UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

História evolutiva de *Cnesterodon* Garman, 1895: Padrões de diversificação em ambientes campestres do Bioma Pampa e da Floresta Atlântica revelados por estudos filogenéticos e filogeográficos

ALINE MITCHELI CARVALHO RAMOS FREGONEZI

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Doutor em Ciências.

Orientação: Nelson Jurandi Rosa Fagundes
Coorientação: Luiz Roberto Malabarba

Porto Alegre, outubro de 2015.

Os capítulos 2, 4 e parte do capítulo 3 do presente trabalho foram executados nos Laboratórios de Ictiologia (Departamento de Zoologia) e de Genética Médica e Populacional (Departamento de Genética), Instituto de Biociências, Universidade Federal no Rio Grande do Sul, sob a orientação dos Profs. Doutores Nelson Jurandi Rosa Fagundes e Luiz Roberto Malabarba. Parte do capítulo 3 deste trabalho foi desenvolvida no Center for Comparative Genomics - California Academy of Sciences (San Francisco, Califórnia, EUA), sob a orientação do Doutor Luiz Rocha, com financiamento do programa PDSE - Doutorado Sanduíche da CAPES.

Fontes Financiadoras:
- Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)
- Programa de Pós-Graduação em Genética e Biologia Molecular (PPGBM-UFRGS)
- Programa de Pós-Graduação em Biologia Animal (PPGBAN-UFRGS)

“Do or do not. There is no try.”

(Mestre Yoda)

AGRADECIMENTOS

Este trabalho foi um grande desafio em minha jornada acadêmica e encerra não somente os meus anos de doutorado, como também os meus mais de dez anos de UFRGS. Ao longo deste caminho, tive sorte de encontrar pessoas especiais que contribuíram muito para minha vida pessoal e para os meus conhecimentos científicos. A realização do presente trabalho contou com o apoio de muitas pessoas, às quais devo minha enorme gratidão:

Aos meus orientadores Nelson e Malabarba, por aceitarem a aventura de orientar um aluno com três anos restantes, pela oportunidade de trabalhar com um grupo tão complexo, pelo convívio agradável, por todo apoio e orientação ao longo destes três anos, pela paciência nos momentos finais e principalmente pelos ensinamentos, essenciais à minha formação. Foi um enorme prazer trabalhar com vocês dois!

Ao doutor Luiz Rocha, pela oportunidade de poder trabalhar em um lugar fantástico e ampliar meus conhecimentos sob os pontos de vista pessoal e profissional.

À Universidade Federal do Rio Grande do Sul, pelo suporte na realização deste trabalho.

À California Academy of Sciences pelo espaço e infraestrutura cedidas para a realização de parte deste trabalho.

Ao Programa de Pós-Graduação em Genética e Biologia Molecular, em especial ao secretário administrativo Elmo Cardoso, por resolver absolutamente todos os problemas burocráticos dos alunos, de forma muito eficiente e tranquila.

Ao Programa de Pós-Graduação em Biologia Animal, pelo suporte na realização do projeto.

À professora Doutora Loreta, por ter me iniciado na carreira científica, pelas grandes oportunidades dadas a mim, pelos ensinamentos ao longo dos anos de orientação e por me apresentar ao maravilhoso mundo da filogeografia.

A todos os meus colegas dos Laboratórios de Ictiologia e Lab 118 pelas valiosas discussões, pelas ajudas em campo, pelos momentos de descontração e sessões de terapia em grupo.

Aos amigos Lina, Jay, Minki e Sasha, por me acolherem tão bem, pelas valiosas experiências que me ensinaram tanto sobre a diversidade cultural e principalmente por serem minha família durante minha estadia em SF.

Ao amigo Moisés, pela orientação essencial nas análises e interpretação dos dados de Rad-seq, por ter muita paciência em responder tantas dúvidas e por ser uma ótima companhia de trabalho, sempre animando os colegas com seu humor.

À Anna Sellas, pelos ensinamentos nos procedimentos laboratoriais, pelo 'improve no meu english', e pela companhia nas 12 horas seguidas de trabalho durante os dias de construção das bibliotecas.

Ao Brian Simison, pelo suporte na parte de bioinformática e pela ajuda com a pipelines.

À Esther, pela longa amizade de onze anos e especialmente pela cumplicidade destes anos de doutorado, que claramente nos deixaram mais loucas e mais parecidas com a bruxa do 71.

À Laísa, por sempre ter um chimarrão quentinho (sem o qual a produtividade teria caído) e um ombro amigo ao meu lado.

À Clarice, por sempre ter sábias palavras de conselho e por compartilhar sua risada contagiante que anima qualquer um.

Ao Lucinda, por me informar que em 2015 ainda existem cientistas que não acreditam em moléculas, o que impulsionou ainda mais minha curiosidade.

Por fim, eu gostaria de agradecer às pessoas mais importantes nesse processo todo:

Aos meus pais (Erton e Márcia), por serem o meu porto seguro, por fomentarem a vontade de estudar desde sempre, por nunca terem medido esforços em prover os recursos necessários aos meus estudos e por terem me dado a maior riqueza que poderiam me dar: o acesso ao conhecimento. Sem o apoio de vocês dois eu jamais teria chegado até aqui. Não existem palavras que expressem o tamanho da minha gratidão. Amo vocês dois incondicionalmente.

Aos meus filhotes felinos Luke e Leia, que foram a minha fonte de paz durante a escrita. Posso dizer que ninguém acompanhou mais este processo do que estes dois, sempre ao meu lado e, por vezes, em cima dos computadores, acompanhando atentamente a convergência das cadeias de Markov.

Ao Jéferson, pela parceria na vida pessoal e acadêmica, por não me deixar desistir de tudo em vários momentos, pela paciência e compreensão, pelo apoio imenso (mesmo quando a gente estava a 11 mil km de distância), pelo companheirismo e por tentar cuidar da minha saúde mental nos últimos meses! Te amo muito!

SUMÁRIO

RESUMO

Cnesterodon Garman, 1895 (Poeciliidae) é um gênero sul-americano de peixes de água doce composto por onze espécies reconhecidas, sendo dez delas já descritas formalmente. A maioria dessas espécies ocorre no sul do Brasil nos Biomas Floresta Atlântica e Pampa. Acredita-se que o gênero possua um elevado grau de endemismo, com oito espécies associadas a regiões específicas de distribuição muito restrita. As espécies do gênero habitam ambientes de água doce com baixa correnteza e estão tipicamente associadas à vegetação campestre. Este trabalho teve como objetivo principal estudar e entender os padrões de diversificação em *Cnesterodon* através da utilização de ferramentas moleculares, e os resultados foram divididos em três manuscritos. No primeiro, foram utilizadas sequências de DNA dos genomas mitocondrial e nuclear para avaliar a estrutura genética e geográfica de *C. decemmaculatus*, uma espécie amplamente difundida no Bioma Pampa. Os resultados demonstraram que *C. decemmaculatus* possui uma estrutura genética associada aos amplos sistemas de drenagem do Sul da América do Sul. Esta espécie possui uma história evolutiva recente, com estimativas de divergência entre grupos genéticos entre 0.8 e 0.02 milhões de anos. Reconstruções filogeográficas apontam para um ancestral de *C. decemmaculatus* originado na Bacia do Uruguai, com posterior colonização para a Bacia do Rio Negro e bacias costeiras do Uruguai, com evidências de eventos de captura de cabeceiras ao longo de sua diversificação no Bioma Pampa. No segundo estudo, o objetivo foi estabelecer o número de linhagens genéticas relacionadas aos táxons *Cnesterodon* sp. nov. B e *C. brevirostratus* através de sequenciamento de nova geração pela técnica de RAD-seq. Essas espécies, simpátricas, endêmicas do Planalto Sul-Brasileiro (PSB), e restritas a altitudes acima de 750 metros podem ser distinguidas pela morfologia do órgão copulador masculino (gonopódio). Porém, a distinção baseada em outras estruturas morfológicas em fêmeas e machos jovens é extremamente difícil e até o momento não existem limites morfológicos claros entre esses táxons. Os resultados permitiram concluir que *Cnesterodon* sp. nov. B e *C. brevirostratus* constituem um complexo de espécies crípticas formado por quatro linhagens genéticas distintas, associadas ao tamanho do gonopódio e aos sistemas de drenagem Tramandaí-Mampituba, Laguna dos Patos e Rio Uruguai . A origem, diversificação e manutenção destes grupos genéticos estão provavelmente associadas a eventos de especiação alopátrica e seleção sexual. Finalmente, no terceiro estudo, dados de sequência de DNA (mitocondrial e nuclear) foram usados para propor uma hipótese filogenética, incluindo as novas linhagens

identificadas no trabalho anterior. As filogenias moleculares se mostraram discordantes das resultantes a partir de dados morfológicos, e sugeriram que as espécies do Bioma Pampa e do complexo de espécies crípticas do PSB são grupos irmãos. A idade do ancestral comum mais recente do gênero foi estimada em 8 milhões de anos, sugerindo que o início da diversificação das espécies atuais do gênero teria ocorrido no Mioceno. As partes elevadas da bacia do Paraná podem ter sido o centro de origem de *Cnesterodon*, com posterior colonização, através de eventos vicariantes, para drenagens ao Norte e ao Sul. Os resultados obtidos neste trabalho revelaram que o conhecimento sobre a diversidade de *Cnesterodon* era subestimada, e que o grupo possui uma história evolutiva complexa, cuja diversidade genética e distribuição geográfica foram moldadas por processos geológicos e biológicos.

ABSTRACT

Cnesterodon Garman, 1895 (Poeciliidae) is a South American genus of freshwater fishes composed of eleven recognized species, with ten of them formally described. Most species occur in Southern Brazil, in the Atlantic Forest and Pampa biomes. It is believed that this genus has high endemism, with eight species associated to specific regions and a narrow distribution. This genus inhabits freshwater environments with low streamflow, and is typically associated to grassland vegetation. This work had as its main goal to study and understand the diversification patterns of *Cnesterodon* using molecular tools, and the results were divided in three manuscripts. In the first study, DNA sequences from mitochondrial and nuclear genomes were used to evaluate the genetic and geographic structure of *C. decemmaculatus*, a species widely distributed in the Pampa Biome. The results showed that *C. decemmaculatus* has a genetic structure associated to the broad drainage systems in Southern South America. This species has a recent evolutionary history, with divergence time estimates among genetic groups between 0.8 and 0.02 million years. Phylogeographic reconstructions suggest a *C. decemmaculatus* ancestor originated in the Uruguay basin with subsequent colonization of Rio Negro and Coastal basins of Uruguay, showing evidence for headwater capture events along its diversification in the Pampa Biome. In the second study, the goal was to establish the number of genetic lineages related to the taxa *Cnesterodon* sp. nov. B and *C. brevirostratus* using next-generation sequencing (RAD-seq). These sympatric species, endemic to the Southern Brazilian Highlands (SBH) and restricted to elevations above 750m can be distinguished by the morphology of the male copulatory organ (gonopodium). However, the distinction based on other morphological structures in females and young males is extremely difficult, and so far there are no clear morphological limits between these taxa. The results allowed us to conclude that *Cnesterodon* sp. nov. B and *C. brevirostratus* constitute a complex of cryptic species formed by four distinct genetic lineages, associated to gonopodium morphology and to hydrological basins. The origin, diversification and maintenance of these genetic groups are probably associated to events of allopatric speciation and sexual selection. Finally, in the third study, DNA sequence data (mitochondrial and nuclear) were used to propose a phylogenetic hypothesis including the new lineages identified in the previous study. Molecular phylogenies were discordant from those based on morphological data, and suggested that the species from the Pampa Biome and those from SBH represent

sister groups. The time to the most common ancestor of the genus was estimated around 8 million years, suggesting that the initial diversification of the current species in the genus occurred in the Miocene. High elevation areas in the Paraná basin could have been *Cnesterodon* center of origin, with the subsequent colonization, based on vicariant events, to Northern and Southern drainages. The results from this work revealed that knowledge about *Cnesterodon* diversity was underestimated, and that this group has a complex evolutionary history, in which genetic diversity and geographic distribution have been shaped by geological and biological processes.

CAPÍTULO 1

Introdução Geral

1.1 A ictiofauna no Neotrópico Sul

A ictiofauna de água doce neotropical é considerada a mais diversificada do mundo, estimada em aproximadamente 8.000 espécies (Schaefer, 1998; Reis *et al*., 2003), o que representa quase 28% da diversidade global de peixes. Somente durante as duas últimas décadas do século passado, cerca de 800 novas espécies de água doce foram descritas somente para a América do Sul (Vari e Malabarba, 1998). Atualmente, entretanto, a diversidade da ictiofauna Neotropical ainda está longe de ser totalmente conhecida, visto que inventários de biodiversidade são incompletos e o número de táxons existente ainda é subestimado (Lévêque *et al*., 2008).

Dentro da região neotropical, o Brasil é o país de maior área (8.547.404 km^2), possuindo cerca de 22% de todas as espécies mundiais de água doce (Buckup *et al*., 2007). O número exato de espécies em águas continentais brasileiras, além de ser desconhecida, é de difícil determinação porque muitas bacias hidrográficas nunca foram amostradas. O número de pesquisadores e infraestrutura necessários para a amostragem e monitoramento são insuficientes, os inventários aquáticos ainda são escassos e as informações são dispersas e muitas vezes de difícil acesso (Agostinho *et al*., 2005).

Trabalhos com peixes marinhos e de água doce do Sul da América do Sul têm aumentado significativamente nos últimos anos (López *et al*., 2002; Baigún e Ferriz, 2003; Hubert e Renno, 2006; Pascual *et al*., 2007), porém a grande maioria consiste em descrições de espécies e estudos osteológicos, de forma que investigações sobre as relações filogenéticas entre grupos de peixes destas áreas são raras e a informação quanto à diversidade no nível específico da maior parte dos *taxa* é fragmentada (Vari e Malabarba, 1998). A biogeografia da ictiofauna neotropical também é pobremente conhecida e muitas questões evolutivas permanecem não resolvidas em diversos níveis taxonômicos, especialmente os inferiores (famílias e gêneros) (Ribeiro, 2006).

A grande maioria dos peixes neotropicais pertence a cinco grupos: Characiformes (cerca de 1.800 espécies descritas), Siluriformes (pelo menos 2.000 espécies conhecidas), Gymnotiformes (cerca de 200 espécies), Cyprinodontiformes (cerca de 700 espécies) e Ciclídeos (cerca de 500 espécies). Esta ictiofauna é caracterizada pela abundância de peixes pequenos (20-30 mm) entre os Characiformes, Siluriformes e Cyprinodontiformes e alguns peixes muito grandes, como o peixe-gato *Brachyplatystoma goliath* na Amazônia (até 3 m de comprimento e 140 kg), ou o *Arapaima gigas*, (até 4,5 m de comprimento e

200 kg). A diversidade de peixes neotropicais tem uma longa história temporal que, para alguns clados endêmicos mais basais, se reporta ao período Cretáceo (entre 65-145 milhões de anos antes do presente). Alguns clados modernos no nível de gênero diferenciaram-se durante o Paleogeno (entre 23-65 milhões de anos antes do presente) (Vari e Malabarba, 1998).

Peixes de água doce tendem a ser confinados aos sistemas de drenagem e fornecem um sistema relativamente conservado para examinar padrões de distribuição que podem refletir marcas do passado continental e mudanças climáticas (Lévêque *et al*., 2008), visto que a ictiofauna de água doce possui sua história associada principalmente à história geológica dos cursos d´água (Castro, 1999). Os processos de especiação e extinção têm interagido com eventos climáticos e geológicos e ambos têm isolado populações de peixes, fornecendo oportunidades de migração e colonização de novos habitats (Lévêque *et al*., 2008). Conhecer a diversidade dos peixes, sua distribuição e relações evolutivas, e com base nestas informações averiguar possíveis associações pretéritas entre as áreas geográficas constituem passos necessários para a detecção de áreas únicas e prioritárias para o estabelecimento de planos de conservação. Informações filogenéticas são de fundamental importância para o entendimento de questões relativas à biogeografia histórica numa escala menor que a continental (Vari e Weitzman, 1990). Estudos de pequenos grupos monofiléticos que estejam limitados em áreas restritas contribuem para aumentar o conhecimento sobre a ictiofauna neotropical (Weitzman e Malabarba, 1998).

1.2 Biomas Sul-Brasileiros

A região sul do Brasil está inserida em biomas ricos em diversidade: Pampa e Mata Atlântica (IBGE, 2012). O bioma Pampa é restrito à metade sul do estado do Rio Grande do Sul e apresenta continuidade florística e faunística nos territórios do Uruguai e Argentina (Boldrini *et al.*, 2010). As fronteiras do Pampa são a região da Patagônia, o planalto brasileiro (Mata Atlântica) e na planície do Chaco, no sul, nordeste e noroeste, respectivamente. Esta região é uma das maiores áreas de pastagem no mundo, cobrindo cerca de 500.000 km2 entre as latitudes 29°S e 39°S (Berretta, 2001; Pallarés *et al.*, 2005). A origem dos campos neste bioma é incerta. Segundo Rambo (1954), o fator determinante seria o controle exercido pelo solo, sendo o clima geral propício ao desenvolvimento de florestas subtropicais. Embora a região do Pampa já fosse dominada por campos quando da chegada dos seus primeiros habitantes no final do Pleistoceno (Behling, 2002), foi a partir

de 1800, com a intensa colonização dessa região por colonos europeus, que houve uma alteração mais significativa da composição vegetal da região, principalmente devido ao pastoreio e à implantação de cultivos agrícolas. Mais recentemente, a introdução de espécies exóticas e a implantação da indústria de reflorestamento representam ameaças consideráveis à biota do Pampa (Pillar, 2006). Muito pouco se sabe sobre a fauna do bioma Pampa e sua diversidade ainda está longe de ser completamente conhecida. No âmbito da flora, Boldrini (1997) estimou a ocorrência de 3000 espécies de gramíneas somente para o estado do Rio Grande do Sul. O número de estudos realizados no bioma Pampa ainda é baixo em comparação à outros Biomas. E este fato, aliado à ideia de que áreas dominadas por pastagens não são ricas em diversidade, são consideradas as principais barreiras para esforços de conservação. Em relação às espécies de peixes, o Pampa brasileiro é considerado heterogêneo, devido à peculiaridade geomorfológica das bacias de drenagem neste bioma. Estas bacias têm uma longa história de isolamento, e, portanto, podem servir de base para o estabelecimento de áreas prioritárias para conservação (Roesch *et al*., 2009). Particularmente o Pampa brasileiro tem recebido pouca atenção em comparação com outros biomas reconhecidos, e seu status de risco ainda não é suficientemente documentado (Carvalho e Batello, 2009). Apesar dos avanços na legislação e iniciativas de conservação, menos de 3% da superfície brasileira pampeana está sendo oficialmente protegida em sete unidades de conservação, com cerca de 375.000 ha (Bilenca e Miñarro, 2004).

A porção brasileira do bioma Mata Atlântica é um mosaico de biodiversidade composto por vários tipos de vegetação e por uma série de ecossistemas quase contínuos ao longo da costa litorânea (ocorrendo desde o Rio Grande do Sul até o Rio Grande do Norte) (Pinto e Brito, 2005). De acordo com a lei federal nº 11.428/2006 consideram-se integrantes do bioma Mata Atlântica *sensu lato* as seguintes formações nativas: Floresta Ombrófila Densa, Floresta Ombrófila Mista (mata de araucárias), Floresta Ombrófila aberta, Floresta Estacional Semidecidual, Floresta Estacional Decidual, manguezais, vegetações de restinga, campos de altitude, brejos interioranos e encraves florestais do Nordeste.

No sul do Brasil, o bioma Mata Atlântica é constituído por um mosaico de formações campestres e florestais (Overbeck *et al*., 2006), sendo que as formações campestres são denominados Campos de altitude do planalto das araucárias ou Campos de Cima da Serra (CCS) e se situam em regiões de altitude superiores a 800m. Esta região e

está localizada na metade meridional de Santa Catarina e no extremo nordeste do Rio Grande do Sul. Apesar de ocuparem áreas relativamente pequenas, estas matas e campos de altitude são importantes centros de riqueza de espécies e de endemismos (Safford, 1999), devido à alterações bruscas de relevo que proporcionam diferentes ambientes (Gentry, 1995; Webster, 1995). A hidrografia desta região compreende todas as nascentes dos rios Canoas e Pelotas, que correspondem aos principais formadores iniciais da bacia do rio Uruguai. Ao sul, encontra-se a bacia do rio Taquari e Antas e em menor escala podemos encontrar pequenas porções do rio Maquiné, Três forquilhas e Mampituba, representados por pequenos arroios que descem as escarpas a leste dos CCS. Esta região apresenta rios típicos de terras altas, com corredeiras frias e cristalinas, além de áreas de remansos conhecidas também como regiões de turfeiras. A ictiofauna dos CCS apresenta elevado endemismo e diversidade subestimada, visto que descrições de espécies endêmicas começaram a ser realizadas há menos de duas décadas (Malabarba *et al.*, 2009). Infelizmente, esta fauna já enfrenta algumas ameaças pela introdução de espécies exóticas e práticas agrícolas agressivas que prejudicam as margens dos rios e a qualidade da água.

1.3. Estudos genéticos na região Sul do Brasil

Estudos genéticos com espécies de vertebrados das regiões citadas acima ainda são escassos. Com relação ao sul da América do Sul, a região da Patagônia e subandina e Pampa argentino apresentam o maior número de trabalhos realizados (Morando *et al*., 2004; Braun *et al*., 2005; Palma *et al*., 2005; Fontanella *et al*., 2012 e Mapelli *et al*., 2012). Já para a região dos campos de Cima da Serra a situação é ainda pior, e os poucos trabalhos existentes tratam não de espécies endêmicas, mas de espécies de vertebrados típicas de Mata Atlântica e que também ocorrem nessa região (p.ex. Grazziotin *et al.*, 2006; Cabanne *et al.*, 2007; Gonçalves *et al.*, 2009 e D'Horta *et al.*, 2011). Até o momento, nenhum estudo filogenético ou filogeográfico foi realizado com a ictiofauna endêmica dos CCS, já que mesmo para a região sul do Brasil como um todo, esses estudos são escassos (p. ex. Beheregaray *et al.*, 2002; Thomaz *et al.*, 2010; - espécies da costa Sul-brasileira; Thomaz *et al*, 2015 – drenagens costeiras da Mata Atlântica, Vergara *et al.*, 2008 - peixes de estuário; Iriarte *et al.*, 2011 - peixes marinhos; Dergam *et al.*, 2002 - peixes de água doce continental).

1.4 A família Poeciliidae e o gênero *Cnesterodon* Garman, 1895

A família Poeciliidae (Rosen e Bailey, 1963) compreende 299 espécies válidas de pequenos peixes comprimidos lateralmente e que variam bastante em tamanho. As espécies desta família se distribuem em água doce e salobra nos continentes americano e africano embora sejam nativas do Novo Mundo, onde sua distribuição é principalmente tropical (Rosen e Bailey, 1963). Esta família é uma das quatro dentro da ordem Cyprinodontiformes que apresentam fertilização interna e uma das três que apresentam viviparidade (Parenti, 1981; Meyer e Lydeard, 1993). Sua classificação mais atual a divide em 3 subfamilias: Poeciliinae (220 espécies), Procatopodinae (78 espécies) e Aplocheilichthynae (1 espécie).

A subfamília Poeciliinae é amplamente distribuída através das Américas e Caribe, e compreende espécies de peixes de aquário bem conhecidas como os *guppies*. Peixes machos desta subfamília possuem um órgão copulador exclusivo denominado gonopódio, formado a partir de modificações nos raios 3, 4 e 5 da nadadeira anal (Parenti, 1981; Lucinda e Reis, 2005). Algumas espécies desta subfamília já foram utilizadas para estudos experimentais sobre seleção natural e sexual (Endler, 1983; Houde, 1997; Schluter *et al*., 1998; Hamilton, 2001) e estudos comparativos de história de vida e evolução (Grove e Wourms, 1991, 1994; Arias e Reznick, 2000; Reznick *et al*., 2002, 2007). Porém, o conhecimento acerca das relações filogenéticas, biogeografia e o grau de diversidade intra e intergenérica é escasso (Lucinda, 2003; Lucinda e Reis, 2005). Dentro da subfamília Poeciliinae podemos destacar a tribo Cnesterodontini, composta por três gêneros (*Phallotorynus*, *Phalloceros* e *Cnesterodon*) e autóctone Sul-americana, com distribuição restrita à região sudeste da América do Sul (Argentina, Brasil, Paraguai e Uruguai) (Lucinda e Reis, 2005). Membros desta tribo parecem não ter parentes próximos, são altamente especializados e podem ser muito antigos (Rosen e Bailey, 1963).

O gênero-tipo da tribo Cnesterodontini é *Cnesterodon* Garman 1895, composto por 11 táxons validos e 10 espécies descritas de peixes pequenos (4-5 cm) que se distribuem no Sul da América do Sul, nas bacias do Alto Araguaia, sistema Paraná-Paraguai, Bacia do Uruguai e ao longo das drenagens costeiras de São Paulo até a Argentina, bem como pequenas drenagens do Oeste Argentino (Lucinda, 2005). Embora a taxonomia deste grupo pareça estar bem resolvida, muito pouco se sabe sobre a história das espécies além de suas descrições. Popularmente são conhecidos como barrigudinhos, pois a espécie é vivípara e as fêmeas ficam "barrigudas" no período de gestação. Este gênero apresenta um elevado

grau de endemismo por possuir ao menos oito espécies associadas a regiões específicas e com distribuições muito restritas (Figura 1).

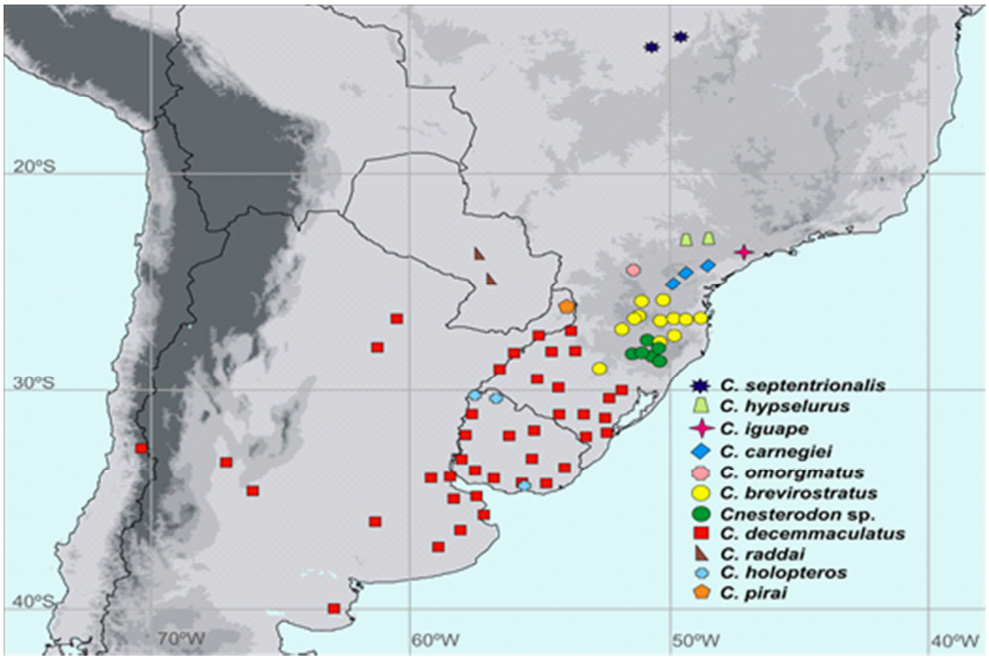


Figura 1: Distribuição geográfica das espécies do gênero *Cnesterodon*, com base nos estudos de Lucinda (2005) e Aguilera *et al*. (2009).

Estudos filogenéticos moleculares com a família Poeciliidae sugerem que a família seja antiga, cujas origens precedem o início do Cretáceo (Hrbek *et al*., 2007). A tribo Cnesterodontini ocupa uma posição basal dentro da família, corroborando a hipótese de Rosen e Bailey (1963) sobre a antiguidade da tribo. Por sua vez, o gênero *Cnesterodon* também ocupa uma posição basal dentro da tribo e estes dados, aliados com a distribuição fragmentada do gênero, podem ser um indicio de que a diversificação e a distribuição atual do gênero foram moldadas por eventos climáticos e geológicos muito antigos.

Ao longo da presente tese, serão abordados em maior detalhe três dos 11 táxons válidos para o gênero (Lucinda e Reis, 2005). A espécie *C. decemmaculatus* (Jenyns, 1842) ocorre no bioma Pampa e apresenta a distribuição mais ampla dentro do gênero, ocorrendo no rio Uruguai, sistema da Laguna dos Patos, Rio Negro, Rio Salado, drenagens oeste da Argentina e pequenas bacias costeiras do Uruguai e Argentina. Na região dos Campos de Cima da Serra podemos encontrar duas espécies endêmicas: *C. brevirostratus*

Rosa & Costa, 1993 e *Cnesterodon. sp*. B, ainda não descrita (Juan Anza, Paulo H. F. Lucinda, Luiz R. Malabarba, dados não publicados). Ambas ocorrem em ambientes sem correnteza, geralmente em banhados e/ou turfeiras que tenham vegetação aquática. Estas espécies são restritas a altitudes acima de 750 m e a ambientes aquáticos cercados por vegetação típica de campo. Podem ser encontradas na porção superior dos rios Pelotas e Canoas (drenagem superior do rio Uruguai; do rio Jacuí; nas cabeceiras do rio Maquiné, no sistema Tramandaí; e nas cabeceiras da drenagem do rio Itajaí-Açu) (Lucinda, 2005). Não existem muitas informações acerca da biologia e relações filogenéticas de *Cnesterodon* sp. B. Além de ocorrer em simpatria com *C. brevirostratus*, seu focinho é visualmente mais comprido e estreito, possuindo também gonopódio mais delgado e curvado (Figura 2). Porém, não existem limites morfológicos descritos na literatura que tornem possível a separação destes dois táxons. Além disso, estas diferenças morfológicas relacionadas ao tamanho do gonopódio são mais pronunciadas em populações simpátricas e a distinção de morfótipos em populações alopátricas não é clara. Na ausência de gonopódio formado, caso das fêmeas e machos juvenis, a separação se torna ainda mais difícil.

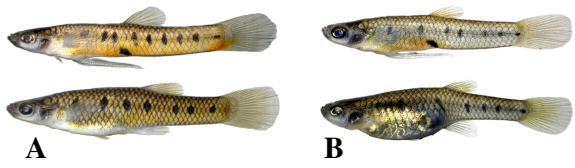


Figura 2: Morfótipos das espécies de *Cnesterodon* do Planalto Sul-Brasileiro, machos acima e fêmeas abaixo: A) Exemplares de focinho longo (Macho: comprimento total 46,2mm e comprimento padrão 38,5 mm. Fêmea: comprimento total 30,2 mm e comprimento padrão 26,5 mm). B) Exemplares de focinho curto (Macho: comprimento total 23,9 mm e comprimento padrão 19 mm. Fêmea: comprimento total 33,5 mm e comprimento padrão 27,7 mm).

1.5 A abordagem molecular como uma ferramenta em estudos evolutivos

Inúmeros avanços em estudos evolutivos têm sido feitos nos últimos anos com o aprimoramento das técnicas de sequenciamento. Marcadores moleculares são uma ferramenta poderosa nos estudos que envolvem questões taxonômicas e populacionais, pois permitem acessar a história evolutiva dos organismos em muitas escalas geográficas (Avise, 2000). Até recentemente, a maioria dos estudos de populações naturais com espécies não modelos utilizavam amplamente marcadores como: sequências de DNA (especialmente mtDNA), ALFPs (*amplified fragment length polymorphisms*) e SSR (microssatélites). Porém, estas metodologias são restritas a um número moderado de marcadores moleculares (Luikart *et al.*, 2003). A maior parte da produção de sequências de DNA gerada até o momento foi através da primeira geração de tecnologias de sequenciamento, desenvolvida por Sanger *et al.* (1977). Este método proporcionou grandes avanços nas áreas biomédicas, na compreensão das relações evolutivas representadas por filogenias, nos estudos populacionais e nos processos biogeográficos. Aprimorada ao longo do tempo, atualmente permite a leitura de fragmentos de até 1000 pares com precisão elevada de 99,9% (Shendure e Ji, 2008).

Desde as décadas de 70-80, informações provenientes do DNA organelar vêm sendo amplamente utilizadas em diversos estudos evolutivos. Em especial, a molécula do DNA mitocondrial (mtDNA), passou a fazer parte de muitos, senão da maioria dos estudos envolvendo estrutura populacional, relações filogenéticas e o entendimento de vários aspectos biológicos e evolutivos de uma grande variedade de organismos (Moritz *et al*., 1987; Avise, 1994; Avise, 2000). Características genéticas e estruturais extremamente peculiares e únicas tornam esse genoma interessante para diversos estudos, dentre elas podemos citar: herança materna de uma única molécula (haploide); não está sujeito à recombinação e fornece marcadores homólogos; ocorre em várias cópias em cada célula; replicação é contínua, unidirecional e simétrico sem qualquer edição ou aparente mecanismo de reparo; não possui introns presentes; é um genoma pequeno (com aproximadamente 16 kb nos animais) e circular, com raras exceções e a taxa evolutiva, ou seja, de substituições de base é muito alta, quando comparada a do genoma nuclear (Billington, 2003).

Nos últimos 15 anos, o mtDNA tornou-se um marcador muito popular nos estudos genéticos destinados a responder questões filogenéticas e populacionais em peixes. Particularmente a utilização do mtDNA pode: fornecer evidências sobre a hibridação e

introgressão entre peixes, servir como um marcador genético em análise forense e fornecer informações críticas para uso nos programas de conservação e reabilitação (Avise, 1994; Billington, 2003). Porém, estudos filogeográficos/populacionais, ou ainda, filogenias de difícil resolução, requerem conjuntos de dados com um grande número de *loci* e maior cobertura taxonômica para aumentar a confiabilidade na reconstrução da história evolutiva de espécies e resolver os limites entre táxons (Dupuis *et al*., 2012; Hohenlohe *et al*., 2012). Entretanto, o desenvolvimento de marcadores de DNA e o sequenciamento de grandes conjuntos de dados sob a tecnologia de Sanger ainda é trabalhoso e caro. Por consequência, a maioria dos estudos evolutivos cujos dados foram gerados por essa tecnologia de sequenciamento são baseadas na análise de um pequeno número de *loci* (Shendure e Ji, 2008).

1.5.1 O Sequenciamento de nova geração

O sequenciamento de Nova Geração (*Next Generation Sequencing* – NGS) surgiu por volta do ano de 2005, como uma técnica revolucionária para a investigação evolutiva. Enquanto um sequenciador de eletroforese capilar processa, em média, 96 fragmentos por vez, os sequenciadores de nova geração podem ler até bilhões de fragmentos ao mesmo tempo, produzindo gigabases de sequências de DNA em um curto espaço de tempo e com um custo mínimo por par de base. Esta nova tecnologia permite o re-sequenciamento para organismos com um genoma de referência existente a um custo relativamente baixo. Nos últimos anos, novas técnicas de NGS têm sido desenvolvidas com intuito de genotipar marcadores SNPs em larga escala para organismos não modelo (sem genoma de referência), tornando-se um método atraente para estudos genéticos populacionais em todo nível de genoma (Baird et al., 2008). As vantagens da utilização de marcadores SNPs em relação à outros marcadores (AFLP, SSR, RFLP) é que eles representam a fonte mais abundante de variação no genoma na maioria dos organismos e a distribuição ao longo de todo genoma ocorre em alta densidade (Narum *et al*., 2008).

O NGS utiliza-se de plataformas como Illumina (Bentley *et al.*, 2008), Roche 454 (Rothberg e Leamon, 2008) e AB SOLiD (Pandey *et al.*, 2008) e baseia-se num processo inicial de fragmentação do DNA genômico por um processo químico, mecânico ou enzimático, gerando pedaços de DNA chamados de *template*. O objetivo é quebrar o DNA da maneira o mais uniforme possível para que todo o genoma seja coberto. Após a

fragmentação são incorporados adaptadores (sequências artificiais conhecidas) ao *template*, os quais são utilizados em etapas posteriores de isolamento dos fragmentos, tornando possível combinar diferentes amostras em uma mesma reação de sequenciamento. Depois de ligadas, as amostras são misturadas, amplificadas e sequenciadas juntas. Após isso, no processo de sequenciamento, essa parte do adaptador é lida e as amostras são separadas computacionalmente. Os resultados do sequenciamento são chamados de *reads* (leituras curtas), por terem cerca de 100 pares de base. Esta técnica resulta em sinais luminosos que são decodificados de modo a determinar a base com um score de qualidade com uma precisão de acerto acima de 99,5% (Nowrousian, 2010; Zhang *et al*., 2011). O número de aplicações do NGS é bastante amplo, basta modificar a etapa de preparo da biblioteca e de análise para obter-se uma nova aplicação. Diversos métodos de NGS foram desenvolvidos nos últimos anos: *Target amplicon*, *Reduced representation*, *Transcriptome*, *UCEs, Anchored Enrichment* e *EPIC*. Embora todos eles sejam capazes de identificar e marcar milhares de marcadores genéticos distribuídos aleatoriamente em todo o genoma alvo, sua utilidade pode variar conforme os dados a serem analisados.

O advento do sequenciamento de nova geração tornou possível um maior conhecimento sobre espécies crípticas, a diferenciação de linhagens genéticas únicas dentro de espécies morfologicamente muito semelhantes, ou mesmo idênticas (Bickford *et al.*, 2007). Estas novas plataformas de sequenciamento apresentam a grande vantagem de permitir um sequenciamento altamente representativo de genomas em um único passo, o que é extremamente relevante, em razão da grande redução de custo alcançada com essas metodologias. Até o momento tais técnicas ainda são escassas na pesquisa brasileira e, por consequência, há uma grande lacuna no conhecimento técnico e na formação de pesquisadores que tenham o domínio sobre a geração, tratamento, análise e interpretação dos dados de NGS.

1.5.2 A técnica RAD-seq

O método RAD-seq (*restriction site-associated DNA sequencing*) pode ser definido como uma representação reduzida do genoma, o qual é fragmentado com enzimas de restrição específicas, cortado e, em seguida sequenciado. Os fragmentos curtos (chamados de *tags*) do DNA genômico que flanqueiam cada sítio da enzima de restrição são, então, investigados para a presença de SNPs. Essa técnica é capaz de gerar dezenas de milhares

deles a partir de uma amostra de vários indivíduos em uma única corrida. Sua principal vantagem é o delineamento amostral sem o conhecimento prévio do genoma do grupo de interesse. Isto, aliado ao baixo custo do método e à rapidez na geração de grande quantidade de dados tornam o método interessante para aplicação em organismos não-modelo (Hohenlohe *et al.*, 2012). Nestes casos, é necessário realizar a montagem do genoma *de novo* a partir dos fragmentos obtidos, o qual demanda uma cobertura maior e mais trabalhosa, bem como a utilização de computadores com quantidades massivas de memória RAM, quando comparada a processos onde um genoma de referência está disponível.

Em geral, e de maneira um pouco mais detalhada, o método de RAD-seq baseia-se nas seguintes etapas: 1) DNA genômico é fragmentado com uma enzima de restrição; 2) os fragmentos de DNA são ligados a um adaptador P1. A sequencia de P1 é constituída por um identificador molecular (*barcode*), sequências complementares aos *forward primers* de sequenciamento e uma extremidade complementar ao sítio de corte da enzima; 3) os fragmentos de diferentes indivíduos são colocados juntos e fragmentados aleatoriamente; 4) Estes novos fragmentos são ligados, então, à um adaptador P2, formado pela sequencia complementar aos r*everse primers* de sequenciamento; 5) amplificação por PCR com os *primers* P1 e P2; 6) fragmentos agrupados serão separados computacionalmente através dos identificadores moleculares.

Diversas técnicas foram propostas para simplificar, aprimorar as análises e reduzir custos associados ao método (Baird *et al.*, 2008; Davey e Blaxter, 2010; Elshire *et al.*, 2011). A técnica utilizada no presente trabalho foi proposta por Peterson *et al.* (2012) e difere das outras por utilizar simultaneamente duas enzimas de restrição (uma de reconhecimento comum, de 4-6 pb para o sítio alvo, e outra de reconhecimento raro, de 8 pb para o sítio alvo), eliminando qualquer fragmentação aleatória e sendo mais precisa na seleção de tamanho dos fragmentos, como demonstrado no esquema da Figura 3. Estudos bem-sucedidos utilizando RAD-seq são cada vez mais comuns em muitos tipos de organismos não-modelo. Muitos trabalhos com espécies que possuem um número limitado de sequências conhecidas empregaram o uso desta metodologia para gerar grandes quantidades de SNPs: salmão (Houston *et al.*, 2012), truta arco-íris (Amish *et al.*, 2012), alcachofra (Scaglione *et al.*, 2012), berinjela (Barchi *et al.*, 2011), peixes marinhos (Gaither *et al*., 2015) e corais (Combosch e Vollmer, 2015).

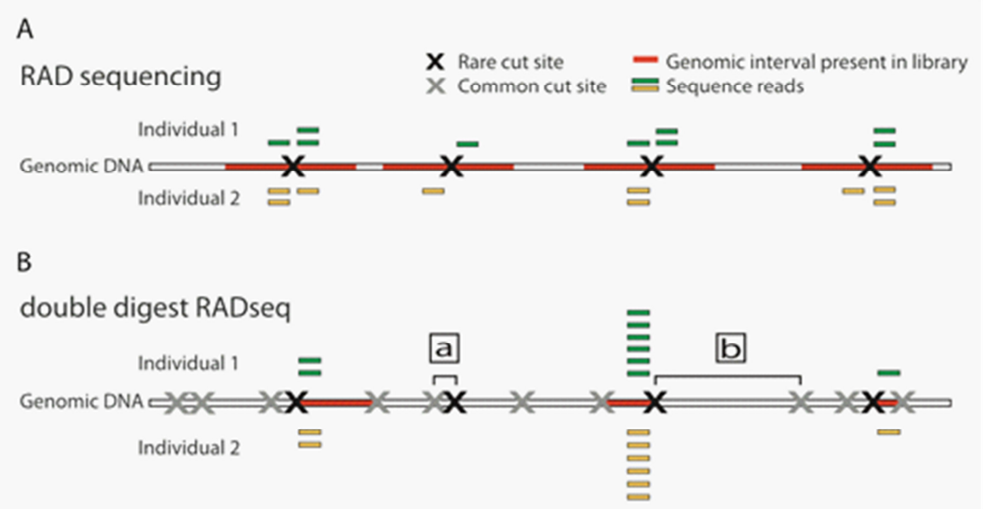


Figura 3: (A) Métodos tradicionais de RAD-seq usam uma única enzima de restrição juntamente com uma fragmentação aleatória secundária e uma seleção de tamanho de fragmentos ampla para gerar bibliotecas de representação reduzida, que consistem de todas as regiões adjacentes ao sítio de corte da enzima (segmentos vermelhos). (B) Método de dupla digestão (ddRADseq) usa duas enzimas de restrição, seguido pela seleção precisa de tamanho, que exclui regiões flanqueadas por [a] locais de reconhecimento muito próximos ou [b] muito distantes, recuperando uma biblioteca composta apenas de fragmentos de tamanho semelhantes (segmentos vermelhos). Nesta biblioteca é esperado que os tamanhos dos fragmentos sejam correlacionados entre os indivíduos (segmentos verdes e amarelos). Retirado de Peterson *et al*. (2012).

Esta metodologia é uma ferramenta útil para testes de atribuição individual entre unidades populacionais de organismos que têm baixos níveis de diferenciação com marcadores tradicionais. Por exemplo, populações selvagens de mosquitos da espécie *Wyeomyia smithii* estudadas pelo método tradicional de sequenciamento (DNA mitocondrial) apresentaram uma filogenia com dois clados atribuídos às regiões Norte e Sul, porém sem resolução entre populações dentro dos clados. Com o uso do RAD-seq foi possível identificar 3741 SNPs que permitiram resolver completamente as relações filogeográficas das 21 populações selvagens (Emerson *et al.*, 2010). Em outro estudo, populações de salmão do leste do Canadá foram divididas em grupos: interior e exterior à Baía de Fundy, com base em seu comportamento migratório. Estudos anteriores com base

no DNA mitocondrial não foram capazes de encontrar evidências de diferenciação genética entre estes dois locais (Fraser *et al.*, 2007). Entretanto, usando 320 SNPs, Freamo *et al.* (2011) encontraram pequenas, embora significativas, diferenças entre estas populações, e testes independentes demonstraram que oito dos loci observados estavam sob seleção divergente. Considerando que, nos últimos anos, tem havido uma diminuição significativa no número de adultos que retornam para desovar na Baía, esta ferramenta é muito útil para a conservação, uma vez que permite a determinação da origem dos poucos adultos que chegam ao local de desova. O relacionamento filogenético de espécies intimamente relacionadas tem sido solucionado com sucesso por este método, como demonstrado no trabalho de Wagner *et al.* (2013) com ciclídeos do lago Vitória. A análise de aproximadamente 90 mil *loci* permitiu a obtenção de uma filogenia com altos valores de suporte, fornecendo a primeira evidencia conclusiva da monofilia destas espécies.

O método de RAD-seq também permite fazer inferências de hibridação e fluxo gênico em uma escala temporal. Eventos de introgressão e substituição de DNA mitocondrial foram observados em ursos marrons e ursos polares, e esta evidência sugere um tempo do ancestral comum mais recente de 150 Kya entre populações híbridas (Lindqvist *et al.*, 2010; Edwards *et al.*, 2011). A análise de mais de um milhão de SNPs confirmou este evento de hibridação, mas mostrou que ursos polares e ursos marrons tiveram sua divergência inicial entre 4 e 5 milhões de anos atrás e experimentaram um longo intervalo de tempo com pouco ou nenhum fluxo gênico. A hibridização teria ocorrido só recentemente, provavelmente por causa dos efeitos de flutuações climáticas históricas (Miller *et al.*, 2012). Outro exemplo de detecção de híbridos é o trabalho de Catchen *et al.* (2013) que estudaram um complexo de espécies de peixes *stickleback* (*Gasterosteus aculeatus*). Este grupo experimentou diferenciação fenotípica e genética repetidas vezes quando populações panmíticas de peixes oceânicos invadiram habitats de água doce. Com aproximadamente 25 mil *loci* encontrados, foi possível detectar hibridização introgressiva persistente entre as populações oceânicas e de água doce ao longo da costa, além de expansão populacional recente para populações de água doce.

2. Objetivos

Este trabalho teve por objetivo geral estudar padrões de diversificação dentro do gênero *Cnesterodon* em ambientes campestres do sul do Brasil através do uso de marcadores moleculares de forma que este conhecimento possa servir como base para programas de conservação, principalmente para os cursos d'água das regiões estudadas.

Objetivos específicos:

a. Estabelecer um cenário filogeográfico para a espécie *C. decemmaculatus*, visando compreender: a distribuição espacial da estrutura genética e a escala de tempo da divergência genética entre populações;

b. Relacionar os resultados obtidos aos efeitos potenciais de barreiras ao fluxo gênico histórico no Bioma Pampa;

c. Estabelecer o número de linhagens genéticas envolvidas sob as nomenclaturas *Cnesterodon.* sp. nov. B e *C. brevirostratus* e compreender como estas linhagens estão geograficamente distribuídas;

d. Avaliar se as diferentes morfologias do gonopódio estão envolvidas no processo de separação e manutenção da estrutura genética dos indivíduos que habitam o Planalto Sul-Brasileiro;

e. Entender as relações filogenéticas entre as linhagens presentes no Planalto Sul-Brasileiro bem como a relação filogenética destas com outras espécies de *Cnesterodon*;

f. Determinar a escala de tempo para a história evolutiva dos táxons presentes no Planalto Sul-Brasileiro.

CAPÍTULO 2

Phylogeography of *Cnesterodon decemmaculatus* (Poeciliidae): a freshwater look at the Pampa biome in Southern South America

Manuscrito submetido para a revista *Zoologica Scripta*

Nelson Jurandi Rosa Fagundes
Telephone number: 55 51 3308 9826
Department of Genetics, Av. Bento Gonçalves, 9500, prédio 43312, sala 113, Federal University of Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brazil.
E-mail: nelson.fagundes@ufrgs.br

Phylogeography of *Cnesterodon decemmaculatus* (Poeciliidae): a freshwater look at the Pampa biome in Southern South America

ALINE M. C. RAMOS-FREGONEZI, LUIZ R. MALABARBA and NELSON J. R. FAGUNDES

Running Title: Phylogeography of *Cnesterodon decemmaculatus*
Ramos-Fregonezi *et al.*

Ramos-Fregonezi A.M.C., Malabarba L.R., Fagundes N.J.R. (2015) Phylogeography of *Cnesterodon decemmaculatus* (Poeciliidae): a freshwater look at the Pampa biome in Southern South America. *Zoologica Scripta*, 00, 000-000.

The Pampa is a South American Neotropical Biome considered rich in endemicity with unique biological features. Despite its importance, the phylogeographic patterns of this biome are still poorly understood and few studies are available. The formation of the Pampas hydrological system is the result of a combination of tectonism, climate and sea level changes. The processes involved in its evolution affected aquatic, semi-aquatic, and terrestrial communities that inhabited these environments. In this study, we access edpopulation genetic structure of *Cnesterodon decemmaculatus* from Southern Brazil and Uruguay, to examine the historical consequences of putative barriers in the Pampa Biome and temporal estimates, using mitochondrial and nuclear markers. We found genetic lineages overall correspondent to the major Pampean drainage systems. Estimates of divergence times for them ranged between 0.8 and 0.02 million years before the present, showing a recent history for this group. Bayesian phylogeographical reconstruction suggested that the lineage ancestral to *C. decemmaculatus* probably came from the upper Uruguay or the Parana Basin. Our results showed little and shallow genetic structure for *C. decemmaculatus*, which is in line with patterns found for other Pampean species. While our results showed a general genetic structure correspondent to drainage systems, some vicariant processes had effect on determining the current distribution of this freshwater taxa.

Aline M. C. Ramos-Fregonezi, Departament of Genetics, Federal University of Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brazil, E-mail: alinemitcheli@gmail.com

Luiz R. Malabarba, Departament of Zoology, Federal University of Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brazil, E-mail: malabarb@ufrgs.br

Nelson J. R. Fagundes, Departament of Genetics, Federal University of Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brazil, E-mail: nelson.fagundes@ufrgs.br

INTRODUCTION

The Pampa is a Neotropical biome dominated by natural grasslands spreading over plains of Uruguay, Northern Argentina, Southern Brazil, and part of Paraguay (Pallarés *et al*. 2005). However, this Biome is far from homogeneous, and several smaller biological "provinces" have been suggested, mainly based on the different plant communities across the Pampa (Cabrera & Willink 1980; Overbeck *et al*. 2007). Together, the unique biological features of this Biome and its heterogeneity have led to endemic species and to spatial and genetic structure in the Pampa (Freitas *et al*. 2012; Turchetto *et al*. 2014; Felappi *et al*. 2015), but similarly to other Neotropical biomes, the Pampas are underrepresented in studies of conservation and phylogeography (Lawler *et al*. 2006; Beheregaray 2008; Turchetto-Zolet *et al*. 2013). For example, several studies have described the influence of climatic changes during the Quaternary glacial cycles in the Northern Hemisphere (Hewitt 2000; Bennett 2004). On the other hand, little information exist on the relevance and impact of Quaternary glacial cycles over the distribution of Neotropical species (see Turchetto-Zolet *et al*. 2013 for a review).

The Pampas hydrological systems result from a long history of tectonism, climate and sea level changes since the Neogene (Casciotta *et al*. 1999). During the Late Miocene (11.8–10.0 million years ago, Mya), an extensive marine transgression resulted in the "Paranean Sea", flooding a significant portion of the Pampas in eastern Argentina, western Uruguay, and part of Brazil (Lundberg *et al*. 1998). In the early Pliocene (around 5.0 Mya) the sea-level retreated giving rise to extensive plains (Ortiz 1998). The final regressive stages of the ''Paranean Sea'' were linked to the formation of the pre-Paraná river, which drained into this sea, or into marginal lagoons or marshes (Aceñolaza 2004). Maack (1968) and Bossi (1969), based on indirect estimates, suggested that the Uruguay River was formed in the Pliocene with a probable migration of the Paraná River to the Uruguay Basin, even though other authors suggest its origin in lower Pleistocene (Iriondo 1996; 1999). In the Quaternary, new marine regressive/transgressive cycles and erosion continued to shape Pampean river

morphology (Tonni & Cione 1997; Quattrocchio *et al*. 2008). As such, current streams could be tributaries of other drainages in the past (Casciotta *et al*. 1999), and sea-level fluctuations may have altered the connection of isolated drainages in lagoons or estuaries (Martin & Dominguez 1994; Thomaz *et al*. 2015). The geomorphological process involved in the formation of the Pampas hydrological system surely affected aquatic, semi-aquatic, and terrestrial communities that inhabited these environments (e.g. Candela *et al*. 2012). While at the species level it is well accepted that vicariance processes have a major effect on determining the distribution freshwater taxa (Rosen 1978), dispersal of fish species may be related to the history of connections among basins (Bermingham & Martin 1998), since temporal or permanent land barriers will isolate populations within specific drainages (Vari 1988). Therefore, the distribution of isolated populations may be mostly attributed to drainage rearrangements (Hurwood & Hughes 1998; McGlashan & Hughes 2000; Zemlak *et al*. 2008; Burridge *et al*. 2008) and sea-level regressions (Ketmaier *et al*. 2004; Swartz *et al*. 2007, 2009 ; Thomaz *et al*. 2015).

It is likely that Pampean drainages have been associated with grassland ecosystems for a long period, since studies of past climate in the Pampas indicate that grasslands were dominant in both glacial and interglacial periods (Tonni *et al*. 1999; Behling *et al*. 2005). In agreement with such stability, some phylogeographic studies in the Pampas have found little geographic structure and shallow (<100,000 years) gene trees in this biome (Speranza *et al*. 2007; Turchetto *et al*. 2014). Nonetheless, in spite of grassland stability, changes in precipitation and temperature during glacial/interglacial periods may have affected aquatic and terrestrial communities, especially for species with very stringent ecological requirements. Indeed, a recent phylogeographic study of an endemic Pampean gecko found a strong geographic structure and a deep mitochondrial gene tree probably associated with its ecological habits (Felappi *et al*. 2015). However, there are no phylogeographic studies of freshwater taxa in the Pampas so far. To fill this gap, in this study we used mitochondrial and nuclear genetic markers to evaluate the genetic structure of *Cnesterodon decemmaculatus* (Jenyns 1842) in the Pampas. This species is one of the most widespread freshwater fish species in Pampean region, inhabiting rivers, ponds and shallow wetlands with vegetation. Since it is well accepted that the distribution of freshwater fish lineages mainly reflects the

paleogeography of a region providing historical information about a specific geographical region (Bermingham & Avise 1986; Bernatchez & Wilson 1998; Avise 2000), we aim to answer the following questions: 1) How well does genetic structure in this species correspond to the current river basins? 2) How old are the gene trees for *C. decemmaculatus* and how old is the genetic divergence among populations? 3) How these results relate to the potential effects of putative geological and biological barriers to historical gene flow in the Pampa Biome?

MATERIAL AND METHODS

We sampled 99 individuals from 38 localities, covering almost all of the species' range. Our sampling ranged from one to eight individuals per locality (Table 1). Individuals included in our analyses have been fixed in alcohol 96% and deposited at Universidade Federal do Rio Grande do Sul (UFRGS) scientific collection. All collections were performed under the approval of the government authorities of both countries (Ministério do Meio Ambiente, Brazil – SISBIO 12038-2 and Dirección de Recursos Naturales del Ministerio de Ganadería, Agricultura y Pesca, Uruguay).

Total genomic DNA was isolated from each sampled tissue with cetyltrimethyl ammonium bromide (CTAB) as described by Doyle & Doyle (1987). We tested two mitochondrial (mtDNA) genes: cytochrome oxidase subunit I (COI) and NADH dehydrogenase 2 (ND2); and two nuclear (nDNA) genes: SH3 and PX domain containing 3-like protein (SH3PX3) and myosin heavy chain 6 (Myh6). Unfortunately, two of them (SH3PX3 and COI) showed insufficient variation and were excluded from further analysis. Primers and amplification protocols for ND2 and Myh6 were described by Sorenson (2003) and Li *et al*. (2007), respectively. PCR products were checked on a 1% agarose gel, purified with Exonuclease I and Shrimp Alcaline Phosphatase (GE Healthcare®) and sequenced using the Sanger method in Macrogen Inc. (Seoul, South Korea). Nucleotide sequences were determined on both strands, checked and aligned with GeneiousPro v. 4.8 (Drummond *et al*. 2007). All sequences obtained in the present study were deposited in GenBank (XXXXXX-XXXXXX).

Geographical groups were defined according to broad systems of watersheds (Figure 1): Uruguay River Basin – URU, Negro River Basin – NEG, Mirim Lagoon

Basin – MIR and Southern coastal Basins of Uruguay – SOU (Figure 1, Table 1). Individuals collected in different localities within each basin were grouped and considered as a single population in the analysis. Descriptive statistics such as nucleotide (π) and haplotype diversity (H) as well as Tajima's D (Tajima 1983) and Fu's (Fu 1997) neutrality tests, analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) and pairwise Φ_{ST} were carried out in Arlequin 3.5 (Excoffier & Lischer 2010). We performed neutrality tests under two schemes: considering all individuals for a given basin, or excluding localities for which there was evidence of headwater capture. The evolutionary relationship between haplotypes was inferred using a median-joining network (Bandelt *et al.* 1999) estimated in NETWORK 4.1.0.9 (available at www.fluxux-engineering.com). The mitochondrial haplotypes were determined using DNAsp (Rozas *et al.* 2003). Haplotypes for the Myh6 gene were estimated in PHASE 2.1 (Stephens *et al.* 2001) using 10000 steps sampling every 10 steps and discarding the first 1000 steps as burnin. We ran PHASE several times using different starting seeds to ensure the reliability of the final estimate. We tested the correlation between geographic and genetic distances using Mantel and Spatial-autocorrelation tests, implemented in Alleles in Space program (AIS; Miller 2005).

Molecular divergence times among mitochondrial sequences were estimated by a phylogenetic analysis using Bayesian inference in BEAST 1.7.5 (Drummond & Rambaut 2007). Input .xml files were generated in BEAUti v. 1.7.5 (Drummond & Rambaut 2007), using a coalescent constant size tree prior, a random starting tree, TN93+G (ND2) model of sequences evolution, determined by Akaike Information Criterion in jModelTest2 (Darriba *et al.* 2012). The MCMC was run twice for 10 million generations each, and a 10% burn-in. We used a normal strict molecular clock calibrated with a normal distribution for molecular substitution rate of $8.6 \pm 0.1 \times 10^{-9}$ substitutions per base pair per year. This substitution rate is described in literature for the mitochondrial genome in Cyprinodontiformes (Hrbek & Meyer 2003). Since we had some evidence for headwater capture for some localities, we also used a Bayesian phylogeographic approach to estimate the most likely location of the ancestors in the mtDNA tree. We used the discrete phylogeographic diffusion model and used the Bayesian stochastic search variable selection (BSSVS) implemented in Beast. We adopted the same geographical groups used in another analysis classified as four

discrete states: URU, NEG, MIR and SOU. Furthermore, we also performed a coalescent-based species-tree analyses (Bayesian Estimation of Species Trees - BEAST) excluding localities with evidence for headwater capture in order to infer the colonization history of each river basin. For this analysis we used both markers (mitochondrial - ND2 and nuclear - Myh6), and the mtDNA rate was used as a reference for calibrating the nuclear molecular clock for Myh6. Remaining parameters and priors were the same used in phylogenetic analysis mentioned above. For all Bayesian analysis, sampling sufficiency was evaluated by monitoring effective sample size (ESS) and ensuring that all values were >200.

Estimates for effective population size (N_E) for each geographical group (URU, NEG, MIR and SOU) without the localities with evidence for headwater capture were performed in the program LAMARC v. 2.1.6 (Kuhner 2006). We used analysis of likelihood with a search strategy composed of three replicates of 10 initial chains and two long final chains. The initial chains were performed with 250 samples and a sampling interval of 20, using a burn-in of 1000 samples for each chain. The two final chains were carried out with the same burn-in and interval sampling, but with 10000 samples. We assumed one year as the generation time based on estimates of sexual maturity for the poeciliid *Poecilia reticulata* (Reznick & Bryga 1987)

RESULTS

Ninety-nine sequences were obtained for the mtDNA ND2 gene (980 bp) yielding 31 different haplotypes defined by 42 variable sites (28 parsimoniously informative). For the nDNA Myh6 gene we had 188 sequences (681 bp) resulting in 25 different haplotypes defined by 21 variable sites (16 parsimoniously informative). Haplotype diversities were high for both markers (H = 0.94 for ND2 and H = 0.79 for Myh6) considering the complete data set. Among geographic groups the highest values were found in URU and SOU Groups (Table 2). The groups with lower H values (NEG and MIR) also presented lower nucleotide diversity values (Table 2). Pairwise Φ_{ST} comparisons indicate that geographic groups are significantly different from each other except for the pair MIR vs. SOU (Table 3). However, genetic variation within groups was also high and significant, as evidenced by AMOVA (Φ_{ST} = 0.25 for ND2 and Φ_{ST} =

0.29 for Myh6; $P < 0.001$ for both), (Table 4). We found significant negative Fu's *Fs* for both markers in SOU, suggesting recent population expansion. Fu's *Fs* tests were also significant negative considering all samples and URU (Myh6). Tajima's *D* values were significant and negative considering all samples (ND2 data). Excluding localities with evidence for headwater capture (see below) resulted in significant and negative values for both Fu's *Fs* and Tajima's *D* in URU for the Myh6 marker.

The network of mtDNA haplotypes showed a general structure associated to basin and a close relationship among haplotypes (Figure 2A). The four most common haplotypes (H3, H8, H11 and H14) represented almost 45% of the sample, were the most centrally located in the network, and the most widely distributed geographically (URU, MIR and NEG). Populations of MIR and SOU shared similar haplotypes, except for two localities in SOU, which shared haplotypes with URU. As in the mtDNA network, Myh6 haplotypes were closely related, but we could not observe any trivial structure based on basin (Figure 2B). URU, MIR and SOU shared the most frequent haplotypes (H5 and H1), and H9 was the only haplotype composed by populations of all drainage systems. The general weak genetic structure and close relationship among haplotypes for both markers are suggestive of a shallow genealogical history for this species (see below). Mantel and spatial-autocorrelation tests were not significant for both markers, indicating no correlation between geographic and genetic distances.

According to the ND2 mitochondrial tree (Figure 3), *Cnesterodon decemmaculatus* had a most recent common ancestor by approximately 0.8 million years ago (Mya) (95% Credible Interval (CI) 0.5 – 1.2 Mya). The phylogeographical reconstruction suggests (PP=0.58) that URU represents the most probable location of all mtDNA diversity in this species. Even though the mtDNA tree was not fully resolved, a few well supported clades could be identified. A first clade (posterior probability PP=1.00), which was sister to all other lineages, occurred in populations inhabiting elevated plains including three basins: URU, MIR and NEG, and was formed by haplotypes H11, H12, H14 and H19. The location of the ancestral lineage of this clade was uncertain given that all locations received PP<0.5, but URU and NEG received most support (PP=0.44, PP=0.24, respectively). We found a very divergent lineage represented by a single haplotype (H25), sampled in URU (Fig. S1 in Appendix S1). The next well-supported clade (PP=1.00) contains mtDNA haplotypes (H5, H6, H15,

H16 and H17) from populations inhabiting the Queguay drainage, a sub-basin from URU (Fig. S1 in Appendix S1). Another well supported clade (PP = 0.90) was mainly formed by individuals from SOU and MIR, with the presence of two individuals from NEG. The most likely location of the ancestral lineage of this clade is SOU (PP=0.46), even though MIR also had a relatively high support (PP=0.30). Finally, several widely distributed haplotypes formed a weakly supported group, having haplotypes present essentially in URU basin, but also in individuals from NEG and SOU. Although this group had a low support value, an internal clade formed by haplotypes H13, H4, H3, H9, H18 and H2 had high support (PP=0.99), being exclusive from URU.

The combined analysis of mtDNA and nDNA markers for *C. decemmaculatus* suggests that the split between URU and NEG, which was also the root of the tree, dates back from ~0.6 Mya (95% CI 0.27-0.95) (Figure 4). Around 0.12 Mya (95% CI 0.04-0.19) there is a split between URU and MIR+SOU, with the subsequent split among MIR and SOU more recent than 30,000 years ago (95% CI 23,000 – 57,000). Estimates for historical effective population size showed larger effective sizes in URU and SOU (>106 effective individuals), with intermediate values in MIR (~600,000 effective individuals) and the lowest values in NEG (~200,000 effective individuals; Table 5).

DISCUSSION

Our results show a general genetic structure based on broad watersheds systems that is more evident considering mtDNA sequences (Figures 2A and 3). However, there is no simple relationship between clades and basins, such that a single basin may contain more than one mtDNA clade, and different basins may share the same mtDNA haplotype (Figure 2A and 3). These observations may result from three alternative scenarios: allopatric differentiation followed by secondary contact due to either active dispersal, or headwater capture, or through sharing of ancestral polymorphism. Active long distance dispersal seems unlikely since *C. decemmaculatus* has demonstrated a poor swimming capacity in high-speed currents (Trenti *et al.* 1999). *Cnesterodon* species, however, are commonly found in shallow wetland habitats (Malthik *et al.* 2014), making it possible the dispersal through wetland habitats connecting very small streams from different river drainages or even through the capture of very small

headwater streams in the Pampa Biome. Our Bayesian phylogeographical reconstruction suggested URU as the ancestral location for all haplotypes (Figure 3). This is in agreement with the distribution of other species in the genus, which are mainly distributed in the Paraná-Uruguay river systems (Lucinda 2005).

Some interesting cases of haplotype (or haplogroup) sharing in this species deserve further discussion. The first involves H11 (typically from NEG) shared with MIR, H14 and H19, found in URU and MIR, respectively, but that are derived from H8 (Figure 2A). These haplotypes occur in the first diverging clade in mtDNA phylogeny (Figure 3), for which URU had higher probabilistic support as location of the ancestral lineage. However, there are three reasons that make NEG a better candidate for the ancestral location of this clade. First, the sampling points in URU and MIR containing these lineages are adjacent to NEG. Second, the seven exclusive substitutions found in this clade (Figure 2A) suggest it has evolved in relative isolation. Third, H14 and H19 are probably descendent from H11, which has been found in NEG. Therefore, this clade shows evidence of three events of headwater capture: NEG would have been colonized from URU between 0.2-0.8 Mya, and the recent period from 0.2 Mya would be associated with the colonization of MIR and URU from NEG. Even in the alternative scenario in which URU is the true location for the ancestor of this clade, two headwater capture events would have occurred, since NEG would have been colonized from URU by ~0.2 Mya, dispersing to MIR later on. Other examples of haplotype sharing include H8 (typically from URU, but shared with NEG) and H10 (typically from SOU, but shared with NEG). In both cases, the fact that these haplotypes are in a clear genealogical context within its respective basin, and that they occur in NEG in sampling sites neighbor to URU and SOU, respectively, lend support for the hypothesis that they also represent events of headwater capture, even though further sampling effort would be required to test this hypothesis more conclusively.

Headwater capture (or, more broady, “river capture”) is very common biogeographic process for freshwater fish taxa (Howard & Morgan 1993; Bishop 1995; Lundberg *et al*. 1998; Waters *et al*. 2001; Burridge *et al*. 2006; Ribeiro 2006; Hubert & Renno 2006; Waters *et al*. 2006). These events can occur due to erosion of the current landscape or as the result of tectonic or volcanic activity. These mechanisms have been evoked to explain some of the biogeographic patterns of freshwater fishes in South

America, including the presence of the same species in upland rivers and coastal drainages among isolated hydrographic basins exemplified by *Hypostomus*, *Bryconamericus* and *Cnesterodon* genera (Ribeiro 2006). Indeed, the tectonic activity in the Atlantic margin of South America may be as young as 1.6 Mya (Saadi *et al*. 2002). The presence of H11 and H19 in MIR and H8 in NEG are in agreement whit this pattern in which Atlantic drainages capture upland shield drainages. However, the presence of H14 in URU and H10 in NEG indicate that these events could also occur in the opposite direction, suggesting a highly dynamic scenario in the Pampas, possibly because there are not steep differences in elevation in this biome. Interestingly, Loureiro *et al*. (2011) also reported drainage rearrangements in both directions based on distribution pattern of *Austrolebias* species in Uruguay. Likewise in *Cnesterodon*, the species of *Austrolebias* inhabit shallow wetland habitats (Malthik *et al*. 2014), reinforcing the hypothesis of dispersal through very small streams by headwater capture of by temporary connections through wetland habitats.

On the other hand, the most parsimonious explanation for the presence of haplotypes H21 and H30 in SOU basin is incomplete lineage sorting, given the low phylogenetic signal associated with location and clade structure, and because the sampling points in which these haplotypes were found are distant from the border of the URU basin.

The contrast found between mitochondrial and nuclear haplotypes are probably related to differences in substitution rates and effective population size between these markers. Autosomal genes have four times larger effective size compared to mtDNA genes (Hedrick 2011), and usually have slower substitution rates, as a result has an improved power to detect recent population and genealogical structure mtDNA (Brown 1979). Hence, for our data it is difficult to discuss scenarios of headwater capture or shared polymorphism based on Mhy6 haplotypes (Figure 2B). However, it is possible that Myh6 haplotypes H23/H4 reflect the headwater capture event affecting NEG/MIR, while H9 distribution may be associated to the event connecting SOU/NEG discussed previously based on mtDNA data.

The evolutionary history of *C. decemmaculatus* dates back from late Pleistocene (± 0.6 Mya - Figure 4). While the Pampean region has been always dominated by grasslands during the whole Pleistocene (Behling *et al*. 2005), changes in precipitation

regimes may have promoted the formation of new wetland areas and the modification in the connection of different drainages (Zemlak *et al*. 2008; Jones & Johnson 2009). Since *C. decemmaculatus* is typically associated with grasslands, the Pleistocene could have been a period of great opportunity for occupation of new areas following rearrangements in the river systems. From a broader biogeographic perspective, the lineage ancestral to *C. decemmaculatus* probably came from the upper Uruguay or Parana basin, reaching the lower Uruguay before 0.6 Mya. (Figure 4). From URU, *C. decemmaculatus* colonized NEG basin, which is the most genetically isolated, and has the smallest effective population size (Table 5). The high genetic diversity in URU is consistent with its role as the ancestral location. Even though we have found a case of headwater capture from NEG to MIR, the most likely colonization route leading to MIR is through SOU. SOU also had a large N_E estimate, which suggests, in agreement to the our genetic data (Figure 2A, 2B, 3) that the genetic diversity in MIR is closely related to lineages also found in SOU (but not in NEG). The high genetic diversity in SOU is surprising, since this area suffered at least three transgression marine events during the Pleistocene and Holocene (Villwock & Tomazelli 1995) that could have led to declines in population size and local extinction. Pleistocenic changes in sea level, however, and consequent connection through the extended coastal plain during sea retraction were found associated with the genetic structure shown by several populations of *Hollandichthys* along the coastal rivers of southern and southeastern Brazil, and this may have affected in the same manner genetic diversity in SOU populations of *Cnesterodon*. The high N_E and the significant negative values of Fu's neutrality test indicates that if local extinction occurred in SOU populations, they have also experienced recent and strong population growth. Our results suggest that MIR and SOU diverged around 27,000 years ago. The different genetic diversity and N_E estimates for MIR and SOU may thus be related to differences in growth rate after the marine transgressions or to founder effect during the colonization of MIR (from SOU). On the other hand, the Pleistocene marine transgressions did not affect the NEG area (Panario & Gutiérrez 1999). Therefore, NEG may have been isolated from other populations during marine transgressions, leading to the accumulation of new mutations and fixation of different haplotypes (Figures 2A, 2B, 3). Taken together, our results showed a complex evolutionary scenario for *C. decemmaculatus*. Overall, river basins

provide a good relation to population structure (Φ_{ST}=0.25 and Φ_{ST}=0.29 for tDNA and Myh6, respectively), but several instances of headwater capture have occurred along its history. The genealogical structure is recent (less than 1.0 Mya), and some divergences between populations may be as recent as ~30,000 years ago, in agreement to other studies suggesting little and shallow genetic structure in this biome (Speranza *et al*. 2007; Turchetto *et al*. 2014).

ACKNOWLEDGEMENTS

The authors would like to thank Marcelo Loureiro Barrela (Universidad de la República, Uruguay), Juliana Wingert, Alice Hirschmann and Juliano Ferrer which provided or helped to collect samples for this study. We also thank Jéferson Fregonezi for contributions in this manuscript. This work was supported by Conselho Nacional de desenvolvimento Científico (CNPq), Programa de Pós Graduação em Biologia Animal (PPGBAN-UFRGS) and Programa de Pós Graduação em Genética e Biologia Molecular (PPGBM-UFRGS).

REFERENCES

Aceñolaza, F. G. (2004). Paleobiogeografía de la región Mesopotámica. In F. G. Aceñolaza (Ed.), *Temas de la biodiversidad del litoral fluvial argentino* (pp. 25-30). San Miguel de Tucumán: INSUGEO.

Avise, J. C. (2000). *Phylogeography*: *The history and formation of species.* Cambridge, MA: Harvard University Press

Bandelt, H. J., Forster, P. & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16, 37-48.

Beheregaray, L. B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology*, 17, 3754-3774.

Behling, H., Pillar, V., & Bauermann, S. G. (2005). Late Quaternary grassland (Campos), gallery forest, fire and climate dynamics, studied by pollen, charcoal and multivariate analysis of the São Francisco de Assis core in western Rio Grande do Sul (southern Brazil). *Review of Palaeobotany and Palynology*, 133, 235-248.

Bennett, K. D. (2004). Continuing the debate on the role of Quaternary environmental change for macroevolution. *Philosophical Transactions of the Royal Society B*, 359, 295-303.

Bermingham, E. & Avise, J. C. (1986). Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics*, 113, 939-965.

Bermingham, E. & Martin, A. P. (1998). Comparative mtDNA phylogeography of neotropical freshwater fishes: Testing shared history to infer the evolutionary landscape of lower Central America. *Molecular Ecology*, 7, 499-517.

Bernatchez, L. & Wilson, C.C. (1998). Comparative phylogeography of nearctic and palearctic freshwater fishes. *Molecular Ecology*, 7, 431-452.

Bishop, P. (1995). Drainage rearrangement by river capture, beheading and diversion *Progress in Physical Geography*, 19, 449-473.

Bossi, J. (1969). *Geologia del Uruguay*. Montevideo: Universidad de La Republica.

Brown, W. M., George, M., & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences, 76, 1967-1971.

Burridge, C. P., Craw, D. & Waters, J. M. (2006). River capture, range expansion, and cladogenesis: the genetic signature of freshwater vicariance. *Evolution*, 60, 1038-1049.

Burridge, C. P., Craw, D., Jack, D. C., King, T. M., & Waters, J. M. (2008). Does fish ecology predict dispersal across a river drainage divide? *Evolution*, 62, 1484-1499.

Cabrera, A. L. & Willink, A. (1980). *Biogeografia de America Latina*. Washington DC: OEA.

Candela, A. M., Bonini, R. A. & Noriega, J. I. (2012). First continental vertebrates from the marine Parana Formation (Late Miocene, Mesopotamia, Argentina): Chronology, biogeography, and paleoenvironments. *Geobios*, 45, 515-526.

Casciotta, J., Almirón, A., Cione, A. & Azpelicueta., M. (1999). Brazilian freshwater fish assemblages from southern Pampean area, Argentina." *Biogeographica*, 75, 67-78.

Doyle, J. J. & Doyle, J. L. (1987). A rapid isolation procedure for small quantities of fresh leaf tissue. *Phitochemical Bulletin*, 9, 11-15.

Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9, 772.

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

Drummond, A. J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T. & Wilson, A. (2007). Geneious v4.8. Available via: http://www.geneious.com.

Excoffier, L., Smouse, P. E., Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes – Application to human mitochondrial-DNA restriction data. *Genetics*, 131, 479-491.

Excoffier, L. & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564-567.

Felappi, J. F., Vieira, R. C., Fagundes, N. J. & Verrastro, L. V. (2015). So Far Away, Yet So Close: Strong Genetic Structure in *Homonota uruguayensis* (Squamata, Phyllodactylidae), a Species with Restricted Geographic Distribution in the Brazilian and Uruguayan Pampas. *PloS one*, 10, 1-19.

Freitas, T. R. O., Fernandes, F. A., Fornel, R. & Roratto, P. A. (2012). An endemic new species of tuco-tuco, genus *Ctenomys* (Rodentia: Ctenomyidae), with a restricted geographic distribution. *Southern Brazil Journal of Mammalogy*, 93, 1355-1367.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147, 915-925.

Hedrick , P. W. (2011) Genetics of Populations. Sudbury: Jones & Bartlett Publishers.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907-913.

Howard, J. H. & Morgan, R. P. (1993). Allozyme variation in the mottled sculpin (Cottus bairdi): a test of stream capture hypotheses. *Copeia*, 3, 870-875.

Hrbek, T. & Meyer, A. (2003). Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *Journal of Evolutionary Biology*, 16, 17-36.

Hubert, N. & Renno, F. (2006). Historical biogeography of South American freshwater fishes. *Journal of Biogeography*, 8, 1414-1436.

Hurwood, D. A. & Hughes, J. M. (1998). Phylogeography of the freshwater fish, Mogurnda adspersa, in streams of northeastern Queensland, Australia: evidence for altered drainage patterns. *Molecular Ecology*, 7, 1507-1517.

Iriondo, M. (1996). Estratigrafía del Cuaternario de la cuenca del río Uruguay. In *Actas del XIII Congreso Geológico Argentino y III Congreso de Exploración de Hidrocarburos* (pp. 15-25). Buenos Aires: Asociación Geológica Argentina.

Iriondo, M. H. (1999). Climatic changes in the South American plains: records of a continent-scale oscillation. *Quaternary International*, 57-58, 93-112.

Jones, C. P., & Johnson, J. B. (2009). Phylogeography of the livebearer Xenophallus umbratilis (Teleostei: Poeciliidae): glacial cycles and sea level change predict diversification of a freshwater tropical fish. *Molecular ecology*, 18, 1640-1653.

Ketmaier, V., Bianco, P. G., Cobolli, M., Krivokapic, M., Caniglia, R. & Matthaeis, E. (2004). Molecular phylogeny of two lineages of Leuciscinae cyprinids (Telestes and Scardinius) from the peri-Mediterranean area based on cytochrome b data. *Molecular Phylogenetics and Evolution*, 32, 1061-1071.

Kuhner, M. K. (2006). LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters." *Bioinformatics*, 6, 768-770.

Lawler, J. J., Aukema, J. E., Grant, J. B., Halpern, B. S., Kareiva, P., Nelson, C. R., Ohleth, k., Olden, J. D., Schlaepfer, M. A., Silliman, B. R. & Zaradic, P. (2006). Conservation science: a 20-year report card. *Frontiers in Ecology Environment*, 4, 473-480.

Li, C., Ortí, G., Zhang, G., & Lu, G. (2007). A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evolutionary Biology*, 7, 1-11.

Loureiro, M., Duarte, A. & Zarucki, M. (2011). A new species of Austrolebias Costa (Cyprinodontiformes: Rivulidae) from northeastern Uruguay, with comments on distribution patterns. *Neotropical Ichthyology*, 9, 335-342.

Lucinda, P. H. (2005). Systematics of the genus Cnesterodon Garman, 1895 (Cyprinodontiformes: Poeciliidae: Poeciliinae). *Neotropical Ichthyology*, 3, 259-270.

Lundberg, J.G., Marshall, L.G., Guerrero, J., Horton, B., Malabarba, M.C.S.L. & Wesselingh, F. (1998) The stage for Neotropical fish diversification: a history of tropical South American rivers. In L. R. Malabarba, R. E. Reis, R. P. Vari, Z. M. S. Lucena & C. A. S. Lucena (Eds) *Phylogeny and Classification of Neotropical Fishes* (pp. 13–48). Porto Alegre: Edipucrs.

Maack, R. (1968). *Geografia física do Estado do Paraná*. Curitiba: Banco de desenvolvimento do Paraná.

Maltchik, L., Lanés, L. E. K., Keppeler, F. W., Rolon, A. S., & Stenert, C. (2014). Landscape and habitat characteristics associated with fish occurrence and richness in southern Brazil palustrine wetland systems. *Environmental biology of fishes*, 97, 297-308.

Martin, L. & Dominguez, J. M. L. (1994). Geological history of coastal lagoons. *Elsevier Oceanography Series*, 60, 41-68.

McGlashan, D. J., & Hughes, J. M. (2000). Reconciling patterns of genetic variation with stream structure, earth history and biology in the Australian freshwater fish Craterocephalus stercusmuscarum (Atherinidae). *Molecular Ecology*, 9, 1737-1751.

Miller, M. (2005). Alleles in space: computer software for the joint analysis of interindividual spatial and genetic information. *Journal of Heredity*, 96, 722-724.

Ortiz-Jaureguizar, E. (1998). Paleoecología y evolución de la fauna de mamíferos de América del Sur durante la Edad de las Planicies Australes (Mioceno superior-Plioceno superior). *Estudios Geológicos*, 54, 161-169.

Overbeck, G. E., Muller, S. C., Fidelis, A., Pfadenhauer, J. Pillar, V. D., Blanco, C. C., Boldrini, I. I., Both, R., Forneck, E. D. (2007). Brazil's neglected biome: the South Brazilian campos. *Perspectives in Plant Ecology, Evolution and Systematics*, 9, 101-116.

Pallarés, O. R., Beretta, E. J. & Maraschin, G. E. (2005). The south american campos ecosystem. In J. Suttie, S. G. Reynolds & C. Batello (Eds) *Grasslands of the world* (pp. 171-219). Rome: Food and agriculture organization of the United Nations.

Panario, D. & Gutiérrez, O. (1999). The continental Uruguayan Cenozoic: an overview. *Quaternary International*, 62, 75-84.

Quattrocchio, M. E., Borromei, A. M., Deschamps, C. M., Grill, S. C. & Zavala, C. A. (2008). Landscape evolution and climate changes in the Late Pleistocene–Holocene, southern Pampa (Argentina): evidence from palynology, mammals and sedimentology. *Quaternary International*, 181, 123-138.

Reznick, D. N., & Bryga, H. (1987). Life-history evolution in guppies (*Poecilia reticulata*): 1. Phenotypic and genetic changes in an introduction experiment. *Evolution*, 41, 1370-1385.

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

Rozas J, Sánchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, 19, 2496-2497.

Rosen, D. E. (1978). Vicariant patterns and historical explanation in biogeography. *Systematic Biology*, 27, 159-188.

Saadi, A., Machette, M. N., Haller, K. M., Dart, R. L., Bradley, L. & de Souza, A. M. P. D. (2002). Map and database of Quaternary faults and lineaments in Brazil. *U.S. Geological Survey Open-File Report 02-230*, 1-59.

Stephens, M., Smith, N. J. & Donnelly P. (2001). A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics*, 68, 978-989.

Sorenson, M. D. (2003, May) Avian mtDNA primers. Avaiable via http://people.bu.edu/msoren/ (Bird.mt.Primers.pdf.)

Speranza, P. R., Seijo, J. G., Grela, I. A. & Neffa, V. G. S. (2007). Chloroplast DNA variation in the Turnera sidoides L. complex (Turneraceae): biogeographical implications. *Journal of Biogeography*, 34, 427-436.

Swartz, E. R., Skelton, P. H. & Bloomer, P. (2007). Sea-level changes, river capture and the evolution of populations of the Eastern Cape and fiery redfins (*Pseudobarbus afer* and *Pseudobarbus phlegethon*, Cyprinidae) across multiple river systems in South Africa. *Journal of Biogeography*, 34, 2086-2099.

Swartz, E. R., Skelton, P. H. & Bloomer, P. (2009). Phylogeny and biogeography of the genus *Pseudobarbus* (Cyprinidae): shedding light on the drainage history of rivers

associated with the Cape Floristic Region. *Molecular Phylogenetics and Evolution*, 51, 75–84.

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, 105, 437–460.

Thomaz, A. T., Malabarba, L. R., Bonatto, S. L. & Knowles, L. L. (2015). Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: study of a Neotropical fish of the Brazilian coastal Atlantic Forest. *Journal of Biogeography*, *in press*, 1-13.

Tonni, E. P. & Cione, A. L. (1997). Did the Argentine Pampean ecosystem exist in the Pleistocene. *Current research in the Pleistocene*, 14, 131-133.

Tonni, E. P., Cione, A. L. & Figini, A. J. (1999). Predominance of arid climates indicated by mammals in the pampas of Argentina during the Late Pleistocene and Holocene. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 147, 257-281

Trenti, P. S., Gómez, S. E. & Ferriz, R. A. (1999) Capacidad de natación en tres peces pampásicos. *Aprona*, 38, 2-9.

Turchetto, C., Fagundes, N. J. R., Segatto, A. L. A., Kuhlemeier, C., Neffa, V. G. S., Speranza, P. R., Bonatto, S. L. & Freitas. L. B. (2014). Diversification in the South American Pampas: the genetic and morphological variation of the widespread Petunia axillaris complex (Solanaceae). *Molecular ecology*, 23, 374-389.

Turchetto-Zolet, A. C., Pinheiro, F., Salgueiro, F. & Palma-Silva, C. (2013). Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology*, 22, 1193-1213.

Vari, R. P. (1988) The Curimatidae, a lowland Noetropical fish family (Pisces: Characiformes); distribution, endemism, and phylogenetic biogeography. In P. E. Vanzolini & W. R. Heyer (Eds), Proceedings of a Workshop on Neotropical Distribution Patterns (pp. 343-377). Rio de Janeiro, Brazil: Academia Brasileira de Ciências.

Villwock, J. A. & Tomazelli, L. J. (1995). Geologia costeira do Rio Grande do Sul. *Notas Técnicas*, 8, 1-45.

Waters, J. M., Craw, D., Youngson, J. H. & Wallis, G. P. (2001). Genes meet geology: fish phylogeographic pattern reflects ancient, rather than modern, drainage connections. *Evolution*, 55, 1844-1851.

Waters, J. M., Allibone, R. M. & Wallis, G. P. (2006). Geological subsidence, river capture, and cladogenesis of galaxiid fish lineages in central New Zealand. *Biological Journal of the Linnean Society*, 88, 367-376.

Zemlak, T. S., Habit, E. M., Walde, S. J., Battini, M. A., Adams, E. D., & Ruzzante, D. E. (2008). Across the southern Andes on fin: glacial refugia, drainage reversals and a secondary contact zone revealed by the phylogeographical signal of Galaxias platei in Patagonia. *Molecular Ecology*, 17, 5049-5061.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Additional figure (S1).

Table 1. Voucher specimens, drainage and broad drainage systems for samples used in this study. All specimens belong to UFRGS zoological collection (Universidade Federal do Rio Grande do Sul).

Voucher	Drainage/ Population	Latitude	Longitude
UFRGS 13774/TEC 0001	Uruguay/URU	-30.6308	-57.6883
UFRGS 13783/TEC 0022 A, B, D, E, F	Uruguay/URU	-29.8597	-56.9369
UFRGS 13784/TEC 0023 B, C, D, E	Uruguay/URU	-30.2119	-55.0550
UFRGS 13811/TEC 0087 A, B, C, D, E	Uruguay/URU	-32.0527	-57.6719
UFRGS 13812/TEC 0095 A, C, D, E	Uruguay/URU	-31.6222	-57.8808
UFRGS 13912/TEC 0138 A, D, F	Uruguay/URU	-30.5319	-57.6658
UFRGS 13927/TEC 0175 A, B	Negro River/NEG	-33.4152	-56.2016
UFRGS 13933/TEC 0184 A, B, C, D	Negro River/NEG	-31.9083	-56.0177
UFRGS 14531/TEC 0328 A, B, C, D	Negro River/NEG	-31.3905	-55.2538
UFRGS 14538/TEC 0367 A, B, C, D, E	Uruguay/URU	-31.3858	-57.5627
UFRGS 14549/TEC 0395 A, B, E	Uruguay/URU	-31.4761	-57.9016
UFRGS 14550/TEC 0396 A, C, D, E, F	Uruguay/URU	-31.2755	-57.1561
UFRGS 14552/TEC 0406 A	Uruguay/URU	-30.4377	-57.2961
UFRGS 14561/TEC 0425 A, B, C, D, E	Uruguay/URU	-32.2069	-57.2130
UFRGS 14566/TEC 0436 A,B	Uruguay/URU	-31.7346	-57.8669
UFRGS 14571/TEC 0456 A, B, G, H, I	Uruguay/URU	-30.6586	-56.6741
UFRGS 14576/TEC 0466 A, B, C, D	Mirim Lagoon/MIR	-32.3661	-54.1997
UFRGS 14590/TEC 0496 A	Uruguay/URU	-30.6308	-57.6883
UFRGS 12372/TEC 0583 G, M, R, S, U	Uruguay/URU	-31.3225	-57.2825
UFRGS 14637/TEC 0593 B	Negro River/NEG	-31.9758	-55.4702
UFRGS 14640/TEC 0601 A, C, G, H, M, O	Uruguay/URU	-30.4688	-57.5122
UFRGS 17842/TEC 3520 A, B, C, D	Patos/MIR	-31.9497	-51.9600

UFRGS 17843/TEC 3521	Rocha Lagoon/SOU	-34.5133	-54.2956
UFRGS 17845/TEC 3523	Atlantic Ocean/SOU	-34.2061	-53.7766
UFRGS 17846/TEC 3524	Atlantic Ocean/SOU	-33.9204	-53.5420
UFRGS 17852/TEC 3530 B	Mirim Lagoon/MIR	-33.8295	-54.7662
UFRGS 17854/TEC 3532 A	La Plata River/SOU	-34.4228	-57.0112
UFRGS 17856/TEC 3534	Santa Lucia River/SOU	-34.2812	-55.2784
UFRGS 17859/TEC 3537	Uruguay/URU	-31.0354	-56.8972
UFRGS 17861/TEC 3539	Mirim Lagoon/MIR	-32.5190	-53.4698
UFRGS 17863/TEC 3541 A, B	Mirim Lagoon/MIR	-33.5734	-54.4987
UFRGS 17865/TEC 3543	Santa Lucia River/SOU	-34.0989	-56.2032
UFRGS 17866/TEC 3544 A B	Santa Lucia River/SOU	-34.5349	-56.5765
UFRGS 17869/TEC 3547	Santa Lucia River/SOU	-34.0118	-56.9440
UFRGS 17870/TEC 3548 A, B	La Plata River/SOU	-34.8415	-55.0960
UFRGS 17873/TEC 3551 B	Mirim Lagoon/MIR	-33.6136	-54.3295
UFRGS 17874/TEC 3552	La Plata River/URU	-33.9394	-58.3663
UFRGS 18003/TEC 3675 A, B, C, D	Patos/MIR	-33.4996	-53.4310

Codes after vouchers numbers are individual identification.

Table 2. Diversity indexes obtained for the mitochondrial and nuclear markers respectively (ND2 values/Myh6 values) for the four populations considered in the study.

Population	**N**	***H***	**π**	**Fu's Fs**	**Tajima's D**	**Fu's Fs‡**	**Tajima's D‡**
URU	62/60	0.89/0.74	0.006/0.001	0.17/-9.36**	-0.48/-1.40	-1.90/-10.17**	-1.43/-1.49*
NEG	11/9	0.60/0.46	0.004/0.001	2.93/0.93	0.78/-0.62	-0.18/0.55	-1.05/0.15
MIR	15/14	0.86/0.71	0.005/0.002	2.27/-1.67	1.06/-0.29	0.41/-1.94	0.46/-0.71
SOU	11/11	0.96/0.79	0.003/0.002	-4.39**/-5.24**	-0.63/-1.13	-	-
All samples	99/94	0.94/0.79	0.007/0.002	-6.40/-16.76**	-0.75/-1.46*	-	-

N – number of individuals; *H* – haplotype diversity; π – nucleotide diversity; *$P \leq 0.05$; **$P \leq 0.02$; ‡ Excluding river capture localities.

Table 3. Pairwise Φ_{ST} values for mtDNA and nDNA data (ND2/Myh6).

Population	**1-URU**	**2- NEG**	**3-MIR**	**4-SOU**
1-URU	0.00/0.00	-	-	-
2-NEG	0.32*/0.58*	0.00/0.00	-	-
3-MIR	0.19*/0.10*	0.32*/0.36*	0.00/0.00	-
4-SOU	0.17*/0.04*	0.57*/0.55*	0.11*/0.03	0.00/0.00

*P<0.05

Table 4. Analysis of molecular variance (AMOVA) for mtDNA (ND2) and nDNA (Myh6) data.

	ND2		Myh	
Source of Variation	d.f.	Variation (%)	d.f.	Variation (%)
Among populations	3	25.39	3	29.54
Within populations	95	74.61	184	70.46
Φ_{ST}		0.25*		0.29*

d.f. – Degrees of freedom; *P<0.05;

Table 5. Estimates of Effective Population Size (N_E).

Population	**Theta (Θ)**	**CI 95%**	**Mean N_E**	**CI 95%**
URU	18.05×10^{-3}	15.44×10^{-3} - 21.28×10^{-3}	1.38×10^{6}	1.18×10^{6} - 1.63×10^{6}
NEG	2.38×10^{-3}	1.60×10^{-3} - 3.74×10^{-3}	1.82×10^{5}	1.22×10^{5} - 2.86×10^{5}
MIR	7.66×10^{-3}	5.18×10^{-3} - 11.53×10^{-3}	5.88×10^{5}	3.96×10^{5} - 8.81×10^{5}
SOU	23.57×10^{-3}	16.71×10^{-3} - 34.47×10^{-3}	1.80×10^{6}	1.28×10^{6} - 2.63×10^{6}

95% CI – 95% Credible intervals for the estimates.

Figure 1 Geographical distribution of sampled sites showing the limits of the major drainage systems (populations). For more details about localities, see Table 1.

Figure 2 Median-joining networks for (A) mtDNA ND2 haplotypes, and (B) nDNA Myh6 haplotypes. Circles size are proportional to the observed frequency of each haplotypes. Cross marks represent mutational differences between haplotypes.

Figure 3 Maximum clade credibility tree among ND2 lineages found in *Cnesterodon decemmaculatus* individuals. Posterior probabilities (PP) are shown above the branches for selected nodes with PP>0.8; Pie charts beside selected ancestral node represents the PP of ancestral location among each of the four populations according to the color scheme shown in the inlet.

Figure 4 Maximum clade credibility tree considering ND2 and Myh6 under a coalescent approach. The thickness of the branches are proportional to effective population size (N_E) estimates for each geographical group (see Table 5 for more details)**.**

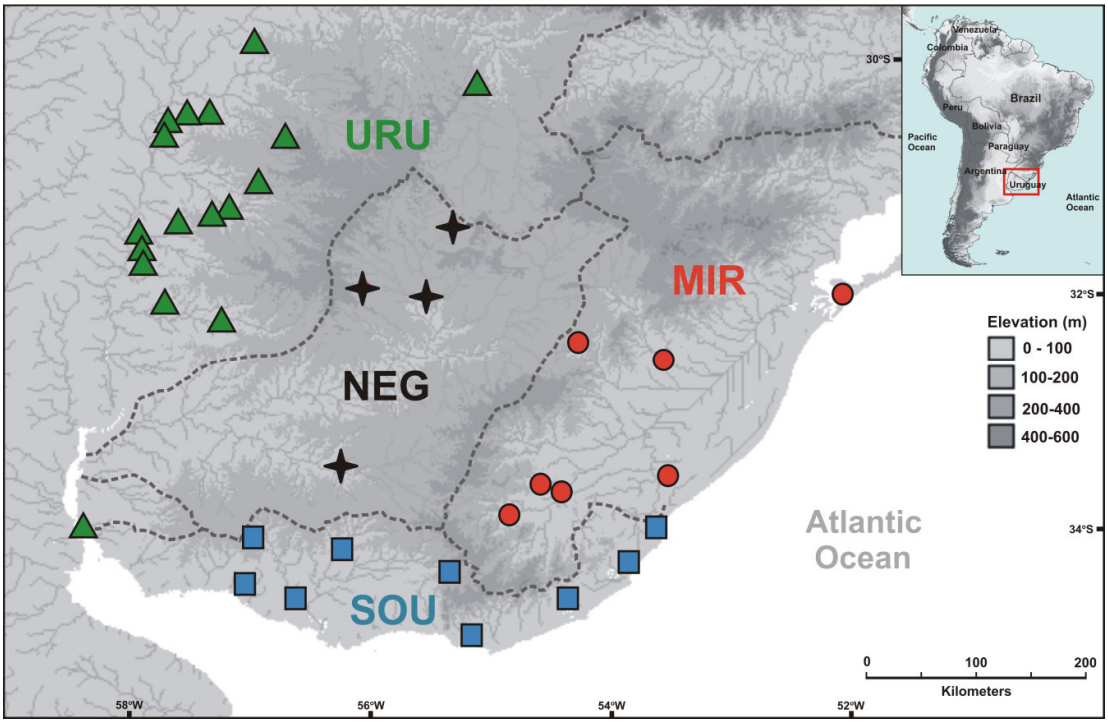


Figure 1

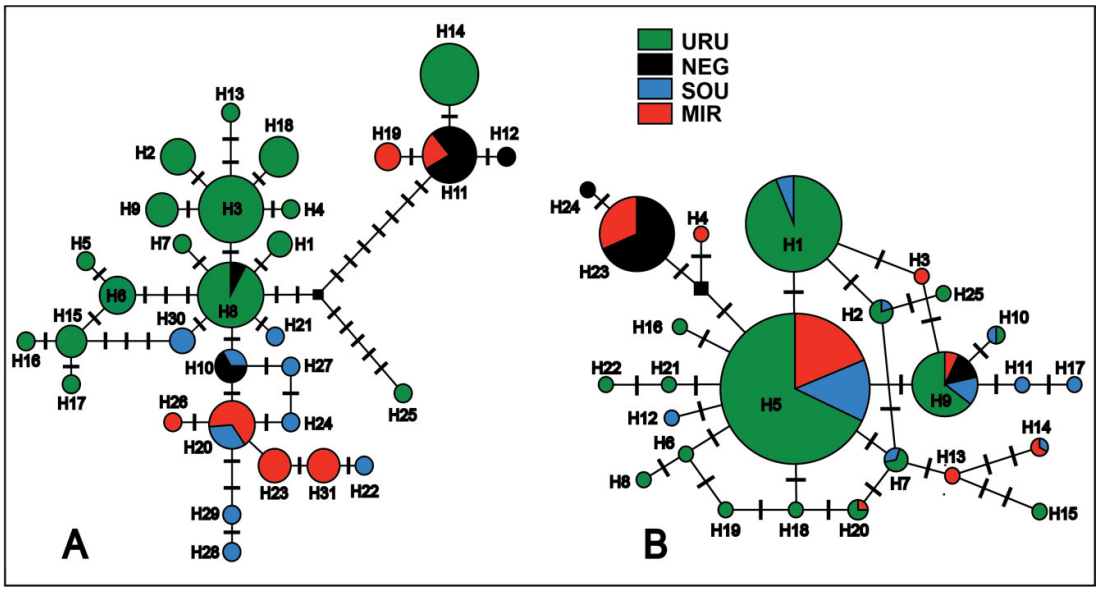


Figure 2

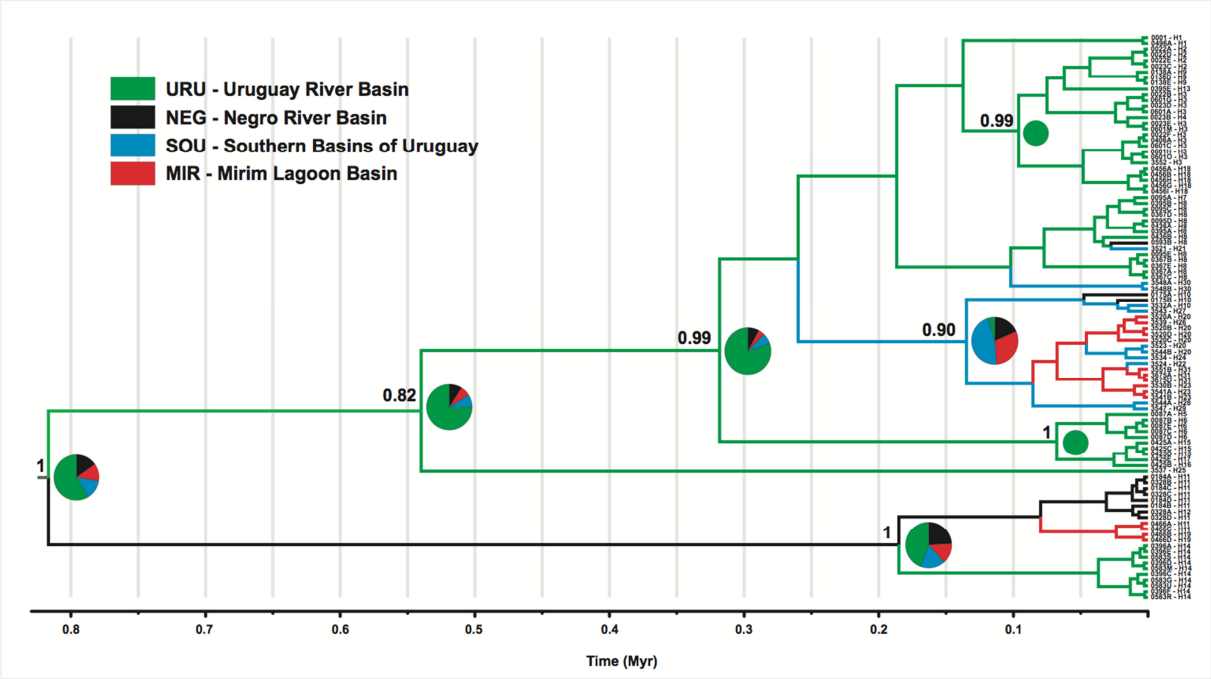


Figure 3

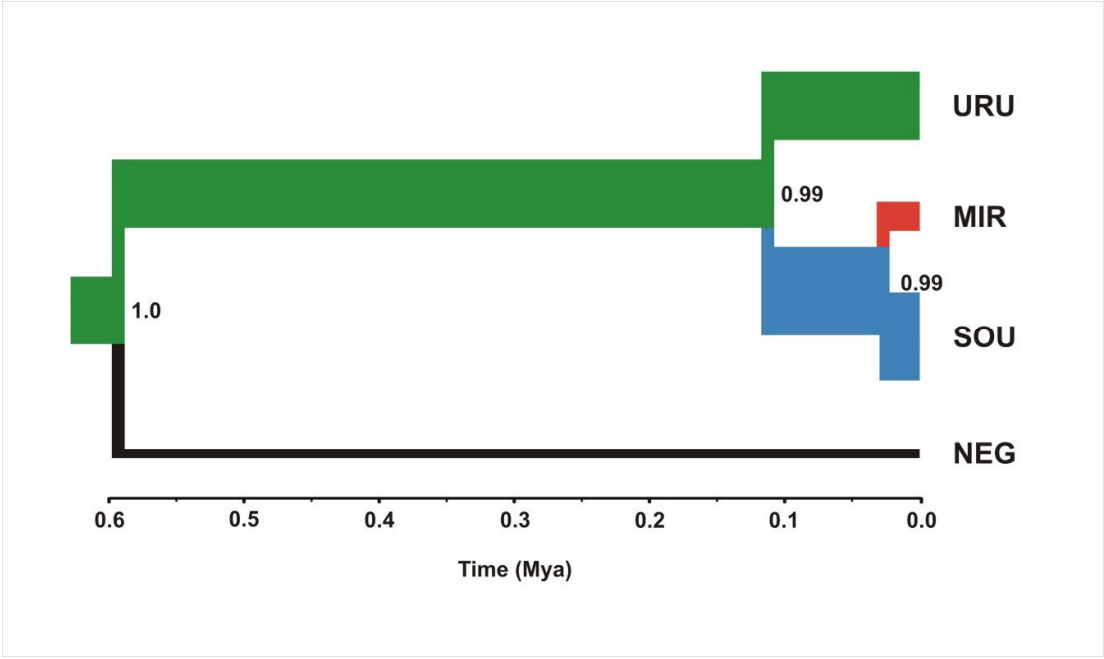


Figure 4

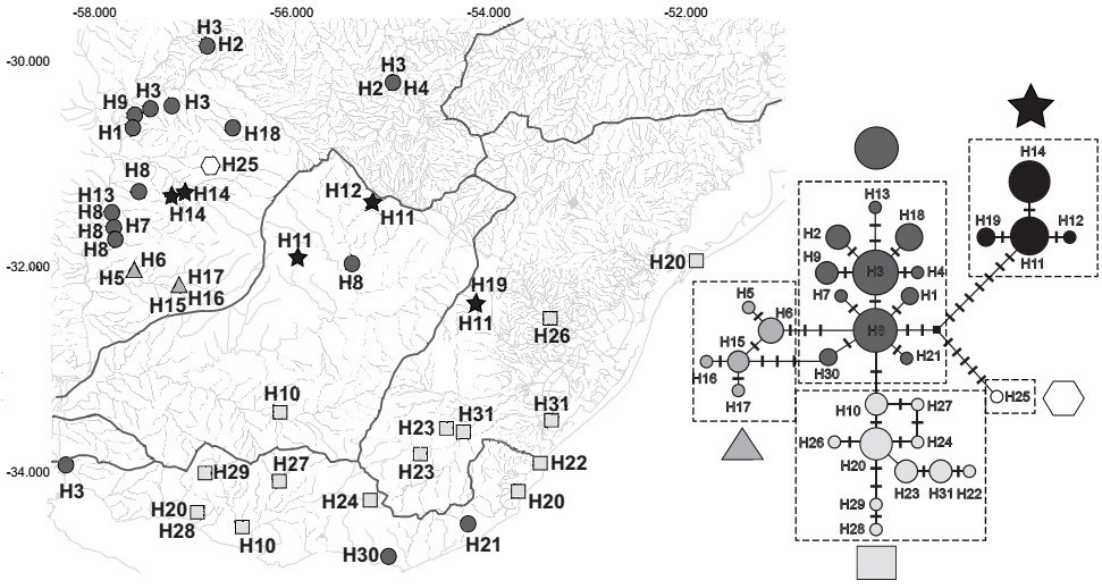


Figure S1. MtDNA haplotype distribution along sampled sites.

CAPÍTULO 3

Unveiling the cryptic diversity of freshwater fishes in Southern South America by the use of RAD-seq sequencing: the case of *Cnesteorodon* species

Manuscrito a ser submetido para a revista *Molecular Ecology*

Unveiling the cryptic diversity of freshwater fishes in Southern Brazil by the use of RAD-seq sequencing: the case of *Cnesteorodon* species

Aline Ramos-Fregonezi,[1] Moisés A. Bernal, [3,4] Luiz R. Malabarba, [2] Luiz A Rocha, [4] Nelson Fagundes [1*]

[1]Departament of Genetics, Institute of Biosciences, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil.
[2]Departament of Zoology, Institute of Biosciences, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil.
[3]University of Texas Marine Science Institute, 750 Channel View Drive, Port Aransas, TX 78373, USA.
[4]Section of Ichthyology, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, CA 94118, USA.

Key-words: Freshwater fish, *RAD-seq*, species-complex

*Corresponding Author
Nelson J.R. Fagundes,
Departament of Genetics – UFRGS
P.O. Box 15053
91501-970 Porto Alegre, RS
Brazil
nelson.fagundes@ufrgs.br

Abstract

Species concepts has been one of the most controversial issues among biologists in the last century. Different evolutionary processes such as geographical isolation or sexual selection may culminate on reproductive isolation among populations and the origin of independent evolutionary lineages. Another related problem is the identification of independent evolutionary lineages, which is a fundamental first step to identify the drivers of speciation. The study of sympatric species-complexes might provide new insights into speciation mechanisms in terms of space, tempo and mode of speciation. Here, we used a genomic dataset based on RAD-seq method coupled with mitochondrial DNA sequence data to identify evolutionary lineages in a sympatric species-complex from the South American freshwater fish poeciliid genus *Cnesterodon*. In this genus, males display a copulatory organ (gonopodium), whose morphological variation within this species-complex may be a key factor for evolutionary isolation. Our results showed that there are four different genetic lineages in this system, with two allopatric lineages associated with long gonopodium and two sympatric lineages associated with short gonopodium. Divergence among lineages dates back to ~3.48 (2.90 - 4.14) million years ago. Even though the gonopodium morphology have certainly played an important role in reproductive isolation, we found at least one case of mtDNA introgression, suggesting that such isolation among different lineages is not complete, and that rare events of secondary sexual contact may occur. The two lineages having long gonopodium are allopatric, suggesting that this divergence may have occurred by factors unrelated to natural/sexual selection operating on gonopodium morphology. On the other hand, the strong genetic isolation between the two sympatric lineages having short gonopodium suggests that there are other isolation mechanisms maintaining reproductive isolation beyond the short/long gonopodium dichotomy.

Introduction

What constitutes a species remains a controversial issue in evolutionary biology and has provided heated discussion among biologists in the last 100 years (e.g. Darwin 1859; Dobzhansky 1937; Mayr 1942; Sneath & Sokal 1973; Donoghue 1985; Coyne *et al.* 1988; Cracraft *et al.* 1989; Templeton 1989; Hey 2001; de Queiroz 2007, 2011). From the evolutionary point of view, a species can be broadly defined as separately evolving metapopulation lineages (de Queiroz 2007), a concept that ultimately traces back to Darwin's view (de Queiroz 2011). In this sense, most of the controversy between different species concepts lies at emphasizing different properties of the speciation process as diagnostic to define, or circumscribe, individuals to a given species (Hey 2001; de Queiroz 2007. For example, the speciation process occurs when populations have limited interbreeding such that they start to follow independent evolutionary paths. Sexual isolation can result from geographic isolation alone (Coyne & Orr 2004), or from phenotypic differences among populations that arise via direct selection on mate preferences (Servedio 2001), or, alternatively, via adaptation to different ecological environments (Funk 1998; Jiggins *et al.* 2001; Rundle & Nosil 2005; Funk *et al.* 2006; Nosil *et al.* 2007). On its turn, these processes may lead to diagnosable phenotypic differences among populations or the structuring of these population in identifiable clades (Cracraft 1983). Many different reproductive barriers may play a role in the evolution of reproductive isolation such as gametic incompatibility, habitat isolation, mechanical isolation, hybrid sterility, etc. Some of these barriers have been shown to evolve in sympatry, but is not easy determine which of these were determinant for speciation (Coyne & Orr 2004).

The idea that species are evolutionary lineages but that we need some diagnostic properties (criteria) implies that many animal species (i.e. evolutionary lineages) may

remain undescribed because they are too similar to be separated based on morphology by traditional taxonomy. Modern molecular genetic methods have enable biologists to distinguish between morphologically similar species that are genetically distinct from each other (Avise 2004). In fishes, evolutionary and taxonomic studies have contributed for understanding the nature of species and speciation (Meyer *et al.*, 1990; Vrijenhoeck, 1994; Schluter, 1996; Seehausen & Wagner 2014; Martinez-Takeshita *et al.* 2015; Smith *et al.* 2015). Importantly, the occurrence of sympatric species-complex has led to the concern that biodiversity may be underestimated (and perhaps undervalued in conservation programs) because different members of a complex will share the same scientific name (Bickford *et al.* 2007). The study of sympatric species complexes might provide new insights into the mechanisms of speciation particularly in terms of space, tempo and mode of speciation.

Under the model of ecological speciation, divergent natural selection among environments might lead to speciation. Adaptation to different selective regimes results in reproductive isolation that can occur in different geographical contexts and does not require selection directly favoring reproductive isolation (van Valen 1976; Schluter 2001; Nosil *et al.* 2003; Rundle & Nosil 2005; Vines & Schluter 2006). This theory is well accepted in a context of allopatric populations, in which interbreeding opportunities are rare. Thus, natural selection cannot act directly on reproductive isolation, but it has an incidental role in speciation, as demonstrated in some studies (Rundle *et al.* 2000; Jiggins *et al.* 2001; Boughman *et al.* 2005). However, in zones of sympatry, prezygotic isolation barriers evolve as a response of selection against hybridization, a process also known as reinforcement (Howard 1993a). A potential outcome of reinforcement is character displacement, in which there is greater divergence among sympatric rather than in allopatric populations stemming from the selection favoring less competition and/or

reproductive interactions (Howard 1993a; Howard 1993b; Pfenning & Pfenning 2009). Therefore, character displacement drives divergence among conspecific populations and both initiate and finalize the process of speciation (Pfenning & Pfenning, 2009).

Character displacement has been demonstrated as a usual mechanism driving diversification in poeciliid fishes possibly because internal fertilization, which is typical in this group, provides many opportunities for sexual and natural selection to occur, since males inseminate females using a non-retractable, modified anal fin called a gonopodium and therefore mate selection and anatomic features in the copulatory organs of both sexes may be determinant of reproductive success (Ryan *et al.* 1996; Poeser 1998; Espinedo *et al*. 2010; Scott & Johnson 2010; Langerhans 2011; Swenton 2011). In the South-Brazilian Highlands (SBH), between 700-1,850m above sea level, two sympatric species have been recognized for the genus *Cnesterodon*. This genus includes small freshwater fishes, which inhabit shallow environments surrounded by grasslands in SBH, in the Pampa Biome, with a single species being described for Central Brazil (Lucinda 2005). In SBH, a species complex of the genus can be found formed by *Cnesterodon brevirostratus* and at least one undescribed *taxon* (called *Cnesterodon* sp. nov. B in Lucinda 2005). Taxonomic identification of *Cnesterodon* individuals found in this region is unclear, even though adult males show a variation in size and in the morphology of the gonopodium: typically, *Cnesterodon* sp. nov. B shows a longer gonopodium compared to *C. brevirostratus.* Morphological variation in male gonopodium is well-known for Poeciliidae at both inter- and intraspecific level (Rosen & Bailey 1963; Ptacek & Travis 1998; Greven 2005; Evans *et al.* 2011). In spite of these differences, there are no sufficient morphological differences for unambiguous distinction between *C. brevirostratus* and *Cnesterodon* sp. nov. B, and this is further complicated in male juveniles and females.

In this study, we generated mtDNA and RAD-seq data for individuals belonging to both species from SBH to allow a high-resolution assessment of population structure among these individuals. Based on this, our goals were: (1) to find out how many genetic lineages occur in SBH; (2) to understand how these evolutionary lineages are geographically distributed; and (3) to assess the role of different gonopodium morphologies in generating the observed genetic structure.

Material and Methods

Sample and DNA extraction

Whole genomic DNA was extracted with the DNeasy Tissue Kit (Quiagen) from frozen muscle tissues of specimens representing two putative taxonomic entities: *Cnesterodon* sp B and *C. brevirostratus*. All individuals were collected from natural populations in Southern South America (Figure 1), fixed in ethanol 96% and stored at the scientific collection UFRGS (Universidade Federal do Rio Grande do Sul) Details on vouchers and collection locals are provided in Table S1. All collections were performed under the approval of the government authorities (Ministério do Meio Ambiente, Brazil – SISBIO 12038-2).

RAD-Seq Data

RAD library construction was carried out according to the double digest protocol described by Peterson *et al.* (2012). Genomic DNA was digested using Sph1 and MluCl restriction enzymes. Resulting fragments were ligated to P1 (individual barcoded adapter) and P2 (index that will be read off in a separate multiplexing read per the standard Illumina multiplexed paired-end sequencing protocol) onto the ends of digested DNA. Samples were pooled with the same index and size-selected 300 bp fragments by gel excision or by

Pippin Prep (Sage Science, Beverly, MA). We then amplified size-selected fragments by Real Time PCR for 09 cycles (at 98 °C for 15s, 65 °C for 30 s and 72 °C for 30s with an initial denaturation step at 98°C for 45s and a final extension step at 72°C for 1min) in order to increase libraries concentration. Reactions were pooled together and final concentrations were measured on an Agilent BioAnalyzer 2100 (Agilent Technologies) and a Qubit Fluorometer (Invitrogen).

Libraries were sequenced on an Ilumina HiSeq 2000 (Vincent J. Coats Genomics Sequencing Laboratory of UC Berkeley) generating single end reads of 95 bp. Loci assembly and quality filtering were conducted in STACKS (Catchen *et al.* 2011). Samples were de-multiplexed and reads with low quality (Phred score <30) were discarded. Reads per individual ranged from 0.7 to 1 million. Loci were identified by aligning the sequence reads from each individual to the assembled contigs resulting in 960158 'stacks'. A stack is a set of identical sequences in the terminology of this pipeline (Catchen *et al.* 2011); several of these stacks may then be merged to form putative loci. Single nucleotide polymorphisms (SNPs) were identified through restrictive filters: a deleveraging algorithm was used to identify and remove significantly highly repetitive stacks (that might represent sequencing errors and/or repetitive genomic regions); a 'populations' component was used to filter only those loci with ≥ 8x coverage ('m' command) and that aligned in ≥ 50% of individuals ('r' command) in each population ('p' command). Input files for other programs were produced from STACKS using the option 'write single snp'. Therefore, we retained only the first single nucleotide polymorphism (SNP) in each locus to avoid having two SNPs highly linked in the same RAD-locus. After filtering, our data set consisted of 6,840 loci. Data file conversion for programs was performed using PGDSPIDER version 2.0 (Lischer & Excoffier 2012). Because STACKS emphasizes analysis on a population

level-variation, we separated our samples in 4 categories (i.e. populations): females, male juveniles, males with large gonopodium and males with short gonopodium.

We used the Bayesian analysis software STRUCTURE 2.3.4 to determine the number of genetic groups (Pritchard *et al.* 2000). Groups (*K*) were tested from 1 to 7, each replicated 05 times, to determine the maximum value of the posterior likelihood [lnP(D)] and the best number of *K*. Each run was performed using $5x10^4$ burn-in periods and $5x10^5$ Markov chain Monte Carlo (MCMC) repetitions after burn-in. We determined the most likely number of populations (*K*) using the ΔK method (Evanno *et al.* 2005) implemented in STRUCTURE HARVESTER 0.6.93 (Earl & von Holdt 2012). The results of the best *K* value were summarized with the CLUMPP 1.1.2 program based on the average pairwise similarity of individual assignments across runs using Greedy's method and the statistic G′ (Jakobsson & Rosenberg 2007). The final output was visualized using the software DISTRUCT (Rosenberg 2004). Since principal coordinates analysis (PCA) has similar power to detect population structure as STRUCTURE (Patterson *et al.* 2006), we performed it as an alternative method using the software package GENODIVE (Meirmans & Van Tienderen 2004). Maximum Likelihood (ML) analyses were conducted with RAxML 1.31 (Silvestro & Michalak 2011). We used the GTR-GAMMA model, which includes a parameter for site heterogeneity, and 1000 rapid bootstrap replicates to estimate clade confidence. We used a matrix containing all SNPs per sample and IUPAC ambiguity base to code for heterozygous SNPs.

Loci with high levels of genetic differentiation among populations were identified using the F_{ST} outlier detection approach implemented in BayeScan v.2.1 (Foll & Gaggiotti 2008). Two models were tested: the first tested outliers among genetics lineages detected in clustering analyses, while in the second we excluded females and juveniles males in attempt to identify locus under selection related to the length of the gonopodium (short or

long). Prior odds of the neutral model were initially set in 10. This prior for two models were further tested by changing it to different values, but this did not influence the results. Results were viewed in R version 3.0.0 (R Development Core Team 2013) using a false discovery rate (FDR) of 0.05 and markers with evidence of selection were subjected to a BLAST search via NCBI.

Mitochondrial Sequence Data

We amplified and sequenced two mitochondrial regions (cytochrome oxidase subunit I - COI) and NADH dehydrogenase 2 - ND2) from 56 individuals using primers and amplification protocols described by Herbert *et al*. (2003) and Sorenson (2003), respectively. PCR products were checked on a 1% agarose gel, purified with Exonuclease I e Shrimp Alcaline Phosphatase (GE Healthcare®) and sequenced using the Sanger method in Macrogen Inc. (Seoul, South Korea). Nucleotide sequences were determined on both strands, checked and aligned with GeneiousPro v. 4.8 (Drummond *et al.* 2007). All sequences obtained in the present study were deposited in GenBank (XXXXXX-XXXXXX).

Both mitochondrial genes were concatenated and analyzed as one sequence in all analyses. Nucleotide (π) and haplotype diversity (H) estimates, as well as Tajima's D (Tajima 1983) and Fu's (Fu 1997) neutrality tests, were estimated using Arlequin 3.5.1.3 (Excoffier & Lischer 2010) for each lineage detected. The mitochondrial sequences were used to estimate a yule-tree prior and a species-tree approach in the software BEAST 1.7.5 (Drummond & Rambaut 2007). All Bayesian analyses were run using a strict molecular clock model, using a normally distributed rate of $8.6x10^{-9}\pm0.1x10^{-9}$ substitutions per site per year (Hrbek & Meyer 2003), and the TN93+I evolutionary model of sequences, suggested by Akaike Information Criterion in jModelTest2 (Darriba *et al.* 2012). Two independent runs of $5x10^{7}$ chains were performed for each analysis, sampling every $5x10^{3}$

generations. The software TRACER v1.4 (available at http://beast.bio.ed.ac.uk/Tracer) was used to check for convergence of the MCMC and adequate effective sample sizes (>200) after delete the first 10% of generations as burn-in. The species-tree approach was also performed under the same parameters.

Results

RAD-seq Data

Over 56 million high-quality RAD-Seq reads were generated, with an average of 41.5 million reads per pooled sample. We obtained 6840 RADSeq loci in 133 individuals across 31 locations (Figure 1) producing a robust data matrix that was 76% complete. Using the 'write_single_snp' command in STACKS (as described above), only the first single nucleotide polymorphism (SNP) in each locus was taken into account for the analyses.

Samples with different sizes of gonopodium were, in general, separated in four distinct clusters by STRUCTURE (figure 2) and PCA (Figure 3) analyses. By examining the change in *LnP(D)*, and using the ΔK approach of Evanno *et al*. (2005), we found that a model with K=4 best fits the data (Figure S1). Clustering results for other K-values are shown in Figure S2. PCA analysis indicated the same clusters evidenced by STRUCTURE. Axis one separated short and long gonopodium and axes two indicated that individuals with longer gonopodium can be attributed to two clusters, as well as individuals with short gonopodium (Figure 3).

Phylogenetic analyses consistently returned four major clades, with three of them (yellow, blue and red in Figure 4) having high support values, and the other (green in Figure 4) with lower support. These clades match the four genetic groups previously

identified by STRUCTURE and PCA. Most adult male individuals in the yellow clade had short gonopodium, although one individual had long gonopodium. The phylogenetic analysis suggests that this clade could be further divided in two subclades, which are strongly associated to the geographical distributions. One subclade is from the Tramandaí-Mampituba drainage system (yellow triangles in Figure 4), occurring exclusively inside an environmental protection area (Parque Nacional do Aparados de Serra – PNAS, Figure 4B), while the other is from the Laguna dos Patos drainage (yellow squares in Figure 4). On its turn, the blue clade is formed exclusively by individuals having short gonopodium, and can be divided in two subclades with a similar geographical distribution compared to the yellow clade. One subclade (blue squares in Figure 4) consists of populations from the Laguna dos Patos drainage and one population from the Tramandaí-Mampituba system, and the second subclade (blue triangles in Figure 4) consists mostly of populations from the Tramandaí-Mampituba drainage plus three populations from the Laguna dos Patos drainage. This subclade is found exclusively in PNAS. The red clade is formed exclusively by individuals having long gonopodium and occurs in Tramandaí-Mampituba and Laguna dos Patos drainages, being found entirely inside the PNAS (red circles in Figure 4). Finally, the clade with lower support value is also the most geographically widespread, occurring in the drainages mentioned above plus the Uruguay drainage, with a single population occurring inside PNAS (green star in Figure 4). Even though this clade is formed mostly by individuals with long gonopodium, there are a few locations that include individuals with short gonopodium.

We found expected heterozygosity levels ranging from 0.034 to 0.055 in the red and green lineages, respectively, with no clear relationship with geographic range. The same occurred for G_{IS} values, which ranged from 0.155 in the red lineage to 0.344 in the yellow lineage (Table 1). All lineages had significant genetic structure distinguishing

lineages (Table 2), with an average F_{ST} = 0.2885 among lineages. We did not detect any F_{ST}-outlier *locus* under divergent selection either among lineages or between gonopodium sizes. On the other hand, we detected over 30 *loci* under balancing selection (Figure S3).

Mitochondrial Sequence Data

The mitochondrial concatenated dataset resulted in 1631 bp for 56 individuals. Overall, nucleotide diversity within lineages was low (π = 0.002-0.032), while the corresponding haplotype diversity was high (*H* = 0.84-0.94; Table 3). Neutrality tests were not significant (Table 3). As for RAD-Seq data, we found significant genetic structure among lineages (Table 2), with an average F_{ST}=0.53 among lineages based on AMOVA. A time-calibrated mtDNA phylogeny suggests a time to most recent common ancestor (TMRCA) for the *Cnesterodon* species complex around 3.1 Ma (million years ago) (95% HPD = 2.7 - 3.7 Ma) (Figure 5), even though most of the early diverged mtDNA haplotypes have been retained only by the green lineage. The red lineage was the only one for which monophyly of its mtDNA haplotypes could be recovered. Nonetheless, for the yellow and blue lineages there was a single individual from the blue lineage carrying a mtDNA haplotype belonging to the yellow clade, as well as one individual from the yellow lineage having a mtDNA haplotype from the blue clade (Figure 5). Bayesian species-tree approach indicated a not clear phylogenetic structure, resulting in a polytomy. Estimated time for root tree suggests that radiation of those lineages was around 3.48 Ma (95% HPD = 2.90 - 4.14 Ma) (data no shown).

Discussion

Using 6,840 RAD-seq SNPs, we find clear evidence for at least four sympatric distinct genetic groups. This highlights the power of large SNP data sets to assign

individuals to genetic clusters in closely related or taxonomically difficult groups, as demonstrated by recent studies (Lindqvist *et al.* 2010; Edwards *et al.* 2011; Freamo *et al.* 2011; Catchen *et al.* 2013; Wagner *et al*. 2013). In general, the lineages identified in SBH had a strong association with gonopodium size. Furthermore, both mtDNA and RAD-seq data suggest that these four main lineages can be further subdivided based on geographical distribution**.** Both lineages associated to the short gonopodium (yellow and blue) are sympatric and show a similar internal structure, giving rise to subclades strongly associated with two drainage systems: the Laguna dos Patos and Tramandaí-Mampituba drainages. On the other hand, clades associated to the long gonopodium (red and green) are allopatric, but its distribution is similar to that found for the two subclades in the short gonopodium lineages (Figure 4, Figure 5). The red lineage is restricted to the Tramandaí-Mampituba drainage while the green lineage is the most widespread. While geographic isolation would be sufficient for explaining the divergence between the two long gonopodium lineages, it is possible that other factors such as natural or sexual selection associated with genital morphology have played a role not only in the divergence between long vs. short lineages but also between the two short gonopodium lineages, which are currently found in sympatry but with no compelling evidence of gene flow.

Geographic isolation may be favored in this region, as *Cnesterodon* species inhabits shallow flooded areas with no permanent contact with each other, and has demonstrated a poor swimming capacity in high-speed water current (Trenti *et al.* 1999). Climatic oscillations during the past 3 Ma (millions of years) are known have influenced species distribution (Hewitt 1996; Hewitt 2000), and the global variation in water regime caused by Milankovitch cycles (Berger 1988) may have been an important factor by changing freshwater fauna, in which wet periods would favor population connectivity and genetic mixing, while dry periods would favor population isolation (Vila *et al.* 2013). Divergence

times of *Cnesterodon* lineages are compatible with this scenario (Figure 5). Five out of six major clades had a very restrict distribution, possibly reflecting that population isolation during dry periods. During wet periods, a higher connectivity would lead to an increase in genetic homogeneity, except if counteracted by other processes, such as natural or sexual selection.

The large number of *Cnesterodon* mtDNA lineages (Table 3) concentrated in a restricted geographic region might indicate that different haplotypes evolved in different subpopulations belonging to each major lineage. Indeed, the large G_{IS} values (Table 1) may be due to a strong subpopulation structure that causes a Wahlund effect when different subpopulations are analyzed together, producing a heterozygote deficiency regarding Hardy-Weinberg expectations. The effects of past climate changes, which affected all evolutionary lineages may be related to the inference of some *loci* evolving under balancing selection (Figure S3).

Another effective reproductive barrier in SBH species is gonopodium size. This is illustrated by the fact that intermediary morphologies are not typically found, and because there is no strong signal of genetic admixture among lineages. In some poeciliid species, premating sexual selection might influence genital morphology through female choice (Bischoff *et al.* 1985; Basolo 1995; Ptacek 1998; Gould *et al.* 1999; Langerhans *et al.* 2005). Female preference for males with longer gonopodium has been demonstrated in two species of poeciliid fishes (Langerhans et al. 2005; Kahn et al 2009). Differences in male genitals of poeciliid fishes has been shown associated with changes in the use of different mating tactics and predation risk (Kelly *et al.* 2000; Jennies & Kelly 2002; Evans *et al.* 2011). For example, Bisazza *et al.* (2001) and Kahn *et al.* (2009) have demonstrated a strong preference of females for males having larger body size, rather than for males with longer gonopodium in some *Gambusia* species. Gonopodium morphology may also change

due to natural selection to match female genitalia in a "lock-and-key" type of mechanism (Langerhans 2011). Therefore, genetic isolation between the blue and yellow lineages, which have short gonopodium and are sympatric, may be explained due to divergent sexual selection by female choice over a trait other than gonopodium size, or due to a coevolved "lock-and-key" male and female genitalia that prevents hybridization. However, mate choice by males may also contribute to genetic isolation, since sexual isolation among two closely related sympatric live-bearing fish species was reported in the literature as a result of male mate choice rather than female choice (Hugues 1985; Langerhans *et al.* 2005; Espinedo *et al.* 2010). A better characterization of courtship and mate choice patterns in *Cnesterodon* would be extremely important to solve this issue.

Despite the possible role of gonopodium morphology driving population divergence, we did not find evidence of loci evolving under divergent selection among lineages or among adult males having short or long gonopodium. The inability to detect loci under selection in this species system may be attributed to the relative small genomic coverage of our dataset relative to the number of functional elements in *Cnesterodon* genome, or to specific features concerning the traits responsible for reproductive isolation or gonopodium development. Reports from literature have so far identified few genes regulating gonopodium development in poeciliid fishes (Zauner *et al.* 2003; Offner *et al.* 2008; 2009), but how these or other factors affect the differential development of the gonopodium among distinct lineages is unknown. Therefore, it is possible that many more genomic loci and large sample sizes would be required to pinpoint the genetic basis of gonopodium differences among species and allow of these loci show any signature of divergent selection.

On the other hand, the strong association between lineage and gonopodium morphology is not strict. One individual from the yellow lineage has long gonopodium,

and four individuals from the green lineage have short gonopodium (Figure 2). Recent admixture seems an unlikely explanation for this finding, as only one out of these five individuals have an ancestry proportion smaller than 0.998 for a given genetic lineage. However, if mate choice operates primarily over other characters than gonopodium morphology, genetic cohesion of distinct evolutionary lineages would be maintained even though the selective pressure on gonopodium size would be more relaxed, allowing some phenotypic variance in this trait within an evolutionary lineage. At the mtDNA level, our data also show that one individual from the yellow lineage has a mtDNA haplotype from the blue lineage and vice versa (Figure 5). These individuals come from populations where the yellow and blue lineages are syntopic, and, therefore, the mtDNA data is consistent with limited introgressive hybridization between these lineages. While the genomic proportions of the individual 0548B may be consistent with a recent admixed origin (0.788 vs. 0.204 ancestry on the blue or yellow lineages, respectively), individual 0554A has 0.999 ancestry from the yellow lineage (considering K=4 in both cases), strongly suggesting that recent admixture cannot explain the distribution of these mtDNA lineages. Overall, these results may suggest occasional introgressive events that may occur rarely but constantly along time.

Taken together, our results showed a complex evolutionary history for *Cnesterodon* species-complex in which both allopatric divergence and natural/sexual selection originated four closely related evolutionary lineages. Despite gonopodium size may be a trait highly relevant to natural and sexual selection in these lineages, the evidence of introgressive hybridization revealed by mtDNA suggests that reproductive isolation among different sizes gonopodium lineages is not complete, and that rare events of secondary sexual contact may influence the evolutionary dynamics in this system. Finally, our study also demonstrates the utility of molecular data, especially of new generation sets of

markers, to identify evolutionary lineages, and to unveil the processes underlying the generation and maintenance of current fish biodiversity.

References

Avise JC (2004) *Molecular Markers, Natural History, and Evolution*. Sinauer Associates, Sunderland.

Basolo AL (1995) Phylogenetic evidence for the role of a pre-existing bias in sexual selection. *Proceedings of the Royal Society of London B: Biological Sciences*, **259**, 307-311.

Berger A (1988) Milankovitch theory and climate. *Reviews Geophysics*, **26**, 624-657.

Bickford D, Lohman DJ, Sodhi NS *et al.* (2007) Cryptic species as a window on diversity and conservation. Trends Ecology and Evolution, **22**, 148-155.

Bischoff RJ, Gould JL, Rubenstein DI (1985) Tail size and female choice in the guppy (Poecilia reticulata). *Behavioral Ecology and Sociobiology,***17**, 253-255.

Bisazza A, Vaccari G, Pilastro A (2001) Female mate choice in a mating system dominated by male sexual coercion. *Behavioral Ecology*, **12**, 59-64.

Boughman JW, Rundle HD, Schluter D (2005) Parallel evolution of sexual isolation in sticklebacks. *Evolution,* **59**, 361-373.

Cracraft J (1983) Species concepts and speciation analysis. In: *Current ornithology* (ed Johnston RF), pp. 159-187. Plenum Press, New York.

Cracraft J (1989) Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: *Speciation and its Consequences* (eds Otte D, Endler JA), pp. 28-59. Sinauer, Sunderland.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, **1**, 171-182.

Catchen J, Bassham S, Wilson T, Currey M, O'Brien C, Yeates Q, Cresko WA (2013) The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular Ecology*, **22**, 2864-2883

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland.

Coyne JA, Orr HA, Futuyma DJ (1988) Do we need a new species concept? *Systematic Zoology*, **37**, 190-200.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods,* **9**, 772-772.

Darwin C (1859) *On the origins of species by means of natural selection*. Murray, London.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879-886.

De Queiroz K (2011) Branches in the lines of descent: Charles Darwin and the evolution of the species concept. *Biological Journal of the Linnean Society*, **103**, 19-35.

Dobzhansky T, Dobzhansky TG (1937) *Genetics and the Origin of Species*. Columbia University Press, New York.

Donoghue MJ (1985) A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist*, **88**, 172-181.

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

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2007) Geneious v4.8. Biomatters Ltd., Auckland, New Zealand. Available at: http://www.geneious.com.

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.*Conservation Genetics Resources*, **4**, 359-361.

Edwards CJ, Suchard MA, Lemey P, Welch JJ, Barnes I, *et al.* (2011) Ancient hybridization and an Irish origin for the modern polar bear matriline. Current Biology **21**, 1251-1258

Espinedo CM, Gabor CR, Aspbury AS (2010) Males, but not females, contribute to sexual isolation between two sympatric species of *Gambusia*.*Evolutionary Ecology*, **24**, 865-878.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.

Evans JP, Gasparini C, Holwell GI, Ramnarine IW, Pitcher TE, Pilastro A (2011) Intraspecific evidence from guppies for correlated patterns of male and female genital trait diversification. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 2611-2620.

Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977-

993.

Freamo H, O'Reilly P, Berg PR, Lien, S, Boulding EG (2011) Outlier SNPs show more genetic structure between two Bay of Fundy metapopulations of Atlantic salmon than do neutral SNPs. *Molecular Ecology Resources*, **11**, 254–267

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.

Funk DJ (1998) Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. *Evolution*, **52**, 1744-1759.

Funk DJ, Nosil P, Etges WJ (2006) Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 3209-3213.

Gould JL, Elliott SL, Masters CM, Mukerji J (1999) Female preferences in a fish genus without female mate choice. *Current Biology*, **9**, 497-500.

Greven H (2005) Structural and behavioral traits associated with sperm transfer in Poeciliinae. In: *Viviparous fishes* (eds Uribe MC, GRIER H), pp147-165. New Life Publications, Homestead, Florida.

Herbert PDN, Cywinska A, Ball S, Dewaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the National Academy of Sciences of the United States of America*, **270**, 313-321.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276 .

Hewitt GM (2000) The genetic legacy of the quaternary ages. *Nature,* **405**, 907-913.

Hey J (2001) The mind of the species problem. *Trends in Ecology & Evolution*, **16**, 326-

329.

Howard DJ (1993a) Small populations inbreeding and speciation. In: *The Natural history of inbreeding and outbreeding: theoretical and empirical perspectives* (ed Thornhill NW), pp. 118-142. The University of Chicago Press, Chicago.

Howard DJ (1993b) Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. In: *Hybrid Zones and the Evolutionary Process*. (ed Harrison RG), pp.46-69. Oxford University Press, New York.

Hrbek T, Meyer A (2003) Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *Journal of Evolutionary Biology*, **16**, 17-36.

Hughes AL (1985) Male size, mating success, and mating strategy in the mosquitofish *Gambusia affinis* (Poeciliidae). *Behavioral Ecology and Sociobiology*, **17**, 271-278.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster match- ing and permutation program for dealing with label switch- ing and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jennions MD, Kelly CD (2002). Geographical variation in male genitalia in *Brachyrhaphis episcopi* (Poeciliidae): is it sexually or naturally selected? *Oikos*, **97**, 79-86.

Jiggins CD, Naisbit RE, Coe RL, Mallet J (2001) Reproductive isolation caused by colour pattern mimicry. *Nature*, **411**, 302-305.

Kahn AT, Mautz B, Jennions MD (2009) Females prefer to associate with males with longer intromittent organs in mosquitofish. *Biology Letters*, **11**, 1-4.

Kelly CD, Godin JGJ, Abdallah G (2000) Geographical variation in the male intromittent organ of the Trinidadian guppy (*Poecilia reticulata*). *Canadian Journal of Zoology*, **78**, 1674-1680.

Langerhans RB, Layman CA, DeWitt TJ (2005). Male genital size reflects a tradeoff between attracting mates and avoiding predators in two live-bearing fish species. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 7618-7623.

Langerhans RB (2011) Genital evolution In: *Ecology and evolution of poeciliid fishes* (eds Evans JP, Pilastro A, Schlupp I), pp. 228-240. University of Chicago Press, Chicago.

Lindqvist C, Schuster SC, Sun Y, Talbot SL, Qi J *et al.* (2010) Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. *Proceedings of the National Academy of Sciences of the United States of America,* **107,** 5053-5057.

Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Lucinda PH (2005) Systematics of the genus Cnesterodon Garman, 1895 (Cyprinodontiformes: Poeciliidae: Poeciliinae). *Neotropical Ichthyology*, **3**, 259-270.

Martinez-Takeshita N, Purcell CM, Chabot CL, Craig MT, Paterson CN, *et al.* (2015) A Tale of Three Tails: Cryptic Speciation in a Globally Distributed Marine Fish of the Genus Seriola.*Copeia*, **103**, 357-368.

Mayr E (1942) *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press, London.

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792-794.

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes as suggested by mitochondrial DNA sequences. *Nature,* **347**, 550-553.

Nosil P, Crespi BJ, Sandoval CP (2003) Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *Proceedings of the Royal Society of London B: Biological Sciences*,**270**, 1911-1918.

Nosil P, Crespi BJ, Gries R, Gries G (2007) Natural selection and divergence in mate preference during speciation. *Genetica,* **129**, 309-327.

Offen N, Blum N, Meyer A, Begemann G. (2008) Fgfr1 signalling in the development of a sexually selected trait in vertebrates, the sword of swordtail fish. *BMC developmental biology*, **8**, 98-98.

Offen N, Meyer A, Begemann G. (2009) Identification of novel genes involved in the development of the sword and gonopodium in swordtail fish. *Developmental Dynamics*, **238**, 1674-87.

Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis, *PLoS Genetics*, **2**, 2074-2093.

Peterson, Brant K., Jesse N. Weber, Emily H. *et al.* (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, **7**, e37135.

Pfennig KS, Pfennig DW (2009) Character displacement: ecological and reproductive responses to a common evolutionary problem. *The Quarterly Review of Biology*, **84**, 253-253.

Poeser FN (1998) The role of character displacement in the speciation of Central American members of the genus *Poecilia* (Poeciliidae). *Italian Journal of Zoology*, **65**, 145-147.

Ptacek MB, Travis J (1998) Hierarchical patterns of covariance between morphological and behavioural traits. *Animal Behaviour*, **56**, 1044-1048.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

R Core Team: R: A Language and Environment for Statistical Computing Vienna, Austria: R foundation for Statistical Computing; 2013. Retrieved from http//www.R-project.org.

Rosen DE, Bailey RM (1963) The poeciliid fishes (Cyprinodontiformes): their structure, zoogeography, and systematics. *Bulletin of the American Museum of Natural History*; **126**, 1-176.

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.

Rundle HD, Nagel L, Boughman JW, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science*, **287**, 306-308.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters,* **8**, 336-352.

Ryan MJ, Dries LA, Batra P, Hillis DM (1996) Male mate preferences in a gynogenetic species complex of Amazon mollies. *Animal Behaviour*, **52**, 1225-1226.

Schluter D (1996) Ecological speciation in postglacial fishes. *Proceedings of the National Academy of Sciences of the United States of America*, **351**, 807–814.

Schluter D (2001) Ecology and the origin of species. *Trends in Ecology & Evolution*, **16**, 372-380.

Scott LE, Johnson JB (2010) Does sympatry predict life history and morphological diversification in the Mexican livebearing fish Poeciliopsis baenschi?.*Biological Journal of the Linnean Society*, **100**, 608-618.

Seehausen O, Wagner CE (2014) Speciation in freshwater fishes. *Annual Review of Ecology, Evolution, and Systematics*, **45**, 621-651.

Servedio MR (2001) Beyond reinforcement: the evolution of premating isolation by direct selection on preferences and postmating, prezygotic incompatibilities. *Evolution,* **55**, 1909-1920.

Silvestro D, Michalak I (2012) raxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution*, **12**, 335-337.

Smith AJ, Nelson-Maney N, Parsons KJ, Cooper WJ, Albertson RC (2015) Body Shape Evolution in Sunfishes: Divergent Paths to Accelerated Rates of Speciation in the Centrarchidae. *Evolutionary Biology*, **42**, 283-295.

Sneath PH, Sokal RR (1973). *Numerical taxonomy. The principles and practice of numerical classification*. W. H. Freeman and Company, San Francisco.

Sorenson MD (2003) Avian mtDNA primers. Avaiable at: http://people.bu.edu/msoren/ (Bird.mt.Primers.pdf.).

Swenton DM (2011) Sex differences in mate preference between two hybridizing species of poeciliid fish. *Ethology*, **117**, 208-216.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**, 437-460.

Templeton A. R. (1989). The meaning of species and speciation: a genetic perspective. In: *The units of evolution* (ed Ereshefsky M), pp.159-183. The MIT Prees, London.

Trenti PS, Gómez SE, Ferriz RA (1999) Capacidad de natación en tres peces pampásicos. *Aprona*, **38**, 2-9.

Van Valen L (1976) Ecological species, multispecies, and oaks. *Taxon*, **25**, 233-239.

Vila I, Morales P, Scott S, Poulin E, Véliz D, Harrod C, Méndez MA (2013) Phylogenetic and phylogeographic analysis of the genus *Orestias* (Teleostei: Cyprinodontidae) in the southern Chilean Altiplano: the relevance of ancient and recent divergence processes in speciation. *Journal of Fish Biology*, **82**, 927-943.

Vines TH, Schluter D (2006) Strong assortative mating between allopatric sticklebacks as a by-product of adaptation to different environments. *Proceedings of the Royal Society of London B: Biological Sciences*, **273**, 911-916.

Vrijenhoek RC (1994) Unisexual fish: model systems for studying ecology and evolution. *Annual review of Ecology and Systematics*, **25**, 71-96.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, *et al.* (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology* **22**: 787-798.

Zauner H, Begemann G, Mari-Beffa M, Meyer A. (2003) Differential regulation of msx genes in the development of the gonopodium, an intromittent organ, and of the "sword," a sexually selected trait of swordtail fishes (*Xiphophorus*). *Evolution & development*, **5**, 466–477.

LRM, NF and ARF conceived and designed the project. ARF produced the data. ARF, MAB and NF analysed the data and wrote most of the text. LAR provided reagents and equipment to develop experiments. All authors contributed on preparation of the manuscript and have approved the final version.

Acknowledgements

The authors are very grateful to Anna Sellas for all support, welcoming and help in laboratory procedures, to Brian Simison for support and help in pipelines programming and running softwares, to California Academy of Science for the infrastructure provided. We also thank to Parque Nacional dos Aparados da Serra for help in getting samples for this study. This work was supported by Conselho Nacional de desenvolvimento Científico (CNPq), Programa de Pós Graduação em Biologia Animal (PPGBAN-UFRGS), Programa de Pós Graduação em Genética e Biologia Molecular (PPGBM-UFRGS) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superios (CAPES, Brazil, Proc. nº 3478/14-9), Ministry of Education, Brazil.

Table 1. Genetic diversity among the four lineages for 6,840 RAD-seq *loci*.

Lineage	Sample Size	Effective number of alleles	H_O	H_E	G_{IS}
Red	18	1.048	0.028	0.034	0.155
Yellow	36	1.049	0.023	0.035	0.344
Blue	36	1.077	0.044	0.055	0.196
Green	43	1.046	0.023	0.035	0.343
Overall	133	1.067	0.03	0.05	0.430

H_O – Observed heterozygosity; H_E – Expected heterozygosity; G_{IS} – Inbreeding coefficient.

Table 2. Pairwise F_{ST} values for mtDNA (lower diagonal) and 6,840 RAD-seq *loci* (upper diagonal). All values were statistically significant (P<0.05).

Lineage	1 – Yellow	2 – Blue	3 – Green	4 – Red
1	-	0.2385	0.2465	0.3538
2	0.5477	-	0.1398	0.3097
3	0.4265	0.4338	-	0.2588
4	0.8338	0.8582	0.3977	-

Table 3. Diversity indexes obtained for the mtDNA concatenated set (details in text) for the four genetic lineages.

Lineage	Sample Size	*H*		Fu's Fs	Tajima's D
Red	12	0.84	0.002	-0.26	0.82
Yellow	10	0.91	0.012	2.68	-0.50
Blue	10	0.9	0.009	3.30	-0.95
Green	24	0.94	0.032	6.73	0.56
Overall	57	0.97	0.034	4.22	0.46

H – Haplotype diversity; – nucleotide diversity. All Fs and D values are not statistically significant (*P*>0.05)

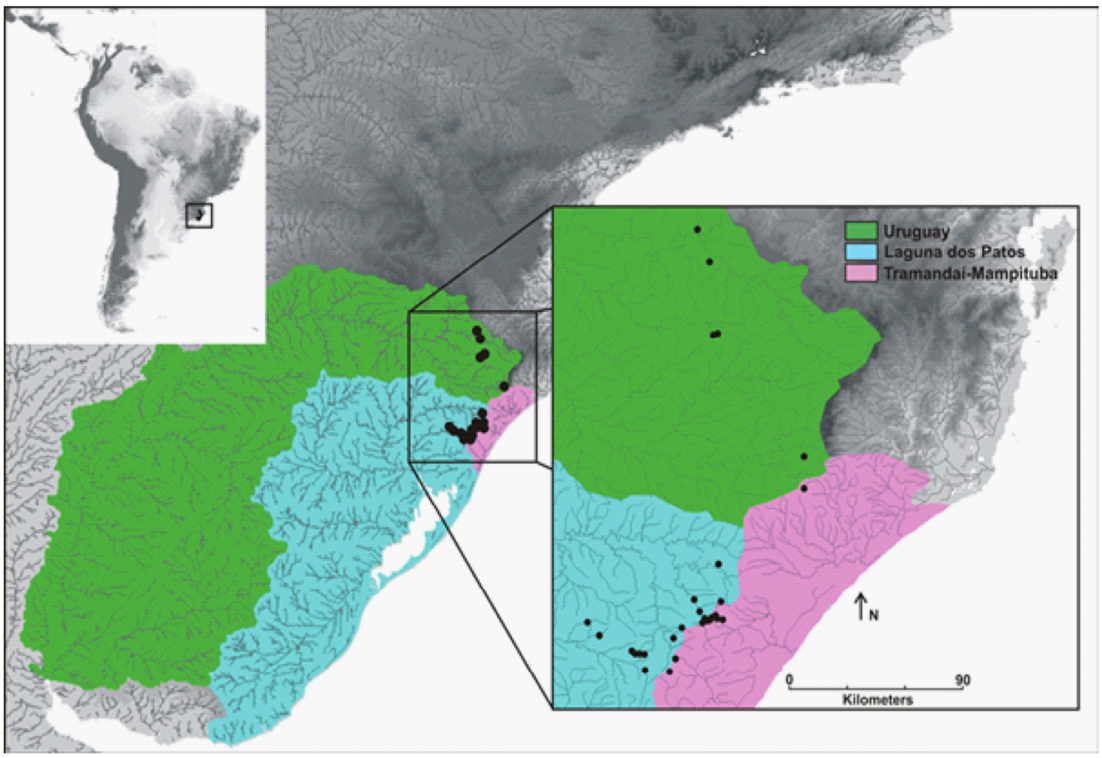


Figure 1. Distribution map of sampled sites showing drainage systems. For more details about localities see Table S1.

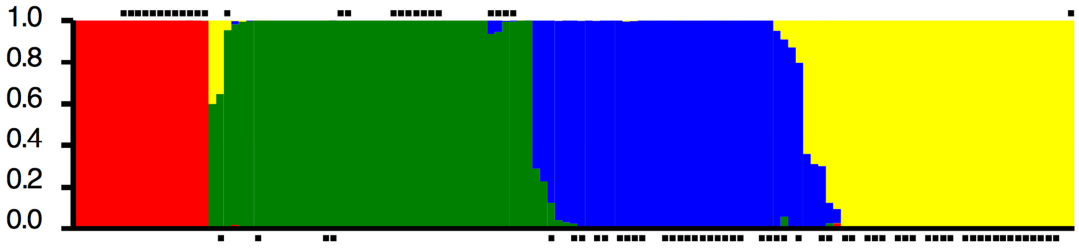


Figure 2. Bayesian ancestry estimates for all 133 *Cnesterodon* individuals based on K=4 genetic clusters. Squares above the ancestry bar represent individuals with large gonopodium, while squares below the ancestry bar represent individuals with short gonopodium.

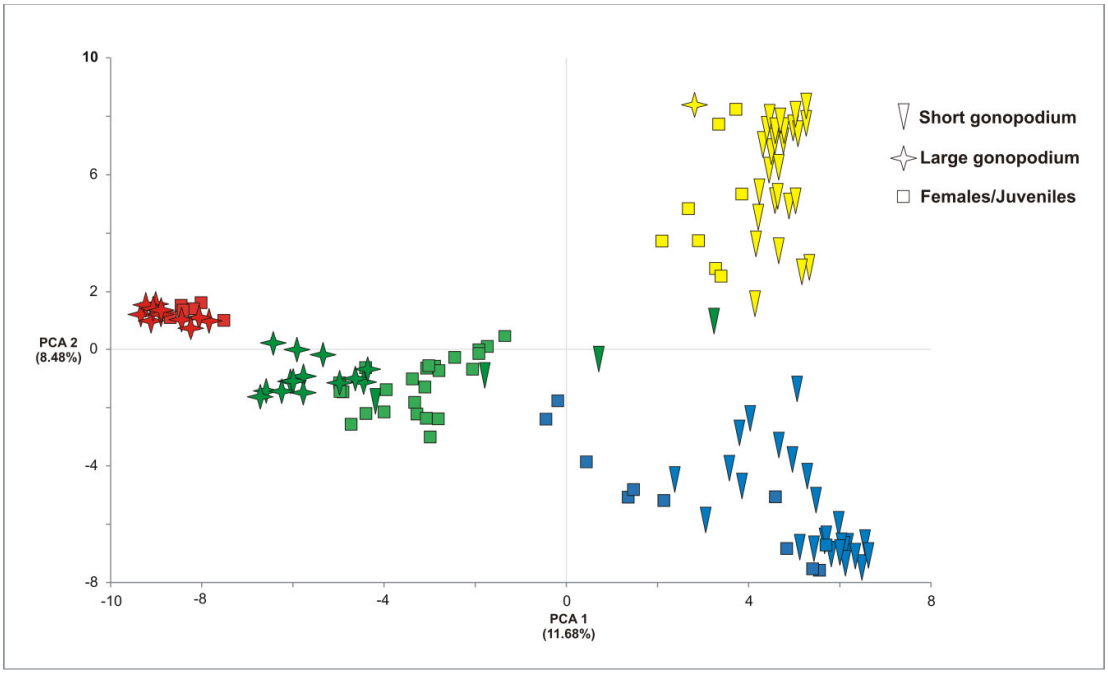


Figure 3. Principal Component Analysis based on RAD-seq genotype data. The first two components (PCA1 and PCA2) are plotted, and the proportion of variation retained by them are shown in the axes. Colors are in agreement to the genetic clusters suggested by STRUCTURE.

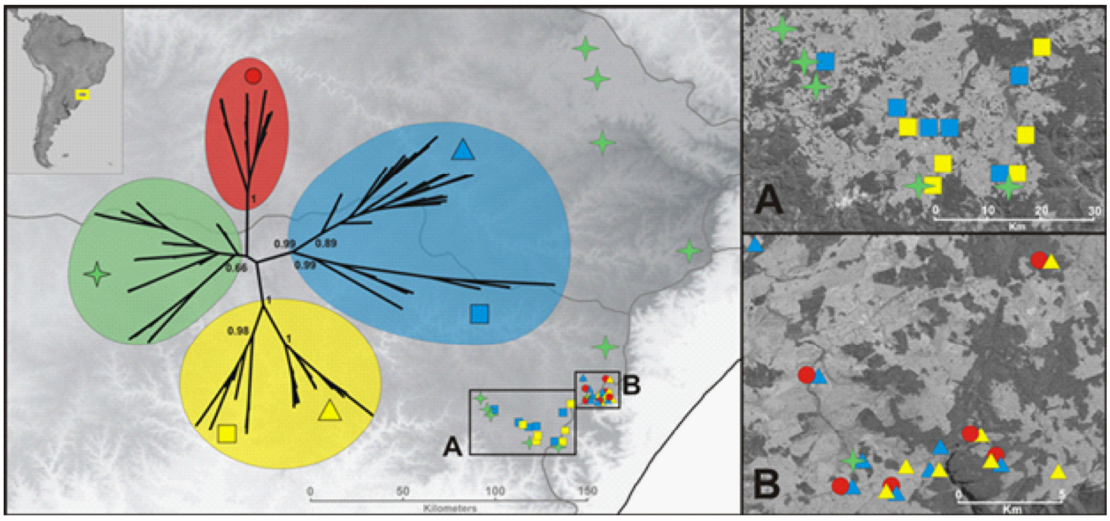


Figure 4. RAD-seq phylogenetic analyses. Tree topology and branch lengths are based on Maximum likelihood reconstructions (RAxML). Bootstrap support (in bold) is based on 1000 pseudo-replicates for RAxML and reported for analyses based on 6,840 loci. Colors are in agreement to the genetic clusters suggested by STRUCTURE. The map shows geographical distribution of the clades. **A** and **B** areas are shown in a larger zoom at right. Gray lines show drainage systems division. Symbols touching each other represent occurrence of more than one clade at the same collection point.

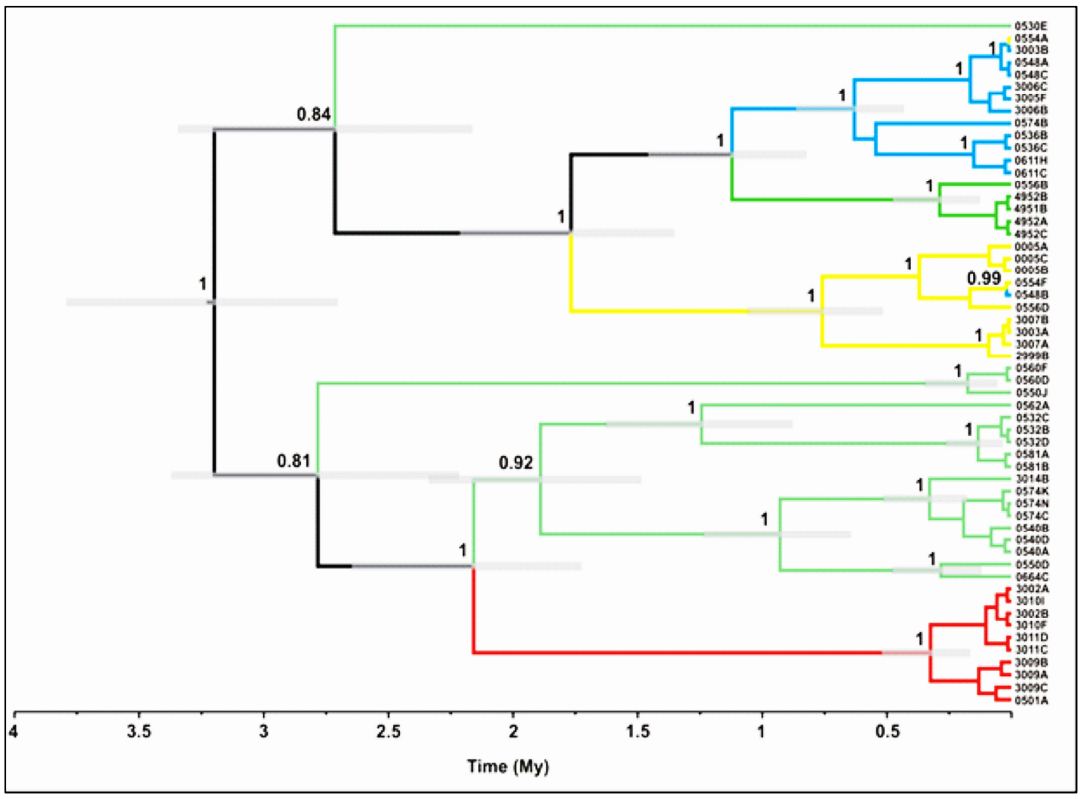


Figure 5. Bayesian phylogenetic tree calibrated based on mitochondrial concatenated dataset (for more details see text), obtained from BEAST for *Cnesterodon* species-complex individuals. Numbers above the branches represent posterior probabilities values (PP). The grey bars represent the 95% credible interval for the time of the most recent ancestor (TMRCA) for each node. Time is given in millions of years. Colors are in agreement to the genetic clusters suggested by STRUCTURE and are in agreement to Figures 3 and 4. For more details see text.

Supplementary Material

Table S1. Voucher specimens and drainage systems for samples used in this study. All specimens belong to the UFRGS zoological collection (Universidade Federal do Rio Grande do Sul).

Voucher	Drainage	Latitude	Longitude
UFRGS 6843/TEC 0005 A, B, C	Laguna dos Patos	-29.2097	-50.2394
UFRGS 14581/TEC 0477 B. C, D	Laguna dos Patos	-29.3425	-50.2666
UFRGS 6849/TEC 0501 A, B, H	Laguna dos Patos	-29.1791	-50.1369
UFRGS 6841/TEC 0525 C,E, H	Laguna dos Patos	-29.2516	-50.2775
UFRGS 14606/TEC 0530 B, C, D, E, F, G	Laguna dos Patos	-29.2450	-50.6294
UFRGS 6839/TEC 0532 A,B, C, D	Laguna dos Patos	-29.9239	-50.0669
UFRGS 12346/TEC 0536 B, C	Laguna dos Patos	-29.3094	-50.4766
UFRGS 14612/TEC 0540 A, B, D, F	Laguna dos Patos	-29.1827	-50.6855
UFRGS 6853/TEC 0548 A, B, C	Laguna dos Patos	-29.0830	-50.1800
UFRGS 14618/TEC 0550 D, H, J	Uruguay	-27.9094	-50.0875
UFRGS 14622/TEC 0554 A, B, F, I	Laguna dos Patos	-29.3938	-50.4138
UFRGS 6846/TEC 0556 B, D, L	Laguna dos Patos	-29.3938	-50.4138
UFRGS 12360/TEC 0558 G, I	Laguna dos Patos	-29.3252	-50.4161
UFRGS 14624/TEC 0560 C, D, F, J	Uruguay	-27.9033	-50.0719
UFRGS 14625/TEC 0562 A, B	Uruguay	-28.4502	-49.6569
UFRGS 14628/TEC 0569 B, D, I	Laguna dos Patos	-29.3227	-50.4619
UFRGS 14629/TEC 0574 B, C, D, E, G, H, I, K, L, N	Laguna dos Patos	-29.2369	-50.6308
UFRGS 6840/TEC 0581 A, B	Uruguay	-27.4383	-50.1644

UFRGS 14646/TEC 0611 C, H	Laguna dos Patos	-29.3219	-50.4408
UFRGS 14663/TEC 0664 C	Uruguay	-27.5877	-50.1080
UFRGS 16682/TEC 2999 A, B	Tramandaí-Mampituba	-29.1586	-50.0803
UFRGS 16684/TEC 3001 A, B	Tramandaí-Mampituba	-29.1586	-50.0803
UFRGS 16685/TEC 3002 A, B	Tramandaí-Mampituba	-29.1728	-50.1032
UFRGS 16686/TEC 3003 A, B, C, D	Tramandaí-Mampituba	-29.1728	-50.1032
UFRGS 16688/TEC 3005 B, C, D, E, F	Tramandaí-Mampituba	-29.1635	-50.0976
UFRGS 16689/TEC 3006 A, B, C, D, E	Laguna dos Patos	-29.1723	-50.1307
UFRGS 16690/TEC 3007 A, B	Tramandaí-Mampituba	-29.1726	-50.1121
UFRGS 16691/TEC 3008 A, B, E	Tramandaí-Mampituba	-29.1742	-50.1167
UFRGS 16692/TEC 3009 A, B, C, G, H	Laguna dos Patos	-29.1723	-50.1307
UFRGS 16693/TEC 3010 A, F, I,	Tramandaí-Mampituba	-29.1742	-50.1167
UFRGS 16694/TEC 3011 C, D, J	Tramandaí-Mampituba	-29.1570	-50.0510
UFRGS 16695/TEC 3012 A, B, C	Tramandaí-Mampituba	-29.1570	-50.0510
UFRGS 16696/TEC 3013 A, B, C	Laguna dos Patos	-29.1693	-50.1332
UFRGS 16697/TEC 3014 A, B, C	Laguna dos Patos	-29.1693	-50.1332
UFRGS 16698/TEC 3015 A, B, C, D, E	Tramandaí-Mampituba	-29.1729	-50.1031

UFRGS 16699/TEC 3016 A, B, C, D, E, F, G, H, I, J	Tramandaí-Mampituba	-29.1728	-50.0454
UFRGS 19258/TEC 4951 A, B, C	Tramandaí-Mampituba	-29.4016	-50.2961
UFRGS 19259/TEC 4952 A, B, C	Tramandaí-Mampituba	-29.4016	-50.2961

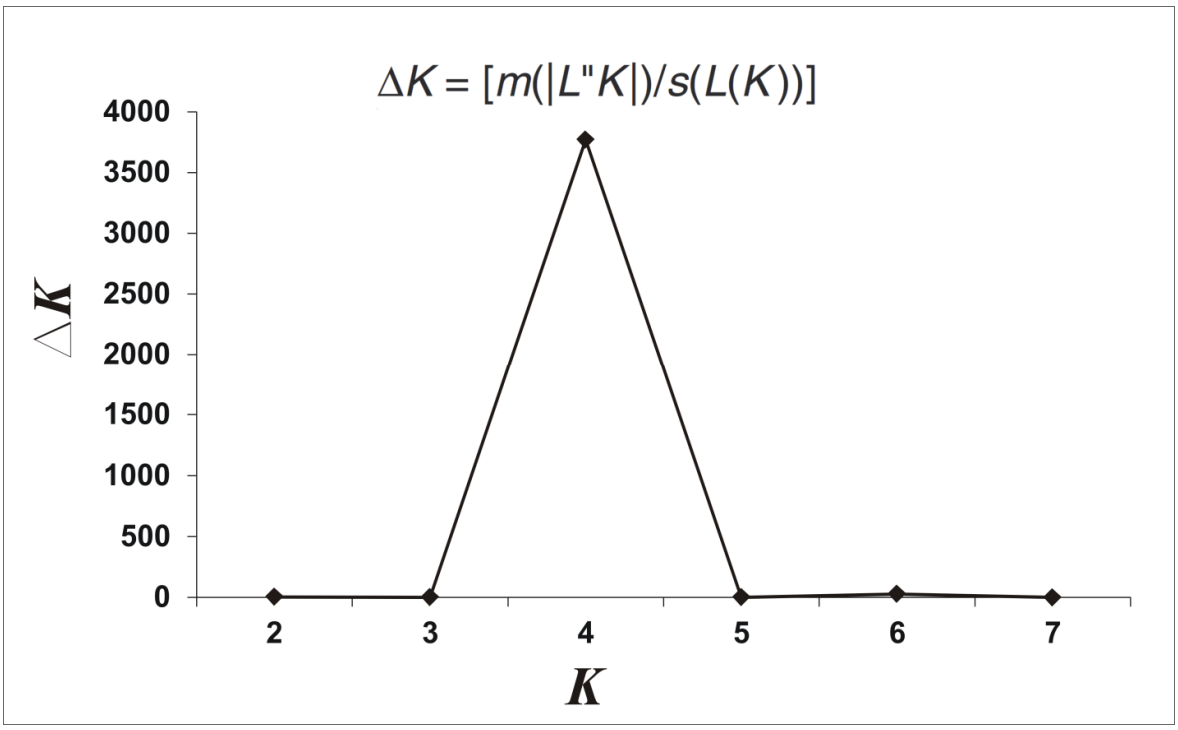


Figure S1. Δ*K* estimate for the best number of genetic clusters in STRUCTURE program. The magnitude of Δ*K* was calculated using the method described by Evanno *et al.* (2005). (see text for details).

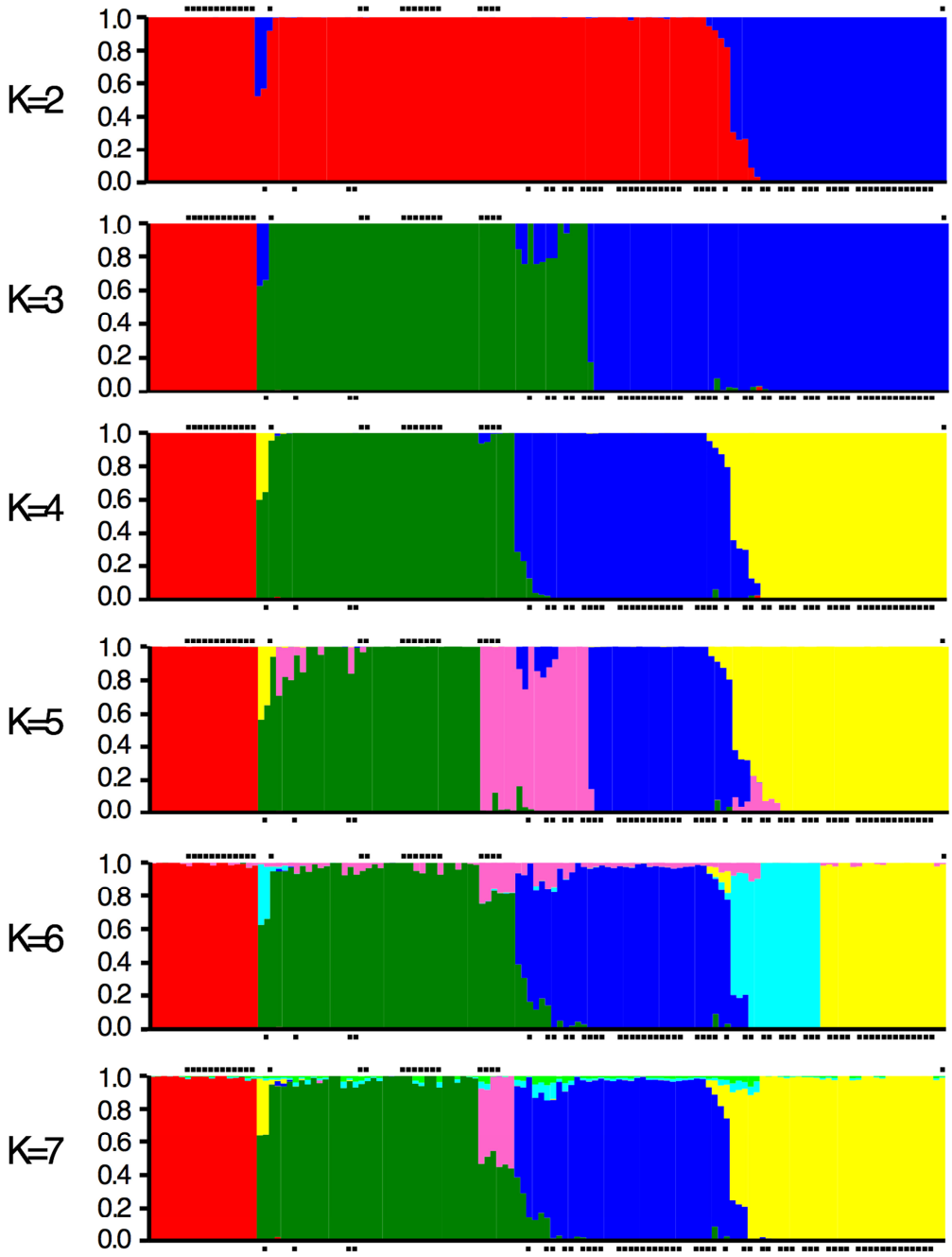


Figure S2. Bayesian ancestry estimates for 133 *Cnesterodon* individuals based on *K* from 1 to 7 genetic clusters. Squares above the ancestry bar represent individuals with large gonopodium, while squares below the ancestry bar represent individuals with short gonopodium.

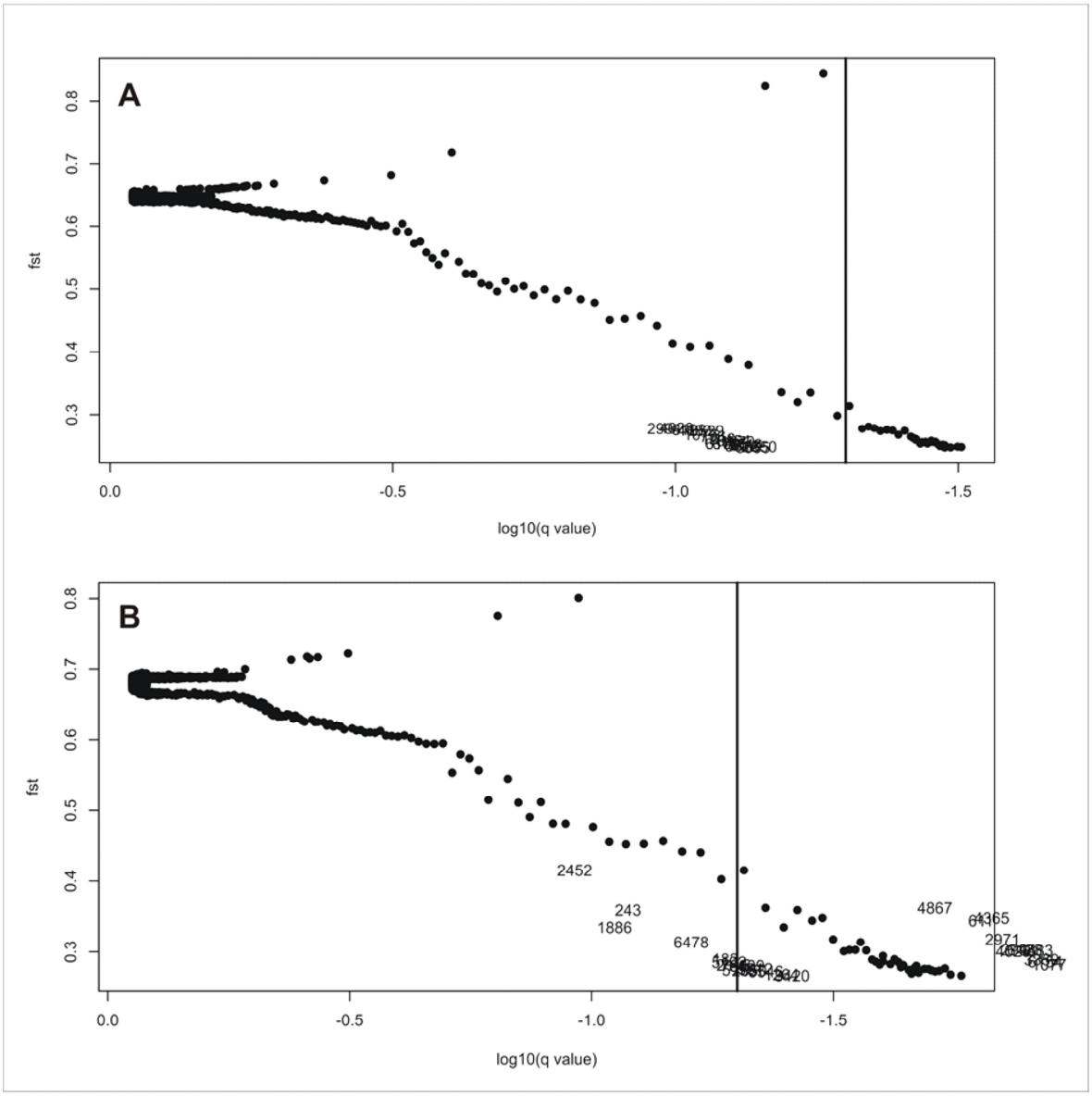


Figure S3. Bayescan plot for 6840 SNP loci sampled among *Cnesterodon* species-complex in sympatry, testing: **(A)** outliers among genetics lineages detected in clustering analyses and **(B)** excluding females and juveniles males in attempt to identify locus related to the type of gonopodium (short or long) under selection. Numbers represent the RAD-locus having the most extreme values.

CAPÍTULO 4

Molecular phylogeny of *Cnesterodon* Garman (1885) "(Cyprinodontiformes: Poeciliidae: Poeciliinae)" with emphasis in new putative species from Southern Brazil

Manuscrito a ser submetido para a revista *Neotropical Ichthyology*

Molecular phylogeny of *Cnesterodon* Garman (1885) "(Cyprinodontiformes: Poeciliidae: Poeciliinae)" with emphasis in new putative species from Southern Brazil

Aline M. C. Ramos-Fregonezi[1], Luiz R. Malabarba[2] and Nelson J. R. Fagundes[1]

[1] Programa de Pós-Graduação em Genética e Biologia Molecular, Departamento de Genética, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, 91501-970 Porto Alegre, RS, Brazil. alinemitcheli@gmail.com (AMCRF), nelson.fagundes@ufrgs.br (NJRF) (corresponding author)

[2] Programa de Pós-Graduação em Biologia Animal, Departamento de Zoologia, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, 91501-970 Porto Alegre, RS, Brazil. malabarb@ufrgs.br (LRM)

Key-words: mtDNA, *Cnesterodon brevirostratus*, freshwater fishes, species tree

Running Head: Molecular phylogeny of *Cnesterodon* (Poecillidae: Poecillinae)

Abstract

The genus *Cnesterodon* Garman (1885) includes small, viviparous live-bearing members of the subfamily Poeciliinae from freshwater ecosystems throughout southeastern South America. The genus currently embraces eleven recognized taxa and ten species described so far. There are several phylogenetic studies for the genus based on morphological characters, but current phylogenetic hypotheses seems to be highly unstable and sensitive to the inclusion of new terminal taxa. In this study, we used data from molecular markers to propose a new phylogenetic hypothesis for the genus including four distinct evolutionary lineages from Southern Brazilian Highlands (SBH) that may represent different species. We analyzed two mitochondrial (COI and ND2) and one nuclear gene (SH3PX3) in a Bayesian framework using both coalescent species tree and concatenated methods. Our results show that the phylogenetic hypothesis based on molecular data is strikingly different from the previous morphological phylogenetic hypotheses. Species from Uruguay/Southern Brazil formed a clade with high support, in which the two Pampean species, *C. decemmaculatus* and *C. holopteros*, are sister to each other, while the four species found in SBH formed another well-supported monophyletic group. The root of our phylogeny dates back to ~8 millions of years ago (Mya), while the origin of Uruguay/Southern Brazilian clade was ~5.5 Mya, and the divergence between Pampean and SBH species occurred ~4 Mya. Our results are compatible with the colonization of Southern drainages from populations occurring in the Paraná basin due to the availability of new habitats following the retreat of the "Paranean Sea".

Resumo

O gênero *Cnesterodon* Garman (1885) inclui membros pequenos e vivíparos da subfamília Poecilinae encontrados em ecossistemas de água doce ao longo do sudeste da América do Sul. Atualmente, o gênero conta com onde taxa reconhecidos, e dez espécies descritas. Diversos estudos filogenéticos para o gênero foram realizados baseados em caracteres morfológicos, mas as hipóteses filogenéticas atuais parecem altamente instáveis e sensíveis à inclusão de novos taxa. Nesse estudo, dados de marcadores moleculares foram utilizados para propor uma nova hipótese filogenética para o gênero incluindo quatro linhagens evolutivas distintas, oriundas do Planalto Sul-Brasileiro (PSB) que podem representar novas espécies. Foram analisados dois genes mitocondriais (COI e ND2) e um nuclear (SH3PX3) através de uma metodologia Bayesiana usando abordagens baseadas em métodos de árvore de espécie coalescente e baseados nos dados concatenados. Nossos resultados mostram que a hipótese filogenética baseada em dados moleculares difere substancialmente das hipóteses filogenéticas morfológicas propostas anteriormente. As espécies do Uruguai/Sul do Brasil formaram um clado com alto suporte, no qual as duas espécies pampeanas, *C. decemmaculatus* e *C. holopteros* são irmãs, enquanto as quatros espécies encontradas no PSB formaram um outro grupo monofilético com alto suporte. A raiz da filogenia data de ~8 milhões de anos atrás (Ma), enquanto a origem do clado do Uruguai/Sul do Brasil foi de ~5,5 (Ma), e a divergência entra as espécies Pampeanas e do PSB ocorreu há ~4 Ma. Os resultados são compatíveis com uma colonização das drenagens do sul a partir de populações que ocorriam na bacia do Paraná devido à disponibilidade de novos habitats subsequente à retração do "Mar Paranaense".

Introduction

The subfamily Poeciliinae comprises a diverse group of small fishes including more than 200 species (Lucinda, 2003). Several species in this subfamily have been used as model organisms in studies of ecology, anatomy, embryology, sexual and natural selection, and in comparative studies of life history and evolution (e.g. Endler, 1983; Grove & Wourms, 1991, 1994; Houde, 1997; Schluter *et al*., 1998; Arias & Reznick, 2000; Hamilton, 2001; Reznick *et al*., 2002, 2007). In addition, one of its members (*Poecilia reticulata* Peters 1859, the guppy), is known worldwide as a popular aquarium species (Nakajima & Taniguchi, 2001; Shen *et al.*, 2007). Despite their importance, many biological features remain poorly understood, including phylogenetic relationships, biogeography and the degree of intra and intergeneric diversity (Lucinda, 2003; Lucinda & Reis, 2005).

The genus *Cnesterodon* Garman (1885) includes small, viviparous live-bearing members of freshwater ecosystems throughout southeastern South America. The genus currently embraces eleven recognized taxa, with ten species described so far (Table 1). There are several phylogenetic studies based on morphological characters (Lucinda, 2005; Lucinda & Reis, 2005; Lucinda *et al*., 2006; Aguilera *et al*., 2009). However, current phylogenetic hypotheses for *Cnesterodon* seems to be highly unstable (Fig. 1). For example, there may be differences among trees after the inclusion of a single species. Another interesting feature of these hypotheses is that taxa such as *C. decemmaculatus* and *C. holopteros*, which may be sympatric in some localities in the Pampa biome, and *C. brevirostratus* and *Cnesterodon.* sp. nov. B, which may be sympatric in the Southern Brazil highlands, are not closely related. This result is important for understanding the biogeographic history of this genus and of the drainage systems in these Biomes. In addition, the current *Cnesterodon* taxonomy probably underrepresents the true species diversity of the genus, given the small number of records available through wide geographical areas, and the number of new species validated or described in the last 15 years. Therefore, the recognition of new *Cnesterodon* species will demand a continuous revisiting of their phylogenetic relationships.

A good example of the yet unrecognized species diversity within *Cnesterodon* occurs in the Campos de Cima da Serra, Southern Brazil, where the specimens occur above 750 meters of elevation in shallow environments surrounded by natural grasslands, draining to the Laguna dos Patos, Uruguay and Tramandaí/Mampituba hydrologic basins (Lucinda, 2005; Malabarba *et al.*, 2009). Based on morphological data, Rosa & Costa (1993) described *C. brevirostratus* but identified other individuals from this region as *C. decemmaculatus*. This was further corrected by other authors (Juan Anza, Paulo H. F. Lucinda, Luiz R. Malabarba, unpublished results), who suggested that these specimens were actually a new species (referred to as *Cnesterodon* sp. nov. B in Lucinda, 2005). However, new collections indicated that *Cnesterodon* sp. nov. B is sympatric, and in many localities syntopic, with *C. brevirostratus* (Ramos-Fregonezi *et al.*, in prep.). The delimitation of distinct evolutionary lineages based on genomic data further suggested that *C. brevirostratus* and *Cnesterodon* sp. nov. B correspond to distinct evolutionary groups, and that these two species actually represent four distinct *Cnesterodon* lineages in this region (Ramos-Fregonezi *et al.*, in prep.). These findings support the idea that the current taxonomical knowledge about *Cnesterodon* remains incomplete, affecting our understanding of its systematics and biogeographic history. In this study, we used, for the first time, data from molecular markers to propose a new phylogenetic hypothesis for the genus. More specifically, we aim at resolving the following questions: 1) Which are the phylogenetic relationships among the four lineages found by Ramos-Fregonezi *et al.*, (in prep.)? 2) Are species occurring in neighboring regions distantly related, as suggested by morphological data? 3) Which is time-scale for the evolutionary history of these taxa?

Material and Methods

Sampling

We analysed four described species (*C. decemmaculatus, C. holopteros, C. hypselurus,* and *C. septentrionalis*), as well as specimens representative of the four new species found in Southern

Brazil Highlands (SBH). We refer to these species as *Cnesterodon* sp. nov. 1, *Cnesterodon* sp. nov. 2, *Cnesterodon* sp. nov. 3 and *Cnesterodon* sp. nov. 4. These species include both *C. brevirostratus* and *Cnesterodon* sp. B in Lucinda (2005), but neither species can be directly related to any of these four lineages. Ongoing analysis of type specimens of *C. brevirosratus* has proved the species has been described based on more than one of the species recognized herein. Vouchers, geographical coordinates, and GenBank accession numbers for all individuals used in this study are available in Table 2. We conducted field expeditions in 2012-2013 to collect samples of *C. carnegiei*, *C. iguape*, *C. ormorgmatos*, *C. pirai* and *C. raddai* for DNA analyses, but we were unsuccessful.

Laboratory protocols

Total genomic DNA from muscle tissue preserved in 95% ethanol was extracted with cetyltrimethyl ammonium bromide (CTAB), as described in Doyle & Doyle (1987). We used PCR to amplify two mitochondrial genes: cytochrome oxidase subunit I (COI) and NADH dehydrogenase 2 (ND2); and one nuclear gene: SH3 and PX domain containing 3-like protein (SH3PX3). We selected these genes to include both slower and faster evolving regions in order to maximize the likelihood of obtaining phylogenetic resolution of recent as well as more ancient cladogenic events. PCR reactions were conducted following protocols and primers described by Herbert *et al*. (2003) (COI), Sorenson (2003) (ND2) and Li *et al*. (2007) (SH3PX3). The PCR products were verified on a 1% agarose gel, purified with Exonuclease I e Shrimp Alcaline Phosphatase (GE Healthcare®) and sent to Macrogen (Macrogen Inc., Seoul, South Korea) for sequencing.

Data analysis

Nucleotide sequences were determined on both strands, checked and aligned with GeneiousPro v. 4.8 (Drummond *et al*., 2007). Since different mitochondrial segments do not recombine, the two mitochondrial markers (COI and ND2) were combined in a single concatenated

set. We used the Dambe v. 5.6.9 software (Xia, 2013) in order to check the quality of partitions (mitochondrial - mtDNA and nuclear - nDNA) by plotting the saturation of transversions and transitions *versus* genetic distance. The best model of sequence evolution were determined by Akaike Information Criterion in jModelTest 2 (Darriba *et al.,* 2012) for each partition.

Coalescent-based species-tree analyses were performed using *BEAST in BEAST v1.8 (Drummond *et al*., 2012). This method treats gene-trees as random variables from a statistical distribution by estimating population parameters related to the history of the species, at a level above gene, integrating gene-trees as parameters (probability of various species trees candidates, given the genealogy of gene trees). *BEAST estimates separate gene trees while simultaneously estimating the species tree that generated them. Importantly, this method explicitly accounts for uncertainty in the individually estimated gene trees, and allows the separate (but correlated) gene tree estimates to influence each other throughout the analysis (Heled & Drummond, 2010). We also performed a Bayesian tree under a Yule tree-prior concatenating mtDNA and SH3PX3 markers to compare the topology of both methods. We used both strict and relaxed clock (under the uncorrelated lognormal distribution) models to estimate topology and divergence times (TMRCA) simultaneously. All analyses were run with 50 million generations of the Markov-Chain Monte Carlo (MCMC) sampling every 1,000 steps in order to achieve an effective sample size (ESS) of at least 200 for all parameters after removing the first 10% of the chain as burnin. At least three replicate runs were conducted for each configuration to check for consistency. For all analyses, we calibrated the mitochondrial molecular clock with a substitution rate of $8.6\pm0.1\text{x}10^{-9}$ substitutions per site per year (described for mitochondrial genome in Cyprinodontiformes (Hrbek & Meyer, 2003)) and we used the mtDNA rate as a reference for calibrating the nuclear molecular clock for SH3PX3. We used Bayes Factors (Kass & Raftery 1995) to evaluate the best molecular clock model (strict-clock vs relaxed-clock). The Maximum Clade Credibility (MCC) tree was summarized with TreeAnnotator, which is distributed alongside the BEAST package. Only nodes for which the posterior probability was higher than 0.60 (PP>0.6) were annotated for TMRCA information.

Results

A total of 28 individuals representing 8 taxa were sampled (Table 2 and Fig. 2). The mitochondrial concatenated data set resulted in 1634 base pairs (bp), 644bp for COI and 990bp for ND2, totalling 328 variables sites and 212 parsimoniously informative sites. The nuclear data set for the SH3PX3 gene had 712bp, 26 variables sites and 11 parsimoniously informative sites. The amplification of nuclear segment was unsuccessful for some individuals (Table 2). Analysis of transversions/transitions *versus* genetic distance showed no evidence of saturation for both partitions. Based on the Akaike Information Criterion scores, we used the GTR+G and HKY+I substitution models for mtDNA and nDNA, respectively. The best molecular clock model for both species-tree and concatenated analyses was the strict-clock (log10 Bayes Factor ~1.8 against the relaxed clock model for species tree analysis, and log10 Bayes Factor ~1 for the concatenated analysis).

Phylogenetic trees using both the species-tree (Fig. 3) or concatenated (Fig. 4) approaches were very similar to each other and yielded trees with overall strong statistical support. *Cnesterodon hypselurus*, from the Paraná basin, was sister to all remaining species, followed by *C. septentrionalis*, from the Araguaia basin, which is sister to the remaining taxa. Species from Uruguay/Southern Brazil formed a clade in which the two Pampean species, *C. decemmaculatus* and *C. holopteros*, are sister to each other, while the four new species found in SBH formed a well-supported monophyletic group. However, the species-tree analysis resulted in relatively low values for the internal relationships among SBH lineages. The concatenated analysis further suggested that while some of the new species may represent a more limited set of lineages, *Cnesterodon* sp. nov 4, which has a broader geographic distribution, had lineages scattered throughout the SBH clade (Fig. 4) which are paraphyletic in regard to the lineages found for the other species. In general, divergence times were consistent in both analyses. The TMRCA for the genus *Cnesterodon* was around 8 millions of years ago (Table 3). The divergence between *C. septentrionalis* and the

southern species was around 5.5 Mya, while the divergence between Pampean and SBH species occurred around 4 Mya. Considering the species-tree approach (Fig. 3) the divergence time for SBH species was similar to that found between the Pampean species pair (*C. decemmaculatus* and *C. holopteros*), with point estimates of ~1.9 Mya and ~1.6 Mya, respectively. On the other hand, because *Cnesterodon* sp. nov. 4 had lineages paraphyletic in regard to the remaining SBH species, the TMRCA of this group (~3.7 Mya) in the concatenated tree reflects the antiquity of these lineages rather than divergence times between species.

Discussion

Notes on the distribution of Cnesterodon *species*

We conducted field expeditions to collect *C. carnegiei*, *C. iguape*, *C. omorgmatos*, *Cnesterodon pirai* and *C. raddai* during 2010-2013, with no success. These species, mostly from the Paraná/Paraguay basin (with the exception of *C. iguape*), have restricted geographic distributions, occurring in habitats that have been seriously degraded due to pasture, large monoculture areas, soil and water contamination by agricultural pesticides, introduction of exotic faunal and increasing urbanization. Three of these taxa (*C. iguape, C. omorgmatos, and C. pirai*) are known from only one or two localities, and the known distribution of the remaining taxa is also restricted to a few localities (see Lucinda *et al.*, 2006, for example). *Cnesterodon carnegiei, C. iguape and C. omorgmatos* are considered threatened or vulnerable in Brazil (ICMBio, 2014), but the results from our field expedition may indicate that these species may have a rapidly deteriorating status and deserve further attention.

The SBHS clade is distributed in Santa Catarina and Rio Grande do Sul states (Brazil). *Cnesterodon* sp. nov. 1, *Cnesterodon* sp. nov. 2 and *Cnesterodon* sp. nov. 3 occur in both Laguna dos Patos and Tramandaí-Mampituba drainages, but *Cnesterodon* sp. nov. 1 is found exclusively inside an area of environmental protection (Parque Nacional do Aparados da Serra). *Cnesterodon* sp. 4 species is the most widespread, occurring in Laguna dos Patos, Tramandaí-Mampituba and

Uruguay drainages (Fig. 2), which may explain, at least in part, the number of divergent genetic lineages found in this species (Fig. 4).

We found two new occurrences for *C. holopteros*: at rio Tacuarembó (Tacuarembó province - Uruguay) and at Laguna del Sauce (Maldonado province - Uruguay) (Table 2). This species was previously known only at the Cuarein river basin (Artigas province – Uruguay). The presence of this species in Northern (Artigas), Central (Tacuarembó), and Southern Uruguay indicates the species' widespread distribution in Uruguay, where the species may have been misidentified as *C. decemmaculatus*.

Phylogenetic relationships and divergence times

Previous phylogenetic studies in *Cnesterodon* were based on morphological characters (Lucinda, 2005; Lucinda *et al*., 2006, Aguilera, 2009). The present study represents the first phylogenetic hypothesis for this genus based on molecular data (including both mtDNA and nDNA markers), and show a sharp contrast between these two sets of data. For example, in our phylogenetic hypothesis, *C. hypselurus* was sister to all other species. On the other hand, the morphological hypothesis of Lucinda *et al.* (2006) suggested that *C. hypselurus*, *C. brevirostratus*, and *C. septentrionalis* (considering the taxa sampled by us) share a set of four synapomorphies relative to *Cnesterodon* sp. n. B and the two Pampean species. However, only one of these character was unreversed in other branches. This illustrates that there are relatively few phylogenetic informative characters among *Cnesterodon* species compared to the number of morphological synapomorphies defining the genus (13 vs. 38 character states, or 1 vs. 8 unreversed states, respectively, in Lucinda *et al.*, 2006). As a consequence, each clade has been defined based on few characters, and even a few reversals would affect phylogenetic estimation.

Another important difference between datasets was related to the phylogenetic relationship among SBH species. While *Cnesterodon* sp. nov. B and *C. brevirostratus* are unrelated in the morphological analysis (Lucinda *et al.*, 2006), the molecular phylogeny strongly indicates that all

evolutionary lineages occurring in SBH represent a monophyletic group. If this result is correct, it would imply in a rapid rate of morphological evolution in these lineages, which may complicate phylogenetic scenarios based on morphological data. Nonetheless, a thorough morphological characterization of these four lineages is needed to evaluate how this data fits the phylogenetic hypothesis proposed in this study.

Divergence time estimates suggested dates in the Pliocene with exception of the divergences between *C. decemmaculatus* and *C. holopteros*, the two Pampean species, and the radiation of the SBH lineages, which in both cases were between 1.5-2.0 Mya. However, even for the SBH clade, divergence among lineages within species may be as old as 3.7 Mya (Fig. 4). Taken together, these dates suggest that the orogenic events in the Pliocene were more important for the diversification of this genus than the more recent occurred in the Pleistocene processes (Turchetto-Zolet *et al.*, 2013). Our analysis suggested a TMRCA for the whole genus in the Miocene, around 8 Mya (between 5.2 and 10 Mya). This is in agreement with the date suggested Meredith *et al.* (2010), who dated the divergence between *C. decemmaculatus* and *C. hypselurus* around 9.7 Mya (between 7.6 and 11.8 Mya) based on the concatenated analysis of two mitochondrial and seven nuclear markers. The slightly younger TMRCA for *Cnesterodon* found by us may be due to the fact that Meredith et al. (2010) used more slowly evolving markers. The TMRCA for *Cnesterodon* could be even older, since we were unable to include five species in our study (*C. omorgmatus*, *C. carnegiei*, *C. pirai* and *C. raddai* - all of them endemic to drainages from the Paraná-Paraguay system and *C. iguape*, which is endemic to a coastal drainage in Southeastern Brazil neighbour to the Paraná basin). Similarly, the absence of these species could also affect our phylogenetic hypothesis. However, the high congruence found in this study among phylogenetic relationships, geographical distribution and hydrologic basins make this possibility unlikely, especially considering that *Cnesterodon* species are poor swimmers with reduced active long dispersal abilities (Trenti *et al.*, 1999). Nonetheless, it is important to stress that our phylogenetic hypothesis differs from the one based on morphological data even excluding unsampled taxa, including the sister relationship between the

Pampean clade and the SHBS clade, and the sister relationships within each of these clades, even though this latter pattern may be due to incomplete taxon sampling (see next section).

Biogeography and evolution of Cnesterodon

The sister relationship between *C. hypselurus* (from the Paraná basin) and the other *Cnesterodon* species may suggest that the ancestor lineage may have come from the Paraná Basin, diverging in the Late Miocene (around 8 Mya). At this time, two remarkable geologic events affected South America. The first was the retraction of the "Paranean Sea", which was caused by a previous marine transgression that flooded part of Argentina, Uruguay, Southern Brazil, and Southern portions of Bolivia and Paraguay (Lundberg *et al*., 1998; Brea & Zucol, 2011). As a consequence of the increased salinity, freshwater species experienced extirpation due to contractions of habitat that greatly reduced the effective population sizes of species that persisted, thus increasing genetic isolation levels (Albert & Reis, 2011). However, if *Cnesterodon* ancestral populations came from upper regions of the Paraná Basin, they probably were not affected by this marine transgression. Nonetheless, when this intra-continental “seaway” retracted, new habitats were available and could be colonized (Pascual *et al*., 1996; Ortiz-Jaureguizar & Cladera, 2006) permitting these populations to reach northern and southern drainages. The second major geological event was the later phase of the rise of the Andes (Garzione *et al.*, 2008), which affected not only the western portion of South America, but also caused the uplift of the Brazilian shield and, consequently, may have affected river drainages allowing for headwater capture (Simpson, 1979; Hubert & Renno, 2006). Interestingly, Hubert & Renno (2006) have suggested that high elevation portions of the Paraná basin is a major area of endemism that may have served as a refuge area for Characiformes during the Miocene, which is in agreement to the occurrence of several *Cnesterodon* species in this region.

The divergence of *C. septentrionalis* in the Late Miocene/Early Pliocene, around 5.5 Mya, is coherent with this scenario in which both the reconfiguration of the hydrological basins due to

orogenic factors. It seems reasonable to suppose that populations in the Paraná basin would have reached the Tocantins-Araguaia basin, to the north, following some stream capture event and posteriorly experiencing a long term of isolation, giving rise to *C. septentrionalis* (Fig. 5).

According to our phylogeny, Pampean and SBH species form two sister clades. Below, we propose two hypotheses compatible with this phylogenetic arrangement, and a third scenario that accounts for unsampled taxa in our phylogeny. In the first scenario, the origin of the SBH clade would be related to the orogenic events that persisted during the Early Pliocene and allowed SBH colonization across the Brazilian Plateau and the Paraná Basin, up to the upper portion of the Uruguay Basin. Then, an ancestral SBH species would have dispersed to the lower portions of the Uruguay Basin, giving rise to the Pampean clade. Alternatively, the orogenic events before mentioned, as well as the increasing habitat availability following the end of the Paranean sea would explain the dispersal of the ancestral of *C. decemmaculatus* and *C. holopteros* to the Uruguay Basin. The Paraná drainage flows to the southwest until it reaches the La Plata estuary, between Argentina and Uruguay. Therefore, this ancestral Pampean lineage could easily reach the Uruguay basin from the Paraná drainage. Both scenarios presented above are consistent with the sister relationship between the Pampas and SBH clades, differing in the original colonization route. It is likely that Pampean drainages have been associated with grassland ecosystems for a long period, since studies of past climate in the Pampas indicate that grasslands were dominant in both glacial and interglacial periods (Tonni *et al.*, 1999; Behling *et al.*, 2005). A third hypothesis is shown in Fig. 5. In this scenario, two independent dispersals would have occurred from the Paraná basin. One originating the Pampean clade biome and the other originating the SBH. If this scenario is correct, the sister relationship found for these clades results from incomplete taxon sampling, either because we did not include all species in the analysis, or because these taxa are now extinct. Subsequently to the colonization of the Pampas and SBH, more recent speciation events took place during the Pleistocene. Regarding to SBH clade, the great *Cnesterodon* diversity found in this region can be attributed to its physiognomic characteristics, with shallow wetlands scattered through the

landscape and draining to the upper tributaries of four different river drainages (Patos, Uruguay, Tramandaí and Mampituba) that may provide opportunities for population isolation which can cause speciation in the long term.

Acknowledgments

We would like to thank Marcelo Loureiro (Universidad de la República, Uruguay), Cláudio Oliveira (Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP, Brazil), Juliana Wingert, Juliano Ferrer and Rafael Angrizani (Universidade Federal do Rio Grande do Sul, Brazil) for help with sampling. This study was supported by Conselho Nacional de desenvolvimento Científico (CNPq), Programa de Pós-Graduação em Biologia Animal (PPGBAN-UFRGS) and Programa de Pós-Graduação em Genética e Biologia Molecular (PPGBM-UFRGS), Brazil.

References

Aguilera, G., J. M. Mirande & M. M. Azpelicueta. 2009. A new species of *Cnesterodon* (Cyprinodontiformes: Poeciliidae) from a small tributary of arroyo Cuñá-Pirú, río Paraná basin, Misiones, Argentina. *Zootaxa*, *2195*, 34-42.

Albert, J. S. & R. E. Reis. 2011. Introduction to neotropical freshwaters. Pp 3-19. In: Albert, J. S. & R. E. Reis (Eds.). Historical Biogeography of Neotropical Freshwater Fishes. Berkley, University of California Press.

Arias, A. & D. Reznick. 2000. Life history of *Phalloceros caudiomaculatus*: a novel variation on the theme of livebearing in the family Poeciliidae. Copeia, 3: 792–798.

Behling, H., V. Pillar & S. G. Bauermann. 2005. Late Quaternary grassland (Campos), gallery forest, fire and climate dynamics, studied by pollen, charcoal and multivariate analysis of the São Francisco de Assis core in Western Rio Grande do Sul (Southern Brazil). Review of Palaeobotany and Palynology, 133: 235-248.

Brea, M., A. F. Zucol. 2011. The Paraná-Paraguay Basin: Geology and Paleoenviroments. Pp. 69-87. In: Albert, J. S. & R. E. Reis (Eds.). Historical Biogeography of Neotropical Freshwater Fishes. Berkley, University of California Press.

Darriba, D., G. L. Taboada, R. Doallo & D. Posada. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9, 8: 772-772.

Doyle, J. J. & J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull, 19: 11-15.

Drummond, A. J., B. Ashton, M. Cheung, J. Heled, M. Kearse, R. Moir, S. Stones-Havas, T. Thierer & A. Wilson. 2007. Geneious v4.8. Biomatters Ltd., Auckland, New Zealand. Available at: http://www.geneious.com.

Drummond, A. J., M. A. Suchard, D. Xie & A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7 Molecular Biology And Evolution, 29: 1969-1973.

Endler, J. A. 1983. Natural selection on color patterns in poeciliid fishes. Environmental Biology Fishes, 9: 173–190.

Garzione, C. N., G. D. Hoke, J. C. Libarkin, S. Withers, B. MacFadden, J. Eiler, P. Ghosh & A. Mulch. 2008. Rise of the Andes. Science, 320, 1304–1307.

Grove, B. D. & J. P. Wourms. 1991. The follicular placenta of the viviparous fish, *Heterandria formosa*. I. Ultrastructure and development of the embryonic absorptive surface. Journal of Morphology, 209: 265–284.

Grove, B. D. & J. P. Wourms. 1994. Follicular placenta of the viviparous fish, *Heterandria formosa*: II. Ultrastructure and development of the follicular epithelium. Journal of Morphology, 220: 167–184.

Hamilton, A. 2001. Phylogeny of *Limia* (Teleostei:Poeciliidae) based on NADH dehydrogenase subunit 2 sequences. Molecular Phylogeneticis and Evolution, 19: 277–289.

Heled, J. & A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution, 27: 570-580.

Herbert, P. D. N., A. Cywinska, S. Ball, & J. R. Dewaard. 2003. Biological identifications through DNA barcodes. Proceedings Royal of the Society of London B, 270: 313-321.

Houde, A. E. 1997. Sex, Color, and Mate Choice in Guppies. Princeton University Press, Princeton, NJ.

Hrbek, T., & A. Meyer. 2003. Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). Journal of Evolutionary Biology, 16: 17–36.

Hubert, N. & J. F. Renno. 2006. Historical biogeography of South American freshwater fishes. Journal of Biogeography, 33: 1414–1436.

ICMBio. 2014. Instituto Chico Mendes, Ministério do Meio Ambiente, Brazil. Available from: http://www.icmbio.gov.br/portal/biodiversidade/fauna-brasileira/lista-de-especies.html (25 July 2015).

Kass, R. E. & A. E. Raftery. 1995. Bayes Factors. Journal of the American Statistical Association, 90: 773–795.

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

Lucinda, P.H.F., 2003. Family Poeciliidae. In: Reis, R.E., Kullander, S.O., Ferraris, C.J. (Eds.), Check List of the Freshwater Fishes of South and Central America. EDIPUCRS, Porto Alegre, Brazil, pp. 555–581.

Lucinda, P. H. F. 2005. Systematics of the genus *Cnesterodon* Garman, 1895 (Cyprinodontiformes: Poeciliidae: Poeciliinae). Neotropical Ichthyology, 3: 259-270.

Lucinda, P. H. F. & R. E. Reis. 2005. Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. Neotropical Ichthyology, 3: 1–60.

Lucinda, P. H. F., L. Thomas & R. Roberto. 2006. *Cnesterodon holopteros* (Cyprinodontiformes: Poeciliidae: Poeciliinae), a new species from the Republic of Uruguay. Zootaxa, 1350: 21-31.

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

Malabarba, L. R., C. B. Fialho, J. A. Anza, J. F. Santos & G. N. Mendes. 2009. Fauna aquática: peixes. In: Boldrini II (org.) Biodiversidade dos campos no planalto das araucárias. Brasília: Ministério do Meio Ambiente. pp. 131-141.

Meredith, R. W., M. N. Pires, D. N. Reznick, & M. S. Springer. 2010. Molecular phylogenetic relationships and the evolution of the placenta in *Poecilia* (Micropoecilia) (Poeciliidae: Cyprinodontiformes). Molecular Phylogenetics and Evolution, 55:631-639.

Nakajima M., & N. Taniguchi. 2001. Genetics of the guppy as a model for experiment in aquaculture. Genetica 111: 279-289.

Ortiz-Jaureguizar, E. & G. A. Cladera. 2006. Paleoenvironmental evolution of southern South America during the Cenozoic. Journal of Arid Environments, 66: 498-532.

Pascual, R., E. Ortiz-Jaureguizar, J.L. Prado. 1996. Land mammals: paradigm of Cenozoic South American geobiotic evolution. Pp 265-319 In: Arratia, G. (Ed.). Contribution of Southern South America to Vertebrate Paleontology, Müncher Geowissenschaftliche Abhandlungen

Reznick, D. N., M. Mateos & M. S. Springer. 2002. Independent origins and rapid evolution of the placenta in the fish genus *Poeciliopsis*. Science, 298: 1018–1020.

Reznick, D., T. Hrbek, S. Caura, J. de Greef & D. Roff. 2007. Life history of *Xenodexia ctenolepis*: implications for life history evolution in the family Poeciliidae. Biological Journal of the Linnean Society, 92: 77–85.

Rosa, R. S. & W. J. E. M. Costa. 1993. Systematic revision of the genus *Cnesterodon* (Cyprinodontiformes: Poeciliidae) with the description of two new species from Brazil. Copeia, 3: 696-708.

Schluter, A., J. Parzefall & I. Schlupp. 1998. Female preference for symmetrical vertical bars in male sailfin mollies. Animal Behaviour, 56: 147–153.

Shen X., G. Yang, Y. Liu, M. Liao, X. Wang, M. Zhu, W. Song, G. Zou, Q. Wei, D. Wang & D. Chen. 2007. Construction of genetic linkage maps of guppy (*Poecilia reticulata*) based on AFLP and microsatellite DNA markers. Aquaculture, 271: 178-187.

Simpson, B. B. 1979. Quaternary biogeography of the high montane regions of South America . In : The South American Herpetofauna: Its Origin, Evolution, and Dispersal (ed. Duellman, W. E.), pp.157-188. Monograph of the Museum of Natural History, University of Kansas.

Sorenson, M. D. 2003. Avian mtDNA primers. Disponível em: http://people.bu.edu/msoren/ (Bird.mt.Primers.pdf.)

Tonni, E. P., A. L. Cione & A. J. Figini. 1999. Predominance of arid climates indicated by mammals in tha pampas of Argentina during the late Pleistocene and Holocene. Palaeogeography, Palaeoclimatology, Palaeoecology, 147: 257-281.

Trenti, P. S., S. E. Gómez & R. A. Ferriz. 1999. Capacidad de natación en tres peces pampásicos. Aprona, 38: 2-9.

Turchetto-Zolet, A. C., F. Pinheiro, F. Salgueiro & C. Palma-Silva. 2013. Phylogeographical patterns shed light on evolutionary process in South America. Molecular Ecology. 22: 1193–1213.

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

Table 1. *Cnesterodon* species and its geographical distributions (according to Lucinda *et al.*, 2006)

Species[1]	Drainage System	Location
Cnesterodon brevirostratus Rosa & Costa, 1993	Uruguay, Laguna dos Patos, Tramandaí-Mampituba	Rio Grande do Sul and Santa Catarina states, Brazil
Cnesterodon sp. nov. B	Uruguay, Laguna dos Patos, Tramandaí-Mampituba	Rio Grande do Sul state, Brazil
Cnesterodon carnegiei Haseman, 1911	Paraná	Paraná state, Brazil
Cnesterodon decemmaculatus (Jenyns, 1842)	Uruguay, Laguna dos Patos	Rio Grande do Sul state, Brazil; widespread over Uruguay
Cnesterodon holopteros Lucinda, Litz & Recuero, 2006	Uruguay, Laguna dos Patos	Artigas, Paysandú, Tacuarembó and Maldonado departments, Uruguay
Cnesterodon hypselurus Lucinda & Garavello, 2001	Paraná	Paraná state, Brazil
Cnesterodon iguape Lucinda, 2005	Ribeira do Iguape	São Paulo state, Brazil
Cnesterodon omorgmatos Lucinda & Garavello, 2001	Paraná	Paraná state, Brazil
Cnesterodon pirai Aguilera Mirande & Azpelicueta, 2009	Paraná	Misiones province, Argentina
Cnesterodon raddai Meyer & Etzel, 2001	Paraná and Paraguay	Misiones and Ñeembucú departments, Paraguay; Chaco province, Argentina
Cnesterodon septentrionalis Rosa & Costa, 1993	Tocantins-Araguaia	Mato Grosso, Brazil

[1]In this study, we considered that the individuals associated to *C. brevirstratos* and *Cnesterodon* sp. nov. B are actually part of four independent evolutionary lineages (*Cnesterodon* sp. nov. 1, *Cnesterodon* sp. nov. 2, *Cnesterodon* sp. nov. 3, and *Cnesterodon* sp. nov. 4), according to Ramos-Fregonezi *et al.* (in prep.).

Table 2. *Cnesterodon* specimens included in this study

Species	Voucher	Latitude	Longitude	Genbank accession number		
				COI	ND2	SH3PX3
C. decemmaculatus	UFRGS 13811/TEC 0087 B	-32.0528	-57.6719	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. decemmaculatus	UFRGS 13812 /TEC 0095 C	-31.6222	-57.8808	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. decemmaculatus	UFRGS 14640 /TEC 0601 C	-30.4689	-57.5122	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. hypselurus	UFRGS 18402/TEC 3816 A	-24.3327	-49.7908	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. holopteros	UFRGS 13912 /TEC 0138 E	-30.5319	-57.6658	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. holopteros	UFRGS 13912 /TEC 0138 L	-30.5319	-57.6658	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. holopteros	UFRGS 14549/TEC 00395 D	-31.4761	-57.9017	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. holopteros[*]	UFRGS 14637/TEC 0593 A	-31.9758	-55.4703	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. holopteros[*]	UFRGS 17870/TEC 3548 C	-34.8416	-55.0961	XXXXXXXX	XXXXXXXX	XXXXXXXX
C. septentrionalis	LBP2003050602/12563*	-17.5618	-53.3082	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 1	UFRGS 6849/TEC 0501 A	-29.1792	-50.1369	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 1	UFRGS 16685/TEC 3002 A	-29.1729	-50.1033	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 1	UFRGS 16693 /TEC 3010 F	-29.1743	-50.1168	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 1	UFRGS 16694 /TEC 3011 C	-29.1571	-50.0510	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 2	UFRGS 14646/TEC 0611 C	-29.3219	-50.4408	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 2	UFRGS 16686/TEC 3003 B	-29.1729	-50.1033	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 3	UFRGS 6843 /TEC 0005 A	-29.2097	-50.2394	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 3	UFRGS 6846/TEC 0556 D	-29.3939	-50.4139	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 3	UFRGS 16682 /TEC 2999 B	-29.1586	-50.0804	XXXXXXXX	XXXXXXXX	XXXXXXXX

Cnesterodon sp. nov. 3	UFRGS 16690 /TEC 3007 A	-29.1726	-50.1121	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 14606 /TEC 0530 E	-29.2450	-50.6294	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 14606 /TEC 0530 S	-29.2450	-50.6294	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 6839/TEC 0532 B	-28.9239	-50.0669	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 14612 /TEC 0540 B	-29.1828	-50.6856	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 6846 /TEC 0556 B	-29.3939	-50.4139	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 14625/TEC 0562 A	-28.4503	-49.6569	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 14629/TEC 0574 K	-29.2369	-50.6308	XXXXXXXX	XXXXXXXX	XXXXXXXX
Cnesterodon sp. nov. 4	UFRGS 14629/TEC 0574 N	-29.2369	-50.6308	XXXXXXXX	XXXXXXXX	XXXXXXXX

* New records found in this work.

Table 3. Time to the most recent common ancestor (TMRCA) for clades (in millions of years)

Clade	Median	95% HPD[1]
Cnesterodon root	7.79	5.27 – 9.93
C. septentrionalis + SBH[2] + Pampean[3]	5.39	3.97 – 6.80
SBH[2] + Pampean[3]	4.06	2.97 – 5.06
SBH[2]	1.90	1.10 – 2.54
Pampean[3]	1.56	0.85 – 2.29

[1] 95% HPD – 95% Highest Posterior Density; [2]*Cnesterodon* sp. nov. 1 + *Cnesterodon* sp. nov. 2 + *Cnesterodon* sp. nov. 3 + *Cnesterodon* sp. nov. 4; [3] *C. decemmaculatus* + *C. holopteros*;

Figure Legends

Fig. 1. Phylogenetic hypotheses for *Cnesterodon* based on morphologic data. A) Lucinda 2005. B) Lucinda *et al.*, 2006. C) Aguilera *et al.*, 2009 based on equal weighting. D) Aguilera *et al.*, 2009 based on implied weighting (K=3-10). Numbers above and below branches in panels C) and D) represent support values based on GC (groups present/contradicted) or relative Bremmer support, respectively. There are no support values in the phylogenies represented in panels A) and B).

Fig. 2. Distribution of the *Cnesterodon* species sampled in this study. See Table 2 for more details.

Fig. 3. Time-calibrated Bayesian coalescent species-tree. Numbers above the branches represent posterior probabilities values (PP). The grey bars represent the 95% highest posterior density (95% HPD) for the time of the most recent ancestor (TMRCA) for each node. Time is given in millions of years (Mya). Colors of each *taxa* are in agreement to Fig. 1. Names in the right representing the general location of each group.

Fig. 4. Time-calibrated Bayesian concatenated tree. Numbers above the branches represent posterior probabilities values (PP). The grey bars represent the 95% highest posterior density (95% HPD) for the most recent ancestor (TMRCA) for each node. Time is given in millions of years. Colours of each *taxa* are in agreement to Fig. 1. Names in the right representing the general location of each group.

Fig. 5. Biogeographical hypothesis for the distribution of *Cnesterodon* species. Colours indicate the main drainage systems according to the legend. Dashed arrows indicate hypotheses of dispersion inferred in this study. SBHS= Southern Brazilian Highlands species. Note that *Cnesterodon decemmaculatus* distribution is under-represented in this map.to show only the individuals used in this study.

Figure 1

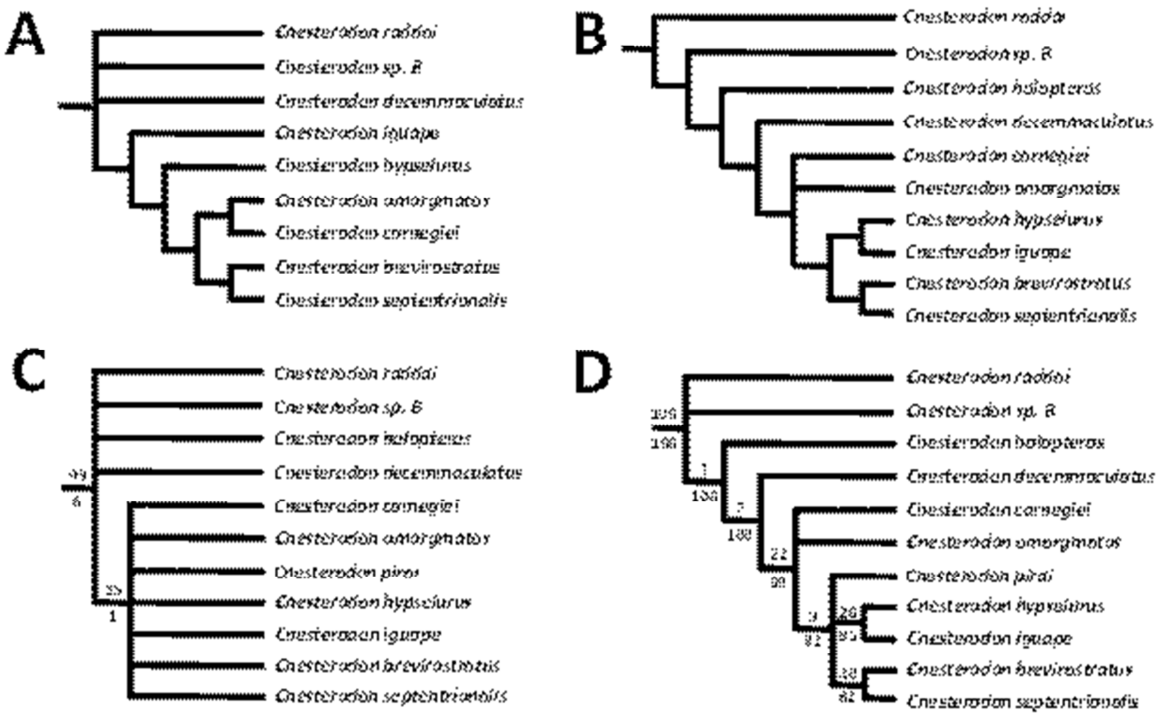

A
Cnesterodon raddai
Cnesterodon sp. B
Cnesterodon decemmaculatus
Cnesterodon iguape
Cnesterodon hypselurus
Cnesterodon emarginatus
Cnesterodon carnegiei
Cnesterodon brevirostratus
Cnesterodon septentrionalis
B
Cnesterodon raddai
Cnesterodon sp. B
Cnesterodon holopteros
Cnesterodon decemmaculatus
Cnesterodon carnegiei
Cnesterodon emarginatus
Cnesterodon hypselurus
Cnesterodon iguape
Cnesterodon brevirostratus
Cnesterodon septentrionalis
C
Cnesterodon raddai
Cnesterodon sp. B
Cnesterodon holopteros
Cnesterodon decemmaculatus
Cnesterodon carnegiei
Cnesterodon emarginatus
Cnesterodon piresi
Cnesterodon hypselurus
Cnesterodon iguape
Cnesterodon brevirostratus
Cnesterodon septentrionalis
D
Cnesterodon raddai
Cnesterodon sp. B
Cnesterodon holopteros
Cnesterodon decemmaculatus
Cnesterodon carnegiei
Cnesterodon emarginatus
Cnesterodon piresi
Cnesterodon hypselurus
Cnesterodon iguape
Cnesterodon brevirostratus
Cnesterodon septentrionalis

Figure 2

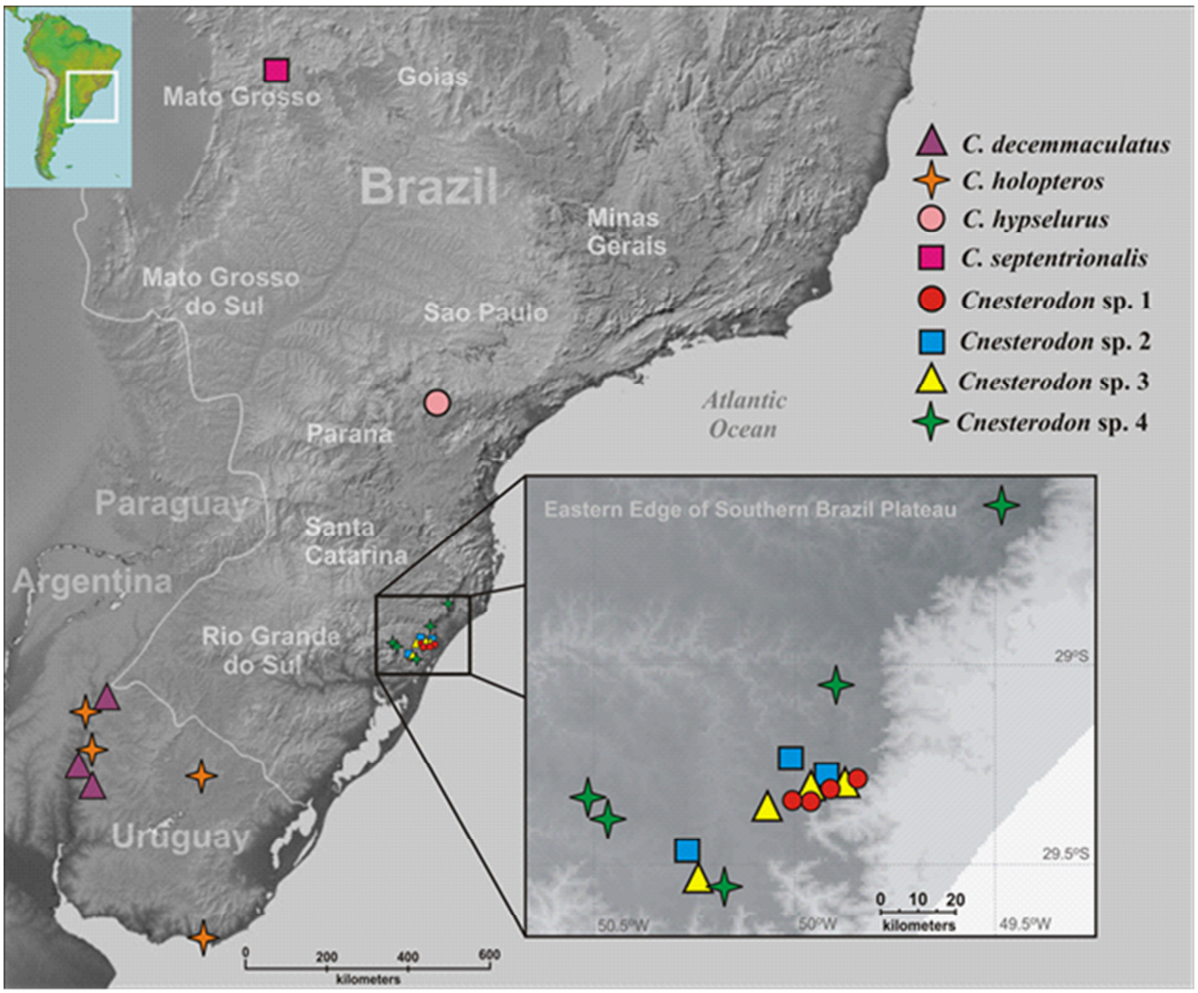

Mato Grosso
Goias
Brazil
Minas Gerais
Mato Grosso do Sul
Sao Paulo
Atlantic Ocean
Parana
Paraguay
Santa Catarina
Argentina
Rio Grande do Sul
Uruguay
C. decemmaculatus
C. holopteros
C. hypselurus
C. septentrionalis
Cnesterodon sp. 1
Cnesterodon sp. 2
Cnesterodon sp. 3
Cnesterodon sp. 4
Eastern Edge of Southern Brazil Plateau
29°S
29.5°S
50.5°W
50°W
49.5°W
0 10 20
kilometers
0 200 400 600
kilometers

Figure 3

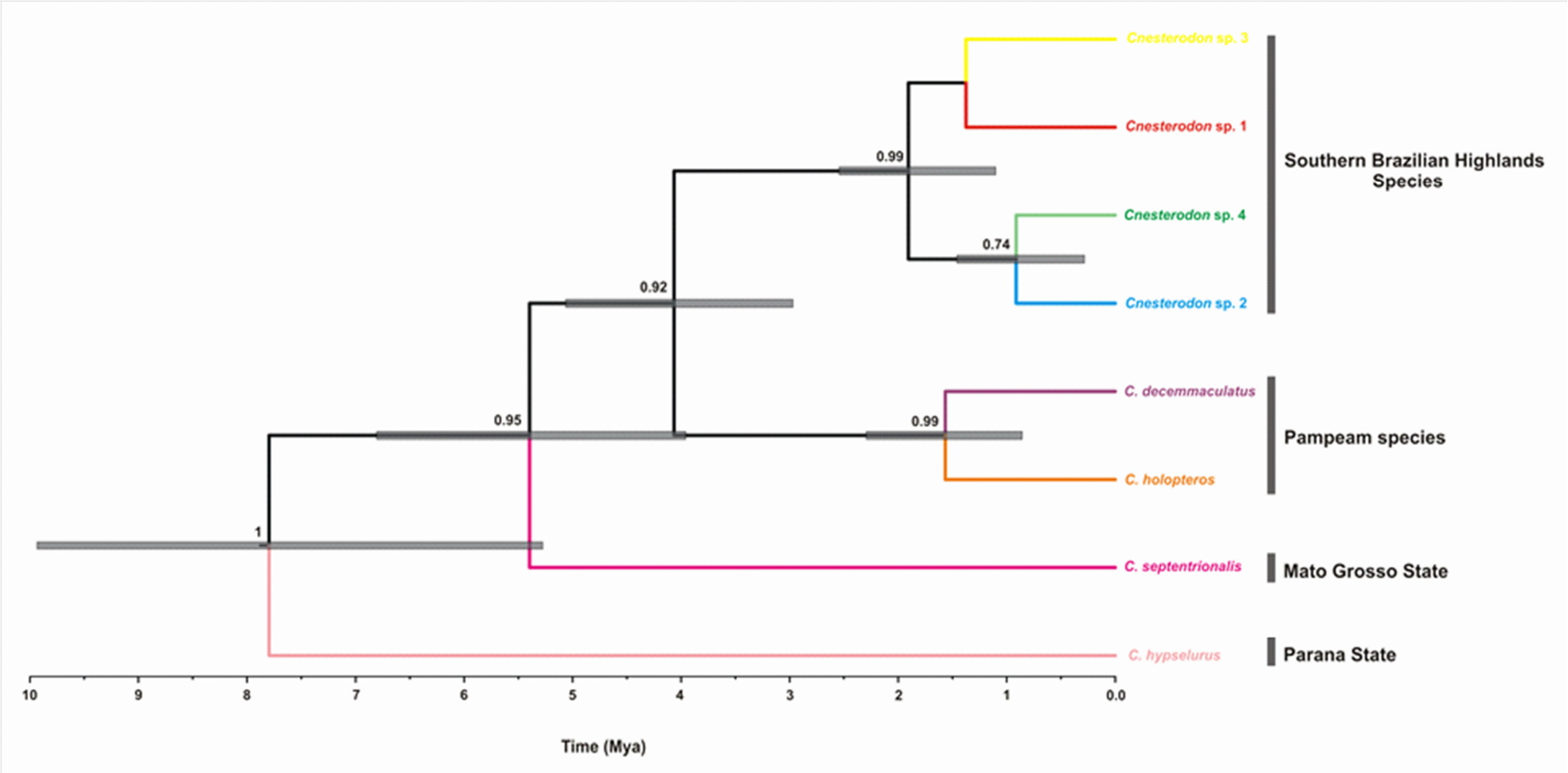

Cnesterodon sp. 3
Cnesterodon sp. 1
Cnesterodon sp. 4
Cnesterodon sp. 2
C. decemmaculatus
C. holopteros
C. septentrionalis
C. hypselurus
0.99
0.74
0.92
0.95
0.99
1
Southern Brazilian Highlands Species
Pampeam species
Mato Grosso State
Parana State
10
9
8
7
6
5
4
3
2
1
0.0
Time (Mya)

Figure 4

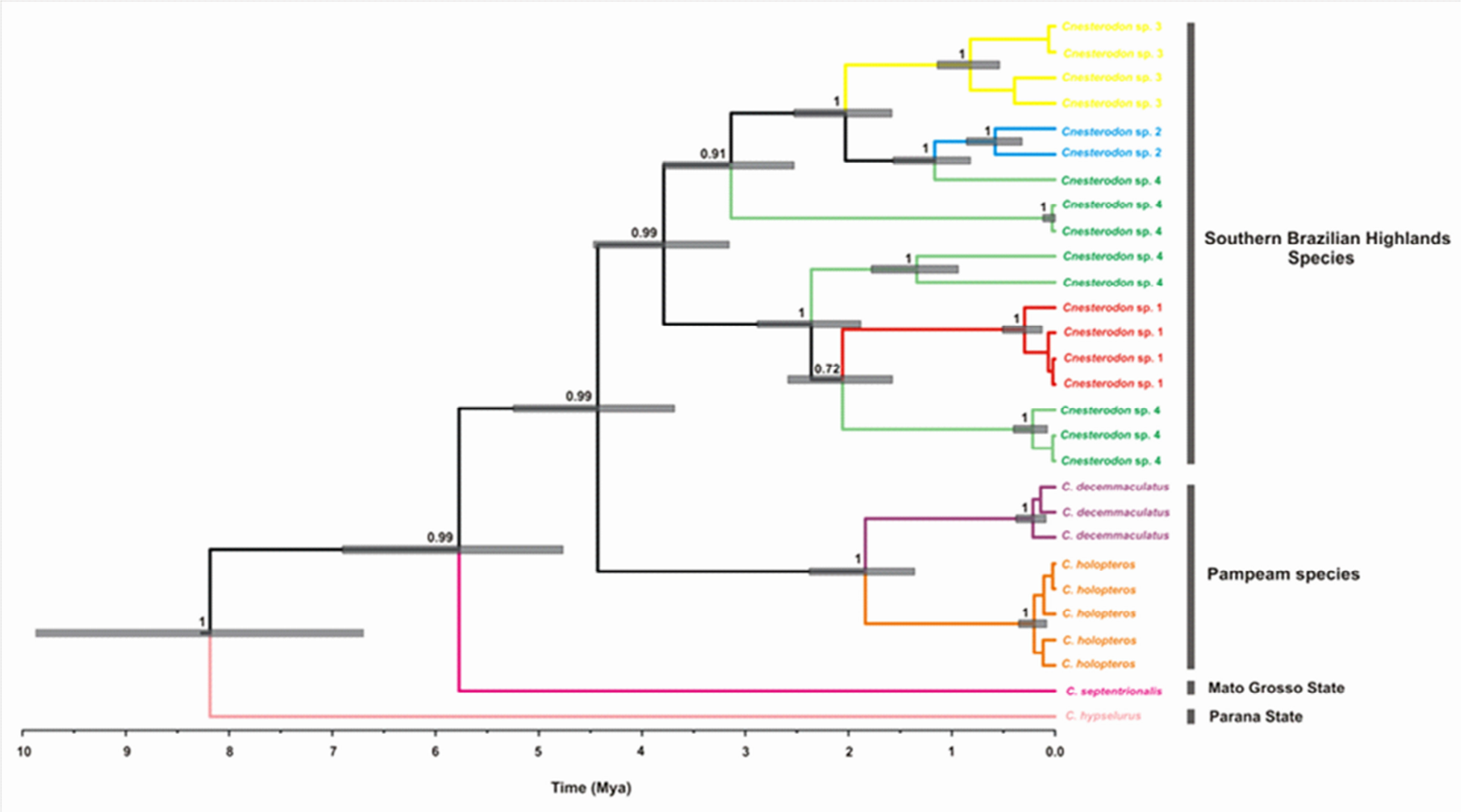
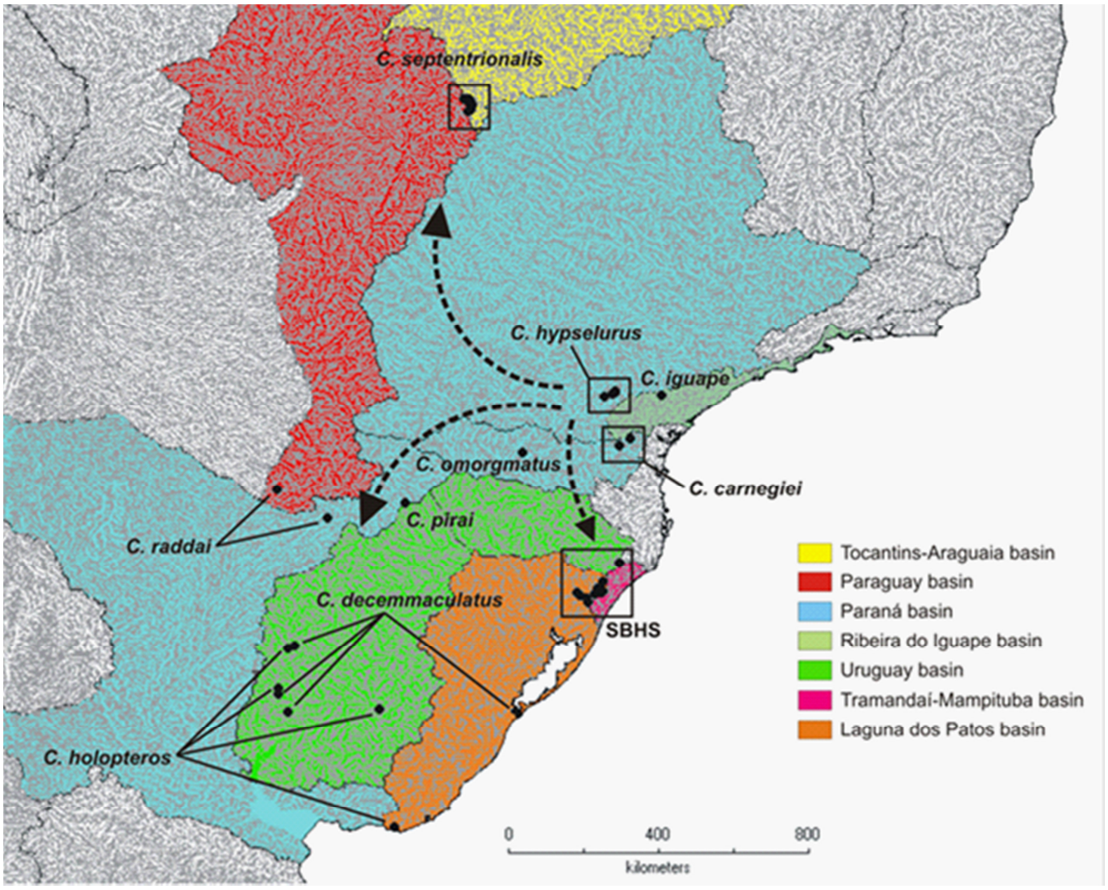

Cnesterodon sp. 3
Cnesterodon sp. 3
Cnesterodon sp. 3
Cnesterodon sp. 3
Cnesterodon sp. 2
Cnesterodon sp. 2
Cnesterodon sp. 4
Cnesterodon sp. 4
Cnesterodon sp. 4
Cnesterodon sp. 4
Cnesterodon sp. 4
Cnesterodon sp. 1
Cnesterodon sp. 1
Cnesterodon sp. 1
Cnesterodon sp. 1
Cnesterodon sp. 4
Cnesterodon sp. 4
Cnesterodon sp. 4
C. decemmaculatus
C. decemmaculatus
C. decemmaculatus
C. holopteros
C. holopteros
C. holopteros
C. holopteros
C. holopteros
C. septentrionalis
C. hypselurus
0.91
0.99
0.99
0.99
0.72
1
Southern Brazilian Highlands Species
Pampeam species
Mato Grosso State
Parana State
10
9
8
7
6
5
4
3
2
1
0.0
Time (Mya)

Figure 5


C. septentrionalis
C. hypselurus
C. iguape
C. omorgmatus
C. carnegiei
C. pirai
C. raddai
C. decemmaculatus
SBHS
C. holopteros
Tocantins-Araguaia basin
Paraguay basin
Paraná basin
Ribeira do Iguape basin
Uruguay basin
Tramandaí-Mampituba basin
Laguna dos Patos basin
0
400
800
kilometers

CAPÍTULO 5

Discussão Geral

A família Poeciliidae ou a subfamília Poeciliinae têm sido por um longo tempo objeto de interesse tanto para pesquisas cientificas quanto para aquaricultura. Estes grupos têm características peculiares que os tornam modelo para estudos de evolução, comportamento, ecologia, entre outros. Dentre estas características podemos citar: variabilidade na morfologia genital, fertilização interna, diferentes comportamentos de corte, e variação na coloração dos machos (Thibault e Schultz, 1978; Meyer e Lydeard, 1993; Espinedo *et al*., 2010; Fraser *et al*., 2007; Langerhans, 2011). No que se refere ao gênero *Cnesterodon*, os últimos anos têm visto a descrição de novas espécies, mas estudos com populações naturais que visam entender processos evolutivos são raros ou inexistentes.

Um dos resultados importantes do presente trabalho é que apesar das novas descrições de espécies feitas recentemente (ver Aguilera *et al.*, 2009), a diversidade de espécies para esse gênero pode permanecer subestimada. Por exemplo, apenas dois táxons eram reconhecidos para o Planalto Sul-Brasileiro: *Cnesterodon* sp. nov. B e *C. brevirostratus*, tendo como principal critério de separação morfológica a forma do gonopódio. O emprego de marcadores moleculares mitocondriais permitiu a descoberta de mais de duas linhagens evolutivas independentes que podem se qualificar como possíveis novas espécies. Para permitir uma avaliação de alta resolução nas relações evolutivas destes indivíduos, buscamos uma metodologia relativamente nova e altamente eficiente na elucidação de relacionamentos filogenéticos de espécies não-modelo intimamente relacionadas. A utilização da metodologia de RAD-seq evidenciou a presença de quatro linhagens evolutivas com forte associação ao tamanho de gonopódio (duas linhagens relacionadas ao gonopódio curto e duas relacionadas ao gonopódio longo). Além disso, estas quatro linhagens ainda podem ser subdivididas com base na distribuição geográfica. O processo de divergência entre estas linhagens envolve divergência alopátrica por isolamento geográfico, embora a seleção sexual também deva ter tido um papel importante no isolamento reprodutivo destas linhagens.

Casos de seleção sexual pré-cruzamento são comuns dentro da família Poecillidae e foram associados à escolha de características de corte e/ou morfologia especificas guiada por fêmeas (Bisazza *et al.*, 2001; Langerhans *et al.*, 2005; Kahn *et al.*, 2009), bem como por escolha de cruzamento coespecífico guiada por machos (Hugues 1985; Langerhans *et al*., 2005; Espinedo *et al.*, 2010). Ainda se desconhece o comportamento de corte em

Cnesterodon. E, embora não tenhamos detectado evidência de *loci* em evolução sob seleção divergente entre as linhagens, os resultados obtidos no Capítulo 3 deste trabalho, aliados ao conhecimento existente dos casos de seleção sexual relatados na literatura para a família, abrem uma nova perspectiva de pesquisa. Neste sentido, o gênero *Cnesterodon* pode servir como um novo organismo modelo para estudos de processos de especiação, seleção sexual e evolução da morfologia genital em peixes. A elucidação do número de linhagens envolvidas sob a nomenclatura de dois táxons auxilia na busca de limites morfológicos que possam validar a separação destas linhagens em novas espécies. Além disso, a descoberta da grande diversidade existente em uma área relativamente pequena apoia a manutenção de áreas de preservação na região do Planalto Sul-Brasileiro. É válido ressaltar a importância do PNAS, o qual abriga representantes das quatro linhagens genéticas do complexo de espécies de *Cnesterodon*, sendo que uma delas é, até o momento, encontrada exclusivamente dentro da área do parque.

Um contraste interessante para a região do Planalto Sul-Brasileiro, onde quatro linhagens muito divergentes ocorrem em uma parea pequena é o Bioma Pampa, onde há poucas espécies com distribuição ampla e uma profundidade genealógica relativamente restrita. Uma análise filogeográfica envolvendo a espécie *C. decemmaculatus* revelou que a bacia do Uruguai seria a localidade mais provável de origem do ancestral comum mais recente desta espécie, corroborando as hipóteses biogeográficas inferidas a partir das filogenias moleculares para o gênero (ver abaixo). Esta espécie é a que possui distribuição mais ampla dentro do gênero. De forma geral, o padrão de estrutura genética de *C. decemmaculatus* está associado, em escala ampla, aos principais sistemas de drenagens do Sul da América do Sul. Foi detectado o compartilhamento de haplótipos entre os sistemas de drenagem, sugerindo não apenas a presença de polimorfismo ancestral compartilhado entre diferentes drenagens, mas também que as linhagens associadas às bacias hidrográficas divergiram alopatricamente com contato secundário posterior ocasionado por eventos de rearranjos de pequenos cursos de rios tanto nos sentidos Norte-Sul, quanto nos sentidos Leste-Oeste e Oeste-Leste. Padrões populacionais diferentes são encontrados dentro da mesa espécie: a bacia do Rio Negro possui uma população com diversidade genética e tamanho efetivo populacional baixo, ao passo que as bacias do Sul do Uruguai possuem uma população com alta diversidade genética, indício de expansão populacional e tamanho efetivo populacional alto. Estas diferenças encontradas refletem a dinâmica

heterogênea do Bioma Pampa, o qual teve a historia dos sistemas de drenagens moldada por mudanças climáticas, alterações drásticas no nível do mar e eventos tectônicos, desde o Neogeno (Casciotta *et al*. 1999).

A estrutura genealógica encontrada para a espécie *C. decemmaculatus* é bem mais recente (menos de 1,0 milhao de anos) do que àquela encontrada no complexo de espécies do Planalto Sul-Brasileiro (3.1 milhões de anos). Trabalhos com enfoque populacional têm demostrado que o Bioma Pampa possui um padrão filogeográfico heterogêneo em termos de estrutura genética e espacial (Freitas *et al*., 2012; Turchetto *et al*., 2014; Felappi *et al*. 2015), Apesar da diversidade de espécies e casos de endemismos, este Bioma tem sido negligenciado quando se compara o número de estudos com Biomas como a Mata Atlântica, por exemplo. O padrão filogeográfico encontrado para a espécie *C. decemmaculatus* acrescenta no entendimento sobre como eventos geológicos e climáticos afetaram a diversidade e a distribuição das espécies que ocupam esta região, especialmente com relação aos ambientes de água doce.

Finalmente, os resultados apresentados na presente tese são importantes no entendimento das relações filogenéticas do gênero *Cnesterodon*. Até o momento, apenas caracteres morfológicos haviam sido utilizados para esse fim (Lucinda, 2005; Lucinda *et al*., 2006; Aguilera *et al.*, 2009). A utilização de ferramentas moleculares permitiu a obtenção de uma hipótese filogenética inédita e contrastante com as existentes na literatura até o momento. Em nossa hipótese filogenética, a espécie *C. hypselurus* é irmã das espécies *C. decemmaculatus*, *C. holopteros*, *C. septentrionalis*, e do complexo de espécies da região do Planalto Sul-Brasileiro. Por outro lado, a hipótese de Lucinda *et al.* (2006), por exemplo, sugere que *C. hypselurus*, *C. brevirostratus*, e *C. septentrionalis* (considerando as espécies amostradas neste trabalho) compartilham um conjunto de quatro sinapomorfias em relação à *Cnesterodon* sp. n. B e às duas espécies que ocorrem no Bioma Pampa. Da mesma forma, a filogenia proposta com base em marcadores moleculares revelou que os *taxa Cnesterodon* sp. nov. B e *C. brevirostratus*, representados agora pelas quatro linhagens evolutivas independentes (identificadas no Capítulo 3 da presente tese), formam um clado monofilético bem suportado, irmão das espécies que ocorrem no Bioma Pampa. Esta nova hipótese modifica completamente o cenário para estas entidades.

A filogenia molecular proposta no presente trabalho permitiu também a inferência de hipóteses biogeográficas que contribuem para um melhor entendimento do surgimento e

dispersão do grupo. O relacionamento irmão entre *C. hypselurus* (da bacia do Paraná) e as outras espécies de *Cnesterodon* pode sugerir que a linhagem ancestral tenha vindo da bacia do Paraná, divergindo no Mioceno (cerca de 8 milhões de anos). Hubert & Renno (2006) sugerem que as partes elevadas da bacia do Paraná formam uma área de endemismo que pode ter servido como uma área de refúgio para Characiformes durante o Mioceno. Considerando que esta região também é uma área de endemismo para *Cnesterodon*, abrigando três espécies conhecidas, é plausível que esta região seja um centro de diversidade para o gênero. O soerguimento dos Andes em sua fase final afetou não somente a parte ocidental da América do Sul, mas também causou a elevação do escudo brasileiro, ocasionando, ainda que de forma indireta, eventos de captura de cabeceiras. A divergência e origem de *C. septentrionalis* no Mioceno tardio/início do Plioceno, pode ser uma consequência destes eventos, nos quais populações da bacia do Paraná teriam alcançado a bacia do Tocantins-Araguaia, ao norte, seguindo algum evento captura e permanecendo isoladas por um longo período.

Três cenários biogeográficos foram propostos para a rota de colonização inicial das espécies que ocorrem na região Sul da América do Sul: 1) A origem do clado formado pelo complexo de espécies do Planalto Sul-Brasileiro estaria relacionada a eventos orogênicos durante o Plioceno, que teriam permitido uma colonização das porções superiores da Bacia do Uruguai, à partir da bacia do Paraná, atingindo posteriormente as porções baixas e dando origem ao clado formado pelas espécies que habitam o Pampa; 2) Eventos orogênicos durante o Plioceno somados à crescente disponibilidade de habitat consequente da retração do "mar do Paraná", que poderiam ter propiciado a chegada da linhagem ancestral do clado composto pelas espécies do Pampa através da própria bacia do Paraná, que drena para o sudoeste até atingir o estuário do Rio da Prata, entre a Argentina e o Uruguai. Posteriormente esse linagem poderia ter colonizado regiões associadas ao Rio Uruguai dando origem à linhagem ancestral do complexo de espécies do Planalto Sul-Brasileiro; 3) Dois eventos independentes de colonização teriam ocorrido a partir da bacia do Paraná, um dando origem ao clado formado pelo complexo de espécies do do Planalto Sul-Brasileiro e outro dando origem ao clado composto pelas espécies do Pampa.

Além dos cenários biogeográficos propostos para o gênero, outra contribuição importante deste trabalho foi quanto ao conhecimento da distribuição geográfica das espécies e ao status de conservação. Novos registros de ocorrência foram encontrados para

C. holopteros no Sul do Uruguai. Esta espécie era conhecida apenas na sua localidade tipo no noroeste do Uruguai. Além disso, expedições à campo ao longo de três anos nos permitiram concluir que os habitats naturais de algumas espécies que possuem distribuição restrita estão seriamente degradados. As espécies *C. carnegiei*, *C. iguape*, *C. omorgmatos* já não são mais encontradas nas suas localidades tipo, bem como em outros locais registrados na literatura. Atividades antrópicas como pastoreio, desvios de rios para irrigação de monoculturas, contaminação de água decorrente do uso de pesticidas e, principalmente, o crescimento urbano sem planejamento estão afetando gravemente as áreas de ocorrência destas três espécies. Uma revisão bibliográfica mostra que *Cnesterodon* ainda é um gênero pouco estudado. Cerca de 60% das descrições de novas espécies foram feitas nos últimos 15 anos. Porém, o ritmo de descoberta de novas espécies e o aumento de conhecimento acerca do gênero ainda é bem menor comparada à velocidade de degradação, fragmentação e extinção dos habitats naturais. As espécies encontradas no Planalto Sul-Brasileiro parecem habitar ambientes menos degradados. As coletas realizadas pelo nosso grupo de trabalho nesta região mostram que indivíduos de *Cnesterodon* ocorrem de forma abundante, especialmente dentro do Parque Nacional dos Aparados da Serra (PNAS) e arredores.

Em resumo, a abordagem molecular executada nessa tese contribuiu com a investigação das relações evolutivas entre espécies e entre populações naturais do gênero *Cnesterodon*, proporcionando um novo panorama com relação à diversidade e os processos envolvidos na diversificação deste gênero. As filogenias moleculares permitiram inferir a origem e dispersão do gênero. As espécies do Planalto Sul-Brasileiro podem agora ser estudadas em profundidade, utilizando amostras populacionais de interesse e o conjunto dos milhares de novos marcadores genéticos SNPs do genoma de *Cnesterodon*, adequando-os às novas perguntas. O padrão filogeográfico encontrado para *C. decemmaculatus* pode servir de modelo e auxiliar a compreensão da estrutura genética de outras espécies de peixes de água doce no Bioma Pampa.

CAPÍTULO 6

Referências Bibliográficas

Agostinho AA, Thomaz SM e Gomes LC (2005) Conservation of the biodiversity of Brazil's inland waters. Cons Biol 19:646-652.

Aguilera G, Mirande JM e Azpelicueta MM (2009) A new species of *Cnesterodon* (Cyprinodontiformes: Poeciliidae) from a small tributary of arroyo Cuñá-Pirú, río Paraná basin, Misiones, Argentina. Zootaxa 2195:34-42.

Amish SJ, Hohenlohe PA, Painter S, Leary RF, Muhlfeld C, Allendorf FW e Luikart G (2012) RAD sequencing yields a high success rate for westslope cutthroat and rainbow trout species-diagnostic SNP assays. Mol Ecol Res 12:653-660.

Arias A e Reznick D (2000) Life history of *Phalloceros caudiomaculatus*: a novel variation on the theme of livebearing in the family Poeciliidae. Copeia:792-798.

Avise JC (1994) Molecular Markers, Natural History and Evolution. 2ª edição. Chapman and Hall, New York, 684 pp.

Avise JC (2000) Phylogeography: The history and formation of species. Harvard University, Londres, 447 pp.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel, JE, Reeb CA e Sauders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu Rev Ecol Syst 18:489-522.

Baigún C e Ferriz R (2003) Distribution patterns of native freshwater fishes in Patagonia, Argentina. Org Divers Evol 3:151-159.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA e Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers, PLoS One 3:e3376.

Barchi L, Lanteri S, Portis E, Acquadro A, Valè G, Toppino L e Rotino GL (2011) Identification of SNP and SSR markers in eggplant using RAD tag sequencing. BMC Genomics 12:304.

Beheregaray LB, Sunnucks P e Briscoe DA (2002) A rapid fish radiation associated with the last sea-level changes in Southern Brazil : the silverside *Odontesthes perugiae* complex. Proc R Soc Lond B Biol Sci 269:65-73.

Behling H (2002) South and Southeast Brazilian grasslands during late Quaternary times: a synthesis. Palaeogeogr Palaeoclimatol Palaeoecol 177:19-27.

Bentley DR, Balasubramanian S, Swerdlow HP *et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456:53–59.

Berretta EJ (2001) Ecophysiology and management response of the subtropical grasslands of Southern America. In: Gomide JA, Mattos WRS e Silva SC (eds) Proceedings of the XIX International Grassland Congress. ACSESS, São Paulo, pp 939-946.

Bickford D, Lohma, DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK e Das I (2007) Cryptic species as a window on diversity and conservation. Trends Ecol Evol 22:148-155.

Bilenca DN e Miñarro FO (2004) Áreas valiosas de pastizal en las pampas y Campos de Argentina, Uruguay y sur de Brasil. Fundación Vida Silvestre Argentina, Buenos Aires, 323pp.

Billington N (2003) Mitochondrial DNA. In: Hallerman EM (ed) Population genetics: principles and applications for fisheries scientists. American Fisheries Society, Maryland, pp 59-100.

Bisazza A, Vaccari G e Pilastro A (2001) Female mate choice in a mating system dominated by male sexual coercion. Behav Ecol 12: 59-64.

Boldrini II (1997) Campos no Rio Grande do Sul. Fisionomia e problemática ocupacional. Bol Inst Biociências UFRGS 56:1-39.

Boldrini II, Ferreira PMA, Andrade BO, Schneider AA, Setubal RB, Trevisan R e Freitas EM. (2010) Bioma Pampa: diversidade florística e fisionômica. Editora Palloti, Porto Alegre, 64pp.

Braun JK, Busscheb RAVD, Mortonb PK e Maresa MA (2005) Phylogenetic and biogeographic relationships of mouse *Opossums thylamys* (Didelphimorphia, Didelphidae) in Southern South America. J Mammal 86:147-159.

Buckup PA, Menezes NA e Ghazzi MS (2007) Catálogo das Espécies de Peixes de Agua Doce do Brasil. Museu Nacional, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 290pp.

Cabanne GS, Santos FR e Miyaki CY (2007) Phylogeography of *Xiphorhynchus fuscus* (Passeriformes, Dendrocolaptidae): vicariance and recent demographic expansion in southern Atlantic forest. Biol J Linn Soc Lond 91:73-84.

Carvalho PCF e Batello C (2009) Access to land, livestock production and ecosystem conservation in the Brazilian Campos biome: the natural grasslands dilemma. Livest Sci 120:158-162.

Casciotta J, Almirón A, Cione A e Azpelicueta, M (1999) Brazilian freshwater fish

assemblages from southern Pampean area, Argentina. Biogeographica 75: 67-78.

Castro RMC (1999) Evolução da ictiofauna de riachos sul-americanos: padrões gerais e possíveis processos causais. In: Caramaschi EP, Mazzoni R, Bizerril CRSF. e PeresNeto PR (eds) Ecologia de peixes de riachos. Série Oecologia Brasiliensis. Volume 7. PPGE-UFRJ, Rio de Janeiro, pp 139-155.

Catchen J, Bassham S, Wilson T, Currey M, O'Brien C, Yeates Q e Cresko WA (2013) The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. Mol Ecol 22:2864-2883.

Combosch DJ e Vollmer SV (2015). Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific *Pocillopora* corals. Mol Phylogenet Evol 88:154-162.

Davey JW e Blaxter ML (2010) RADSeq: next-generation population genetics. Brief Funct Genomics 9:416-423.

Dergam JA, Paiva SR, Schaeffer CE, Godinho AL e Vieira F (2002) Phylogeography and RAPD-PCR variation in *Hoplias malabaricus* (Bloch, 1794) (Pisces, Teleostei) in southeastern Brazil. Genet Mol Biol 25:379-387.

D'Horta FM, Cabanne GS, Meyer D e Miyaki CY (2011) The genetic effects of Late Quaternary climatic changes over a tropical latitudinal gradient: diversification of an Atlantic Forest passerine. Mol Ecol 20:1923-1935.

Dupuis JR, Roe AD e Sperling FAH (2012) Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Mol Ecol 21:4422-4436.

Edwards CJ, Suchard MA, Lemey P, Welch JJ, Barnes I, Fulton TL, Barnett R, O'Connell TC, Coxon P, Monaghan N, Valdiosera C, Lorenzen ED, Willerslev E, Baryshnikov GF, Rambaut A, Thomas MG, Bradley DG e Shapiro B (2011) Ancient hybridization and an Irish origin for the modern polar bear matriline. Curr Biol 21:1251-1258.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, e Mitchell SE (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One 6:e19379.

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE e Holzapfel CM (2010). Resolving postglacial phylogeography using high-throughput sequencing. Proc Natl Acad Sci U S A 107:16196-16200.

Endler JA (1983) Natural selection on color patterns in poeciliid fishes. Environ Biol Fish 9:173-190.

Espinedo CM, Gabor CR, Aspbury AS (2010) Males, but not females, contribute to sexual isolation between two sympatric species of *Gambusia*. Evol Ecol 24: 865-878.

Felappi, J. F., Vieira, R. C., Fagundes, N. J. & Verrastro, L. V. (2015). So Far Away, Yet So Close: Strong Genetic Structure in Homonota uruguayensis (Squamata, Phyllodactylidae), a Species with Restricted Geographic Distribution in the Brazilian and Uruguayan Pampas. PloS One 10: 1-19.

Fontanella FM, Feltrin N, Avila LJ, Sites JW e Morando M (2012) Early stages of divergence: phylogeography, climate modeling, and morphological differentiation in the South American lizard *Liolaemus petrophilus* (Squamata: Liolaemidae). Ecol Evol 4:792-808.

Fraser D, Jones M, McParland T e Hutchings J (2007) Loss of historical immigration and the unsuccessful rehabilitation of extirpated salmon populations. Conserv Genet 8:527-546.

Freamo H, O'Reilly P, Berg PR, Lien, S e Boulding EG (2011) Outlier SNPs show more genetic structure between two Bay of Fundy metapopulations of Atlantic salmon than do neutral SNPs. Mol Ecol Res 11: 254-267.

Freitas TRO, Fernandes FA, Fornel R e Roratto PA (2012) An endemic new species of tuco-tuco, genus *Ctenomys* (Rodentia: Ctenomyidae), with a restricted geographic distribution. S Braz J Mammal 93: 1355-1367.

Gaither MR, Bernal MA, Coleman RR, Bowen BW, Jones SA, Simison BW e Rocha LA (2015) Genomic signatures of geographic isolation and natural selection in coral reef fishes. Mol Ecol 24:1543-1557.

Gentry AH (1995) Patterns of diversity and floristic composition in neotropical montane forests. In: Churchill SP, Balslev H, Forero E e Luteyn JL (eds) Biodiversity and conservation of neotropical montane forests. New York Bot Garden, New York, pp.103-126.

Gonçalves GL, Marinho JR e Freitas TRO (2009) Genetic structure of sigmodontine rodents (Cricetidae) along an altitudinal gradient of the Atlantic Rain Forest in southern Brazil. Genet Mol Biol 32:882-885.

Grazziotin FG, Monzel M, Echeverrigaray S e Bonatto SL (2006) Phylogeography of the

Bothrops jararaca complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic Forest. Mol Ecol 15:3969-3982.

Grove BD e Wourms JP (1991) Follicular placenta of the viviparous fish, *Heterandria formosa*: I. Ultrastructure and development of the embryonic absorptive surface. J Morphol 209:265-284.

Grove BD e Wourms JP (1994) Follicular placenta of the viviparous fish, *Heterandria formosa*: II. Ultrastructure and development of the follicular epithelium. J Morphol 220:167-184.

Hamilton A (2001) Phylogeny of *Limia* (Teleostei:Poeciliidae) based on NADH dehydrogenase subunit 2 sequences. Mol Phylogenet Evol 19:277-289.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA e Cresko WA (2012) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genet 6:e1000862.

Houde AE (1997) Sex, Color, and Mate Choice in Guppies. Princeton University Press, Princeton, 224 pp.

Houston RD, Davey JW, Bishop SC, Lowe NR, Mota-Velasco JC, Hamilton A, Guy DR, Tinch AE, Thomson ML, Blaxter ML, Gharbi K, BronJE e Taggart JB (2012) Characterisation of QTL-linked and genome-wide restriction site-associated DNA (RAD) markers in farmed Atlantic salmon. BMC Genomics 13:244.

Hrbek T, Seckinger J e Meyer A (2007) A phylogenetic and biogeographic perspective on the evolution of poeciliid fishes. Mol Phylogenet Evol 43:986-998.

Hubert N e Renno JF (2006) Historical biogeography of South American freshwater fishes. J Biogeogr 33:1414-1436.

Hughes AL (1985) Male size, mating success, and mating strategy in the mosquitofish *Gambusia affinis* (Poeciliidae). Behav Ecol Sociobiol 17: 271-278.

IBGE - Instituto Brasileiro de Geografia e Estatística (2012) Base de dados electrônica disponível em: http://www.ibge.gov.br/home/ (acessado em junho de 2012).

Iriarte PJF, Alonso MP, Sabadin DE, Arauz PA e Iudica CM (2011) Phylogeography of weakfish *Cynoscion guatucupa* (Perciformes: Sciaenidae) from the southwestern Atlantic. Sci Mar 75:701-706.

Kahn AT, Mautz B, Jennions MD (2009) Females prefer to associate with males with longer intromittent organs in mosquitofish. Biol Lett 11: 1-4..

Langerhans RB, Layman CA, DeWitt TJ (2005). Male genital size reflects a tradeoff between attracting mates and avoiding predators in two live-bearing fish species. PNAS 102: 7618-7623.

Langerhans RB (2011) Genital evolution In: Evans JP, Pilastro A, Schlupp I (eds). Ecology and evolution of poeciliid fishes, University of Chicago Press, Chicago, pp 228-240.

Lévêque C, Oberdorff T, Paugy D, Stiassny MLJ e Tedesco PA (2008) Global diversity of fish (Pisces) in freshwater. Hydrobiologia 595:545-567.

Lindqvist C, Schuster SC, Sun Y, Talbot SL, Qi J, Ratan A, Tomsho LP, Kasson L, Zeyl E, Aars J, Miller W, Ingólfssonf O, Bachmannd L e Wiigd O (2010) Complete mitochondrial genome of a Pleistocene jawbone unveils the origin of polar bear. Proc Natl Acad Sci U S A 107:5053-5057.

López HL, Morgan CC e Montenegro MJ (2002) Ichthyological ecoregions of Argentina. ProBiota, Serie Documentos N01. Universidad Nacional de La Plata, Buenos Aires, 70 pp.

Lucinda PHF (2003) Family Poeciliidae. In: Reis RE, Kullander SO e Ferraris Jr CJ (eds) Check list of the freshwater fishes of South and Central America. EdiPUCRS, Porto Alegre, pp 555-581.

Lucinda PHF (2005) Systematics of the genus *Cnesterodon* Garman, 1895 (Cyprinodontiformes: Poeciliidae: Poeciliinae). Neotrop Ichthyol 3:259-270.

Lucinda PHF e Reis RE (2005) Systematics of the subfamily Poeciliinae Bonaparte (Cyprinodontiformes: Poeciliidae), with an emphasis on the tribe Cnesterodontini Hubbs. Neotrop Ichthyol 3:1-60.

Lucinda PHF, Thomas L e Roberto R (2006) *Cnesterodon holopteros* (Cyprinodontiformes: Poeciliidae: Poeciliinae), a new species from the Republic of Uruguay. Zootaxa 1350: 21-31.

Luikart G, England PR, Tallmon D, Jordan S e Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4:981-994.

Malabarba LR, Fialho CB, Anza JA, Santos JF e Mendes GN (2009) Fauna aquática: peixes. In: Boldrini II (ed) Biodiversidade dos campos no planalto das araucárias. Ministério do Meio Ambiente, Brasília, pp 131-141.

Mapelli FJ, Mora MS, Mirol PM e Kittlein MJ (2012) Effects of Quaternary climatic changes on the phylogeography and historical demography of the subterranean rodent

Ctenomys porteousi. J Zool 286:48-57.

Meyer A e Lydeard C (1993) The evolution of copulatory organs, internal fertilization, placentae and viviparity in killifishes (Cyprinodontiformes) inferred from a DNA phylogeny of the tyrosine kinase gene *X-src*. Proc R Soc Lond B Biol Sci 254:153-162.

Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-Reina OC, Zhao F, Kim HL, Burhans RC, Drautz DI, Wittekindt NE, Tomsho LP, Ibarra-Laclettee E, Herrera-Estrella L, Peacock E, Farley S, Sage GK, Rode K, Obbard M, Montiel R, Bachmann L, Ingólfsson O, Aars J, Mailund T, Wiig O, Talbot SL e Lindqvist C (2012) Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. Proc Natl Acad Sci U S A 109:2382-2390.

Morando M, Avila FJ, Baker J e Sites JW (2004) Phylogeny and phylogeography of the *Liolaemus darwinii* complex (Squamata: Liolaemidae): Evidence for introgression and incomplete lineage sorting. Evolution 58:842-861.

Moritz C, Dowling TE e Brown WM (1987) Evolution of animal mitochondrial DNA. Relevance for population biology and systematics. Annu Rev Ecol Syst 18:269-292.

Moritz C e Faith DP (1998) Comparative phylogeography and the identification of genetically divergent areas for conservation. Mol Ecol 7:419-429.

Narum SR, Banks M, Beacham TD, Bellinger MR, Campbell MR, Dekoning J, Elz A, Guthrie CM, Kozfkay C, Miller KM, Moran P, Phillips R, Seeb LW, Smith CT, Warheit K, Young SF e Garza JC (2008) Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. Mol Ecol 17:3464-3477.

Nowrousian M (2010) Next-Generation Sequencing Techniques for Eukaryotic Micro-organisms: Sequencing-Based Solutions to Biological Problems. Eukaryot Cell 9:1300-1310.

Overbeck GE, Müller SC, Pillar VD e Pfadenhauer J (2006) Floristic composition, environmental variation and species distribution patterns in burned grassland in southern Brazil. Braz J Biol 66:1073-1090.

Pallarés OR, Berretta EJ e Maraschin GE (2005) The South American Campos ecosystem. In: Suttie J, Reynolds SG e Batello C (eds) Grasslands of the World. Food and Agriculture Organization of the United Nations, Rome, pp 171–219.

Palma RE, Rivera-Milla E, Salazar-Bravo J, Torres-Pérez F, Pardiñas UFJ, Marquet PA,

Spotorno AE, Meynard ASP e Yates TL (2005) Phylogeography of *Oligoryzomys longicaudatus* (Rodentia: Sigmodontinae). J Mammal 86:191-200.

Pandey V, Nutter RC e Prediger E. (2008) Applied biosystems SOLID system: ligation-based sequencing. In: Janitz M (ed) Next Generation Genome Sequencing: Towards Personalized Medicine. Wiley-VCH, Weinheim, pp 431-444.

Parenti LR (1981) A phylogenetic and biogeographical analysis of Cyprinodontiform fishes (Teleostei, Atherinomorpha). Bull Am Mus Nat Hist 168:341-557.

Pascual MA, Cussac V e Dyer B (2007) Freshwater fishes of Patagonia in the 21st century after a hundred years of human settlement, species introductions, and environmental change. Aquat Ecosyst Health Manag 10:212-227.

Peterson BK, Weber JN, Kay EH, Fisher HS e Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS One 7:e37135.

Pillar VP (2006) Estado atual e desafios para a conservação dos campos. Workshop Estado Atual e desafios para a conservação dos campos. UFRGS, Porto Alegre, http://www.ecologia.ufrgs.br/ecologia/campos/finalcampos.htm (acessado em abril de 2015)

Pinto LP e Brito MCW (2005) Dinâmica da perda da biodivresidade na Mata Atlântica brasileira: uma introdução. In: Galindo-Leal C e Câmara IG (eds) Mata Atlântica biodiversidade, ameaças e perspectivas. Fundação SOS Mata Atlântica, Belo Horizonte, pp 27-30.

Rambo B (1954) Análise histórica da flora de Porto Alegre. Sellowia 6:9-112.

Reis RO, Kullander SO e Ferraris CJ (2003) Check list of the freshwater fishes of South and Central America. EdiPUCRS, Porto Alegre, 729 pp.

Reznick DN, Mateos M e Springer MS (2002) Independent origins and rapid evolution of the placenta in the fish genus *Poeciliopsis*. Science 298:1018-1020.

Reznick D, Hrbek T, Caura S, Greef J e Roff D (2007) Life history of *Xenodexia ctenolepis*: implications for life history evolution in the family Poeciliidae. Biol J Linn Soc Lond 92:77-85.

Ribeiro AC (2006) Tectonic history and the biogeography of the freshwater fishes from the coastal drainages of eastern Brazil: an example of faunal evolution associated with a divergent continental margin. Neotrop Ichthyol 4:225-246.

Roesch LFW, Vieira FCB, Pereira VA, Schünemann AL, Teixeira IF, Senna AJT e Stefenon VM (2009) The Brazilian Pampa: A fragile biome. Diversity 1:182-198.

Rosen DE e Bailey RM (1963) The poeciliid fishes (Cyprinodontiformes), their structure, zoogeography, and systematics. Bull Am Mus Nat Hist 126:1-176.

Rothberg J e Leamon J (2008) The development and impact of 454 sequencing. Nat Biotechnol 26:1117-1124.

Safford HD (1999) Brazilian paramos I: Introduction to the physical environment and vegetation of the campos de altitude. J Biogeogr 26:693-712.

Sanger F, Nicklen S e Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74:5463-5467.

Scaglione D, Acquadro A, Portis E, Tirone M, Knapp SJ e Lanteri S (2012) RAD tag sequencing as a source of SNP markers in *Cynara cardunculus* L. BMC Genomics 13:3.

Schaefer SA (1998) Conflict and resolution: impact of new taxa on phylogenetic studies of the neotropical Cascudinhos (Siluroidei: Loricariidae). In: Malabarba LR, Reis RE, Vari RP, Lucena ZMS e Lucena CAS (eds) Phylogeny and classification of neotropical fishes. EdiPUCRS, Porto Alegre, pp 363–374.

Schluter A, Parzefall J e Schlupp I (1998) Female preference for symmetrical vertical bars in male sailfin mollies. Anim Behav 56:147-153.

Shendure J e Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135-1145.

Thibault RE e Schultz RJ (1978) Reproductive adaptations among viviparous fishes (Cyprinodontiformes: Poeciliidae). Evol 32: 320-333.

Thomaz AT, Malabarba LR e Bonnato SL (2010) The phylogenetic placement of *Hollandichthys* Eigenmann 1909 (Teleostei: Characidae) and related genera. Mol Phylogenet Evol 57:1347-1352.

Thomaz AT, Malabarba LR, Bonatto SL e Knowles LL (2015) Testing the effect of palaeodrainages versus habitat stability on genetic divergence in riverine systems: study of a Neotropical fish of the Brazilian coastal Atlantic Forest. J Biogeogr, *in press*: 1-13.

Turchetto, C., Fagundes, N. J. R., Segatto, A. L. A., Kuhlemeier, C., Neffa, V. G. S., Speranza, P. R., Bonatto, S. L. & Freitas. L. B. (2014). Diversification in the South American Pampas: the genetic and morphological variation of the widespread *Petunia axillaris* complex (Solanaceae). Mol Ecol 23: 374-389.

Vari RP e Malabarba LR (1998) Neotropical Ichthyology: an overview. In: Malabarba LR,

Reis RE, Vari RP, Lucena ZMS, Lucena CAS (eds) Phylogeny and classification of Neotropical fishes. EdiPUCRS, Porto Alegre, pp 1-11.

Vari RP e Weitzman SH (1990) A review of the phylogenetic biogeography of freshwater fishes of South America. In: Peters G e Hutterer R (eds) Vertebrates in tropics. Museum Alexander Koening, Bonn, pp 381-392.

Vergara J, Azpelicueta MM e Garcia G (2008) Phylogeography of the Neotropical catfish *Pimelodus albicans* (Siluriformes: Pimelodidae) from río de la Plata basin, South America, and conservation remarks. Neotrop Ichtyol 6:75-85.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A e Seehausen O (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol Ecol 22:787-798.

Webster GL (1995) The panorama of neotropical clouds forests. In: Churchill SP, Balslev H, Forero E e Luteyn JL (eds) Biodiversity and conservation of neotropical montane forests. New York Botanical Garden, New York, pp 53-78.

Weitzman SH e LR Malabarba (1998) Perspectives about the phylogeny and classification of Characidae (Teleostei: Characiformes). In: Malabarba LR, Reis RE, Vari RP, Lucena ZMS e Lucena CAS (eds) Phylogeny and classification of neotropical fishes. EdiPUCRS, Porto Alegre, pp. 161-170.

Zhang J, Chiodini R, Badr A e Zhang G (2011) The impact of next-generation sequencing on genomics. J Genet Genomics 38:95-109.