for all fishes studied until now.

On the other hand, previous studies involving a low number of *Astyanax* samples notice the low genetic distance among some nominal species, uncommon among fishes (Ward et al., 2009). These low genetic divergence values were observed between close related species from a restrict area, as in the São Francisco Basin with values ranging from 0 between *A. bimaculatus lacustris* and *A. bimaculatus* to 0.93% between *A. cf. fasciatus* and *Astyanax rivularis* (Carvalho et al., 2011). In this case the low genetic distance between members of A. bimaculatus complex probably occurred since them correspond

to the same species (*A. lacustris*, as found for cluster *A. bimaculatus/ A. lacustris* in Clade 3), but the data from Carvalho et al. (2011) study also point to genetic similarity of barcode value among members A. fasciatus and A. scabripinnis complexes. In a more recent work, the analysis of DNA barcoding of fishes from Argentina showed that between *A. eigenmanniorum* and *A. cf. pampa* the genetic distance is also very low (0.62%) (Rosso et al., 2012).

In the present study species identified as *Astyanax anterior*, *A. bifasciatus*, *A. bimaculatus*, *A. laticeps* and *A. scabripinnis* are present in two clades and *A. fasciatus* in three clades. *A. bimaculatus*, *A. scabripinnis* and *A. fasciatus* are recognized species complex (Moreira-Filho & Bertollo, 1991; Bertaco and Lucena, 2006; Melo and Buckup, 2006; Garutti, 1995; Garutti and Langeani, 2009). *A. anterior* was described by Eigenmann (1908) and *A. laticeps* was described as *Tetragonopterus laticeps* by Cope (1894). These two species represent wide distributed forms that need a review based on the actual known distribution of specimens signed by these species. *A. bifasciatus* was described for the Rio Iguaçu Basin by Garavello & Sampaio (2010) in a review of the *Astyanax* species of this basin. In the present study samples identified as *A. bifasciatus* from close sites at Iguaçu River are in distinct groups, but our morphological and morphometric data do not permit the distinction between fishes, indicating the need of a review.

Astyanax Clade 1

In this clade we have 772 individuals and we found 25 nominal species and the 20 clusters identified at genus level. The use of different methodologies permitted the identification from 17 (BIN) to 82 clusters (GMYC) (Table 2) and

the NJ analysis with a 2% cutoff shows 25 groups (Figure 3). In the Clade 1 more than one nominal species were observed in 44% of the NJ clusters.

Species of this clade are distributed in Brazil and southern neighbor countries. The large number of clusters identified and the overall genetic distance of about 4% make impossible to separate the species by DNA barcoding. In this clade the number of species observed in the 2% groups were composed from 2 to 11 species (nominal species), most of them belonging to the A. scabripinnis and A. fasciatus complexes. In some cases the low divergence among them can suggest the description of different species only based on specimens distributions as the case of *A. paranae* and *A. rivularis* that are from Rio Paraná and Rio São Francisco Basins respectively. Today we know that exist a shared fauna between the two basins (Buckup, 2011). Eigenmann (1921) already pointed that fishes from different rivers forms a "labyrinth" (*A. fasciatus, A. taeniatus, A. scabripinnis* and *A. intermedius*), in which there would be intermediate forms between them and also with overlaps in characters and a reduction in the number of species would be justified.

There are some possible explanations in the literature for cases of low divergence (<2%), as found in the Clade 1. One relates to the phenotypic plasticity of the species that difficult proper identification and it is a factor that should not be disregarded (Hebert et al., 2003). In addition there is the fact that COI have different rates of evolution in different groups leading to different species cluster together (Frézal and Leblois, 2008; Ward, 2009). Other point is related to the evolutionary history, with a possible recent radiation of some groups (Ward, 2009). Particularly for *Astyanax*, Ornelas-Garcia et al. (2009) using three mitochondrial (16S, Cytb and COI) and one nuclear markers

(RAG1), found different groups within the genus in Central America, a fact that was attributed to a possible recent colonization and rapid expansion of the group. The invasion of Central America for *Astyanax* ancestors may be occurred between 3.1 and 8.1 million years ago thus confirming its recent radiation (Ornelas-Garcia et al, 2009; Gross, 2012).

Specimens identified as *A. fasciatus* are present in several groups and in three of the four clades found in the present study. Morphological studies with the A. fasciatus complex suggest that the name *A. fasciatus* should only be given to the specimens from São Francisco river basin as described by Cuvier (1819), and other specimens identified as *A. fasciatus* from Paraná river basin, eastern Brazil and Central America could belong to different species (Melo and Buckup, 2006). Independent studies, as the cytogenetic investigations, have suggested the existence of many different forms in *A. fasciatus* (Pazza et al, 2008, Medrado et al., 2008) and these data could be used in a further revision of this species complex. In the present study the sample that better match the morphology and distribution of the type specimens of Cuvier are that found in the Clade 1 (Figure 3). Interestingly, the genetic distance between specimens from this form of *A. fasciatus* Cuvier and others from A. scabripinnis complex (*A. paranae*, collected near type locality) are less than 1.3%, reinforcing the idea of recent radiation from this fishes.

In the Clade 1, *A. xavante*, *A. goyanensis* and *A. anterior* were the only three nominal species that could be unequivocally identified by DNA barcode (Figure 3). Considering that specimens identified as *A. anterior* are also found in the Clade 4, this number is reduced to two. In some cases, as for example, *A. sp.*/*A. erytroptherus* and *A. sp.*/*A. fasciatus*/*A.schubarti*, a redefinition of the

species limits (morphology and distribution) could taxonomically solve the barcode problems. In many other cases, taxonomy studies should involve the description of new species and/or redescription of previously described ones and molecular studies which involve fast evolving molecular markers.

Astyanax Clade 2

In this clade we have 100 individuals and we found 10 nominal species and the 9 clusters identified at genus level (Figure 4). The use of different methodologies permitted the identification from 2 (BIN) to 19 clusters (GMYC) (Table 2) and the NJ analysis with a 2% cutoff show 17 groups. Of these, three groups (*A. sp.* 2, *A. sp.* 9 and *A. petenensis*) were combined because shows less than 2% of genetic distance between them (mean 1.4%, see figure 4). In the Clade 2 more than one species were observed in 17.6% of the NJ clusters.

This group includes all trans-Andean species, mainly data from Central America, less *A. festae*. The average distance of this clade was 3.72%, with the largest distance between clusters (*A. sp.* 8 and *A. fasciatus*) of 10%.

However, the cluster with specimens of *A. mexicanus* also involves cave specimens identified as *A. jordani* suggesting that they may be the same species (distance within the cluster of 0.73%). These data from Central America were derived from public records, which Ornelas-Garcia et al. (2008) also points out similarities between *Astyanax* and *Bramocharax* (less than 1% divergence) and proposes *Astyanax* as polyphyletic. Another study also demonstrates that DNA barcoding fail to separate species of *Bramocharax* and *Astyanax* (Valdez-Moreno et al., 2009).

In this Clade 2 a group of individuals from Maracaibo Lake (Venezuela) identified as *A. fasciatus* represent the only sample that came from outside Central America, but can still be considered a trans-Andean species. Valenciennes in Cuvier and Valenciennes (1848) described *Tetragonopterus viejita* from Maracaibo Lake that was synonymized as *A. fasciatus* by Buckup (in Reis et al., 2003). Considering the present data and those from Melo (2005 – unpublished PhD Thesis) it is possible to recognize *A. viejita* as a valid species. In the Clade 2, *A. orthodus*, *A. belizanus*, *A. nasutus* were the only three nominal species that could be unequivocally identified by DNA barcode (Figure 4). In many other cases, taxonomy studies should involve the description of new species and/or redescription of previously described ones and molecular studies which involve fast evolving molecular markers.

Astyanax Clade 3

In this clade we have 565 individuals and we found 10 nominal species and the 5 clusters with specimens identified at genus level (Figure 5). The use of different methodologies permitted the identification from 19 (ABGD) to 44 clusters (GMYC) (Table 2) and the NJ analysis with a 2% cutoff show 20 groups (Figure 5). In the Clade 3 more than one species were observed in 35% of the NJ clusters.

Interesting, in this clade we basically found the species of the A. bimaculatus complex as proposed by Garutti (1995), which encompass the *Astyanax* species in which the humeral spot has an oval shape and is horizontally positioned, with a spot on the caudal peduncle extending to the edge of the median caudal rays. Only one cluster (A. sp./*A. elachylepis*) and a

species (*A. varzeae)* does not fit to *A. bimaculatus* species group external body color as proposed by Garutti (1995). On the other hand, species with the same color pattern as that of A. bimaculatus species complex, according to Garutti (1995): *A. bimaculatus* L, *A. bimaculatus M, A. bimaculatus N, A. bimaculatus O, A. bimaculatus P, A. bimaculatus Q and A. goyacensis* were allocated in the Clade 4 of the present study.

One cluster was identified with three nominal species (*A. altiparanae*, *A. jacuhiensis* and *A. bimaculatus*) and one non-identified species (*Astyanax sp.*). The mean genetic distance among specimens in this cluster was only 0.35%, and the minimum distance of the closest cluster with other samples identified as *A. altiparanae* was 2.8%. The distribution of these two clusters includes coastal and continental basins in the states of São Paulo, Goiás, Mato Grosso, Minas Gerais, Paraná and Rio Grande do Sul. The existence of two clusters with specimens identified as *A. altiparanae* reflects what is already present in the cytogenetic literature (Fernandes and Santos-Martins, 2004) and DNA barcoding (Pereira et al., 2013). Thus, these data point to the existence of two species, so far treated as a single taxonomic entity (*A. altiparanae*). The type locality of *A. altiparanae* is the Rio Grande at Volta Grande Dam, Miguelópolis, São Paulo (Garutti and Britski, 2000) and in the present study we analyzed a sample from Colômbia, São Paulo that it is our closest point to the type locality. Individuals from this collection point are mixed with fishes in the group composed by *A. sp./A. abramis/A. altiparanae/ A. asuncionensis/A. lacustris* suggesting that the sample identified as *A. altiparanae* and grouped with *A. jacuhiensis*, *A. bimaculatus* and *Astyanax sp.* belongs to a different species. It

is interesting to notice that *A. altiparanae* apparently has a wide chromosomal plasticity with 22 cytotypes described (Fernandes and Martins-Santos, 2004).

Astyanax lacustris was described as *Tetragonopterus lacustris* by t e n (1875) from São Francisco Basin and this name has been currently used for *Astyanax* from A. bimaculatus complex in this basin. In the present study we found three clusters with specimens identified as *A. lacustris* (2.2, 2.8 and 2.9% genetic distance between them) (Figure 5). According to Buckup (2011) *A. lacustris* it is rarely compared to *A. altiparanae* since these two species are normally taken as distinct because the separation between the São Francisco and Paraná rivers where these species occur. However, the author draws attention to the fact that all characters used to differentiate these species are variable and overlap. Thus, our results reinforce the need of a wide revision of *A. altiparanae* including samples of *A. lacustris*. Moreover, a consistent group of *A. lacustris* from Bahia State seems to belong to other species, grouping with other *A. bimaculatus* from coastal rivers, indicating a possible cryptic species.

In the Clade 3 only *A. argyrimarginatus* could be unequivocally identified by DNA barcode (Figure 5) and in all other cases, taxonomy studies should involve the description of new species and/or redescription of previously described ones.

Astyanax Clade 4

In this clade we have 247 individuals and we found 27 nominal species and the 32 clusters with specimens identified at genus level (Figure 6). The use of different methodologies permitted the identification from 57 (GMYC) to 61 (ABGD) clusters (Table 2) and the NJ analysis with a 2% cutoff show 60

groups. In the Clade 4 more than one species were observed only in 18.3% of the NJ clusters, being thus the best resolved clade.

Besides the high number of unidentified species (32) in the Clade 4 we also identified specimens belonging to A. scabripinnis, A. bimaculatus and A. fasciatus species complex in this clade. Species are widely distributed throughout Brazil, but there are also species from Colombia, Guyana and Venezuela (as *A. metae*, *A. magdalenae*, *A. mutator* and *A. venezuelae*) and the only trans-Anden species that does not belong to Clade 2: *A. festae.*

Although COI gene is very resolute at the species level, little can be said about the relationships between organisms analyzed because there is little phylogenetic signal (Hebert & Gregory, 2005). However, the groups formed by this analysis provide some insights into groups and similarities between species. The presence of several clusters in this clade with high genetic distances (up 30.9%) between them and among other clades reinforces the idea of *Astyanax* is even more complex.

In the Clade 4, *A. pirapuan*, *A. metae*, *A. venezuelae*, *A. magdalenae*, *A. guaporensis, A. festae*, *A. marionae, A. vermilion*, *A. pelecus*, *A. hamatilis, A. burgerai*, *A. taeniatus* and *A. hastatus* could be unequivocally identified by DNA barcode (Figure 6). *A. eigenmanniorum* was also identified by barcode but other samples of this nominal species was also found in the Clade 1, similar for *A. scabripinnis* and *A. fasciatus*. In many other cases, taxonomy studies should involve the description of new species and/or redescription of previously described ones.

Conclusions

This study represent the deeper analyses conducted in *Astyanax* since the revisionary study of Eigenmann (1921). The analysis of more than a 1,600 samples, including 60 named species and 65 species only identified at genus level, showed the complexity of the genus and the difficult task to identify species in it. As the main result we found four artificial species clades with a very high genetic divergence among them (from 13.5 to 20.7). In one of the clades (Clade 2) we found the Central America *Astyanax* forms and the species *A. mexicanus*, the type species of the genus, thus this clade may represents the real *Astyanax* and the remaining groups may correspond to other genera. This hyphothesis need to be tested with the construction of a phylogeny for the genus, however, Eigenmann (1921) was already recognized three different genera and subgenera inside *Astyanax*: *Astyanax strictu sensu* (which correspond to *A. mexicanus*, Clade 2), *Poecilurichthys* and *Zygogaster*.

Three of identified clades (Clades 1 to 3) exhibited a very low intragroup genetic divergence (between 3.36 to 5.82) and the Clade 4 exhibited a very high intragroup genetic divergence (18.72) suggesting that many species could have suffered a very fast speciation process, which make much more difficult the species recognition.

As a final result, among the 60 nominal species analyzed only 18 (30%) could be identified by DNA barcode (*A. xavante*, *A. goyanensis*, *A. orthodus*, *A. belizanus*, *A. nasutus, A. argyrimarginatus, A. pirapuan*, *A. metae*, *A. venezuelae*, *A. magdalenae*, *A. guaporensis, A. festae*, *A. marionae, A. vermilion*, *A. pelecus*, *A. hamatilis*, *A. taeniatus* and *A. hastatus*). This very poor result may be related to a combination of fast speciation, species with broad

distribution, the absence of a complete revision of the group, the absence of a phylogeny for the group and related species and the actual preference for the description of local forms without a deep revision of all available data (including specimens deposited in many fish collections). All these points should be taken in account in the future description or revision of *Astyanax*.

Acknowledgements

We thank all the individuals who assisted us in the collection and identification of the specimens that served as the basis for this study as well as those that contributed with the donation of fishes and tissues. The last author is a CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico do Brasil) researcher (CNPq grant number 303854/2009-0). This investigation was granted by FAPESP (00264-2/2011 and 2010/17009-2), CNPq and CAPES.

References

Albert JS, Reis RE. (2011). Historical biogeography of Neotropical Freshwater Fishes. Univ. of California Press, Berkeley, Los Angeles.

Bertaco, V.A. e Lucena, C.A.S. (2006). Two new species of Astyanax (Ostariophysi: Characiformes: Characidae) from eastern Brazil, with a synopsis of the Astyanaxscabripinnis species complex. Neotropical Ichthyology, 4(1):53-60.

Buckup, P.A. (2011). The Eastern Brazilian Shield. In: James S. Albert, Roberto E. Reis (Editors). Historical Biogeography of Neotropical Freshwater Fishes. University of California Press. 408p.

Carvalho, D.C.; Oliveira, D.A.A.; Pompeu, S.A.; 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(S1): 80–86.

Cope, E.D.; On the fishes obtained by the naturalist expedition in Rio Grande do Sul. Proceedings of the American Philosophical Society, Vol. 33:144(Jan., 1894), pp. 84-108.

Cuvier, L.B. (1819). Memoirs du Museum D'Historie Naturelle. Paris, Tomme 5.352-353.

Cuvier, L.B., Valenciennes, A. (1848). Memoirs du Museum D'Historie Naturelle Des Poissons. Paris, 22: 1-149.

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

Eigenmann, C.H. (1908). Preliminary descriptions of new genera and species of tetragonopterid characins (Zoölogical Results of Thayer Brazilian expedition). Bulletin of the Museum of Comparative Zoology 52(6):91-106.

Eigenmann, CH. (1921). The American Characidae. Memoirs Museum of Comparative Zoology, vol. 43, no. 3, p. 209-310.

Eigenmann, CH. (1927). The American Characidae. Memoirs Museum of Comparative Zoology, vol. 43, no. 4, p. 311-428.

Eschmeyer, W. N. & Fong, J. D. (2014). SPECIES BY FAMILY/SUBFAMILY. (http://research.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version accessed 01 June 2014.

Fernandes, C.A.; Martins-Santos, I.C. (2004). Cytogenetic studies in two populations of Astyanax altiparanae (Pisces, Characifromes). Hereditas, 141: 328-332.

Frézal, L. E.; Leblois, R. (2008). Four years of DNA barcoding: current advances and prospects. Infection, Genetics and Evolution, v. 8, n. 5, p. 727-736.

Froese, R., e Pauly, D. (Eds.) (2014). FishBase, World Wide Web electronic publication:www.fishbase.org,ver. Accessed April 2014.

Fujisawa, T., Barraclough, T.G. 2013. Delimiting species using single locus data and the generalized mixed Yule coalescent approach. Systematic Biology, doi: 10.1093/sysbio/syt033.

Garavello, J.C. e Sampaio, F.A.A. (2010). Five new species of genus Astyanax Baird & Girard, 1854 from Rio Iguaçu, Paraná, Brazil (Ostariophysi, Characiformes, Characidae).Braz. J. Biol., 70(3): 847-865.

Garutti, V. & H. A. Britski. 2000. Descrição de uma espécie nova de Astyanax (Teleostei: Characidae) da bacia do Alto rio Paraná e considerações sobre as demais espécies do gênero na bacia. Comunicações do Museu de Ciências e Tecnologia da PUCRS, Série Zoologia, 13: 65-88.

Garutti, V. (1995). Revisão taxonômica dos Astyanax (Pisces, Characidae), com mancha umeral ovalada e mancha no pedúnculo caudal, estendendo-se à extremidade dos raios caudais medianos, das Bacias do Paraná, São Francisco e Amazônica. Tese (livre docência) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. São José do Rio Preto : [s.n.]. viii, 286 f. : il.

Garutti, V. e Langeani, F. (2009). Redescription of Astyanax goyacensis Eigenmann, 1908 (Ostariophysi: Characiformes: Characidae). Neotropical Ichthyology, 7(3):371-376.

Géry, J., 1977. Characoids of the World. Neptune City: T. F. H. Publications. 672 p.

Gross, J.B. 2012. The complex origin of Astyanax cavefish. BMC Evolutionary Biology, 12:105.

Hall, T.A. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." Nucl. Acids Symp. Ser. 41: 95-98.

Hebert P. D. N., Cywinska A., Ball S. L., deWaard J. R. (2003) Biological identification through DNA barcodes. Proc. R. Soc. Lond. B.;270:313-321.

Hebert, P. D. N.; Penton, E. H.; Burns, J. M.; Janzen, D. H.; Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astrapes fulgerator. Proc. Natl. Acad. Sci. USA, v. 101, p. 14812-14817.

Hebert, P.D.N. and T.R. Gregory. 2005. The promise of DNA barcoding for taxonomy. Systematic Biology 54: 852-859.

Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burridge M, Watkinson D, Dumont P, Curry A, Bentzen P, Zhang J, April J, Bernatchez L. 2008. Identifying Canadian freshwater fishes through DNA barcodes. PLoS One, 3:e2490.

Ivanova, N. V.; Zemlak, T. S.; Hanner, R. H.; Hebert, P. D. N. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes, v. 7, p. 544–548, 2007.

Kavalco, K.F.; Brandao, K.O.; Pazza, R.; Almeida-Toledo L.F. (2009). *Astyanax hastatus* Myers, 1928 (Teleostei, Characidae): A new species complex within the genus *Astyanax*?. Genetics and Molecular Biology, 32, 3, 477-483.

Kimura M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution, v. 16, p. 111-120.

Kingman, J.F.C. (1982) On the genealogy of large populations. Journal of Applied Probability, 19, 27–43.

Lakra, et al. (2011). DNA barcoding Indian marine fishes. Molecular Ecology Resources, 11: 60–71.

Lara, A., J. L. P. León, R. Rodríguez, D. Casane, G. Côté, L. Bernatchez & E. Machado. 2010. DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts. Molecular Ecology Resources, 10: 421-430.

Lima, F.C.T. et al. (2003). Genera Incertae Sedis in Characidae. Pp. 106-169. In: Reis, R. E., Kullander, S. O., & Ferraris, C. J. (Eds.). Check list of the freshwater fishes of South and Central America (CLOFFSCA). Porto Alegre, EDIPUCRS, 729 p.

Mabragaña E, Astarloa JMD, Hanner R, Zhang J, Castro MG. 2011. DNA Barcoding Identifies Argentine Fishes from Marine and Brackish Waters. PLoS One, 6:e28655.

Medrado AS, Figueiredo AVA, Waldschmidt AM, Affonso PRAM, Carneiro PLS. Cytogenetic and morphological diversity in populations of *Astyanax fasciatus* (Teleostei, Characidae) from Brazilian northeastern river basins. Genetics and Molecular Biology. 2008;31:208–214.

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(1):49-56.

Melo, F.A.G. (2005). Revisão taxonômica do complexo de espécies *Astyanax fasciatus* (Cuvier, 1819) (Teleostei: Characiformes: Characidae). Rio de Janeiro: UFRJ/MN. 293pp. Phd Thesis.

Melo, F.A.G. e Buckup, P.A. (2006). Astyanax heseli, a new name for Tetragonopterus aeneus Hensel, 1870 from southern Brazil (Teleostei: Characiformes). Neotropical Ichthyology, 4(1):45-52.

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biology, 3, 2229–2238.

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

Miya M, Nishida M (2000) Use of mitogenomic information in teleostean molecular phylogenetics: a tree-based exploration under the maximum-parsimony optimality criterion. Molecular Phylogenetics and Evolution, 17, 437–455.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP. 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol 58:298–311.

Moreira-Filho, O. e Bertollo, L.A. (1991). Astyanax scabripinnis (Pisces, Characidae): a species complex. Revista Brasileira de Genética, 14(2): 331-357.

Oliveira, C.; Avelino, G.S.; Abe, K.T.; Maariguela, T.C.; Beine, R.C.; Ortí, G.; Vari, R.P.; Correa 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 2011, 11:275..

Ornelas-Garcia, C. P.; Dominguez-Dominguez, O.; Doadrio, I. (2008). Evolutionary history of the fish genus Astyanax Baird & Girard (1854) (Actynopterigii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. BMC Evolutionary Biology, v. 8, p. 340.

Pazza, R.; Kavalco, S.A.F.; Penteado, P.R.; Kavalco, K.F.; Almeida-Toledo, L.F. (2008). The species complex *Astyanax fasciatus* Cuvier (Teleostei, Characiformes) - a multidisciplinary approach. Journal of Fish Biology, 72: 2002–2010. doi:10.1111/j.1095-8649.2008.01837.x

Pereira LHG, Hanner R, Foresti F, Oliveira C (2013) Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics 14: 20 doi:10.1186/1471-2156-14-20.

Pereira, L.H.G.; Maia, G.M.G.; Hanner, R.; Foresti, F.; Oliveira, C. (2011a). DNA barcodes discriminate freshwater fishes from the Paraíba do Sul River Basin, São Paulo, Brazil. Mitochondrial DNA, 22(S1): 71–79.

Pereira LHG, Pazian MF, Hanner R, Foresti F, Oliveira C. (2011b) DNA barcoding reveals hidden diversity in the Neotropical freshwater fish Piabina argentea (Characiformes:

Characidae) from the Upper Paraná Basin of Brazil. Mitochondrial DNA 2011, 22(Suppl 1):87–96.

Pie, M.R. et al. (2009). Molecular data reveal a diverse Astyanax species complex in the upper Iguaçu River. Journal of Fish Biology, 75:2357–2362.

Pons J., Barraclough T. G., Gomez-Zurita J., Cardoso A., Duran D. P., Hazell S., Kamoun S., Sumlin W. D., Vogler A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 2006;55:595-609.

Powell, J.R. (2012). Accounting for uncertainty in species delineation during the analysis of environmental DNA sequence data. Methods in Ecology and Evolution 2012, 3, 1–11 doi: 10.1111/j.2041-210X.2011.00122.x

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012a. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol Ecol, 21(8):1864-1877.

Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier MC, Cruaud C, Holford M, Samadi S. 2012b. Large-scale species delimitation method for hyperdiverse groups. Mol Ecol, 21(11):2671–2691.

Rambaut, A., Suchard, M.A., Drummond, A.J. 2013. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/tracer

Ratnasingham, S.; Hebert, P.D.N. (2007). BOLD: The barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes, 7:355–364.

Ratnasingham, S.; Hebert, P.D.N. (2013). A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. Plos One. DOI: 10.1371/journal.pone.0066213

Reis, R.E., Kullander, S.O., Ferraris, C. (2003). Check List of the Freshwater Fishes of South and Central America (CLOFFSCA), EDIPUCRS, Porto Alegre, 729 pp.

Rosso, J.J.; Mabragaña, E.; Gonzalez Castro, M.; Díaz de Astarloa, J.M. (2012). DNA barcoding Neotropical fishes: recent advances from the Pampa Plain, Argentina. Molecular Ecology Resources 12, 999–1011. doi: 10.1111/1755-0998.12010.

Saitou, M. e Nei, N. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. Mol. Biol. Evol. 4(4):406-425.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28: 2731-2739.

Váldez-Moreno, M. et al. (2009). Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes. Journal of Fish Biology, v. 74, p. 377-402.

Ward, R. D. et al. (2005). DNA barcoding Australia´s fish species. Phil. Trans. R. Soc. B, v. 360, p. 1847-1857.

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. Molecular Ecology Resources, v. 9, p. 1077-1085..

Yule, G.U. (1924) A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character, 213,21–87.

Zhang, Jiajie, Kapli, P., Pavlidis, P., and Stamatakis, A. 2013. A General Species Delimitation Method with Applications to Phylogenetic Placements. Bioinformatics (Oxford, England), 29 (22): 2869-2876.

Tables

Table 1. K2P distance within the four major clades identified (main diagonal, marked with asterisk) and average K2P divergence between these clades (below diagonal average values and above diagonal standard error).

	Clade 1	Clade 2	Clade 3	Clade 4
Clade 1	4.08*	1.4	1.5	1.5
Clade 2	14.5	3.36*	1.4	1.4
Clade 3	16.0	13.5	5.82*	1.4
Clade 4	21.1	20.4	20.7	18.72*

Table 2. Number of clusters identified on the different analysis.

	Cutoff value	Number of clusters	X (relative gap width) - K2P	Number of clusters	Number of clusters	Number of clusters
Method		**NJ**		**ABGD**	**BIN**	**GMYC**
All dataset	2%	122	0.1	120	*	205 CI (201-228) threshold time: - 0.01135164
Clade 1		25		25	17*	82
Clade 2		17		15	2*	19
Clade 3		20		19	24	44
Clade 4		60†		61	58*	57

† Excluding Tetragonopterus carvalhoi

* Incomplete sampling (See text for details)

Figures

Figure 1. Distribution of *Astyanax* samples deposited in the fish collection of Laboratório de Biologia e Genética de Peixes (LBGP-UNESP). Samples from Central America and Argentina are not showed.

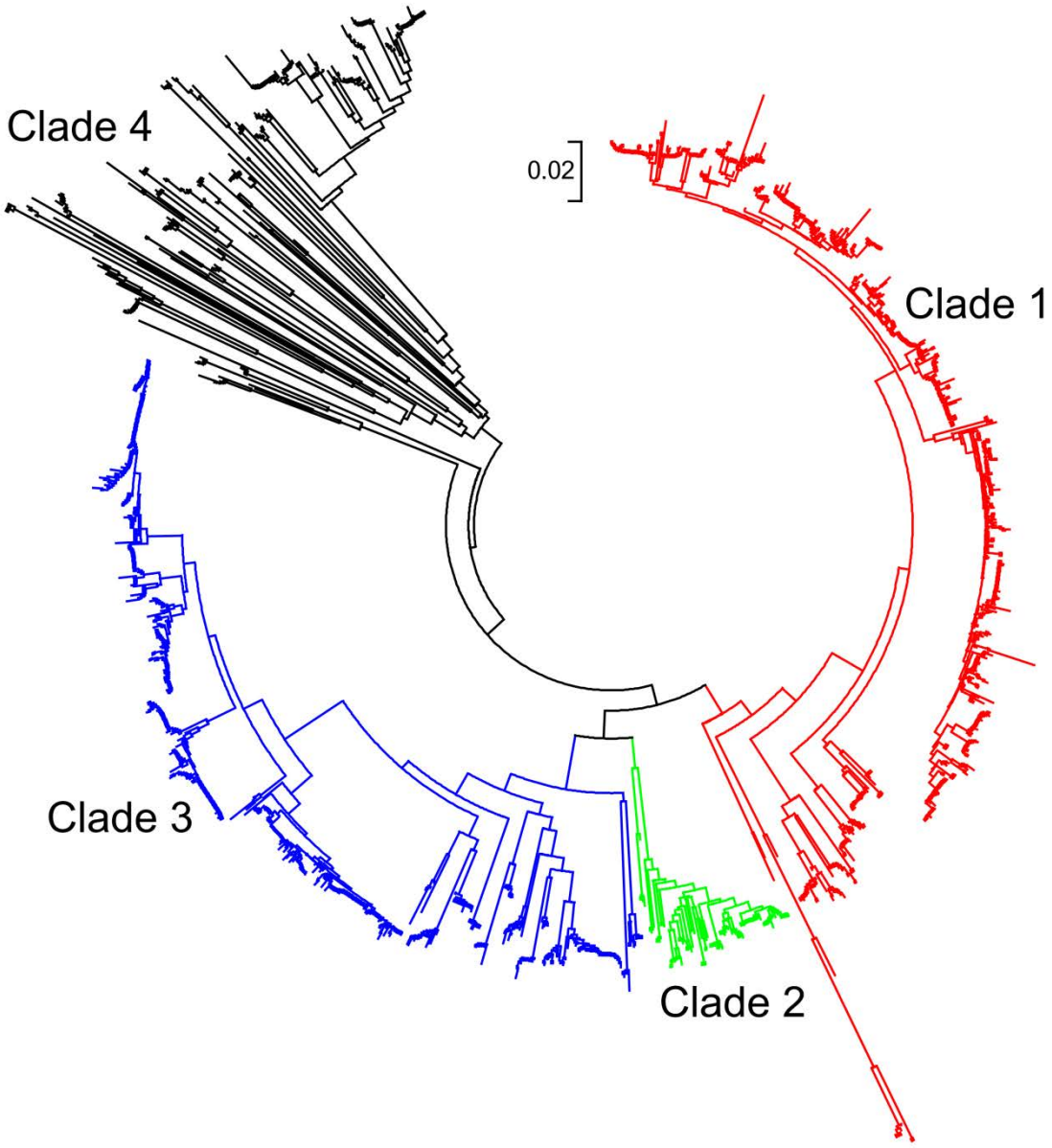

Figure 2. NJ dendrogram of all specimens analyzed. There are at least four major clades identified as Clade 1 to Clade 4 (see text for details).

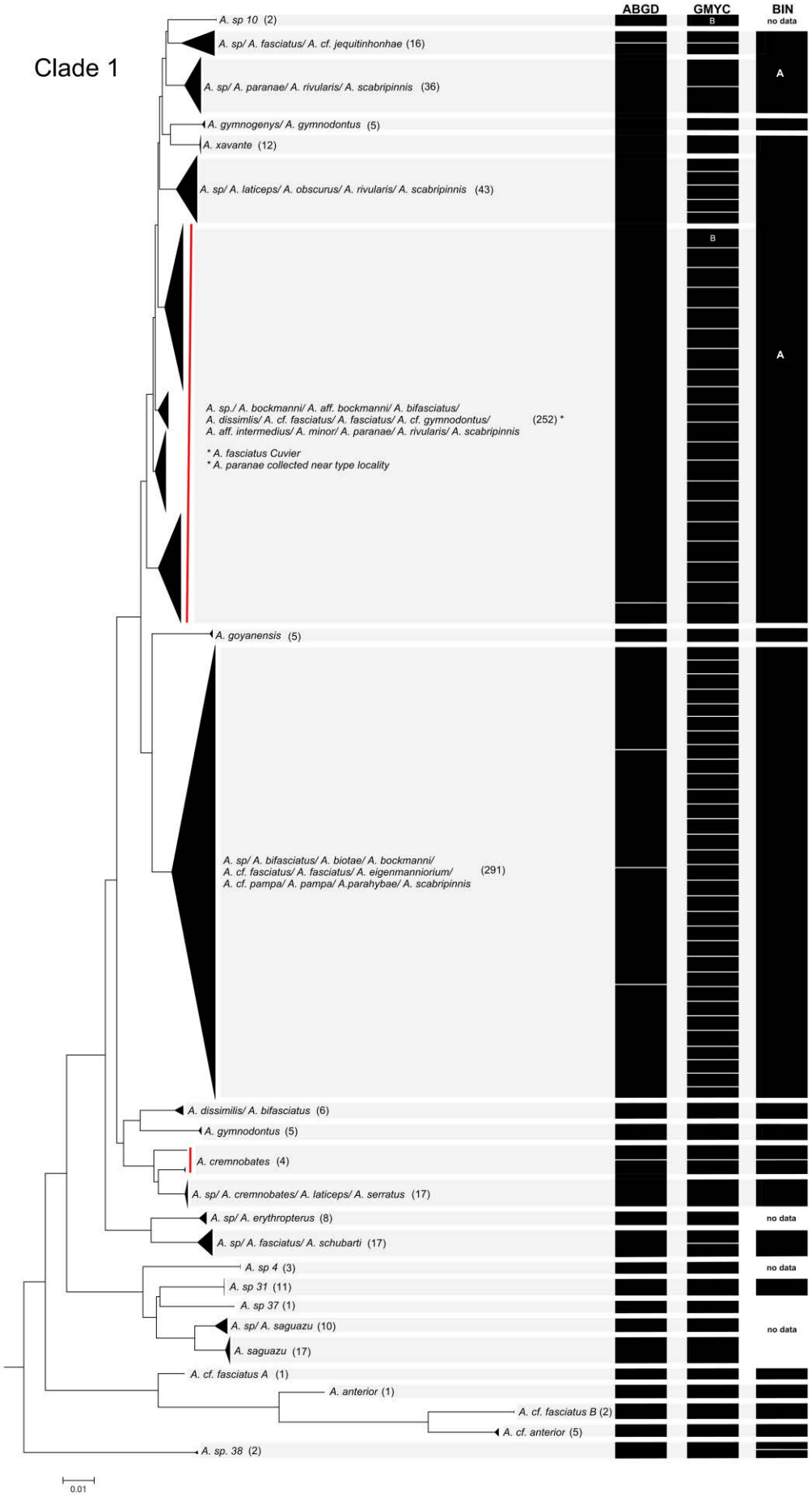
Clade 1
ABGD
GMYC
BIN
no data
A. sp 10 (2)
A. sp/ A. fasciatus/ A. cf. jequitinhonhae (16)
A. sp/ A. paranae/ A. rivularis/ A. scabripinnis (36)
A. gymnogenys/ A. gymnodontus (5)
A. xavante (12)
A. sp/ A. laticeps/ A. obscurus/ A. rivularis/ A. scabripinnis (43)
A. sp./ A. bockmanni/ A. aff. bockmanni/ A. bifasciatus/
A. dissimlis/ A. cf. fasciatus/ A. fasciatus/ A. cf. gymnodontus/ (252) *
A. aff. intermedius/ A. minor/ A. paranae/ A. rivularis/ A. scabripinnis
* A. fasciatus Cuvier
* A. paranae collected near type locality
A. goyanensis (5)
A. sp/ A. bifasciatus/ A. biotae/ A. bockmanni/
A. cf. fasciatus/ A. fasciatus/ A. eigenmanniorium/ (291)
A. cf. pampa/ A. pampa/ A.parahybae/ A. scabripinnis
A. dissimilis/ A. bifasciatus (6)
A. gymnodontus (5)
A. cremnobates (4)
A. sp/ A. cremnobates/ A. laticeps/ A. serratus (17)
A. sp/ A. erythropterus (8)
A. sp/ A. fasciatus/ A. schubarti (17)
A. sp 4 (3)
A. sp 31 (11)
A. sp 37 (1)
A. sp/ A. saguazu (10)
A. saguazu (17)
A. cf. fasciatus A (1)
A. anterior (1)
A. cf. fasciatus B (2)
A. cf. anterior (5)
A. sp. 38 (2)
0.01

Figure 3. NJ dendrogram showing the groups of Clade 1. Results from BIN are partially presented because lack of data. Gray shaded areas represents groups in each cluster delimited by K2P cutoff 2%. Letters A and B represents that groups with same letter are clustered together by GMYC analysis. Vertically bar indicates that besides there are a separation between clusters, the genetic divergence are less than 2%.

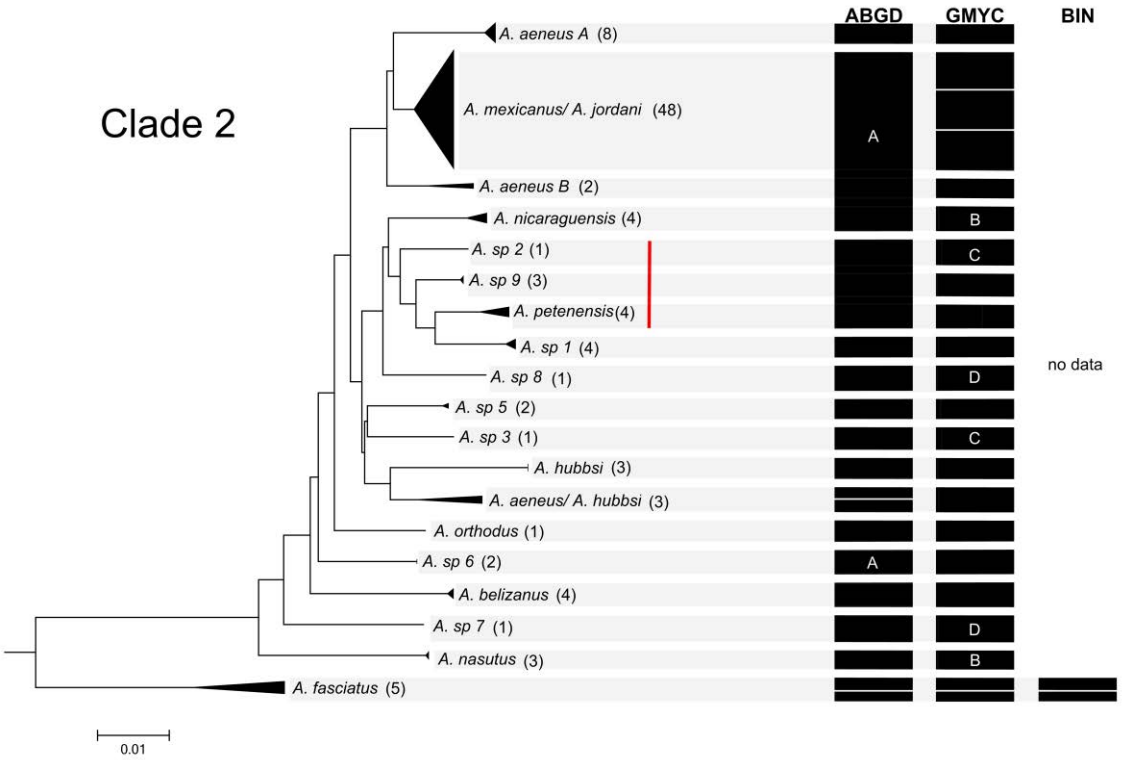

Figure 4. NJ dendrogram showing the groups of Clade 2. Results from BIN are partially presented because lack of data. Gray shaded areas represents groups in each cluster delimited by K2P cutoff 2%. Letters A to D represents that groups with same letter are clustered together by GMYC analysis. Vertically bar indicates that besides there are a clear separation between clusters, the genetic divergence are less than 2%.

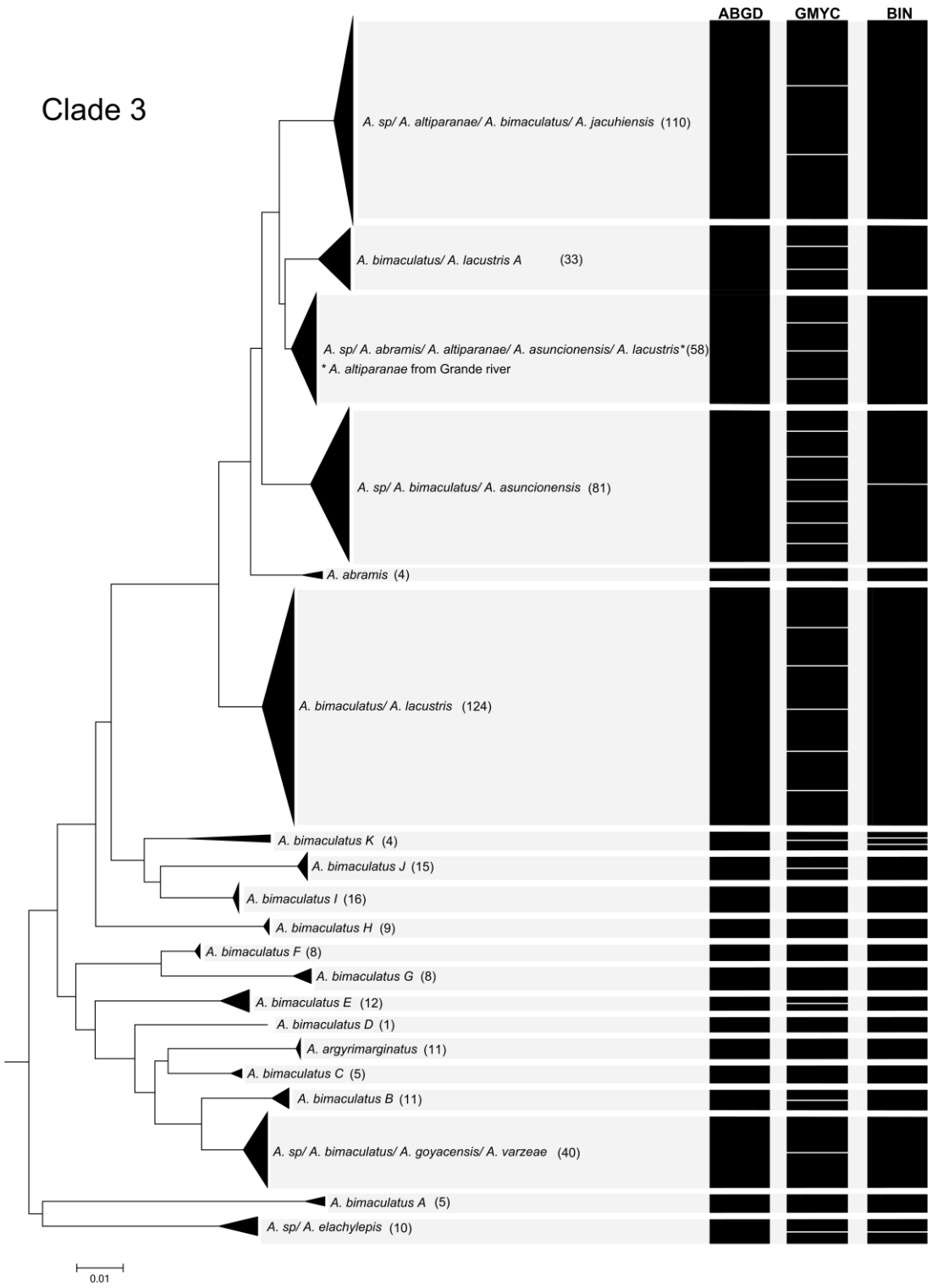

Figure 5. NJ dendrogram showing the groups of Clade 3. Gray shaded areas represents groups in each cluster delimited by K2P cutoff 2%.

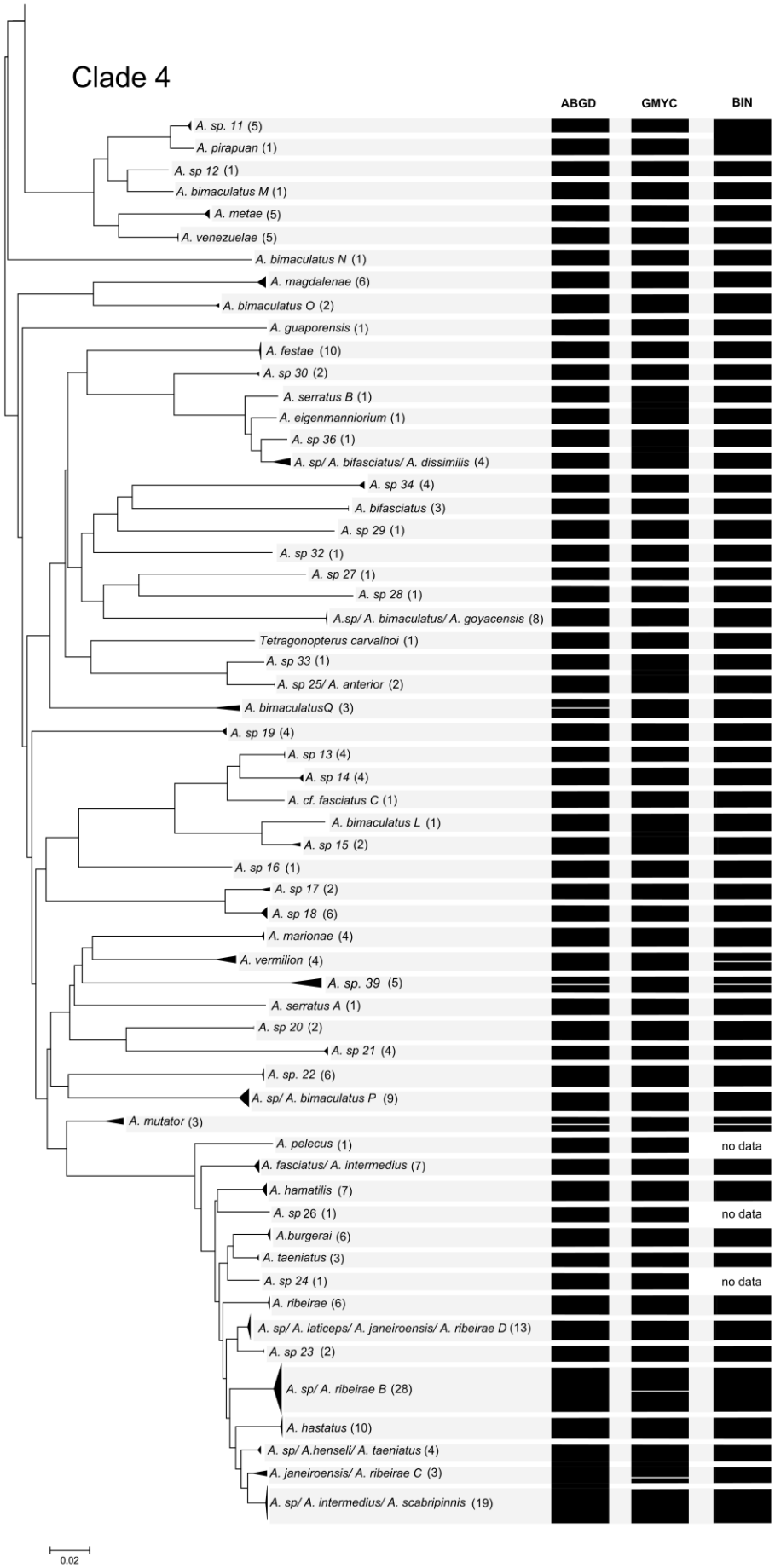
Clade 4
ABGD
GMYC
BIN
A. sp. 11 (5)
A. pirapuan (1)
A. sp 12 (1)
A. bimaculatus M (1)
A. metae (5)
A. venezuelae (5)
A. bimaculatus N (1)
A. magdalenae (6)
A. bimaculatus O (2)
A. guaporensis (1)
A. festae (10)
A. sp 30 (2)
A. serratus B (1)
A. eigenmanniorium (1)
A. sp 36 (1)
A. sp/ A. bifasciatus/ A. dissimilis (4)
A. sp 34 (4)
A. bifasciatus (3)
A. sp 29 (1)
A. sp 32 (1)
A. sp 27 (1)
A. sp 28 (1)
A.sp/ A. bimaculatus/ A. goyacensis (8)
Tetragonopterus carvalhoi (1)
A. sp 33 (1)
A. sp 25/ A. anterior (2)
A. bimaculatusQ (3)
A. sp 19 (4)
A. sp 13 (4)
A. sp 14 (4)
A. cf. fasciatus C (1)
A. bimaculatus L (1)
A. sp 15 (2)
A. sp 16 (1)
A. sp 17 (2)
A. sp 18 (6)
A. marionae (4)
A. vermilion (4)
A. sp. 39 (5)
A. serratus A (1)
A. sp 20 (2)
A. sp 21 (4)
A. sp. 22 (6)
A. sp/ A. bimaculatus P (9)
A. mutator (3)
A. pelecus (1)
no data
A. fasciatus/ A. intermedius (7)
A. hamatilis (7)
A. sp 26 (1)
no data
A.burgerai (6)
A. taeniatus (3)
A. sp 24 (1)
no data
A. ribeirae (6)
A. sp/ A. laticeps/ A. janeiroensis/ A. ribeirae D (13)
A. sp 23 (2)
A. sp/ A. ribeirae B (28)
A. hastatus (10)
A. sp/ A.henseli/ A. taeniatus (4)
A. janeiroensis/ A. ribeirae C (3)
A. sp/ A. intermedius/ A. scabripinnis (19)
0.02

Figure 6. NJ dendrogram showing the groups of Clade 3. Gray shaded areas represents groups in each cluster delimited by K2P cutoff 2%. Results from BIN are partially presented because lack of data.

Supplemental files

These files are in digital format.

Capítulo 2

Multilocus phylogenetic analysis of *Astyanax* (Characiformes: Characidae) reveals multiple lineages

Running title: Phylogeny *of Astyanax*

Keywords: Freshwater fish, Phylogenetics, Neotropical, Systematic, Biodiversity

Abstract

Characidae is the most species rich family of Characiformes. Morphological and molecular studies have been conducted in this family, but some relationships still remain uncertain. One particular species rich genus and scarcely studied extensively is *Astyanax*. This genus is characterized by the presence of species complexes and considered polyphyletic, but this was not tested over a broad representative sample of the genus. Here, we present an effort of sampling including almost the entire area of occurrence of the genus (Central and South America) used in a phylogenetic analysis based on multilocus data from mitochondrial and nuclear genes. The data supports that all *Astyanax* belong to Clade C, with just one exception: *A. festae.* The results suggests that are at least four lineages within *Astyanax*, comprising the three major species complexes and specimens from Central America, plus one formed by species from South America coastal rivers. Also, *Astyanax* can be found closed allied with *Moenkhausia*, *Jupiaba* and *Deuterodon*, all morphological similar genera. These findings helps to better understand the complex origin of this very complex and controversial group of fishes.

Introduction

Fishes represents the major group of vertebrates (Pough, 2008). Among them, the Neotropical region has more than 7000 species currently described, occupying a vast area from the portion of Central America and extending to the southern portion of South America (Albert & Reis, 2011). The diversity of Neotropical fish is associated with the geomorphological changes in the South American continent. Since the separation of the South American/African plate about 118 My, several events such the separation of watersheds, uplift of the Andes (starting ~90 My), the rise of Eastern Cordillera (separation cis and transandines regions ~12 My), Orinoco and Maracaibo split (~8 My) and Orinoco and Amazon (~10 and 8 My). In addition, marine incursions and regressions in the low lands of the continent, as well as fluctuations in sea level and glaciation periods caused many vicariant species culminating in the great biodiversity of fish (Albert et al., 2006; Albert & Reis, 2011; Lundberg et al, 1998).

A species rich and complex genus is *Astyanax*, which was described by Baird & Girard (1854) and has 142 valid species (Froese and Pauly, 2015). Eigenmann (1921, 1927) made the first morphological revisions of the genus, describing new species and including some subgenus in *Astyanax*, but many of these today are allocated in other genera. After Eigenmann, only Gery (1975) studied the genus as a whole. Garutti (1995) also revised *Astyanax*, but restricted to A. bimaculatus complex. At least two other complexes are described for the genus, A. fasciatus and A. scabripinnis (Moreira-Filho & Bertollo, 1991; Melo & Buckup, 2006).

The monophyly of the genus is questioned, and several authors (Rosen, 1972; Weitzman & Fink, 1983; Weitzman & Malabarba, 1998; Mirande, 2010; Oliveira et al., 2011) suggest that possibly there are several independent lineages within *Astyanax*. In addition, from the cytogenetic point of view, the genus is very variable, with 2n = 36 to 2n = 50 chromosomes, with most of the species with chromosome number conserved 2n = 50 (Pazza & Kavalco, 2007).

Despite efforts in an attempt to elucidate the phylogenetic relationships of the group, little is known about the relationships within the genus and between other members of Characidae. The first study based on molecular data of partial sequences of mtDNA (12S and 16S) trying to understand the relationship of characiforms was performed by Ortí and Meyer (1997). After that, a study involving analysis within Characidae and other groups, based on two mitochondrial and four nuclear genes, the alignment of 3660 base pairs revealed the presence of a well-supported clade formed by *Astyanax, Astyanacinus, Moenkhausia, Inpaichthys , Hemigrammus* and *Hyphessobrycon* (Calcagnotto et al., 2005).

Javonillo et al. (2010) based on molecular data from three mitochondrial genes (12S, 16S and COI) and one nuclear (RAG2) in a total of 2940 base pairs analyzed two species of the genus *Astyanax* (*A. cremnobates* and *A. mexicanus*) and considered the group polyphyletic. More recently, in a more extensive DNA analysis of Characidae, with two mitochondrial genes (16S rRNA and cytochrome b) and three nuclear (Myh6, RAG1 and RAG2, total of 4680pb), also examined the position of three species of A*styanax* (*A. aeneus, A.* mexicanus and *A. jordani*) between the characids. The results include *Astyanax* in Clade C (as proposed by Javonillo et al., 2010), along with other

specious genera such as *Hemigrammus, Hyphessobrycon, Moenkhausia, Knodus* and *Jupiaba* and two subfamilies Stethaprioninae and Rhoadsiinae. Although this work has involved a large number of family Characidae species, only a small portion of the genus *Astyanax* was analyzed.

Regarding morphological studies, in a phylogenetic analysis of Characidae involving 160 species (17 from *Astyanax*) and based on the study of 360 morphological characters, Mirande (2010) proposed that *Astyanax* is polyphyletic. According to the author the species of *Astyanax* analyzed belong to three groups identified as (1) "*Hyphessobrycon luetkenii clade*" with a *Astyanax latens* more related to two others species of *Hyphessobrycon*, (2) "*Astyanax paris clade*" composed only by *A. paris*, and (3) "*Astyanax clade*" formed by several species of *Astyanax* and species of *Hyphessobrycon, Psellogrammus* and *Markiana*.

The resolution of internal groups (genera and subfamilies) of Characidae is fundamental to understanding the evolutionary relationships within the family. The few molecular studies involving *Astyanax*, problems of identification of their species, the uncertain number of species and the polyphyletic hypothesis, makes the study of this group necessary, so that the internal relationships can be established, the relationship of this with other genera within the family and also to the knowledge of the diversity of species group. Thus, we present here the phylogeny of *Astyanax* based on molecular characters.

Material and methods

Taxon sampling

We included 101 *Astyanax* taxa based on the sampling of DNA barcoding data proposed by Rossini et al. (unpublished) comprising almost the entire area of occurrence of the genus (Central and South America). Data of some *Astyanax* specimens from Central America were obtained from GenBank. All other taxa from clades A, B and C, studied by Oliveira et al. (2011), were included in the analyzes. *Chalceus erythrurus* was used to root the trees, since this genus represent the sister group of Neotropical characids (Calcagnotto et al, 2005; Oliveira et al, 2011) (Table 1). Tissue samples were preserved in 95% ethanol and the specimens are deposited in the collection of the Laboratory of Biology and Genetic of Fish (LBP) of Universidade Estadual Paulista, Brazil.

DNA extraction and sequencing

Total DNA was extracted from muscle fragments following the Canadian Center for DNA-Barcoding (CCDB) protocol (available http://www.ccdb.ca) or following the protocol NaCl adapted from Aljanabi & Martinez (1997). Partial sequences were obtained from the 16S rRNA (16S), *cytochrome C oxidase subunit I* (COI), *ATP synthase 6 and 8* (ATPase 6/8) and *cytochrome b* (Cytb). Additionally, gene sequences from *myosin heavy chain gene 6* (Myh6), *recombination activating gene 1* (RAG1), *recombination activating gene 2* (RAG2) were obtained by nested PCR procedure according to Oliveira et al. (2011). Polymerase chain reactions (PCR) were perfomed in 12.5 µl,

containing: 1µl DNA (concentration 10-50 ng/µl), 0.25µl of each of the Forward and Reverse primers (concentration 10 mM), 1,25µl reaction buffer, 0.2µl of 200mM dNTPs mix, 0.37µl $MgCl_2$ and 0.0625µl (5 units/µl) Platinum Taq DNA polymerase (Invitrogen). The amplifications were performed in a thermocycler (Veriti® 96-Well Thermal Cycler, Applied Biosystems) following: initial denaturation of 5 minutes at 96°C, 35 cycles, 45 seconds at 96°C, 45s at 48-58°C (according to primer, table 2), 60-90s at 72ºC, with a finally extension step at 72°C for 5 minutes. PCR amplified products were cleanup with ExoSAP-IT (USB Corporation) and sequenced in both directions using the BigDye Terminator v3.1 Cycle Sequencing Kit (Life Technologies) in ABI3130 Genetic Analyzer automated sequencer (Applied Biosystems).

Sequence and phylogenetic analysis

The sequences of each gene were aligned using the default parameters of MUSCLE algorithm (Edgar, 2004) implemented in the MEGA software 5 (Tamura et al., 2011). The final alignment of concatenated sequences was visually inspected to correct any alignment errors. The occurrence of nucleotide substitutions saturation was estimated by saturation index (Iss) as described by Xia et al. (2003) and Xia & Lemey (2009) in DAMBE5 program (Xia, 2013).

Maximum Likelihood analyses (ML) were generated with the partitioned data set (19 partitions, one for gene 16S and one for each position of the codon for other genes, table 3) in RAxML (Stamatakis 2006) through the CIPRES portal (Miller et al., 2010). All parameters were used in the default, with the analysis starting from a random tree. The nucleotide substitution model was

used GTR + G, since this is the only model implemented in RAxML (Stamatakis et al., 2008). The robustness of analysis was investigated using 1000 pseudoreplicates. Bayesian analysis was performed using the program MrBayes 3.2.3 (Ronquist et al., 2012) through the CIPRES portal (Miller et al., 2010). The choice of nucleotide substitution model was under the BIC (Bayesian Information Criterion) for each partition obtained by the PartitionFinder v1.1.1 program (Lanfear et al, 2012; table 3). Two independent analyzes with four MCMC chains were run each under 30,000,000 replicas, sampling every 10,000 steps. The distribution of log likelihoods was analyzed in TRACER v1.6 program (Rambaut et al., 2013). Consensus tree (majority rule of 50%) was built with software 4.0b10 PAUP * (Swofford, 2003).

Results

A total of 6060bp were obtained after sequence concatenation (16S, 645bp; COI, 633bp; ATPase 6/8, 719bp; Cytb, 992bp; Myh6, 755bp; RAG1, 1265bp; RAG2, 1051bp). Substituition saturation analysis did not indicate occurrence of saturation in transitions or tranversions in any tests (Iss.cSym and Iss.cAsym). Data of variable sites, number of sequences for each gene and base composition are shown in table 2. The nucleotide substitution models obtained for each partition used in Bayesian analysis are shown in Table 3.

The combined data from the majority consensus trees of Bayesian analysis and ML are shown in figures 2-9. The values of the nodes with high support (> 0.9 posterior probability,> 50% bootstrap values) are shown with a black circle for the Bayesian analysis and a white diamond for the ML analysis.

The topology of the trees between the two analyzes were similar, but the support of the nodes was lower in the ML analysis. The complete trees are presented in the Supplemental Material (Supplemental F1 and F2).

The two analyzes returned similar topologies, with the nomenclature of clades A, B and C are the same proposed by Javonillo et al. (2010) and maintained by Oliveira et al. (2011) (Figure 1). *Astyanax* is not monophyletic. In both analyses *A. festae* was included in Clade A (figure 2), but all the others *Astyanax* are in Clade C (figures 4-9), however distributed within several others species from different genera. None *Astyanax* species can be found in Clade B (figure 3). At least four major clades of A*styanax* inside Clade C can be identified, with representatives from A. scabripinnis and A. fasciatus species complexes, one with A. bimaculatus species complex, and other with *Astyanax* from Central America and a last one with *Astyanax* from South American coastal region together with other genera. The remaining *Astyanax* are in groups together genera such as *Moenkhausia* and *Jupiaba*.

In the figure 9, representatives of A. scabripinnis and A. fasciatus species complexes are related with *Oligosarcus*. The figure 8 has the specimens of A. bimaculatus complex. Of these, the sister group of all other *A. bimaculatus* is *A. elachylepis*, a species that not belongs to the A. bimaculatus complex. Regarding *Astyanax* from Central America, these also do not represent a monophyletic group, since *Bramocharax* is present between them. A well supported node contain the genera *Stygichthys* (cave fish), *Myxiops*, *Probolodus* and *Deuterodon* together *Astyanax* from coastal rivers of Brazil (*A. taeniatus, A. ribeirae, A. hastatus, A. janeiroensis, A. intermedius, A. pelecus, A. hamatilis* and *A. burgerai*) and coastal specimens from Guyana (*A. mutator*

and *A. cf. wappi*). Another fact is that *Ctenobrycon* and *Psellogrammus* are sister groups of *A. magdalenae* (figure 7). In addition, *Jupiaba* (figure 6) is sister group of *A. anterior* and others *A. sp.* Also some *Astyanax* are found among the genera *Moenkhausia* and *Jupiaba* (Figure 5; *A. sp '14', A. sp '19 ', A. sp '20'* and *A. sp '21'*).

Discussion

The great anatomical variety and the high number of species of *Astyanax* are possibly the main cause of the absence of more extensive studies of this genus. Problems involving high complexity in the identification and description of new species and the phenotypic plasticity allied to the great variability of the characters used in morphological studies reflects the present results. The genus is shown as polyphyletic according to molecular (Calcagnotto et al, 2005;. Ornelas-Garcia, 2008; Javonillo et al, 2010;. Oliveira et al, 2011.) and morphological studies (Mirande, 2010). Rossini et al. (unpublished) found that are at least four groups inside *Astyanax* with more than 100 species could being recognized based on DNA barcode data, but a great number of species are assigned to be a unique barcode cluster, illustrating the enormous problems of identification and description of the species from this genus.

The first species complex has been described based on cytogenetic data. Moreira-Filho & Bertollo (1991) showed that there are morphological and karyotypic differences between specimens of *A. scabripinnis*, suggesting that are more than one species sharing the same morphological unit (A. scabripinnis complex). Currently there are over 15 species belonging to this complex with

very similar morphological characteristics (Bertaco & Lucena, 2006). Another complex within the genus based on morphological analysis is A. fasciatus (Melo, 2005 - unpublished PhD Thesis). Thus, we can identify a group of several species morphologically related to these complexes in figure 9 (*A. paranae, A. rivularis, A. pirapuan, A. xavante, A. laticeps, A. obscurus, A. scabripinnis, A. eigenmanniorum, A. goyanensis, A. fasciatus, A. schubarti, A. parahybae* and *A. cf. jequitinhonhae*).

Oligosarcus is considered related to *Hyphessobrycon* and closed allied to *Astyanax* by Javonillo et al. (2011). Besides this, the species *A. cremnobates* (A. scabripinnis complex) is the sister group of *Oligosarcus*, as found for *A. paranae, A. rivularis, A. pirapuan, A. xavante, A. laticeps, A. obscurus, A. scabripinnis, A. eigenmanniorum* and *A. goyanensis* (also, from A. scabripinnis complex) in our analysis. About the A. fasciatus complex, the species identified here are *A. fasciatus* (from the supposed type locality), *A. schubarti, A. parahybae* and *A. cf. jequitinhonhae*. Due to the low number of species of *Astyanax* (3 species) analyzed by Oliveira et al. (2011), *Oligosarcus* is considered the sister group of *Astyanax* from Central America. With the sampling number of species expanded in this study, we can say that our results support previously found and confirm that *Oligosarcus* is the sister group of *Astyanax* from Central America and more, they are closely related to A. scabripinnis and A. fasciatus species complexes. Inside this, it also can be found the case of *A. erythropterus* from Argentina, that according to Soneira et al. (2010), is very similar to *A. pelegrini* because it has a large number of rays in anal fins (40-42 to *A. pelegrini* - data taken from Eigenmann, 1921; 38-42 for *A. erythropterus*) and body deep. Lima et al. (2003) include *A. pelegrini* as *species*

inquirendae whitin the genus *Ctenobrycon* but Mirande et al. (2006) believe that this species is different from last genus because it does not have the same pattern of scales as described to *Ctenobrycon*. Despite the similarities between these species, *A. erythropterus* in the present study is placed inside of the group of *Astyanax*.

Further, among the representatives of A. scabripinnis and A. fasciatus complexes, we note the presence of endemic species from Iguaçu River basin, of which *A. minor, A. gymnodontus* and *A. dissimlis* are closely related and more distant are the species *A. varzeae* and *A. serratus*. Unexpected, species from very distant regions of Brazil, as *A. metae* and *A. venezuleae* are also included in this group, with only mention of *A. venezuelae* resembles to *A. fasciatus* in body coloration (Schultz, 1944).

Another species complex that could be identified is A. bimaculatus (figure 8). The only species not recognized to belongs to this complex is *A. elachylepis*, that shows to be a sister group of A. bimaculatus. This species complex was revised by Garutti (1995) for Paraná-Paraguay, São Francisco and Amazon River basins that include species that have a black horizontal elongated humeral spot and a black spot on the caudal peduncle which extends to the tips of the median caudal rays. The number of species proposed by the author reaches almost 30, which had only one species to the Upper Paraná River Basin, but DNA barcode data shows that are in fact two species recognized as a single morphological unit, *A. altiparanae* (Pereira et al., 2013; Rossini et al., unpublished). The only study who included more than one species (*A. asuncionensis* and *A. abramis*) of this complex was Mirande (2010) and also found a closely related relationship between these species as herein.

A third lineage identified is from *Astyanax* of Central America (figure 7). Those *Astyanax* are the type of the genus, proposed by the form of *A. argentatus* (today, *A. mexicanus*). Monophyletic hypothesis of *Astyanax* from Central America was proposed by Strecker et al. (2004) based on Cytb gene using populations from Mexico, Guatemala and Belize, but Ornelas-Garcia et al. (2008) with inclusion of more species and genes, says that this hypothesis only can be supported if *Bramocharax* are considered as belonging to *Astyanax*. Since that basically data analyzed in this work are from this last study, our conclusions follow the previous one. The only species outside Central America and included in this group is *A. fasciatus* (here called *A. fasciatus viejita*). This species was described as *Tetragonopterus viejita* (Cuvier & Valenciennes, 1848) and is considered a synonymy of *A. fasciatus*, but Melo (2005 – unpublished PhD Thesis) consider it as a valid species.

The presence of *Astyanax* together with *Deuterodon* and *Myxiops* it is supported by the hypothesis proposed by Oliveira et al. (2011), which according to the authors these genera inhabits very ancient land formations from southeastern and northeastern of Brazil. The similarities shared between *Deuterodon* and *Astyanax* are described by Eigenmann (1921), in particular *A. taeniatus:* "*In A. taeniatus the teeth of the sides of the lower jaw tend to become graduate, a condition leading to the distinguishing character of Deuterodon*". Géry (1977) also points that *Deuterodon* body shape are very similar with *Astyanax*, were the only differences are in the teeth. Here, *A. taeniatus* are sister group of *Deuterodon* and other species from coastal zones (*Myxiops* and *Probolodus*). In relation to *Astyanax*, Eigenmann (1921) says that *A. taeniatus* are very similar of *A. intermedius*. *Myxiops* was recently described by Zanata &

Akama (2004) with a series of characters, Including the presence of a single series of teeth in the premaxilla, which differentiates it of *Deuterodon* and *Astyanax*. Despite this, the authors indicate that there are very similarities between the shape and cusps of the teeth between *Myxiops* and *Deuterodon*, in addition of several characters shared with *Astyanax*, where the body shape and color pattern are similar of those described to *A. scabripinnis*. Besides similarities in overall body shape and color pattern shared with *A. scabripinnis*, this genus is not related with the other A. scabripinnis complex species as proposed by this work. Another species, *A. burgerai*, which also is morphologically similar to *A. scabripinnis* (Zanata & Camelier, 2009) such *Myxiops*, herein shows that are close related to the coastal species. Our results shows that probably *Astyanax* and the other related genera from coastal zones could be intimately associated by the geologic evolution in the limits of the Eastern Brazilian Shield, with the repeated episodes of sea-level rises, were the different populations of basins became isolated during this events resulting in speciation events (Buckup, 2011).

In the case where *Astyanax* is mixed with *Moenkhausia* and *Jupiaba* it shows clearly the need of a revision of the boundaries of these genera. As discussed by some authors about the weakness of characters used to separate *Moenkhausia* by Eigenmann (1917), where specimens with lateral line interrupted, complete and incomplete are placed together in *Moenkhausia sanctafilomenae* (Benine et al., 2009) and also with the description of *Moenkhausia diktyota* with incomplete lateral line (Lima & Toledo-Piza et al., 2001), shows a overlapping of this characters in these fish groups. Even in *Jupiaba*, with well-established characters: presence of developed pelvic bones

developed spine like vs. absence in other genus (Zanata, 1997), the present analysis shows that there is also relations among some *Jupiaba*, *Astyanax* and *Moenkhausia*.

Other genera related to *Astyanax* are *Ctenobrycon* and *Psellogrammus*, being the two lasts considered morphological very similar between them (Eigenmann, 1921; Mirande, 2010). Morphological studies in Characidae allocates *Psellogrammus* inside of *Astyanax clade* (Mirande, 2010), since that *A. magdalenae* could be related to others genera, we suggests that this species also need a revision.

We found that *A. festae* not belongs to Clade C, but to Clade A (figure 2). This hypothesis is in agreement with the fact that this species has only four teeth in the inner series of premaxilla as described by Eigenmann (1921). Our morphological analysis of the specimens confirmed the presence of this number. Eigenmann (1914) says: *"For instance Astyanax festae and Bryconamericus peruanus of the Pacific slope of Ecuador are more intimately related than festae is to Astyanax anterior of the upper Amazon. And in this case, Astyanax brevirostris or Bryconamericus brevirostris whichever it may be, is intermediate between the two. I am not competent to say whether brevirostris is moving from Bryconamericus to become an Astyanax, or whether it has just completed the reverse process. Certainly festae and brevirostris are more intimately related, have had a common ancestor at a less remote time, than either of them with an Astyanax or Bryconamericus of southeastern Brazil*". In this way, since that this study revealed *A. festae* are sister group of *Bryconamericus emperador* in both analysis, possess only four teeth in the inner row of the premaxillary and are not placed in the Clade C as the remaining

Astyanax, we recommend a revision of this species together of others species of *Bryconamericus*. Other member closed allied to these species into Clade A is *Markiana*. Mirande (2010) found a relationship among *Astyanax* and *Markiana*, but, as in the Oliveira et al. (2011) analysis, the presente data refute this hypothesis. The remaining relationships insides Clades A and B are well discussed previous by Oliveira et al. (2011).

Conclusions

All *Astyanax* are placed in Clade C (with the exception of *A. festae*) as observed in previous molecular studies. Our results indicates and confirms the hypothesis of polyphiletism of *Astyanax*, since it is possible to identify at least four main lineages involving the *Astyanax* from the three major species complexes and a coastal one, plus another species being closed allied to other genera such *Moenkhausia* and *Jupiaba*. This reinforces the idea that the characters used to define *Astyanax* species need to be revised, since that evolution could be as result of morphological convergence to similar ecological factors and habitats.

References

Albert, J., N. R. Lovejoy & W. G. R. Crampton. 2006. Miocene tectonism and the separation of cis- and trans-Andean river basins: Evidence from Neotropical shes. Journal of South American Earth Science, 21:14-27.

Albert JS, Reis R.E. (2011). Historical biogeography of Neotropical Freshwater Fishes. Univ. of California Press, Berkeley, Los Angeles.

Aljanabi SM, Martinez I 1997. Universal and rapid salt-extraction of high genomic DNA for PCR-based techniques. Nucleic Acid Res 25: 4692-4693.

Benine, R.C., Mariguela, T.C., Oliveira, C. (2009). New species of Moenkhausia with comments on the Moenkhausia oligolepis species complex. Neotropical Ichthyology, 7(2): 161-168.

Buckup, P.A. (2011). The Eastern Brazilian Shield. In Albert JS, Reis R.E. (editors). Historical biogeography of Neotropical Freshwater Fishes. Univ. of California Press, Berkeley, Los Angeles.

Calcagnotto, D., Schaefer, S.A., DeSalle, R., 2005. Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. Mol. Phylogenet. Evol. 36, 135-153.

Cuvier, L.B., Valenciennes, A. (1848). Memoirs du Museum D'Historie Naturelle Des Poissons. Paris, 22: 1-149.

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

Eigenmann, C.H. (1914). Some results from studies of South American fishes. IV. New genera and species of South American fishes. Indiana University Studies, no. 20: 20-31.

Eigenmann, C. H. (1917). The American Characidae. Part I. Memoirs of the Museum of Comparative Zoology, 43: 1-102.

Eigenmann, CH. (1921). The American Characidae. Memoirs Museum of Comparative Zoology, vol. 43, no. 3, p. 209-310.

Eigenmann, CH. (1927). The American Characidae. Memoirs Museum of Comparative Zoology, vol. 43, no. 4, p. 311-428.

Froese, R., e Pauly, D. (Eds.) (2015). FishBase, World Wide Web electronic publication:www.fshbase.org,ver. Accessed January 2015.

Garutti, V. (1995). Revisão taxonômica dos Astyanax (Pisces, Characidae), com mancha umeral ovalada e mancha no pedúnculo caudal, estendendo-se à extremidade dos raios caudais medianos, das Bacias do Paraná, São Francisco e Amazônica. Tese (livre docência) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. São José do Rio Preto : [s.n.]. viii, 286 f. : il.

Géry, J., 1977. Characoids of the World. Neptune City: T. F. H. Publications. 672 p.

Ortí, G., Meyer, A. (1997). The radiation of characiform fishes and the limits of resolution of mitochondrial ribosomal DNA sequences. Syst. Biol., 46: 75-100.

Javonillo, R. et al. (2010). Relationships among major lineages of characid fishes (Teleostei: Ostariophysi: Characiformes), based on molecular sequence data. Molecular Phylogenetics and Evolution 54: 498-511.

Lanfear R, Calcott B, Ho SYW, Guindon S (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29 (6): 1695-1701. http://dx.doi.org/10.1093/molbev/mss020.

Lima, F.C.T., Toledo-Piza, M. (2001). New species of Moenkhausia (Characiformes: Characidae) from the rio Negro of Brazil. Copeia, 2004(4): 1058-1063.

Lima, F.C.T. et al. (2003). Genera Incertae Sedis in Characidae. Pp. 106-169. In: Reis, R. E., Kullander, S. O., & Ferraris, C. J. (Eds.). Check list of the freshwater fishes of South and Central America (CLOFFSCA). Porto Alegre, EDIPUCRS, 729 p.

Li C, Ortí G, Zhang G, Lu G: A practical approach to phylogenomics: The phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evol Biol 2007, 7:44.

Li C, Ortí G: Molecular phylogeny of Clupeiformes (Actinopterygii) inferred from nuclear and mitochondrial DNA sequences. Mol Phylogenet Evol 2007, 44:386-398.

Lovejoy NR, Collette BB: Phylogenetic relationships of new world needlefishes (Teleostei: Belonidae) and the biogeography of transitions between marine and freshwater habitats. Copeia 2001, 1: 324-338.

Lundberg JG. (1998). The temporal context for diversification of Neotropical fishes. In: Malabarba LR, Reis RE, Vari RP, Lucena CAS,. Lucena ZMS, eds. Phylogeny and Classification of Neotropical Fishes. Edipurus, Porto Alegre, Brazil. 67-91.

Melo, F.A.G. (2005). Revisão taxonômica do complexo de espécies Astyanax fasciatus (Cuvier, 1819) (Teleostei: Characiformes: Characidae). Rio de Janeiro: UFRJ/MN. 293pp. Phd Thesis.

Melo, F.A.G. e Buckup, P.A. (2006). Astyanax heseli, a new name for Tetragonopterus aeneus Hensel, 1870 from southern Brazil (Teleostei: Characiformes). Neotropical Ichthyology, 4(1):45-52.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. In Gateway Computing Environments Workshop (GCE).

Mirande, J. M., M. Azpelicueta & G. Aguilera. 2006. Redescription of Astyanax correntinus (Holmberg, 1891) (Teleostei: Characiformes: Characidae), more than one hundred years after original description. Zoologische Abhandlungen (Dresden), 55: 9-15.

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

Moreira-Filho, O. e Bertollo, L.A. (1991). Astyanax scabripinnis (Pisces, Characidae): a species complex. Revista Brasileira de Genética, 14(2): 331-357.

Oliveira, C.; Avelino, G.S.; Abe, K.T.; Maariguela, T.C.; Beine, R.C.; Ortí, G.; Vari, R.P.; Correa 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 2011, 11:275.

Ornelas-Garcia, C. P.; Dominguez-Dominguez, O.; Doadrio, I. (2008). Evolutionary history of the fish genus Astyanax Baird & Girard (1854) (Actynopterigii, Characidae) in Mesoamerica reveals multiple morphological homoplasies. BMC Evolutionary Biology, v. 8, p. 340.

Palumbi SR: Nucleic acids II: the polymerase chain reaction. In Molecular Systematics. Edited by Hillis D, Moritz C, Mable B. Massachusetts: Sinauer Associates Inc.; 1996:205-247.

Pazza R and Kavalco KF (2007) Chromosomal evolution in the neotropical characin Astyanax (Teleostei, Characidae). Nucleus, 50:519-543.

Perdices, A., Bermingham, E., Montilla, A., Doadrio, I. 2002. Evolutionary history of the genus Rhamdia(Teleostei: Pimelodidae) in Central America. Mol. Phylogenet. Evol. 25, 172-189.

Pereira LHG, Hanner R, Foresti F, Oliveira C (2013) Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics 14: 20 doi:10.1186/1471-2156-14-20.

Pough, F. H., Heiser, J. B., Janis, C. M. 2008. A vida dos vertebrados. 4ª ed. São Paulo: Atheneu.

Rambaut, A., Suchard, M.A., Drummond, A.J. 2013. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/tracer

Ronquist F., Teslenko M., van der Mark P., Ayres D. L., Darling A., Höhna S., Larget B., Liu L., Suchard M. A., Huelsenbeck J. P. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 2012;61(3):539-542.

Rosen, D.E. (1972). Origin of the Characid Fish Genus Bramocharax and a Description of a Second, More Primitive, Species in Guatemala. American Museum Novitates, No 2500.

Rossini, B.C., Oliveira, C.A.M., Melo, F.A.G, Bertaco, V., Astarloa, J.M.D., Rosso, J.J., Foresti, F., Oliveira, C. (unpublished). Astyanax species diversity through DNA barcoding: molecular data assessment and the species complexes.

Schultz, L.P. 1949. A further contribution to the ichthyology of Venezuela. Proceedings of the United States National Museum, 99: 1-211.

Soneira, P., Casciotta, J., Almirón, A, Ciotek, L., Giorgis, P. (2010). Redescription of Astyanax erythropterus (Holmberg, 1891) (Teleostei: Characiformes: Characidae) from La Plata basin in Argentina. Neotropical Ichthyology, 8(4):779-785.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogeneticanalyses with thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for theRAxML web servers. Syst. Biol. 57, 758-771.

Strecker, U., Faúndez, V.H., Wilkens, H. (2004). Phylogeography of surface and cave Astyanax (Teleostei) from Central and North America based on cytochrome b sequence data. Molecular Phylogenetics and Evolution 33, 469-481.

Swofford, D.L., 2003. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Version 4. Sinauer Associates Inc., Sunderland, MA.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution 28: 2731-2739.

Weitzman, S.H. e Fink, W.L. (1983). Relationships of the neon tetras, a group of south american freshwater fishes (Teleostei, Characidae), with comments on the phylogeny of new world characiformes. Bull. Mus. Comp. Zool., 150(6): 339-395.

Weitzman, S.H., Malabarba, L.R. (1998). Perspectives about the phylogeny and classification of the Characidae (Teleostei: Characiformes). In: L.R. Malabarba, R.E. Reis, R.P. Vari, Z.M.S. Lucena, C.A.S. Lucena (Eds.), Phylogeny and Classification of Neotropical Fishes, EDIPUCRS, Porto Alegre (1998), pp. 161-170.

Xia, X., Z. Xie, M. Salemi, L. Chen, Y. Wang. 2003. An index of substitution saturation and its application. Molecular Phylogenetics and Evolution 26:1-7.

Xia, X. and Lemey, P. 2009. Assessing substitution saturation with DAMBE. Pp. 615-630 in Philippe Lemey, Marco Salemi and Anne-Mieke Vandamme, eds. The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny. 2nd edition Cambridge University Press.

Xia X (2013) DAMBE5: A Comprehensive Software Package for Data Analysis in Molecular Biology and Evolution. Mol Biol Evol 30: 1720-1728.

Zanata, A.M. (1997). Jupiaba, A New Genus Of Tetragonopterinae With Spine like Pelvic Bones. Iheringia Serie Zoologia 83:99-136.

Zanata, A.M., Akama, A., 2004. Myxiops aphos, new characid genus and species (Characiformes: Characidae) from the rio Lencois, Bahia, Brazil. Neotrop. Ichthyol. 2, 45-54.

Zanata, A.M, Camelier, P. (2009). Astyanax vermilion and Astyanax burgerai: new characid fishes (Ostariophysi: Characiformes) from Northeastern Bahia, Brazil. Neotropical Ichthyology, 7(2):175-184.

Tables

Table 1. Specimens list used in this work.

Group/species	Voucher	Specimen	Locality	Latitude/longitude
*Acestrocephalus sardina**	LBP 6876	33172	Rio Negro/São Gabriel da Cachoeira/AM/Brazil	S 00°08.156' W 67°05.057'
*Aphyocharacidium bolivianum**	LBP 9055	42219	Arara/Arara/RO/Brazil	S 09°36'39.5" W 64°55'38.9"
*Aphyocharax alburnus**	LBP 1587	11774	Rio das Garças/Barra do Garça/MT/Brazil	S 15°54'18.1" W 2°19'24.2"
*Aphyocharax anisitsi**	LBP 3764	22190	Rio Negro/Aquidauana/MS/Brazil	S 19°34'33.7' W 56°14'49.5"
*Aphyocharax pusillus**	LBP 4046	22920	Rio Moa/Cruzeiro do Sul/AC/Brazil	S 7°37'20.0" W 2°47'42.2"
*Aphyocheirodon hemigrammus**	LBP 8306	40025	Rio Araras/Araras/SP/Brazil	S 22°22'42.4' W 47°25'37.9"
*Aphyodite grammica**	LBP 9050	42214	Rio Madeira/Pacaás/ Mutumparaná/ RO/Brazil	S 09°37'05.3" W 64°56'01.3"
*Astyanacinus moorii**	LBP 5783	28195	Rio Muzambinho/Muzambinho/ MG/Brazil	S 21°21'03.1" W 46°29'33.2"
Astyanax abramis	LBP 8427	41612	Rio Sepotuba/Rio Paraguai/La Plata Basin/Tangará da Serra/MT/Brasil	S 14°30'04.4' W 57°34'38.1"
*Astyanax aeneus**	LBP 8938	42019	Chichancanab lagoon/Quintana Roo/México	19,882999420166/-88,8710021972656
Astyanax aeneus[1]	-	MNCN/E:3606Tux	San Joaquin **R**	-
Astyanax aeneus[1]	-	MNCN/E:374GU	Rio Mopan	-
Astyanax altiparanae	LBP 5005	25912	Córrego do Batata/Rio Grande/Rio Paraná/Colômbia/SP/ Brasil	S 20°14'10' W 48°40'42'
Astyanax anterior	LBP 13790	57143	Igarapé na BR 230/Rio Amazonas/Rio Tapajós/Itaituba/PA/Brasil	S 04°30'20.6" W 56°16'58.7"
Astyanax argyrimarginatus	LBP 1547	11885	Córrego das Mulas/Rio Araguaia/Aragarças/GO/Brasil	S 15°54'23.4'' W 52°05'39.4''
Astyanax assuncionensis	LBP 8541	43262	Rio Sepotuba/Rio Paraguai/La Plata Basin/Nova Fernandópolis/MT/Brasil	S 14°56'54.6' W 57°44'11.3"
Astyanax belizanus[1]	-	MNCN/E:485GU	Izabal **L**	-
Astyanax bimaculatus	LBP 2278	15790	Afluente do Rio Orinoco/Rio Orinoco/Caicara del Orinoco /Bolivar/Venezuela	N 07º30'01,2'' W 66º08'14,4''
Astyanax bimaculatus	LBP 3440	20304	Rio Matipó/Rio Doce/Matipó/MG/Brasil	S 20°18'09.3" W 42°20'04.6"

Astyanax bimaculatus	LBP 3962	22792	Rio Mutuca/Rio Paraguai/Cuiabá/MT/Brasil	S 15°21'20' W 56°06'54"
Astyanax bimaculatus	LBP 5368	27009	Igarapé Piumquara/Rio Amazonas/Laranjal do Jari/AP/Brasil	S 00°34'54'' W 52°31'46''
Astyanax bimaculatus	LBP 9417	42622	Rio Guamá/Rio Guamá/Amazonas/Ourém/PA/Brasil	S 01°34'00.5" W 47°09'51.4"
Astyanax bimaculatus	LBP 8591	43405	Corixo Rio Arinos/Rio Tapajós/Amazônia/Diamantino/MT/Brasil	S 14°09'14.9' W 56°05'41.6"
Astyanax bimaculatus	LBP 9514	44681	Riacho sem nome/Rio Tocantins/Amazônia/Formosa/GO/Brasil	S 15°28'54.8' W 47°20'31.4"
Astyanax bimaculatus	LBP 10599	49361	Igarapé São Francisco/Amazonas/Rio Branco/AC/Brasil	S 09°56'16.6'' W 67°52'48.6"
Astyanax bimaculatus	LBP 10809	49957	Afluente Rio Guaporé/Amazonas/Madeira/Vila Bela da Santíssima Trindade/MT/Brasil	S 14°58'08.5" W 59°58'59.5"
Astyanax bimaculatus	LBP 10856	50093	Riacho Fazenda Lagoa Azul/Amazonas/Madeira/Ariquemes/RO/Brasil	S 09°58'59.4" W 63°08'15.4"
Astyanax bimaculatus	LBP 11006	50543	Rio Lajeado/Amazonas/Madeira/Guajará Mirim/RO/Brasil	S 10°26'23.5" W 65°20'34.1"
Astyanax bimaculatus	LBP 15851	64114	Tributário rio Toguro/Rio Xingu/Querência/MT/Brasil	S 13°00'26.6" W 52°11'27.0"
Astyanax bimaculatus	LBP 15595	64219	Lagoa sem nome/Rio Amazonas/Bonfim/RR/Brasil	N 03°19'16.1" W 59°56'46.1"
Astyanax bimaculatus	LBP 15595	64220	Lagoa sem nome/Rio Amazonas/Bonfim/RR/Brasil	N 03°19'16.1" W 59°56'46.1"
Astyanax bimaculatus	LBP 17767	70233	Quebrada sem nome/Rio Ucayali/Amazonas/Coronel Portillo/Pucallpa/Peru	S 08°30'14.4" W 74°48'21.8"
Astyanax bimaculatus	LBP 18712	74618	Caño La Union/Rio Meta/Orinoco/Villavicencio/Meta/Colombia	-
Astyanax bockmanni	LBP 7264	35818	Rio Arapuca/Rio Paranaíba/La Plata Basin/Bela Vista de Goias/GO/Brasil	S 17°05'58.0" W 48°45'41.5"
Astyanax burgerai	LBP 8328	40078	Rio do Braço/Costeira/Ihéus/BA/Brasil	S 14°41'11.7' W 39°16'28.0"
Astyanax burgerai	LBP 18840	61848	Riacho Palmeirá/rio Almada/Almadina/BA/Brasil	-14.42376 39.37332
Astyanax cf. bimaculatus	LBP 9417	42619	Rio Guamá/Rio Guamá/Amazonas/Ourém/PA/Brasil	S 01°34'00.5" W 47°09'51.4"
Astyanax cf. jequitinhonhae	LBP 8311	38397	Rio Jequitinhonha/Rio Jequitinhonha/Itaobim/MG/Brasil	S 16°30'35.0' W 41°20'02.0"
Astyanax cf. wappi	LBP 17448	69024	Kuribrong River/rio Potaro/rio Essequibo/Potaro-Siparuni/Guiana	N 05°20'30.2'' W 59°32'30.0''
Astyanax dissimilis	LBP 16173	66550	Riacho sem nome/Tributário Rio Iguaçu/Foz do Iguaçu/PR/Brasil	S 25°32'13.2" W 54°21'11.2"
Astyanax eigenmanniorum	LBP 13093	54684	Rio Cará/Rio Uruguai/Cacique Duble/RS/Brasil	S 27°46'44.8" W 51°39'31.7"

Astyanax elachylepis	LBP 1811	13115	Córrego Água Funda/Rio Araguaia/Barra do Garças/MT/Brasil	S 15°52'40.1" W 52°18'15.0"
Astyanax erythropterus	UNMDP	UNMDP-T 0557	Parana River (San Nicolas)/Buenos Aires/Argentina	-
Astyanax fasciatus	LBP 10302	42092	Rio São Francisco/Rio São Francisco/São Roque de Minas/MG/Brasil	S 20°20'53.1' W 46°04'10.7"
Astyanax fasciatus viejtita	LBP 6116	29588	Rio Apon/Lago Maracaibo/Machiques de Perijá/Zulia/Venezuela	N 10°01'42.0" W 72°25'58.0"
Astyanax festae	LBP 9351	43911	Rio Zarumilla/Pacífico/Papayal/Tumbes/Peru	S 03°32'00.1" W 80°13'41.1"
Astyanax goyacensis	LBP 4013	22836	Córrego Taquaralzinho/ Rio Araguaia/Barra do Garça/MT/Brasil	S 15°42'43.4" W 52°15'32.1"
Astyanax goyanensis	LBP 17143	68317	Rio São Miguel/Rio Tocantins/Alto Paraíso de Goiás/GO/Brasil	S 14°10'38.4" W 47°46'27.1"
Astyanax gymnodontus	LBP 17561	69179	Rio Iguaçu/Porto Barreiro/PR/Brasil	S 25°35'15.6" W 52°17'53.3"
Astyanax gymnogenys	LBP 17570	69188	Rio Iguaçu/São Jorge d'Oeste/PR/Brasil	S 25°32'34.3" W 53°01'44.2"
Astyanax hamatilis	LBP 7187	34870	Rio Lapão/Rio Paraguaçu/Atlântico/Lençóis/BA/Brasil	S 12°32'33.6" W 41°22'51.5"
Astyanax hastatus	LBP 14400	60521	Riacho sem nome/Costeira/Ubatuba/SP/Brasil	S 23°23.345' W 45°01.405'
Astyanax hubbsi1	-	MNCN/DNA:32446	El Ahuaje **S**	-
Astyanax hubbsi[1]	-	MNCN/DNA:33368	Peñon Blanco	-
Astyanax intermedius	LBP 6447	29066	Rio Paraíba do Sul/Bacia Costeira Ocidental/Guararema/SP/Brasil	S 23°22'26.2" W 46°03'10.6"
Astyanax jacuhiensis	LBP 4744	25508	Rio Guaíba/Atlântico/Barra do Ribeiro/RS/Brasil	S 30°17'07.0" W 51°18'01.1"
Astyanax janeiroensis	LBP 14419	60585	Riacho sem nome/Costeira/Paraty/RJ/Brasil	S 23°06.816' W 44°43.443'
Astyanax jordani	LBP 4527	24599	Sem procedência/Brasil	
Astyanax lacustris	LBP 10295	47246	Baia da Inhuma/Rio São Francisco/São Roque de Minas/MG/Brasil	S 20°11'03.4' W 45°50'57.9"
Astyanax laticeps	LBP 726	8292	Rio Itapucu/Jaraguá do Sul/SC/Brasil	S 26°26,812' W 49°09.908'
Astyanax magdalenae	LBP 6115	29582	Rio Apon/Lago Maracaibo/Machiques de Perijá/Zulia/Venezuela	N 10°01'42.0" W 72°25'58.0"
Astyanax metae	LBP 18700	74628	Caño Río Negro/Rio Meta/Orinoco/Villavicencio/Meta/Colombia	-
*Astyanax mexicanus**	LBP 8937	42016	Ojo San Bernabe/Spring/Nuevo Leon/México	-

Astyanax mexicanus[1]	-	MNCN/DNA:32262-32263	El Limón **R**	-
Astyanax minor	LBP 17565	69201	Rio Iguaçu/Saudade do Iguaçu/PR/Brasil	S 25°36'05.2" W 52°34'17.8"
Astyanax mutator	LBP 17436	69006	Kuribrong River/rio Potaro/rio Essequibo/Potaro-Siparuni/Guiana	N 05°21'07.0'' W 59°32'45.0''
Astyanax nasutus[1]	-	MNCN:230073	Telica **R**	-
Astyanax nicaraguensis[1]	-	MNCN:225082	Compazague **R**	-
Astyanax obscurus	LBP 3624	21640	Rio Itapucu/Atlântico/Jaraguá do Sul/SC/Brasil	S 25°26'49.3" W 49°09'37.6"
Astyanax orthodus[1]	-	JJ28	Sixaola **R**	-
Astyanax parahybae	LBP 639	7072	Afluente do Riacho Borboleta/Pindamonhangaba/SP/Brasil	S 22°47' W 45°28'
Astyanax pelecus	LBP 18841		rio Panelão/Rio Pardo/Camacan/BA/Brasil	-
Astyanax petenensis[1]	-	MNCN/E:941GU	Candelaria-Yalicar **R**	-
Astyanax pirapuan	LBP 5067	26070	Afluente Rio Aricá Mirim/Rio Cuiabá/Chapada dos Guimarães/MT/Brasil	S 15º46'03.83" W 55º30'44.54"
Astyanax ribeirae	LBP 7387	35493	Rio Água Doce/Ribeira de Iguape/Atlântico/Tapiraí/SP/Brasil	S 22°27'02.2" W 49°14'26.8"
Astyanax rivularis	LBP 10256	47854	Nascente rio São Franciso/Rio São Francisco/São Roque de Minas/MG/Brasil	S 20°14'34.5' W 46°26'47.6"
Astyanax saguazu	UNMDP	UNMDP-T 0646	Ramos Stream/Misiones/Argentina	-
Astyanax scabripinnis	LBP 1205	10629	Afluente do rio do Salto/Rio Tibagi/Ponta Grossa/PR/Brasil	S 25°22.674 W 49°48.322'
Astyanax scabripinnis	LBP 8744	38110	Rio Claro/Rio Grande/La Plata Basin/Delfinópolis/MG/Brasil	S 20°20'32.2' W 46°47'12.2"
Astyanax scabripinnis	LBP 9548	44785	Riacho sem nome/Rio Paraná/Brasília/DF/Brasil	S 15°43'31.7' W 47°56'24.6"
Astyanax schubarti	LBP 5077	26028	Lagoa do Diogo/Rio Mogi-Guaçu/Luis Antônio/SP/Brasil	S 21°37'26.7" W 47°48'22.6"
Astyanax serratus	LBP 13028	54580	Córrego de Inguca/Rio Iguaçu/São Mateus do Sul/PR/Brasil	S 25°45'37.6" W 50°12'07.5"
Astyanax sp.	LBP 3442	20310	Rio Jucu/Atlântico/Domingos Martins/ES/Brasil	S 20°24'30.3" W 40°54'55.5"
Astyanax sp.	LBP 3914	22551	Rio Novo/Rio Paranapanema/Rio Paraná/Avaré/SP/Brasil	S 23°01'27.4" W 48°49'41.0"
Astyanax sp.	LBP 6880	33201	Rio Negro/Rio Amazonas/São Gabriel da Cachoeira/AM/Brasil	S 00°08.156' W 67°05.057'

Astyanax sp.	LBP 7059	34175	Igarapé do 20/Rio Negro/Rio Amazonas/São Gabriel da Cachoeira/AM/Brasil	N 00°02.975' W 66°38.365'
Astyanax sp.	LBP 7083	34615	Igarapé margem direita rio Negro/Rio Negro/Rio Amazonas/São Gabriel da Cachoeira/AM/Brasil	S 00°08.625' W 67°05.605'
Astyanax sp.	LBP 10184	47680	Rio Mucuri/Costeira/Carlos Chagas/MG/Brasil	S 17°41'42.4' W 40°46'11.3"
Astyanax sp.	LBP 10805	49940	Afluente Rio Guaporé/Amazonas/Madeira/Vila Bela da Santíssima Trindade/MT/Brasil	S 14°58'08.5" W 59°58'59.5"
Astyanax sp.	LBP 14197	59340	Igarapé sem nome/Rio Amazonas/Rio Tapajós/Itaituba/PA/Brasil	S 04°27'59.1" W 56°08'42.8"
Astyanax sp.	LBP 18834	61847	Rio Piabinha/Rio Paraguaçu/Mucugê/BA/Brasil	
Astyanax sp.	LBP 17108	66422	Córrego Grande/Rio Araguaia/Araguaiana/MT/Brasil	S15°46'05" W52°05'18"
Astyanax sp.	LBP 17785	70291	Quebrada sem nome/Rio Ucayali/Amazonas/Huánuco/Huánuco/Peru	S 08°39'57.2" W 74°48'08.7"
Astyanax sp.	LBP 17736	70564	Quebrada sem nome/Rio Ucayali/Amazonas/Coronel Portillo/Pucallpa/Peru	S 08°34'30.1" W 74°48'04.7"
Astyanax sp.	UNMDP	UNMDP-T 2564	Misiones/Argentina	-
Astyanax sp. Novo 1[1]	-	MNCN:184817-184818	La Guija **A**	-
Astyanax sp. Novo 2[1]	-	MNCN:2338-2339-MNCN/E:PO70	Maquinas **R**	-
Astyanax sp. Novo 3[1]	-	MNCN:174626	Montebello **A**	-
Astyanax sp. Novo 5[1]	-	MNCN:243147	Ciruelas **R**	-
Astyanax sp. Novo 6[1]	-	MNCN:243156	Colorado **R**	-
Astyanax sp. Novo 7[1]	-	MNCN/DNA:30003	Lagarto **R**	-
Astyanax sp. Novo 8[1]	-	MNCN/DNA:34422	Chagres **R**	-
Astyanax sp.5	UNMDP	UNMDP-T 2700	Adeg Oveja Negra Stream/Misiones/Argentina	-
Astyanax taeniatus	LBP 3411	20359	Rio São Manoel/Rio Paraíba do Sul/Rio Pombas/MG/Brasil	S 21°15'16.2" W 43°11'55.7"
Astyanax varzeae	LBP 1043	9102	Rio dos Patos/Rio Iguaçu/Lapa/PR/Brasil	S 25°52.477' W 49°43.409'
Astyanax venezuela	LBP 18701	74623	Caño Río Negro/Rio Meta/Orinoco/Villavicencio/Meta/Colombia	
Astyanax vermilion	LBP 8319	40053	Rio Almada/Costeira/Ilhéus/BA/Brasil	S 14°39'52.1' W 39°13'26.7"

Astyanax xavante	LBP 1432	12558	Córrego do Sapo/Rio Araguaia/Alto Araguaia/MT/Brasil	S 17°33'42.4" W 53°18'29.7"
*Bario steindachneri**	LBP 4389	24187	Rio Branco/Porto Velho/RO/ Brazil	S 09º34'10.7" W 63º58'08.2"
*Brachychalcinus copei**	LBP 192	8853	Igarapé São Francisco/Rio Branco/AC/Brazil	S 9°56,271' W 67°52,923'
*Bramocharax baileyi**	LBP 8940	42025	Chisec/Chajmaic/Alta Verapaz/Guatemala	15,72 -89,94
*Bramocharax caballeroi**	LBP 8939	42022	Catemaco/Veracruz/México	18,877 -95,292
Bryconadenos tanaothoros[2]	MCP 40399	-	-	-
*Bryconamericus emperador**	LBP 2754	18528	Río Llano Sucio/Santa Rita Arriba/Colón/Panamá	N 09°19'26.2" W 79°46'08.2"
*Bryconamericus exodon**	LBP 7123	34200	Córrego Lageadinho/Sapopema/ PR/Brazil	S 23°55'29.0" W 50°37'27.3"
*Bryconella pallidifrons**	LBP 4646	24696	Aquarium	-
*Carlana eigenmanni**	LBP 3300	19864	Rio Mandinga/Panama	-
*Carlana eigenmanni**	LBP 3301	19865	Rio Playon Chico/Panama	-
*Ceratobranchia cf. delotaenia**	LBP 3257	20042	Rio Chontabamba/Oxapampa/ Pasco/Peru	S 10º36'06,6'' W 075º29'10,8''
*Chalceus erythrurus**	LBP 4211	22727	Rio Juruá/Cruzeiro do Sul/AC/Brazil	S 07°09'49.6' W 73°43'29.7"
*Charax leticiae**	LBP 1480	12700	Rio Taquari - Pesqueiro Recnato Alegre/Coxim/MS/Brazil	S 18°25'42.5" W 54°50'02.8"
*Cheirodon ibicuhiensis**	LBP 4777	25598	Rio Guaíba/Barra do Ribeiro/RS/Brazil	S 30°18'03.9" W 51°20'40.8"
*Cheirodon killiani**	LBP 3115	19803	Río La Laja/Monte Aguila/VIII Region/Chile	S 37°12'54.8" W 72°26'49.1"
*Compsura heterura**	LBP 4733	24984	Rio Ceará-Mirim/Natal/RN/Brazil	S 05°37'47'' W 35°37'09"
*Coptobrycon bilineatus**	LBP 3809	33169	Afluente rio Itatinga/Bertioga/SP/Brazil	S 23°45'01.2' W 46°09'52.9"
Corynopoma riisei[2]	-	-	-	-
*Creagrutus peruanus**	LBP 3267	20057	Rio Santa Cruz/Pozuzo/Pasco/Peru	S 10º02'20,4'' W 075º34'55,4''
Ctenobrycon hauwelllianus	LBP 1696	12776	Lago do Vanico/Rio Solimões/Carero/AM/Brasil	S 03°09'17.3" W 59°53'12.3"
*Ctenobrycon hauxwellianus**	LBP 4095	23538	Rio Japiim/Mâncio Lima/AC/Brazil	S 07°34'28.8' W 72°55'24.9"

*Cyanocharax alburnus**	LBP 4746	25516	Rio Guaíba/Barra do Ribeiro/RS/Brazil	S 30°17'07.0" W 51°18'01.1"
*Cynopotamus kincaidi**	LBP 3225	19449	Lagoa marginal/Nobres/MT/Brazil	S 14°40'32,8" W 56°13'14,0"
*Cynopotamus venezuelae**	LBP 6132	29515	Rio Santa Rosa/Machiques de Perijá/Zulia/Venezuela	N 09°38'53.8' W 72°34'56.4"
*Deuterodon iguape**	LBP 6827	33065	Rio Fau/Miracatu/SP/Brazil	S 24°12,441' W 47°28,616'
Deuterodon parahybae	LBP 10738	49742	Rio Macabu/Paraíba do Sul/Atlântico/Conceição do Macabu/RJ/Brasil	S 22°04'07.8" W41°54'36.2"
Deuterodon pedri	LBP 17098	40125	Rio Mucuri/Costeira/Carlos Chagas/MG/Brasil	S 17°41'42.4' W 40°46'11.3"
Deuterodon supparis	LBP 17091	21712	Rio Garuva/Atlântico/Garuva/SC/Brasil	S 26°00'10.3" W 48°52'22.7"
*Exodon paradoxus**	LBP 4006	23040	Lago Morto/São Félix do Araguaia/MT/Brazil	S 11°40'9" W 50°51'0.30"
*Galeocharax knerii**	LBP 3496	20164	Rio Tietê/Birigui/SP/Brazil	S 21°06'25.2" W 50°15'52.7"
*Gen. & sp. nov.**	LBP 5699	27603	Córrego Taquaral/Barra do Garças/MT/Brazil	S 15°40.678' W 52°17.863"
*Gen. and sp. new **	LBP 7243	33196	Rio Uberaba/Ponte Alta/MG/ Brazil	S 19°40'59.8" W 48°40'08.6"
*Gephyrocharax atracaudatus**	LBP 2753	18519	Río Llano Sucio/Santa Rita Arriba/Colón/Panamá	N 09°19'26.2" W 79°46'08.2"
*Glandulocauda melanogenys**	LBP 4507	24538	Rio Paranapiacaba/Santo André/SP/Brazil	S 23°46'13.2" W 46°18'39.6"
*Gymnocorymbus ternetzi**	LBP 3737	21989	Lagoa Marginal Rio Negro/ Aquidauana/MS/Brazil	S 19°34'54.6' W 56°15'16.5"
*Hasemania sp.**	LBP5967	28455	Rio Paraibuna/Comendador Levy Gasparian/RJ/Brazil	S 22°01'24.3" W 43°10'08.5"
*Hemibrycon taeniurus**	LBP 6847	33168	Upper Arouca River/Trinidad Tobago	N 10°41.320' W 61°19.499'
*Hemigrammus marginatus**	LBP 6292	29419	Córrego Barbacena/Pontal/SP/ Brazil	S 20°56'49.5" W 48°08'51.9"
*Hemigrammus occellifer**	LBP 10838	50033	Rio Jamari/Amazonas/Madeira/Ariquemes/RO/Brasil	S 10°15'54.4" W 63°18'18.9"
Hemigrammus sp.	LBP 10608	49386	Açude Recanto do Socó/Rio Acre/Amazonas/Rio Branco/AC/Brasil	S 09°56'31.0" W 67°53'46.2"
*Hemigrammus ulreyi**	LBP 7604	36267	Lagoa Margina rio Cuiabá/ Barão de Melgaço/MT/Brazil	S 16°11'39.5" W 55°48'25.1"
*Heterocheirodon yatai**	LBP 4872	24954	Rio Yi/Durazno/Durazno/Uruguai	S 33°23'49'' W 56°24'10"
*Hollandichthys multifasciatus**	LBP 698	8791	Afluente do rio Grande/ Paranapiacaba, SP/Brazil	S 23°46.123' W 46°19.467'

*Hyphessobrycon boulengeri**	LBP 17842	22544	Rio Novo/Rio Paranapanema/Rio Paraná/Avaré/SP/Brasil	S 23°01'27.4" W 48°49'41.0"
*Hyphessobrycon eques**	LBP 7615	36278	Lagoa Margina rio Cuiabá/ Barão de Melgaço/MT/Brazil	S 16°11'39.5" W 55°48'25.1"
*Hyphessobrycon megalopterus**	LBP 7613	36932	Lagoa Margina rio Cuiabá/ Barão de Melgaço/MT/Brazil	S 16°11'39.5" W 55°48'25.1"
*Hyphessobrycon reticulatus**	LBP 1049	8939	Afluente rio São João/ Papandu/ SC/Brazil	S 26°22.049' W 50°07.149'
*Hypobrycon maromba**	LBP 6750	33174	Rio Marombas/Curitibanos/ SC/Brazil	S 27°19'49,6'' W 50°45'05,4''
*Jupiaba acanthogaster**	LBP 8623	43469	Corixo Rio Arinos/Rio Tapajós/Amazônia/Diamantino/MT/Brasil	S 14°08'39.8' W 56°05'48.6"
*Jupiaba anteroides**	LBP 7067	34380	Igarapé Miuá/São Gabriel da Cachoeira/AM/Brazil	S 00°06.308' W 66°52.585'
*Jupiaba cf. acanthogaster**	LBP 7935	37269	Rio dos Patos/Nova Mutum/MT/Brazil	S 13°48'03.1" W 56°01'38.4"
Jupiaba pirana	LBP 9142	43072	Igarapé Açu/Rio Guamá/Amazonas/Capitão Poço/PA/Brasil	S 01°34'28.3" W 47°02'03.5"
Jupiaba scologaster	LBP 6881	33206	Rio Negro/Rio Amazonas/São Gabriel da Cachoeira/AM/Brasil	S 00°08.156' W 67°05.057'
Jupiaba sp.	LBP 8513	41782	Rio Salobra/Rio Paraguai/La Plata Basin/Cáceres/MT/Brasil	S 15°19'53.5' W 57°11'31.1"
Knodus cf. chapadae	LBP 8407	40477	Córrego Águas Claras/Rio Paraguai/La Plata Basin/Tangará da Serra/MT/Brasil	S 14°21'03.2' W 57°33'07.2"
Knodus cf. chapadae	LBP 9516	44686	Riacho sem nome/Rio Tocantins/Amazônia/Formosa/GO/Brasil	S 15°28'54.8' W 47°20'31.4"
Knodus cf. chapadae	LBP 9516	44687	Riacho sem nome/Rio Tocantins/Amazônia/Formosa/GO/Brasil	S 15°28'54.8' W 47°20'31.4"
*Knodus meridae**	LBP 7569	15818	Rio Orinoco/Caicara del Orinoco /Bolivar/Venezuela	N 07º39'06,3'' W 66º10'34,2''
Knodus moenkhausii	LBP 10751	49778	Rio Macucu/Paraíba do Sul/Atlântico/Cordeiro/RJ/Brasil	S 21°59'37.1" W42°16'41.4"
Knodus moenkhausii	LBP 10751	49779	Rio Macucu/Paraíba do Sul/Atlântico/Cordeiro/RJ/Brasil	S 21°59'37.1" W42°16'41.4"
*Kolpotocheirodon theloura**	LBP 5033	25982	Ribeirão Bananal/Distrito Federal/Brazil	S 15°43'42.7' W 47°54'39.4"
*Leptagoniates steindachneri**	LBP 4137	23661	Rio Moa/Mâncio Lima/AC/Brazil	S 07°26'35.5' W 73°03'33.5"
*Lophiobrycon weitzmani**	LBP 1225	38090	Rio Claro/Delfinópolis/MG/Brazil	S 20°20'32.2' W 46°47'12.2"
*Macropsobrycon uruguayanae**	LBP 6039	29061	Rio Piquiri/Cachoeira do Sul/RS/Brazil	30º14'46"S e 52º45'53"W
*Markiana nigripinnis**	LBP 663	8038	Região de Rombado, afluente rio Pirai/Poconé/MT/Brazil	S 16°25,680' W 56°25,143'

*Microschemobrycon casiquiare**	LBP 8161	38058	Rio Tapajós/Pimental/PA/Brazil	S 04°32'25'' W 56°15'15''
*Mimagoniates inequalis**	LBP 3383	21274	Arroio dos Corrientes/Pelotas/RS/Brazil	S 31°28'46.3'' W 52°12'46.9''
*Mimagoniates microlepis**	LBP 1225	11077	Bertioga/SP/Brazil	S 23°57,769' W 46°10,625'
Moenkhausia agnesi	LBP 12449	53748	Afluente Rio Ampiyacu/Amazonas/Pevas/Mariscal Ramon Castilla/Loreto/Peru	S 03°18'46.1'' W 71°50'58.4''
Moenkhausia australe	LBP 3739	22007	Lagoa Marginal Rio Negro/Rio Paraguai/Aquidauana/MS/Brasil	S 19°34'54.6' W 56°15'16.5''
Moenkhausia bipunctiaubicaudalis	LBP 6587	31867	Rio Paraná/Rio Paraná/Bacia do Prata/Marilena/PR/Brasil	S 22°38'59.7' W 53°05'28.4''
Moenkhausia bonita	LBP 3950	22688	Baía do Poço/Rio Cuiabá/Rio Paraguai/Santo Antônio do Leverger/MT/Brasil	S 15°54'03' W 56°01'17''
Moenkhausia bonita	LBP 5796	28231	Baia das Piranhas (ou Jacaré)/Prata/Rio Aricá-Mirim/Santo Antônio do Leverger/M/Brasil	S 15°44'3.60'' W 55°52'48.7''
Moenkhausia celibela	LBP 13737	57012	Rio Tracuá/Rio Amazonas/Rio Tapajós/Itaituba/PA/Brasil	S 04°29'11.1'' W 56°17'22.1''
Moenkhausia ceros	LBP 4504	24526	Igarapé Puxirituba/Rio Negro/Amazonas/Barcelos/AM/Brasil	S 00°53'18.6'' W 62°40'36/1''
Moenkhausia colletti	LBP 4457	24371	Igarapé Zalala/Rio Negro/Amazonas/Barcelos/AM/Brasil	S 00°40'03.1'' W 62°58'23.5''
Moenkhausia comma	LBP 6929	33156	Igarapé BR307/Rio Negro/Amazonas/São Gabriel da Cachoeira/AM/Brasil	N 00°05.610' W 66°49.034''
Moenkhausia copei	LBP 2300	15831	Lagoa marginal do rio Orinoco/Rio Orinoco/Caicara del Orinoco /Bolivar/Venezuela	N 07º31'23,3'' W 66º03'16,2''
Moenkhausia cosmops	LBP 8164	36311	Rio Verde/Rio Paraguai/Campo Novo dos Parecis/MT/Brasil	S 13°37'02' W 58°00'50''
Moenkhausia costae	LBP 7558	36310	Aquário/Brasil	
Moenkhausia coutinho	LBP 4277	23875	Igarapé Boiboi/Rio Negro/Amazonas/Barcelos/AM/Brasil	S 00°49'43.7'' W 62°49'59.8''
Moenkhausia diktiota	LBP 7074	34395	Afluente Igarapé Miuá/Rio Negro/Rio Amazonas/São Gabriel da Cachoeira/AM/Brasil	S 00°06.119' W 66°53.756'
Moenkhausia forestii	LBP 4655	24752	Rio Baia/Rio Paraná/Rio Baia/MS/Brasil	S 22°43'46.2'' W 53°19'04.2''
Moenkhausia gracilima	LBP 9049	42213	Rio Madeira/Pacaás/Rio Amazonas/Pacaás/RO/Brasil	S 10°52'07.7'' W 65°15'42.1''
Moenkhausia gradisquamis	LBP 10906	50234	Rio Jaci-Paraná/Amazonas/Madeira/Porto Velho/RO/Brasil	S 09°15'23.3'' W 64°23'13.6''
Moenkhausia grandisquamis	LBP 5380	27039	Igarapé Uiratapuru/Rio Jari/Rio Amazonas/Laranjal do Jari/AP/Brasil	S 00°33'51'' W 52°34'45''
Moenkhausia hemigrammoides	LBP 7029	34103	Igarapé Ya-Mirim/Rio Negro/Rio Amazonas/São Gabriel da Cachoeira/AM/Brasil	N 00°16.259' W 66°38.365'

Moenkhausia inrai	LBP 10125	47519	Rio Lawa/Atlântico/Sipalawini/Suriname	N 03° 19' 31" W 54° 3' 48"
Moenkhausia intermedia	LBP 4091	23529	Rio Japiim/Rio Moa/Rio Juruá/Mâncio Lima/AC/Brasil	S 07°34'28.8' W 72°55'24.9"
Moenkhausia jamesi	LBP 9048	42212	Rio Madeira/Sotério/Rio Amazonas/Sotério/RO/Brasil	S 11°35'53.1" W 65°13'50.2"
Moenkhausia lata	LBP 13747	57040	Rio Tracuá/Rio Amazonas/Rio Tapajós/Itaituba/PA/Brasil	S 04°29'11.1" W 56°17'22.1"
Moenkhausia lepidura	LBP 7087	34530	Igarapé margem direita rio Negro/Rio Negro/Rio Amazonas/São Gabriel da Cachoeira/AM/Brasil	S 00°08.625' W 67°05.605'
Moenkhausia lopesi	LBP 8550	43297	Córrego Vermelho/Bacia rio Sepotuba/La Plata Basin/Tangará da Serra/MT/Brasil	S 14°20'32.6' W 57°31'22.5"
Moenkhausia nigromarginata	LBP 5156	26296	Rio Verde/Rio Paraguai/Campo Novo dos Parecis/MT/Brasil	S 13°37'02' W 58°00'50"
Moenkhausia oligolepis	LBP 5377	27028	Igarapé Uiratapuru/Rio Jari/Rio Amazonas/Laranjal do Jari/AP/Brasil	S 00°33'51'' W 52°34'45''
Moenkhausia phaenota	LBP 8029	36316	Riacho sem nome/Arinos/Tapajós/Amazonas/Nova Mutum/MT/Brasil	S 13°52'14.7" W 56°11'30.8"
Moenkhausia pirauba	LBP 13319	59827	Riacho Alegre/Rio Tapajós/Diamantino/MT/Brasil	S 13°59'04.1" W 57°04'01.7"
Moenkhausia pyrophthalma	LBP 1521	11557	Ribeirão Ínsula/Rio das Mortes/Araguaia/Barra do Garça/MT/Brasil	S 15°40'57.7'' W 52°13'24.8''
Moenkhausia sanctaefilomenae	LBP 5541	27225	Brejo das Ovelhas/Rio Parnaíba/Santa Filomena/PI/ Brasil	S 09°08'04' W 45°53'48'
Moenkhausia sp.	LBP 5327	26949	Rio Jari/Rio Amazonas/Laranjal do Jari/AP/Brasil	S 00°34'11'' W 52°33'19''
Moenkhausia sp.	LBP 5327	26951	Rio Jari/Rio Amazonas/Laranjal do Jari/AP/Brasil	S 00°34'11'' W 52°33'19''
Moenkhausia xinguensis	LBP 6101	28443	Rio Culuene/Amazonas/Rio Xingu/Paranatinga/MT/Brasil	S 13°49'00.0" W 53°15'00.0"
*Myxiops aphos**	LBP 7184	33173	Rio Lençóis/Lençóis/BA/Brazil	S 12°33'41.8" W 41°24'09.3"
*Nanocheirodon insignis**	LBP 6104	27476	Rio Apon Medio/Machiques de Perijá/Zulia/Venezuela	N 10°09'42.0' W 72°25'58.0'
*Nematobrycon palmeri**	LBP 6124	33165	Aquarium/Brazil	
*Nematocharax venustus**	LBP 8106	37557	Rio Almada/Ilhéus/BA/Brazil	S 14°39'52.1' W 39°13'26.7"
*Odontostilbe fugitiva**	LBP 4052	22932	Rio Moa/Cruzeiro do Sul/AC/Brazil	S 7°37'20.0" W 72°47'42.2"
*Odontostilbe sp.**	LBP 4650	22626	Rio Araquá/Botucatu/SP/Brazil	S 22°47.135' W 48°28.892
*Odontostoechus lethostigmus**	LBP 6752	33171	Arroio Água Parada/Maquiné/ RS/Brazil	29°39'42,8''S / 50°12'37,7''

*Oligosarcus hepsetus**	LBP 2377	16055	Lagoa Feia/Campos dos Goytacazes/RJ/Brazil	S 22°00' W 41°20'
*Oligosarcus paranensis**	LBP 3926	22582	Rio Paraitinguinha/Salesópolis/ SP/Brazil	S 23°31'25.6" W 43°53'22.7"
Oligosarcus sp.	LBP 17554	17386	Córrego sem nome/Rio São José dos Dourados/Paraná/Sebastianópolis/SP/Brasil	S 20º44'43.7" W 49º46'45.3"
Oligosarcus sp.	LBP 17554	17387	Córrego sem nome/Rio São José dos Dourados/Paraná/Sebastianópolis/SP/Brasil	S 20º44'43.7" W 49º46'45.3"
*Orthospinus franciscensis**	LBP 8105	37555	Rio Verde Grande/Jaíba/MG/ Brazil	S 15°19'24.2' W 43°39'52.5"
*Paracheirodon axelrodi**	LBP 4472	24425	Igarapé Zalala/Barcelos/ AM/Brazil	S 00°40'03.1" W 62°58'23.5"
*Paragoniates alburnus**	LBP 9208	43156	Rio Manapire/Cabruta/ Guárico/Venezuela	N 7°52'04.1" W 66°12'40.1"
*Parecbasis cyclolepis**	LBP 9053	42217	Belmont – Foz/Belmont/RO/ Brazil	S 08°37'17.2" W 63°49'25.7"
*Phenacogaster calverti**	LBP 5582	27299	Afluente Parnaíba/Santa Filomena/PI/Brazil	S 09°09'51' W 45°51'15'
*Phenagoniates macrolepis**	LBP 6105	35623	Rio Apon Medio/Machiques de Perijá/Zulia/Venezuela	N 10°09'42.0' W 72°25'58.0"
*Piabarchus analis**	LBP 8514	38382	Rio Salobra/Cáceres/MT/Brazil	S 15°19'53.5' W 57°11'31.1"
*Piabina argentea**	LBP 3509	21306	Córrego da Hortelã/Botucatu/ SP/Brazil	S 22°56'28.9" W 48°35'03.2"
*Planaltina britskii**	LBP 2598	17243	Córrego Boa Vista dos Castilhos/Miraluz/SP/Brazil	S 21º00'46.6" W 49º41'25.1"
*Poptella paraguayensis**	LBP 3732	21986	Lagoa Marginal Rio Negro/Aquidauana/MS/Brazil	S 19°34'54.6' W 56°15'16.5"
*Prionobrama paraguayensis**	LBP 3230	19465	Lagoa marginal/Nobres/ MT/Brazil	S 14°40'32,8" W 56°13'14,0"
*Pristella maxillaris**	LBP 2221	15637	Laguna de Castilleros/Caicara del Orinoco/Bolivar/Venezuela	N 07º30'50,9'' W 66º09'19,8''
*Probolodus heterostomus**	LBP 6454	29141	Rio Paraíba do Sul/ Guararema/SP/Brazil	S 23°21'38.2" W 45°59'69.0"
*Prodontocharax melanotus**	AMNH	233264		
*Psellogrammus kennedyi**	LBP 6578	31813	Lagoa marginal Rio Paraná/Marilena/PR/Brazil	S 22°38'49.4' W 53°04'36.9"
*Pseudocheirodon arnoldi**	STRI	5		
*Pseudocorynopoma heterandria**	LBP 2862	18570	Rio Fau/Miracatu/SP/Brazil	S 24°12,441' W 47°28,616'
*Rachoviscus crassiceps**	LBP 7146	33170	Riacho sem nome/Guaratuba/ PR/Brazil	S 25°55'27.6" W 48°36'39.5"

*Roeboexodon guyanensis**	LBP 5315	26921	Igarapé Uiratapuru/Laranjal do Jari/AP/Brazil	S 00°34'03'' W 52°34'41''
*Roeboides guatemalensis**	LBP 2755	18529	Río Llano Sucio/Santa Rita Arriba/Colón/Panamá	N 09°19'26.2" W 79°46'08.2"
*Saccoderma melanostigma**	LBP 6103	27475	Rio Apon Medio/Machiques de Perijá/Zulia/Venezuela	N 10°09'42.0' W 72°25'58.0"
*Serrapinnus calliurus**	LBP 3731	22121	Lagoa Marginal Rio Negro/Aquidauana/MS/Brazil	S 19°34'54.6' W 56°15'16.5"
*Serrapinnus heterodon**	LBP 9039	37551	Córrego Cachoeira	S 17°08'54.9' W 43°49'32.3"
*Serrapinnus piaba**	LBP 8972	41813	Córrego da Mata/Pedro Leopoldo/MG/Brazil	S 19°37'59.7" W 44°06'25.5"
*Stethaprion crenatum**	LBP 4078	22994	Rio Japiim/Mâncio Lima/AC/Brazil	S 07°34'28.8' W 72°55'24.9"
*Stygichthys typhlops**	LBP 8107	37558	Cacimba fazenda Lajeado/ Jaíba/MG/Brazil	S 15°24'41.7' W 43°45'19.7"
*Tetragonopterus argenteus**	LBP 3758	22029	Rio Negro/Aquidauana/MS/ Brazil	S 19°34'33.7' W 56°14'49.5"
*Tetragonopterus chalceus**	LBP 8268	37556	Rio Verde Grande/Jaíba/MG/ Brazil	S 15°19'24.2' W 43°39'52.5"
*Thayeria obliqua**	LBP 5743	26891	Lagoa Marginal rio Corrente/ Barra do Garças/MT/Brazil	S 15°19'57.6" W 52°12'10.4"
*Tyttocharax madeirae**	LBP 5145	33166	Rio Japiim/Mâncio Lima/AC/Brazil	S 07°34'28.8' W 72°55'24.9"
*Xenagoniates bondi**	LBP 3074	19694	Rio Orinoco/Caicara del Orinoco/Bolivar/Venezuela	N 07°38'11.6" W 66°19'04.2"
*Xenurobrycon pteropus**	LBP 9054	42218	Mutumparaná/Mutumparaná/RO/Brazil	S 09°36'39.5" W 64°55'38.9"

1- data from Ornelas-Garcia (2008); 2- data from Javonillo et al. (2010); * - data from Oliveira et al. (2011)

Table 2. Primer sequences of nuclear and mitochondrial genes used in the phylogeny. The sequences of COI gene for *Astyanax* were obtained from Rossini et al. (unpublished) and for the other genera, the sequences were generated with the same protocol used by the laste work. For primers data references, see Rossini et al. (unpublished).

Gene	Number of sequences	Bp after alignment	Variable sites	Π_T	Π_C	Π_A	Π_G	Primer name	Primer sequence (5'-3')	Reference
COI	177 (70.0%)	633	335	0.317	0.263	0.239	0.181	-	-	-
ATPase 6/8	174 (68.7%)	719	516	0.307	0.281	0.284	0.128	ATP 8.2 - L8331	AAAGCRTYRGCCTTTTAAGC	Perdices et al., 2002
16S	232 (91.7%)	645	361	0.231	0.233	0.318	0.219	CO3.2 - H9236	GTTAGTGGTCABGGCTTGGRTC	Perdices et al., 2002
								16Sa-L	ACGCCTGTTTATCAAAAACAT	Palumbi, 1996
								16Sb-H	CCGGTCTGAACTCAGATCACGT	Palumbi, 1996
Cytb	198 (78.3%)	992	599	0.31	0.282	0.264	0.142	LNF	GACTTGAAAAACCAYCGTTGT	Oliveira et al., 2011
								H08R2	GCTTTGGGAGTTAGDGGTGGGAGTTAGAATC	Oliveira et al., 2011
Myh6	217 (85.7%)	755	306	0.248	0.214	0.304	0.234	F329	CCGCMTGGATGATCTACAC	Li et al., 2007
1stPCR								A3R1	ATTCTCACCACCATCCAGTTGAA	Li et al., 2007
Myh6								A3F2	GGAGAATCARTCKGTGCTCATCA	Li et al., 2007
2ndPCR								A3R2	CTCACCACCATCCAGTTGAACAT	Li et al., 2007
RAG1	225 (89.0%)	1265	761	0.224	0.241	0.248	0.287	Rag1CF1	ACCCTCCGTACTGCTGAGAA	Oliveira et al., 2011
1stPCR								Rag1CR1	CGTCGGAAGAGCTTGTTGCC	Oliveira et al., 2011
RAG1								Rag1CF2	TACCGCTGAGAAGGAGCTTC	Oliveira et al., 2011
2ndPCR								Rag1CR2	TGTTGCCAGACTCATTGCCCTC	Oliveira et al., 2011
RAG2	192 (75.9%)	1051	550	0.228	0.257	0.241	0.274	164F	AGCTCAAGCTGCGYGCCAT	Oliveira et al., 2011
1stPCR								RAG2-R6	TGRTCCARGCAGAAGTACTTG	Lovejoy & Collette, 2001
RAG2								176R	GYGCCATCTCATTCTCCAACA	Oliveira et al., 2011
2ndPCR								Rag2Ri	AGAACAAAAGATCATTGCTGGTCGGG	Oliveira et al., 2011

Table 3. Partitioning scheme used in ML/Bayesian analysis and best models selected for Bayesian analysis.

Gene and position	Partitions	Best-fit model
Coi 1st position	1-633\3	SYM+I+G
Coi 2nd position	2-633\3	F81+I+G
Coi 3rd position	3-633\3	GTR+G
ATPase 1st position	634-1352\3	SYM+I+G
ATPase 2nd position	635-1352\3	GTR+I+G
ATPase 3rd position	636-1352\3	GTR+I+G
16S	1353-1997	GTR+I+G
Cytb 1st position	1998-2989\3	HKY+I+G
Cytb 2nd position	1999-2989\3	GTR+G
Cytb 3rd position	2000-2989\3	SYM+I+G
Myh6 1st position	2990-3744\3	SYM+I+G
Myh6 2nd position	2991-3744\3	GTR+I+G
Myh6 3rd position	2992-3744\3	GTR+G
RAG1 1st position	3745-5009\3	SYM+G
RAG1 2nd position	3746-5009\3	GTR+I+G
RAG1 3rd position	3747-5009\3	SYM+I+G
RAG2 1st position	5010-6060\3	SYM+G
RAG2 2nd position	5011-6060\3	SYM+I+G
RAG2 3rd position	5012-6060\3	SYM+I+G

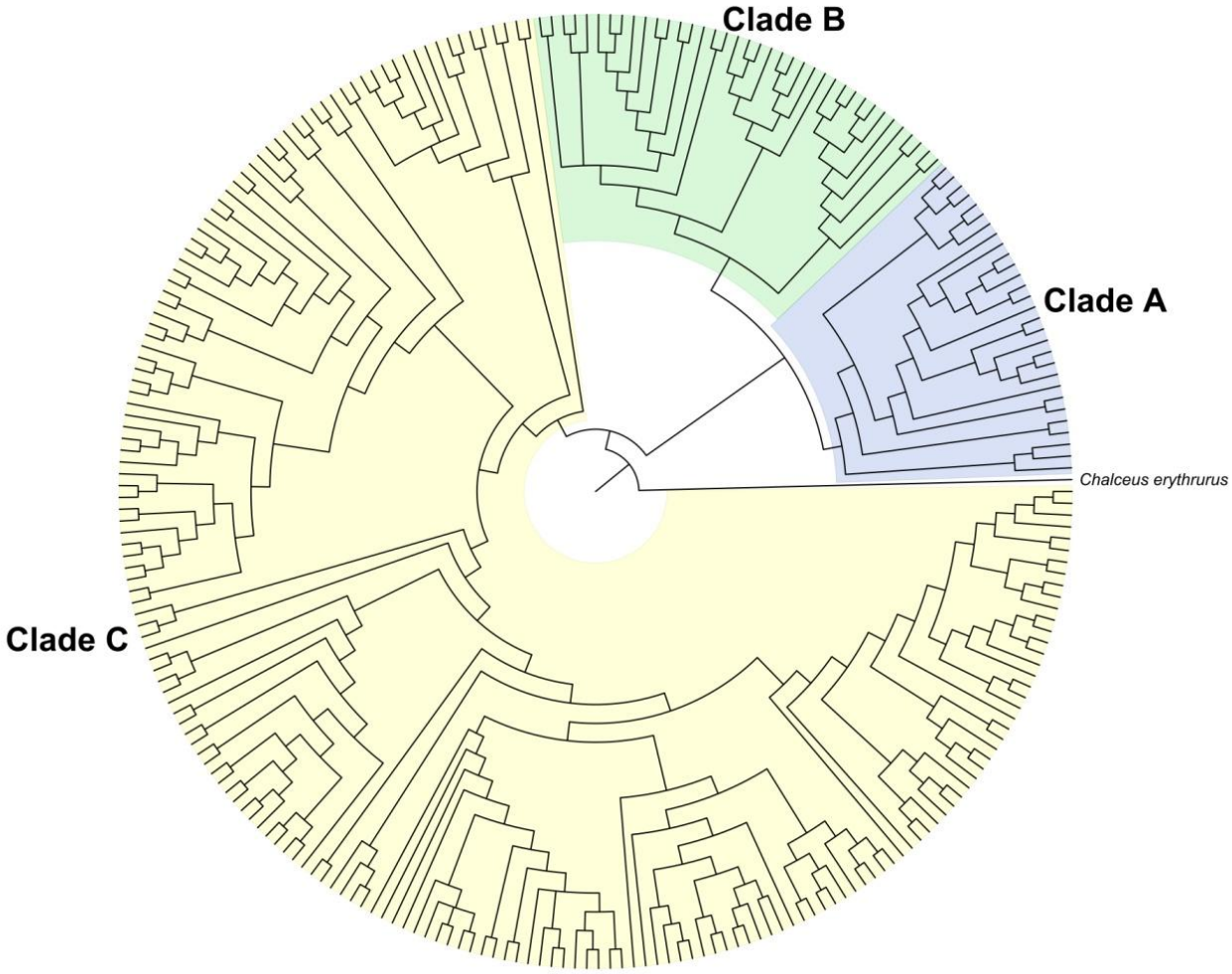

Figure 1. Bayesian majority consensus tree showing the three clades following nomenclature by Javonillo et al. (2010) and Oliveira et al. (2011), rooted in *Chalceus erythrurus*.

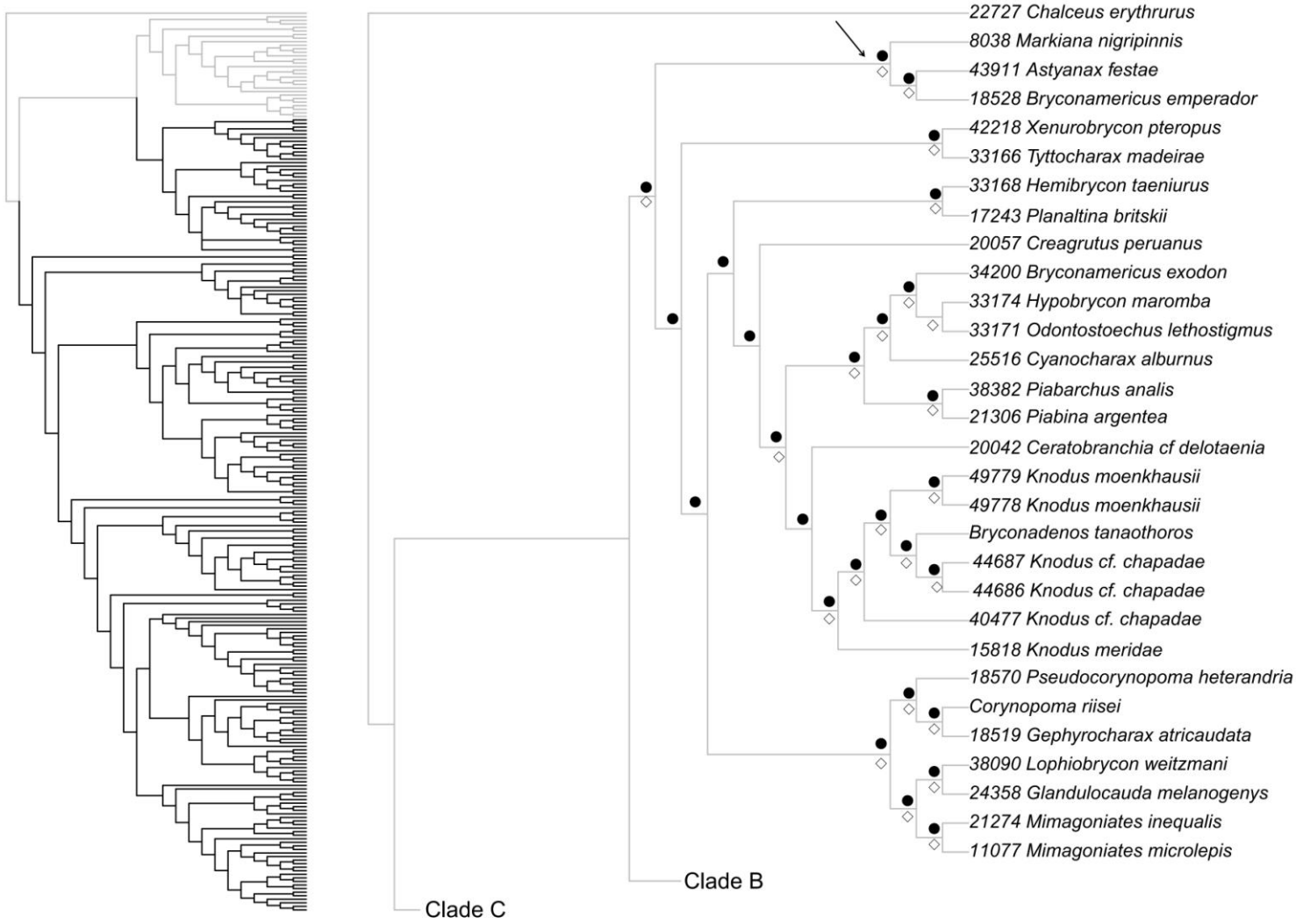

Figure 2. Partially Bayesian majority consensus tree showing Clade A. Black circles correspond posterior probabilities >0.9 for the Bayesian analysis and white diamonds correspond bootstrap percentages greater than 50% in ML analysis. Arrow indicates position of *A. festae*.

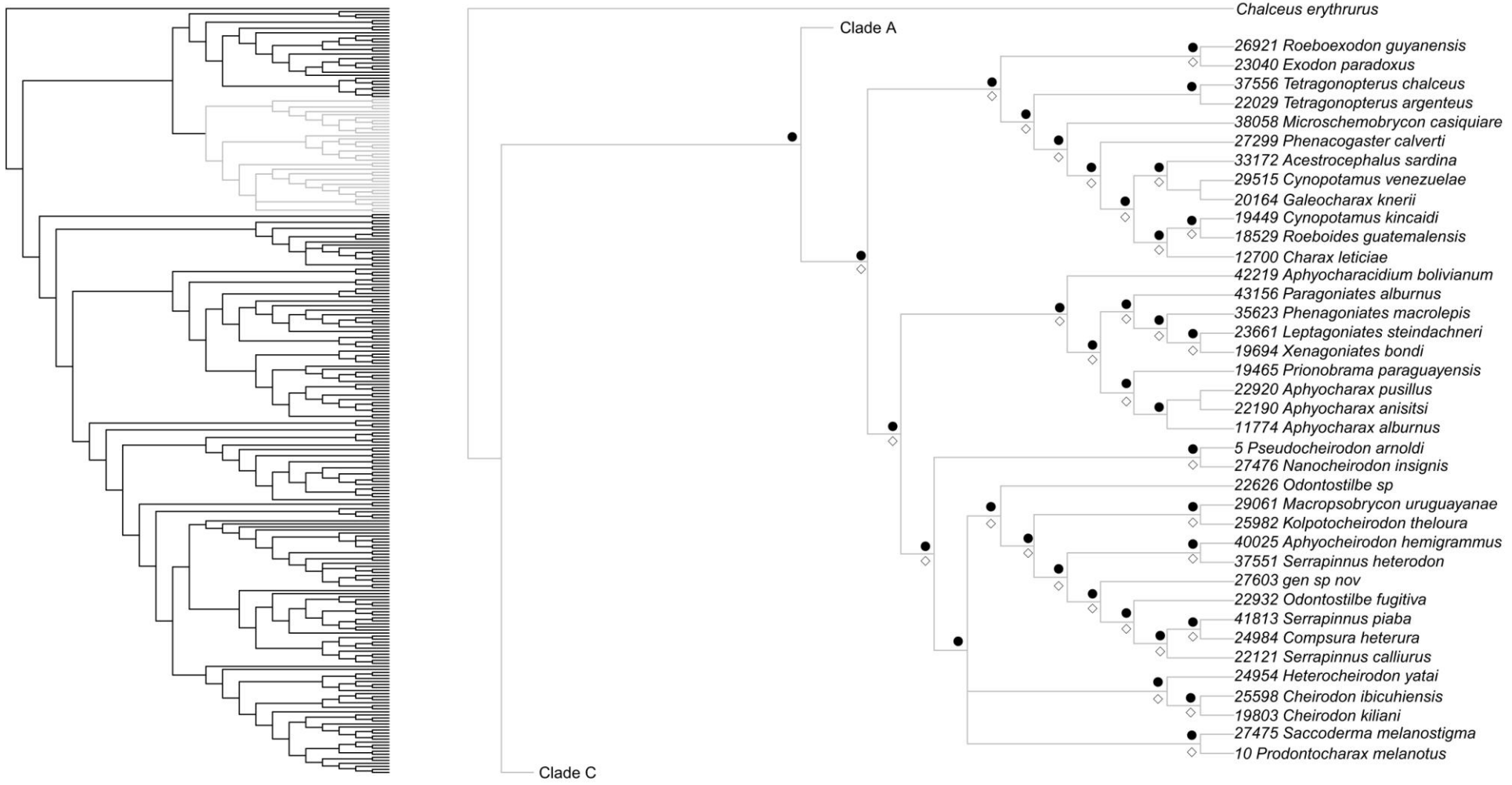

Figure 3. Partially Bayesian majority consensus tree showing Clade B. Black circles correspond posterior probabilities >0.9 for the Bayesian analysis and white diamonds correspond bootstrap percentages greater than 50% in ML analysis.

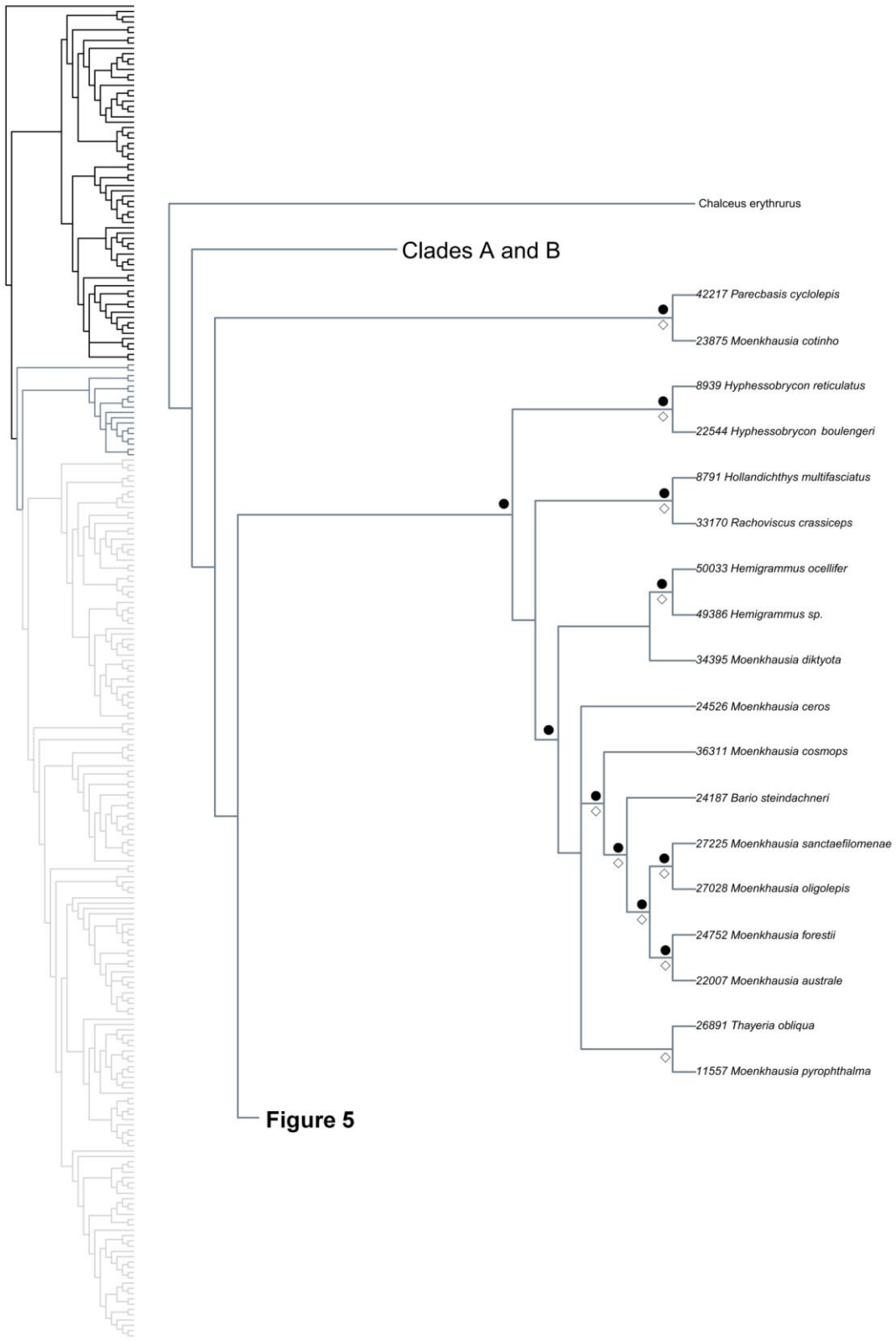

Figure 4. Partially Bayesian majority consensus tree showing partially Clade C. Black circles correspond posterior probabilities >0.9 for the Bayesian analysis (other nodes have posterior probabilities >0.7) and white diamonds correspond bootstrap percentages greater than 50% in ML analysis.

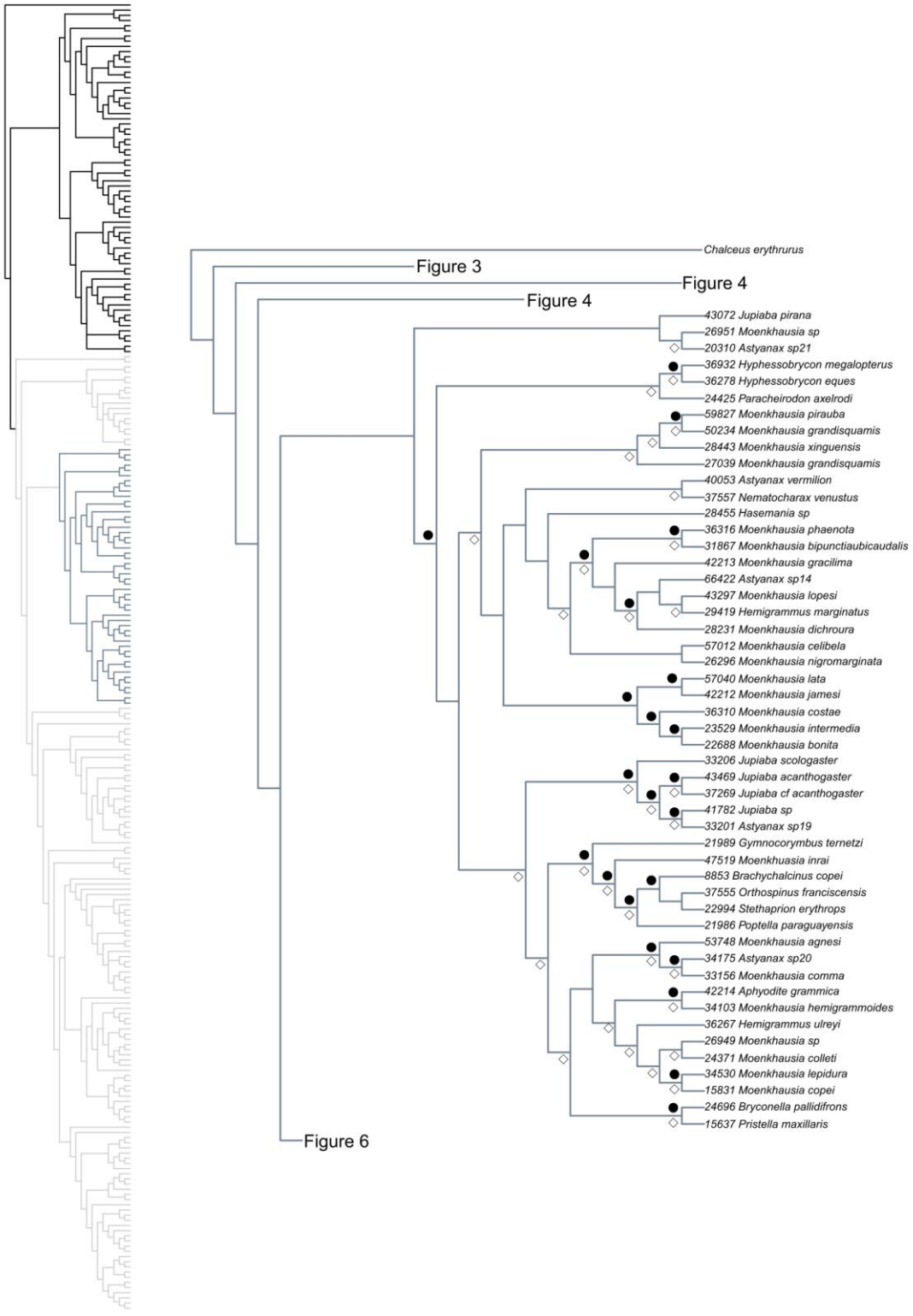

Figure 5. Partially Bayesian majority consensus tree showing partially Clade C. Black circles correspond posterior probabilities >0.9 for the Bayesian analysis (other nodes have posterior probabilities >0.7) and white diamonds correspond bootstrap percentages greater than 50% in ML analysis.

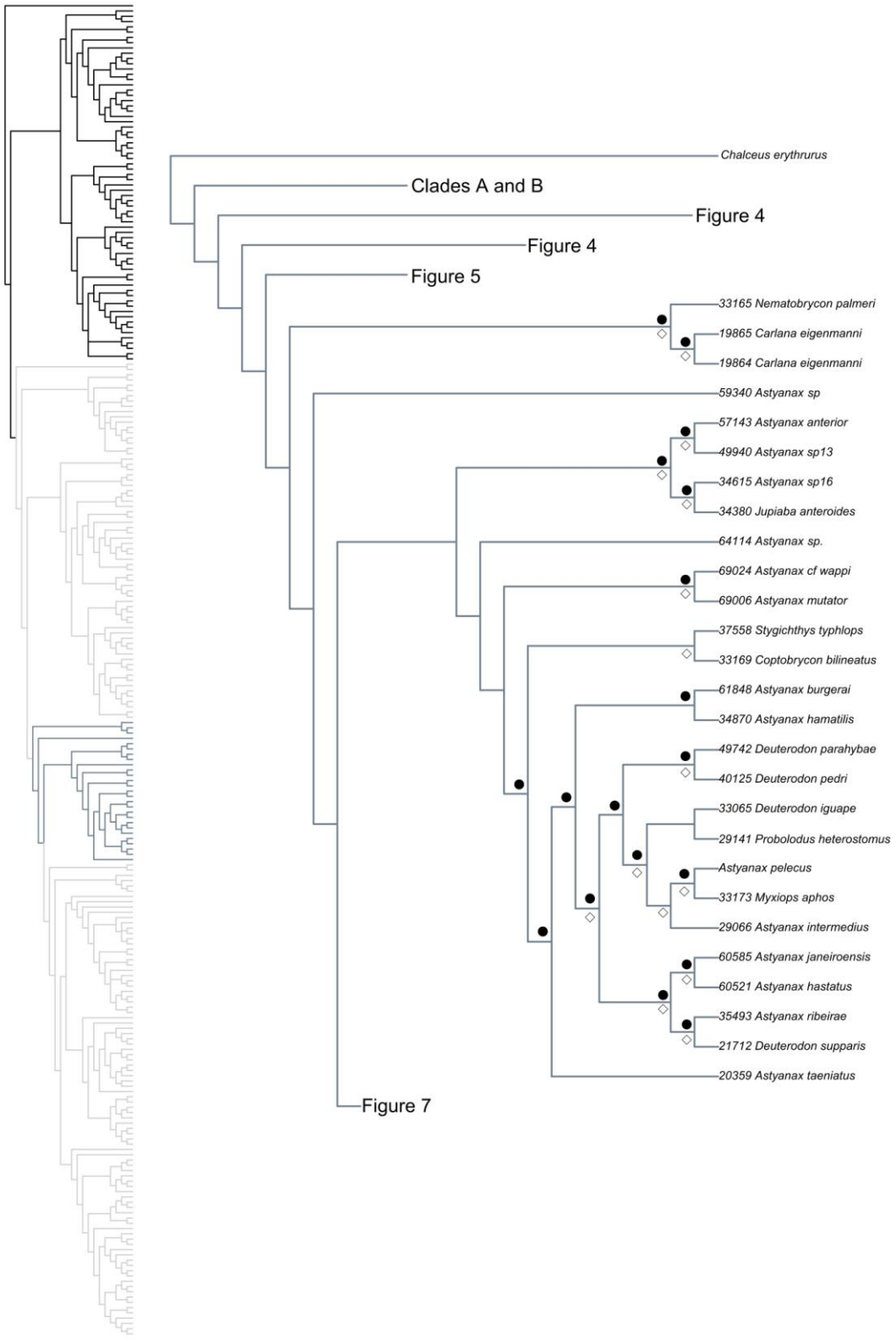

Figure 6. Partially Bayesian majority consensus tree showing partially Clade C. Black circles correspond posterior probabilities >0.9 for the Bayesian analysis (other nodes have posterior probabilities >0.7) and white diamonds correspond bootstrap percentages greater than 50% in ML analysis.

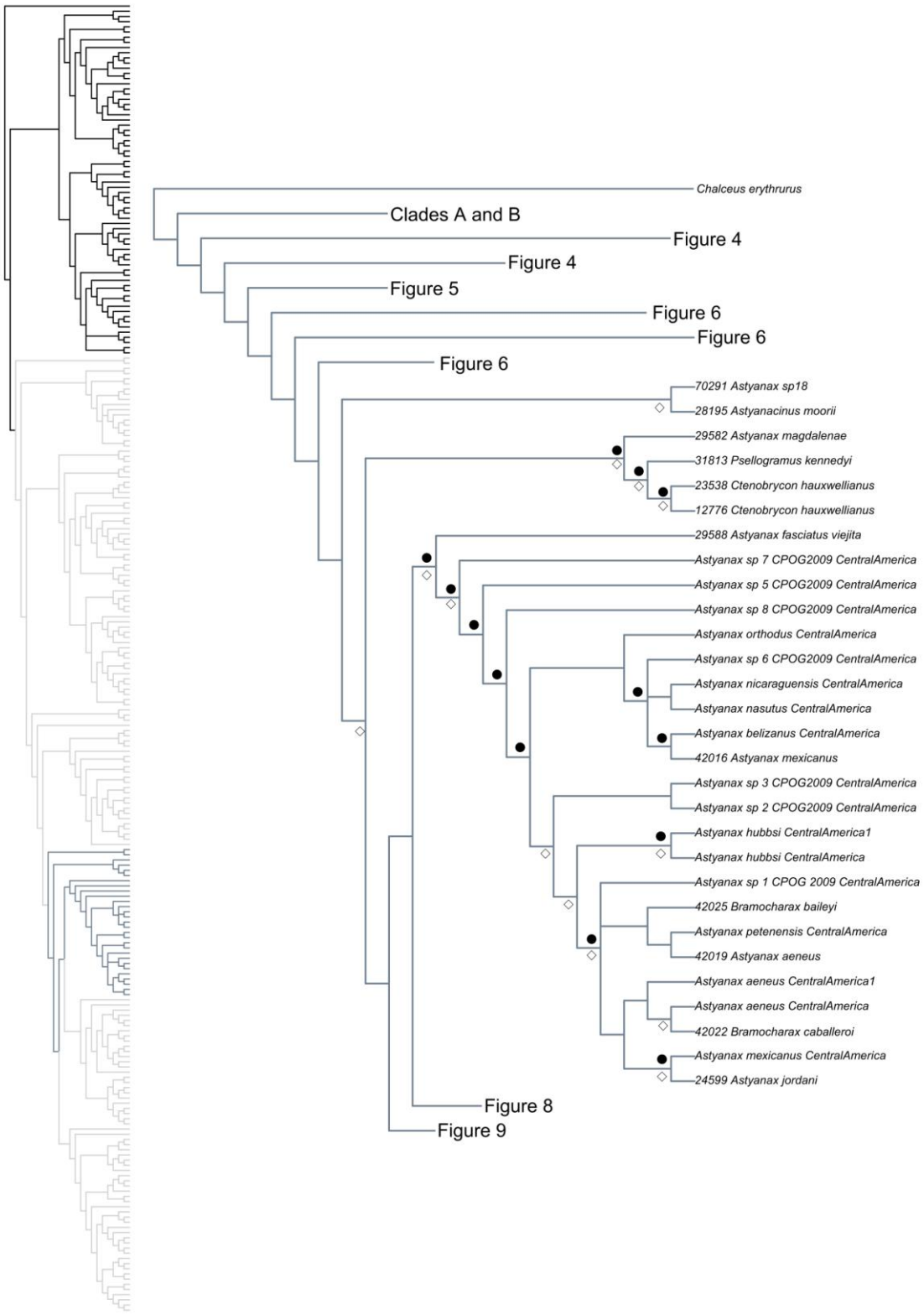

Figure 7. Partially Bayesian majority consensus tree showing partially Clade C. Black circles correspond posterior probabilities >0.9 for the Bayesian analysis (other nodes have posterior probabilities >0.7) and white diamonds correspond bootstrap percentages greater than 50% in ML analysis.

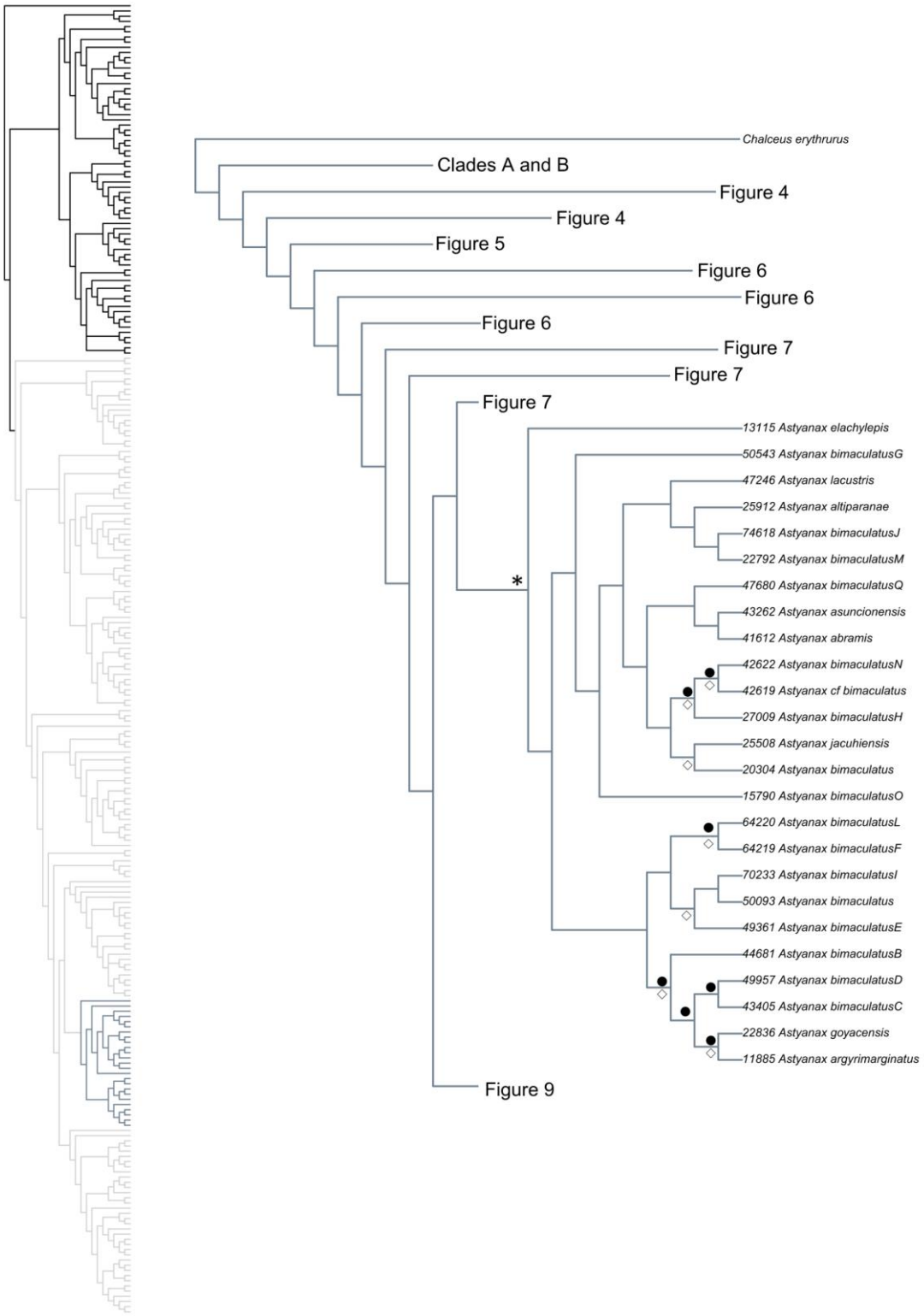

Figure 8. Partially Bayesian majority consensus tree showing partially Clade C. Only its shown posterior probabilities >0.9 for the Bayesian analysis (Black circles, but all nodes have posterior probabilities >0.8, including node support for all 'bimaculatus', marked with an asterisk) and white diamonds correspond bootstrap percentages greater than 50% in ML analysis.

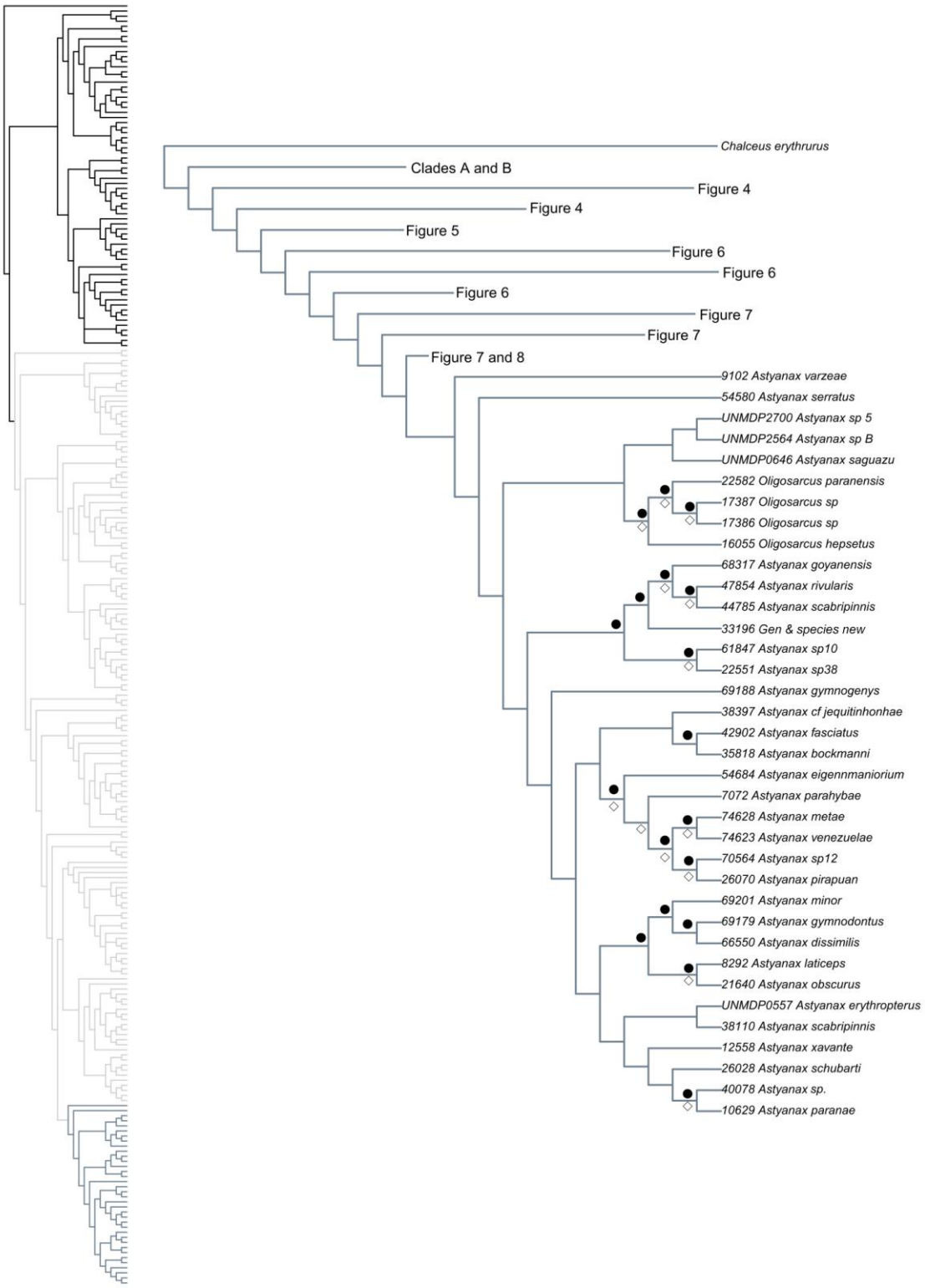

Figure 9. Partially Bayesian majority consensus tree showing partially Clade C. Only its shown posterior probabilities >0.9 for the Bayesian analysis (Black circles, but all nodes have posterior probabilities >0.7) and white diamonds correspond bootstrap percentages greater than 50% in ML analysis.

Supplemental files

These files are in digital format