Marine incursions, cryptic species and ecological diversification in Amazonia: the biogeographic history of the croaker genus *Plagioscion* (Sciaenidae)

Georgina M. Cooke¹, Ning L. Chao² and Luciano B. Beheregaray¹,³*

¹Molecular Ecology Laboratory, Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia, ²Departamento de Ciências Pesqueiras, Universidade Federal do Amazonas, Manaus, Brazil, ³Molecular Ecology Laboratory, School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia

ABSTRACT

Aim We propose a phylogenetic hypothesis for the marine-derived sciaenid genus *Plagioscion* in the context of geomorphology and adaptation to freshwaters of South America, and assess the extent to which contemporary freshwater hydrochemical gradients influence diversification within a widely distributed *Plagioscion* species, *Plagioscion squamosissimus*.

Location Amazon Basin and South America.

Methods Using nuclear and mitochondrial DNA sequence data, phylogenetic analyses were conducted on the five nominal *Plagioscion* species, together with representatives from *Pachyurus* and *Pachypops*, using character and model-based methods. Genealogical relationships and population genetic structure of 152 *P. squamosissimus* specimens sampled from the five major rivers and three hydrochemical settings/‘colours’ (i.e. white, black and clear water) of the Amazon Basin were assessed.

Results Phylogenetic analyses support the monophyly of *Plagioscion* in South America and identify two putative cryptic species of *Plagioscion*. Divergence estimates suggest that the *Plagioscion* ancestor invaded South America via a northern route during the late Oligocene to early Miocene. Within *P. squamosissimus* a strong association of haplotype and water colour was observed, together with significant population structure detected between water colours.

Main conclusions Our analyses of *Plagioscion* are consistent with a biogeographic scenario of early Miocene marine incursions into South America. Based on our phylogenetic results, the fossil record, geomorphological history and distributional data of extant *Plagioscion* species, we propose that marine incursions into western Venezuela between the late Oligocene and early Miocene were responsible for the adaptation to freshwaters in *Plagioscion* species. Following the termination of the marine incursions during the late Miocene and the establishment of the modern Amazon River, *Plagioscion* experienced a rapid diversification. *Plagioscion squamosissimus* arose during that time. The formation of the Amazon River probably facilitated population and range expansions for this species. Further, the large-scale hydrochemical gradients within the Amazon Basin appear to be acting as ecological barriers maintaining population discontinuities in *P. squamosissimus* even in the face of gene flow. Our results highlight the importance of divergent natural selection through time in the generation and maintenance of sciaenid diversity in Amazonia.

Keywords Amazon Basin, divergent natural selection, ecological gradients, marine-derived lineage, phylogeography, *Plagioscion squamosissimus*, South America.
INTRODUCTION

The Amazon drainage basin is home to the world’s richest freshwater fish fauna. Like other tropical systems, the biodiversity of Amazonia has been largely attributed to palaeoecographic and climatic events that occurred during the Tertiary (Moritz et al., 2000; Rull, 2008). Indeed, the majority of Neotropical fishes have been physiologically constrained to freshwater habitats and appear to have a long history of evolution dating back to the final separation of the South American and African continents (e.g. Siluriformes and Characiformes) (Lundberg et al., 1998, 2010; Reis et al., 2003). However, South America also accommodates several species and small clades of freshwater fishes that evolved from predominantly marine lineages. In these cases, the role of ecologically based divergent natural selection across salinity gradients probably facilitated the physiological changes necessary for adaptation to freshwater. Based on taxonomic, phylogenetic, biogeographic and fossil data, it is generally thought that these lineages colonized South America relatively recently, as a product of marine incursions that entered the Amazon during the Oligocene and Miocene (Hoorn, 1993; Webb, 1995; Lovejoy et al., 1998, 2006; Lundberg et al., 1998; Monsch, 1998; Boeger & Kritsky, 2003; Wesselingh & Macrosay, 2006). Marine incursions produced numerous estuarine environments of varying salinity and in concert with significant geological changes may have favoured ecological shifts from saltwater to freshwater. Further, based on fossil data and biogeographic patterns exhibited between marine-derived fish species, it has become apparent that the process of adaptation to Neotropical freshwaters has been repeated many times in parallel by taxonomically diverse species (Lovejoy et al., 2006). Biogeographic reconstructions of invasions by marine-derived lineages can advance our knowledge of spatial and temporal aspects of fish adaptation and speciation and, more broadly, of the biotic consequences of past sea level changes (Lovejoy et al., 1998, 2006; Beheregaray et al., 2002; McCairns & Bernatchez, 2010; Bloom & Lovejoy, 2011).

The Sciaenidae comprises approximately 78 genera and 287 species of fish distributed throughout the warm coastal and estuarine waters of the Atlantic, Indian and Pacific Oceans (Chao, 1978; Casatti & Chao, 2002). In South America, there are some Sciaenidae species that have adapted to freshwaters and become physiologically restricted and speciated in this new environment. These include members of Plagioscion, Pachyurus, Petilipinnis and Pachyops. Current phylogenetic hypotheses based on fossil, distributional and palaeohydrological data (Sasaki, 1989; Monsch, 1998; Boeger & Kritsky, 2003) advocate secondary colonization of sciaenids into the continent from marine derivatives during the marine incursions of the Oligocene and Miocene. Further, these data also suggest independent adaptation and colonization events of Pachyurus + Pachyops and Plagioscion clades, respectively, with the Plagioscion colonizing later (Sasaki, 1989; Monsch, 1998; Boeger & Kritsky, 2003).

Plagioscion Gill, 1861 is a genus of particular importance to subsistence and commercial fisheries of the Amazon Basin (Santos et al., 1984; Casatti, 2005). Further, it is also a biogeographic indicator of key events relating to the marine incursions into South America. According to Casatti (2005) there are five valid nominal species assigned to this genus (Plagioscion squamosissimus, Plagioscion auratus, Plagioscion magdalenae, Plagioscion ternetzi and Plagioscion montez). While the phylogenetic placement of Plagioscion is unclear (Sasaki, 1989; Casatti, 2005), Boeger & Kritsky (2003) suggest a western Atlantic/eastern Pacific sister genus Paralonchurus based on a phylogeny of dactylogyrid parasites of the genus Euryhalitrena. The geographic association of this sister relationship is consistent with the hypothesis that Plagioscion evolved freshwater tolerance as a product of marine incursions into northern South America. Indeed, fossil evidence suggests that the ancestral Plagioscion species colonized the freshwaters of South America sometime between the early Oligocene and middle Miocene when western Venezuela was predominantly an estuarine bay-like environment (Lundberg, 1998; Monsch, 1998). During that time, a Palaeo-Amazon/orinoco Basin flowed northward, contributing freshwater into Venezuela, and enabling a major pathway south for marine transgressions into South America (Lundberg, 1998). Today, native Plagioscion species occur throughout the Magdalena, Amazon, Orinoco and lower Paraná basins, as well as in rivers of the Guianas, while some have also been introduced to the upper Paraná and São Francisco river basins (Casatti, 2003, 2005).

This study has two main objectives. First, to propose a phylogenetic hypothesis for Plagioscion that includes all valid species of the genus. Using this phylogeny we aim to date species origins and assess their biogeographic history in the context of major geological and eustatic events that may have facilitated their adaptive evolution from marine to freshwater habitats. Our second objective is to clarify the intra-specific phylogeographic history of the most widely distributed and commercially important Plagioscion species, P. squamosissimus. In particular, our aim was to assess the influence of freshwater hydrochemistry on the population structure of P. squamosissimus within the Amazon Basin. In the event that population genetic structure is associated with hydrochemistry after controlling for other factors (e.g. riverine distance and geomorphology), ecologically based divergent natural selection may well be a persistent force in the generation of Sciaenidae diversity. Indeed, there is now a mounting body of evidence suggesting that divergent natural selection and ecological speciation might be an important biodiversity driver across multiple taxa and environments (reviewed in Schluter, 2000; Coyne & Orr, 2004; Rundle & Nosil, 2005; Schluter & Conte, 2009), including the tropics (Smith et al., 1997, 2001; Schneider et al., 1999; García-Paris et al., 2000; Ogden & Thorpe, 2002; López-Fernández et al., 2010).

The Amazon Basin actually provides an ideal setting in which to test for divergent natural selection in P. squamosissimus as a result of hydrochemistry. Within the Amazon Basin, there are marked hydrochemical and ecological gradients that
are known to impose physiological constraints upon its aquatic communities (Junk et al., 1983; Henderson & Crampton, 1997; Rodriguez & Lewis, 1997; Saint-Paul et al., 2000; Petry et al., 2003). Differentiated largely by sediment composition, geochemistry and optical characteristics, these aquatic conditions have been divided into three water types or ‘colours’ (Sioli, 1984): (1) white water which has an Andean origin is turbid in nature and characterized by large amounts of dissolved solids; (2) clear water, which is comparatively transparent, contains a low content of dissolved solids; and (3) black water, which is transparent yet stained by tannins and humic acids leached from vegetation, and which differs most dramatically from clear water by its low pH. Unlike other species of Plagioscion, P. squamosissimus can be found abundantly in all three water colours across river systems, making it an ideal candidate with which to test for population structure driven by hydrochemistry. Furthermore, research has shown that, despite showing high larval dispersal from spawning sites, P. squamosissimus larvae are actually quite sensitive to hydrochemical variables such as temperature, pH and dissolved oxygen (Bialetzki et al., 2004).

Here, using data from both nuclear and mitochondrial genes we reconstruct the phylogenetic relationships of the genus Plagioscion and propose a biogeographic scenario for the invasion of freshwaters consistent with the marine incursions of the early Miocene. We also show evidence for the contribution of freshwater hydrochemistry to genetic structure in Amazonian populations of P. squamosissimus. In this way, we highlight the importance of divergent natural selection through time in the maintenance and generation of Sciaenidae diversity in the Neotropics.

MATERIALS AND METHODS

**Sampling**

For phylogenetic analysis, tissue samples from the five nominal Plagioscion species (P. squamosissimus, P. auratus, P. tenetzi, P. magdalenae and P. montei) and three representatives of Pachyurus and Pachyops (Pachyurus junki, Pachyops fourcroi and a currently undescribed Pachyops 'Tapajós' species that is morphologically similar to P. fourcroi but with a gas bladder more typical of Pachyurus species) were collected in the field or obtained from commercially harvested animals during January and February of 2005 and 2008 (Table 1a). Based on molecular analysis of four genes across 53 sciaenid taxa, Pachyurus appears as sister taxon of Plagioscion (W.-J. Chen, National Taiwan University, pers. comm.). For outgroups we included two strictly marine Sciaenidae (Cynoscion guatucupa and Micropogonias furnieri) (Table 1a). While the systematics of the large family Sciaenidae is far from resolved (see Sasaki, 1989; Boeger & Kritsky, 2003; Casatti, 2003), morphological analyses suggest that the likely marine sister group of Plagioscion is Cynoscion (Chao, 1978). The second outgroup used in this study is part of a set of four candidate genera (Micropogonias, Nebris, Otolothodies, Penna, Paralochurus) that could also be potentially related to Plagioscion (N.L.C., unpublished data).

For the phylogeographic analysis of P. squamosissimus, 152 specimens were sampled from 11 sites across a vast area of the Amazon Basin (approximate riverine distance c. 2200 km) (Table 1b). Our study transect included five major river systems (Amazon, Madeira, Branco, Negro and Tapajós rivers) and hydrochemical settings (white, black and clear water) of the Amazon Basin (Fig. 1, Table 1b). From each sample site we obtained up to 25 individuals. Fish were caught using a beach seine net; they were euthanized, muscle tissue was dissected from behind the dorsal fin and preserved in 95% ethanol, and geographic coordinates and hydrochemical variables were recorded for each sampling location (Table 1b). Voucher specimens were deposited at the ichthyology collection of the Museu Nacional (UFRJ), Rio de Janeiro, Brazil.

**Genetic methods**

Genetic data were obtained from both the mitochondrial (mtDNA) and nuclear (nDNA) genomes. DNA was extracted via a modified salting-out protocol (Sunnucks & Hales, 1996). The mtDNA adenine-thymine transferase subunits 6 and 8 (ATPase 6 and 8) were amplified via polymerase chain reaction (PCR) using primers and conditions specified in Corrigan et al. (2008) and sequenced for all samples. The single-copy nDNA fragment of the recombination activating gene 1 (RAG1) was amplified using a nested PCR for each species, and for putative cryptic Plagioscion species identified using mtDNA data (details below). For the first round PCR, RAG1 was amplified using the primers RAG1-2510F (Li & Orti, 2007) and RAG1-4090R (López et al., 2004). For the second round PCR, RAG1 was amplified using the primers RAG1-2533F and RAG1-4078R (López et al., 2004). Each 30 µL reaction contained 0.6 µm of each primer, 3 mM MgCl2, 0.6 mM of each dNTP, 5X buffer (Promega, Madison, WI, USA) and 1 U Taq polymerase (Promega). Amplification was carried out using the following programme: 94 °C for 3 min, followed by 35 cycles at 94 °C for 30 s, 55 °C for 45 s, 72 °C for 45 s, and a final extension of 72 °C for 10 min.

Using 2% TRIS-acetate-EDTA (TAE) agarose gel electrophoresis, PCR products were separated, excised and purified using an ULTRA CLEAN 15 DNA purification kit (Mo Bio Laboratories, West Carlsbad, CA, USA). Purified DNA was sequenced in an ABI3730xl sequencer.

**Phylogenetic analyses**

All sequence data were aligned using SEQUENCHER 4.1 (Gene Codes Corporation, Ann Arbor, MI, USA) and submitted to GenBank (accession numbers JN683723–JN683822). To reconstruct phylogenetic relationships we employed character-based and model-based (maximum parsimony, MP, maximum likelihood, ML, and Bayesian inference, BI) methods of analysis for both ATPase 6 and 8 and RAG1 sequence data. Analyses based on MP were conducted in PAUP* 4.0b10
### Table 1
Sciaenidae species and sample locations: (a) species used in the phylogenetic analysis, their collection locality ('upper', 'middle' or 'lower' Amazon Basin, 'Paraná' Basin or 'marine' for samples from the south coast of Brazil) and markers sequenced, and (b) *Plagioscion squamosissimus* samples used in the phylogeographic analysis, including sampling locations, sample size (*n*) and water colours.

#### (a) Phylogenetic analysis

<table>
<thead>
<tr>
<th>Species</th>
<th>Collection locality</th>
<th>Markers sequenced</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Plagioscion squamosissimus</em></td>
<td>Upper, middle</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Plagioscion auratus</em></td>
<td>Middle</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Plagioscion magdalenae</em></td>
<td>Middle</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Plagioscion ternetzi</em></td>
<td>Paraná</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Plagioscion montei</em></td>
<td>Middle</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Pachyrops foucroi</em></td>
<td>Middle</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Pachyrops 'Tapajós' (new sp.)</em></td>
<td>Middle</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Pachyurus junki</em></td>
<td>Middle</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td><em>Cynoscion guatacupa</em></td>
<td>Marine</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
<tr>
<td>* Micropogonias furnieri*</td>
<td>Marine</td>
<td>ATPase 6 and 8, RAG1</td>
</tr>
</tbody>
</table>

#### (b) Phylogeographic analysis of *P. squamosissimus*

<table>
<thead>
<tr>
<th>River</th>
<th>Site</th>
<th><em>n</em></th>
<th>Water colour</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branco</td>
<td>B2</td>
<td>9</td>
<td>Black</td>
<td>1°03'23.00&quot; S</td>
<td>61°51'29.00&quot; W</td>
</tr>
<tr>
<td>Branco</td>
<td>B1</td>
<td>5</td>
<td>Black</td>
<td>1°15'59.00&quot; S</td>
<td>61°50'55.00&quot; W</td>
</tr>
<tr>
<td>Negro</td>
<td>N1</td>
<td>22</td>
<td>Black</td>
<td>3°4'44.00&quot; S</td>
<td>60°14'44.00&quot; W</td>
</tr>
<tr>
<td>Madeira</td>
<td>M1</td>
<td>20</td>
<td>White</td>
<td>3°28'14.00&quot; S</td>
<td>58°52'5.00&quot; W</td>
</tr>
<tr>
<td>Amazon</td>
<td>A1</td>
<td>20</td>
<td>White</td>
<td>3°20'40.00&quot; S</td>
<td>60°7'10.00&quot; W</td>
</tr>
<tr>
<td>Amazon</td>
<td>A2</td>
<td>12</td>
<td>White</td>
<td>3°65'6.00&quot; S</td>
<td>59°32'19.00&quot; W</td>
</tr>
<tr>
<td>Amazon</td>
<td>A3</td>
<td>18</td>
<td>White</td>
<td>3°47'39.00&quot; S</td>
<td>58°13'13.00&quot; W</td>
</tr>
<tr>
<td>Amazon</td>
<td>A4</td>
<td>18</td>
<td>White</td>
<td>2°33'7.00&quot; S</td>
<td>57°15'59.00&quot; W</td>
</tr>
<tr>
<td>Amazon</td>
<td>A5</td>
<td>5</td>
<td>White</td>
<td>2°10'21.00&quot; S</td>
<td>54°58'21.00&quot; W</td>
</tr>
<tr>
<td>Amazon</td>
<td>A6</td>
<td>4</td>
<td>White</td>
<td>2°28'10.00&quot; S</td>
<td>54°30'5.00&quot; W</td>
</tr>
<tr>
<td>Rio Tapajós</td>
<td>T1</td>
<td>14</td>
<td>Clear</td>
<td>2°52'17.00&quot; S</td>
<td>55°9'38.00&quot; W</td>
</tr>
</tbody>
</table>

#### Figure 1
Sampling localities of *Plagioscion squamosissimus* in the Amazon Basin. Each site is distinguished by a site label shaded according to water colour. Black water sites are labelled B2, B1 and N1; white water sites are labelled A1–A6 and M1; and clear water sites are labelled T1. The inset identifies the study area within northern South America.
(Swofford, 2003) using a heuristic search strategy for the most parsimonious tree. All characters were treated as unordered and unweighted. Bootstrap resampling based on 1000 replicates was used to assess support of relationships.

For ML phylogenetic analyses we used Modeltest 3.06 (Posada & Crandall, 1998) to estimate the most likely model of sequence evolution for our ATPase 6 and 8 and RAG1 data. Based on the Akaike information criterion (AIC), Tamura–Nei (I+G) was selected as the most likely model of sequence evolution for both ATPase 6 and 8 and RAG1. Corrected genetic distances based on 816 bp of ATPase 6 and 8 and 1498 bp of RAG1 were calculated in paup* 4.0b10. ML trees were obtained in paup* 4.0b10 using model parameters specified by Modeltest and also assessed with 1000 bootstrap replicates.

Bayesian inference was conducted using MrBayes 3.1.2 (Huelsenbeck & Ronquist, 2001). Here, MrModeltest 2 (Nylander, 2004) was employed to determine the best fit model of sequence evolution for both ATPase 6 and 8 and RAG1. Based on the AIC, the general time reversible (I+G) model was selected for both datasets. Using Bayesian inference we also constructed a concatenated phylogenetic tree incorporating the appropriate parameters of sequence evolution for each gene. The Bayesian analysis was run using a Metropolis-coupled Markov chain Monte Carlo (MCMC) algorithm from randomly generated starting trees for six million generations with sampling every 1000 generations. Both the standard deviation of splits frequencies and the potential scale reduction factor were used as a convergence diagnostic and the initial 25% of samples were discarded in the burn-in.

**Molecular dating of speciation events**

Divergence times using the concatenated BI phylogeny were estimated using the MCMC approach implemented in the program BEAST 1.6.1 (Drummond & Rambaut, 2007). There is some uncertainty about the correct marine sister taxon of Plagioscion, so date estimates of freshwater invasion should be interpreted with caution. Nonetheless, our tree was calibrated using the oldest known Plagioscion and Pachypops fossils (Monsch, 1998) and the marine transgressions of the middle Miocene. Furthermore, we calibrated isolation of Plagioscion magdalenae with the isolation of the Magdalena Basin (Lundberg et al., 1998). Using these calibrations, we estimated the time to most recent common ancestor (TMRCa) for each species using the relaxed clock method that allows for branch specific variation, drawn from a uncorrelated lognormal distribution (Drummond et al., 2006). Tree priors were modelled according to Yule speciation process, the age at the node of Plagioscion and Pachyrus/Pachypops was set to 22 ± 6 Ma and the age of the node of Plagioscion magdalenae was set to 12 ± 4 Ma. All other priors were set at their default values. The MCMC analyses were run for 60 million generations and sampled every 100th generation with the first 10% of samples discarded as a burn-in. Using Tracer 1.4 (Rambaut & Drummond, 2006), we examined our results and confirmed that sufficient effective sample sizes had been achieved (> 200) and stationarity was reached. The analysis was run three times independently to validate these results.

**Phylogeographic analysis of Plagioscion squamosissimus**

Genealogical relationships between 152 specimens identified as P. squamosissimus were investigated by constructing a haplotype network from mtDNA ATPase 6 and 8 sequence data in TCS 1.21 (Templeton et al., 1992; Clement et al., 2000). Summary statistics including nucleotide diversity (π) and mean number of pairwise differences (PWD) were estimated for each population in Arlequin 3.01 (Excoffier et al., 2005). The demographic history of P. squamosissimus in the Amazon Basin was assessed by comparing a mismatch distribution of the data with a model of demographic growth (Rodgers & Harpending, 1992; Excoffier et al., 2005). Mismatch analysis calculates the estimator of time to growth (τ) and the mutation parameter (θ) (Schneider & Excoffier, 1999). Using the ATPase 6 and 8 mutation rate (μ) of 1.4% per million years after Bermingham et al. (1997), we applied the formula τ = τ/2μ to estimate the timing of the demographic growth. To further assess the signal of demographic growth, we applied Fu’s F_s (1997) test of selective neutrality to dataset. Large negative F_s values are generally observed under demographic growth (Fu, 1997).

To assess levels of genetic differentiation among populations of P. squamosissimus, we calculated the fixation index ΦST in Arlequin. We also used this estimator to test for correlations between population genetic structure and geographic distance (isolation by distance, IBD) (Wright, 1943). The significance of the association was tested using a Mantel permutation test (Mantel, 1967) with 1000 matrix randomizations. The degree of population structure associated with water colour was also explored with an analysis of molecular variance (AMOVA) (Excoffier et al., 1992). AMOVA partitions total variance into covariance components, and tests the significance of the variance components associated with chosen hierarchical levels of genetic structure based on the fixation index, ΦST. Two AMOVAs were conducted. AMOVA 1 partitioned the dataset into regions defined by water colour: (1) black (B2, B1, N1), (2) white (A1–A6, M1), and (3) clear water (T1) populations. AMOVA 2 tested for population subdivision between two true white water rivers: (1) the Madeira River (M1) and (2) the Amazon River (A1–A5).

**RESULTS**

**Sequence characteristics**

We generated 816 bp of the mtDNA ATPase 6 and 8 genes and 1476 bp of the nuclear gene RAG1 for all Plagioscion species, outgroup taxa and major lineages of P. squamosissimus identified in the phylogeographic analyses. The mtDNA dataset were characterized by 344 variable characters, of which 284 were parsimony informative. The base frequencies
were 28.5% A, 38.4% C, 9.40% G and 23.7% T. Tamura–Nei (1+G) corrected pairwise differences ranged from 0.048 between *P. squamosissimus* from Meta River, Colombia and the Amazon River, to 0.650 between *Cynoscion guatucupa* and *Pachyurus fourcrroi*. For the nuclear gene *RAG1*, the base frequencies were 23.3% A, 27.6% C, 23.4% G and 25.7% T.

**Phylogenetic analysis of mtDNA and nuDNA**

For the mtDNA ATPase 6 and 8 dataset all methods of phylogenetic analyses resulted in largely concordant tree topologies. They strongly supported the monophyly of *Plagioscion* in which *P. auratus* appears as the basal lineage, as well as the monophyly of *Pachyurus* + *Pachyops* in which *Pachyurus junki* appears basal. We present only the ML tree with bootstrap values for each method (Fig. 1a and Appendix S1 in Supporting Information). In each case, the *Plagioscion* and *Pachyurus* + *Pachyops* clades were strongly supported. Inter-specific relationships within the *Pachyurus* + *Pachyops* clade were well resolved with strong bootstrap support. Within the *Plagioscion* clade, however, two putative cryptic *Plagioscion* species (these were identified as *P. squamosissimus* in the field) were resolved in the phylogeny, both with strong bootstrap support (herein referred to as *Plagioscion* sp. 1 and sp. 2). *Plagioscion* sp. 1 appears sister to the *P. squamosissimus* clade, whereas *Plagioscion* sp. 2 is closely related to *P. montei*. Additionally, *P. squamosissimus* from the Meta River appeared basal to and relatively divergent from *P. squamosissimus* sampled from the Amazon, Madeira, Negro and Tapajós rivers (based on 152 specimens). Within the *Plagioscion* clade the relationship of *P. magdalenae* with *P. montei* and *P. ternetzi* was poorly supported. In contrast, the placement of *P. squamosissimus* as a recently derived lineage was well supported in all analyses.

For the *RAG1* dataset, only the ML and the BI resulted in consistent tree topologies (Fig. 1b, Appendix S1). While all methods of phylogenetic analysis produced well-supported *Plagioscion* and *Pachyurus* + *Pachyops* clades, inter-specific relationships within *Plagioscion* were not consistent in the MP analysis and the placement of *P. squamosissimus* was poorly resolved. For these reasons we decided to base our discussion on results from the ML and BI analyses. Inter-specific relationships between *Pachyurus* species were consistent with mitochondrial results. However, this was not the case between *Plagioscion* species, which generally showed clades with poor statistical support. This was particularly true regarding the placement of *P. magdalenae*, *P. montei*, *P. ternetzi* and *Plagioscion* sp. 1. Nonetheless, as with mtDNA analyses, *P. auratus* appears basal, *Plagioscion* sp. 2 is closely related to *P. montei* and *P. squamosissimus* is a recently derived lineage. As expected for a slowly evolving sequence marker, *P. squamosissimus* for the Meta River could not be distinguished from other *P. squamosissimus* samples by the *RAG* dataset.

The concatenated ATPase 6, 8 and *RAG1* BI tree was fully resolved regarding relationships among species of *Plagioscion* (Fig. 2a). In this analysis, *P. auratus* appears sister to *P. magdalenae*, followed by two sister clades consisting of the remaining *Plagioscion* species. *Plagioscion ternetzi* appears basal to *Plagioscion* cryptic sp. 1 and *P. squamosissimus*, while *P. montei* and *Plagioscion* cryptic sp. 2 form a separate clade.

**Divergence estimates**

Estimates of TMRCA and their corresponding 95% lower and upper highest probability densities are shown in Fig. 3 and Table 2. Assuming that *Pachyurus* + *Pachyops* are sister to *Plagioscion*, our results indicate a single freshwater lineage composed of *Pachyurus* + *Pachyops* + *Plagioscion* estimated to be 31 Ma (Oligocene). Based on this result, divergence between marine and freshwater lineages probably took place during the early to middle Oligocene. The minimum age estimates of the *Plagioscion* and *Pachyurus* + *Pachyops* clades suggest that *Pachyurus* + *Pachyops* may have radiated before *Plagioscion* during the late Oligocene to early Miocene, while radiation within the *Plagioscion* group began towards the end of the Miocene around 15 Ma.

**Phylogeography and adaptation in *Plagioscion squamosissimus***

Of the 152 *P. squamosissimus* samples sequenced from 11 sites within the Amazon Basin (excluding the Meta River), we identified 74 unique haplotypes. Within these there were 89 variable characters and 56 were parsimony informative. The haplotype network linked all *P. squamosissimus* samples, except those sampled from the Meta River, based on a criterion of 95% statistical parsimony (Fig. 4). tcs identified a haplotype found only in white water (Amazon and Madeira rivers) as ancestral. According to this network, haplotypes from black or clear water rivers are more derived than the majority of white water haplotypes. Within the network, there appears to be little correlation between haplotype and geographic distance. This is because, in many instances, the clear water haplotypes are more closely related to the black water haplotypes that are over 1000 km away than to neighbouring white water haplotypes. Contrary to this result, the Mantel test revealed a significant positive correlation between genetic differentiation and geographic distance (*r* = 0.33, *P* = 0.029). However, when black and clear water sites were excluded from the Mantel test, the correlation between distance and genetic differentiation increased markedly (*r* = 0.62, *P* < 0.000), suggesting that IBD between white water sites contributed substantially to the overall signal of IBD. Based on the association of haplotype with water colour, rather than with tributary or river system, we suggest a more complex model of diversification than a simple IBD model (discussed further below).
Consistent with event(s) of historical population expansion, the haplotype network displays a star-like configuration (excluding Meta River samples). Indeed, based on mismatch analysis, our data did not deviate from a model of demographic expansion (Fig. 5; sum of squared deviations = 0.0046, \( P = 0.619 \); raggedness index = 0.0072, 0.913), a result corroborated by a large negative \( F_S \) value (\( F_S = -25.19 \), \( P < 0.001 \)). Assuming the ATPase mutation rate of Bermingham et al. (1997), estimated \( \tau \) values suggest that the demographic expansion occurred approximately 2.6 Ma, during the Pleistocene (range = 1–3.6 Ma, \( \alpha = 0.05 \)).

On average, genetic diversity was moderate within each population (pairwise distance = 5.34, \( \pi = 0.006 \)). Significant population structure based on \( \theta_{ST} \) analysis was detected between populations sampled from black, white and clear water rivers, indicative of reduced gene flow between water colours (Table 3). In contrast, non-significant population structure was detected between samples from white water rivers, including comparisons between different rivers (e.g. Amazon and Madeira rivers). While there was some population structure between the Branco (seasonally black/white) and Negro rivers (black), this was considerably lower than pairwise \( \theta_{ST} \) comparisons between black and white water populations. In support of this pattern, AMOVA (Table 4) found that 25.07% of the genetic variation within our sample was accounted for by differences between water colours (\( P < 0.0001 \)). On the other hand, there were no detectable differences between river systems of the same water colour (Madeira River versus Amazon River; \( P = 0.41 \)).

**DISCUSSION**

In this study we propose a phylogenetic hypothesis for the marine-derived sciaenid *Plagioscion* based on mitochondrial and nuclear sequences obtained for all valid species of the genus. Our analyses are consistent with a biogeographic scenario of early Miocene incursions into freshwaters of South America and also disclose two putatively new cryptic species of *Plagioscion*. In addition, we present a phylogeographic analysis of *P. squamosissimus* (a commercially important species in northern South America) using a sizeable sample collected from a vast area in the Amazon Basin. The latter analysis provides support for the theory that large-scale hydrochemical gradients found in Amazonia act as ecological barriers, maintaining abrupt population discontinuities in *P. squamosissimus*. In this way, our study highlights the importance of divergent natural selection through time in the generation and maintenance of sciaenid diversity in the Neotropics.
Age and trajectory of sciaenid invasions into freshwater

Multidisciplinary studies involving palaeoenvironmental data are providing an ever increasing array of evidence supporting the marine influence in South America, particularly during the Miocene (Lundberg et al., 1998; Hovikoski et al., 2010). Indeed, it is the Miocene marine incursions that are thought to have contributed most substantially to modern Amazonian biodiversity, especially within the region of focus of this study (e.g. Lovejoy & Collette, 2001; Lovejoy et al., 2006; Bloom & Lovejoy, 2011). Beginning in the late Oligocene (c. 30–20 Ma), these marine transgressions permeated the eastern Orinoco and Magdalena basins as well as eastern Amazonia and the region of Mar del Plata and Patagonia (Lundberg et al., 1998) (Fig. 6a). However, 20–11.8 Ma marked the beginning of more extensive marine transgressions (Fig. 6b). During that time the ‘Pebas system’ in the western Amazon was formed. This relatively long-lived basin consisted of brackish to freshwater salinities that drained northwards towards the Caribbean via the large body of marine water known as the ‘Pebasian Sea’ in western Venezuela (Lundberg et al., 1998; Wesselingh et al., 2002; Hovikoski et al., 2010) (Fig. 6c). Also during that time, an extensive marine incursion known as the Paraná Sea occurred in the Mar del Plata region in the Paraná estuary reaching as far north as the Chapeare buttress in central-eastern Bolivia (Lundberg et al., 1998).

Considering the extent and frequency of marine incursions into South America, one or more of these transgressions and interchanges with freshwater could have facilitated the adaptation of sciaenids from a marine to freshwater environment. While Plagioscion and Pachyurus + Pachypops may or may not have invaded South America separately, our estimates suggest that, regardless, it was probably the less extensive marine

Table 2 Time to most recent common ancestor (TMRCA) statistics and 95% lower and upper highest probability density (HPD) as calculated in BEAST 1.6.1 using the relaxed clock method that allows for branch specific variation, drawn from a lognormal distribution for Plagioscion squamosissimus in the Amazon Basin. Node number is as shown in Fig. 3.

<table>
<thead>
<tr>
<th>Node number</th>
<th>TMRCA (Ma)</th>
<th>95% HPD lower (Ma)</th>
<th>95% HPD higher (Ma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.182</td>
<td>43.6856</td>
<td>19.1322</td>
</tr>
<tr>
<td>2</td>
<td>17.764</td>
<td>22.9288</td>
<td>12.5772</td>
</tr>
<tr>
<td>3</td>
<td>15.445</td>
<td>17.2276</td>
<td>9.6835</td>
</tr>
<tr>
<td>4</td>
<td>12.673</td>
<td>16.4343</td>
<td>8.9678</td>
</tr>
<tr>
<td>5</td>
<td>11.381</td>
<td>15.1374</td>
<td>7.8365</td>
</tr>
<tr>
<td>6</td>
<td>9.216</td>
<td>12.71</td>
<td>5.93</td>
</tr>
<tr>
<td>7</td>
<td>3.329</td>
<td>5.547</td>
<td>1.3299</td>
</tr>
<tr>
<td>8</td>
<td>7.155</td>
<td>10.8677</td>
<td>3.5247</td>
</tr>
<tr>
<td>9</td>
<td>19.48</td>
<td>28.4705</td>
<td>10.8676</td>
</tr>
<tr>
<td>10</td>
<td>11.589</td>
<td>18.0703</td>
<td>5.605</td>
</tr>
</tbody>
</table>
incursions of the late Oligocene and early Miocene that facilitated the invasion(s) (Fig. 4). While the date of freshwater invasion estimated here might be tentative, its range is nonetheless consistent with marine incursions of that time. Because the distributions of *Plagioscion* and *Pachyurus* + *Pachypops* species span both the northern and southern basins of South America, there is debate surrounding the transgression responsible for the freshwater adaptation (i.e. northern or southern). In a northern invasion, the northern-most river basins would be expected to host basal extant lineages. Likewise, in a southern invasion, the most basal lineages should be found in the Paraná Basin. Based on our phylogenetic analysis (Fig. 3), we are able to address both predictions for *Plagioscion* species. Firstly, we can rule out a southern invasion because the two species distributed throughout the Paraná Basin, *P. squamosissimus* and *P. ternetzi*, are both recently derived lineages. In contrast, the most basal extant *Plagioscion* species, *P. auratus*, has a predominantly north-eastern distribution throughout the eastern Amazon and Orinoco basins (Casatti, 2005). Additionally, the second most basal extant species, *P. magdalenae*, which actually possesses the least derived gas bladder morphology, is distributed throughout the Magdalena and eastern Amazon basins (Chao, 1978, 2003; Casatti, 2005). Therefore, phylogenetic history and biogeography are consistent with a north to south colonization and radiation of *Plagioscion* species as a product of northern marine incursions during the late Oligocene.

Our results are consistent with both fossil and geomorphological data. In particular, the presence of *P. magdalenae* in the Magdalena Basin supports an early northern colonization as the Magdalena has been isolated from the remaining northern

---

**Figure 4** Statistical parsimony network for mitochondrial DNA ATPase 6 and 8 haplotypes for *Plagioscion squamosissimus* sampled in the Amazon Basin. Relationships among haplotypes are estimated using the parsimony method of Templeton *et al.* (1992). Each circle denotes a unique haplotype and the area of the circle is proportional to its frequency. The shade(s) of the circle represents the water colour of the sampling locality, as in Fig. 1.

**Figure 5** Mismatch distribution of all individuals of *Plagioscion squamosissimus* sampled from the Amazon Basin (n = 152). The black solid line represents the observed relative frequencies of nucleotide differences between pairs of individuals, the dark grey dashed line represents the distribution fitted to the data under a model of demographic expansion, and the dashed light grey lines represent the 95% confidence interval values of 1000 simulations.
basins since the late Miocene (c. 11.8 Ma) (Lundberg et al., 1998). In addition, the oldest known Plagioscion fossils are from the Pebas Formation in north-western South American dating to the middle Miocene (c. 15 Ma) (Monsch, 1998). These fossils actually pre-date the major marine transgressions that formed the Pebsian and Paranan seas between 11.8 and

<table>
<thead>
<tr>
<th>Table 3 Pairwise $\theta_{ST}$ comparisons for the 11 sampled populations of Plagioscion squamosissimus in the Amazon Basin. Sample sites have been grouped by geographic proximity and water colour (see Table 1 for more details).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>B1</td>
</tr>
<tr>
<td>0</td>
</tr>
</tbody>
</table>

* $P \leq 0.05$.

<table>
<thead>
<tr>
<th>Table 4 Analysis of molecular variance (AMOVA) for mitochondrial DNA genes ATPase 6 and 8 in Plagioscion squamosissimus. In AMOVA 1, regions include: (1) seasonally black/white (B1), (2) black (N1), (3) white (S1, A1, M1, A2, A3, A4, A5) and (4), clear water populations (T1). In AMOVA 2 regions include: (1) the Rio Madeira, and (2) white water sites A1–A5.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black versus white versus clear water</td>
</tr>
<tr>
<td>Source of variation</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Between water colours</td>
</tr>
<tr>
<td>Between populations</td>
</tr>
<tr>
<td>Between individuals</td>
</tr>
</tbody>
</table>

$\Phi$, fixation index; significant results are indicated by *.

Figure 6 Palaeogeography of South America, depicting the location and frequency of marine transgressions since the middle Tertiary (modified from Lundberg et al., 1998).
10 Ma (Fig. 6c) and indicate that adaptation to freshwaters by Plagioscion probably took place earlier (i.e. late Oligocene and early Miocene), as suggested by our Bayesian molecular dating (Fig. 4). Accordingly, it is thought that freshwater stingrays (Potamotrygon) (Lovejoy et al., 1998), freshwater anchovies (Juregnudis and Anchovia) (Lovejoy et al., 2006) and needlefish (Potamorrhaphis) (Lovejoy & Collette, 2001) share similar age estimates in relation to palaeogeographic events between the late Oligocene to early Miocene.

Considering the fossil record, geomorphological history, distributional data of extant Plagioscion species and our phylogenetic results, it is most likely that marine transgressions into western Venezuela between the late Oligocene and early Miocene were responsible for the transition to freshwaters in Plagioscion species. All other marine transgressions occurring within northern South America were either small or geographically isolated from Lake Pebas where fossil Plagioscion species were reported. The Orinoco transgression, for example, was isolated by the Andean Arc, during which time the eastern Amazon transgression was also isolated from the west by the Purus Arch (Lundberg et al., 1998). Thus, by the early Miocene, Plagioscion species had probably adapted to the freshwaters of South America and were establishing a southerly colonization route.

By the middle Miocene both Plagioscion and Pachyurus + Pachyops species were likely to have been well established in the freshwaters of South America. However, it was not until the end of the marine transgressions during the late Miocene that much of the Plagioscion radiation begun. It was during this time that the modern Amazon River and its largest tributaries were becoming established, opening up new available habitat and fusing western and eastern biotas (Hoorn et al., 1995; Lundberg et al., 1998) (Fig. 6d). While the full establishment of the Amazon River did not occur until the late Pliocene, the Amazon River had started to form as a compound response to the breaching of the eastern rim of the sedimentary basin known as the Madre de Dios Formation (Içã Formation in Brazil) and erosion of the proto-Amazon River (Hoorn et al., 1995; Lundberg et al., 1998; Rossetti et al., 2005; Campbell et al., 2006) (Fig. 6e). In fact, based on our phylogeny, each Plagioscion species (including putative cryptic species) that arose during that time is distributed within the modern Amazonian River and its tributaries east of the Madre de Dios formation. What this suggests is that from about 10–15 Ma until recently, conditions for speciation (via adaptive divergence and/or vicariance) have favoured the evolution of the genus Plagioscion. For this reason, we move forward in time and discuss phylogeographic relationships between populations of P. squamosissimus within the Amazon River and its major tributaries to identify recent mechanisms that might have influenced Plagioscion diversity.

Phylogeography and adaptation of Plagioscion squamosissimus

Plagioscion squamosissimus originated within South America sometime between the late Miocene and early Pliocene based on our estimates of TMRCA (Fig. 3, Table 2). During that time, the Amazon River had begun establishing and the marine incursions into South America had terminated (Hoorn et al., 1995; Lundberg et al., 1998). Our results suggest that P. squamosissimus arose in north-western South America prior to the full establishment of the modern Amazon River. For instance, P. squamosissimus sampled from Meta River in Colombia appear basal to those from the Amazon River (TMRCA c. 3.33 Ma). Following the full establishment of the Amazon River system during the late Pliocene to early Pleistocene c. 2.5 Ma (Campbell et al., 2006), subsequent colonization and population expansion within this new available habitat would have occurred. Consistent with this inference is the estimated time of demographic expansion within this region for P. squamosissimus at approximately 2.6 Ma (Table 3).

The final establishment of the modern Amazon River may not only have facilitated population and range expansion, but it may also have promoted adaptive divergence. Our data suggest that the different mineralogical and geochemical compositions (i.e. water colours) of the Amazon Basin may have also driven divergent natural selection between these environments. For instance, the genealogy of P. squamosissimus shows a pronounced association of haplotype with water colour, with black and clear water haplotypes derived from presumably ancestral white water haplotypes. Overall, this genealogy and other summary statistics derived from our phylogeographic dataset (Figs 5 & 6, Table 3) are consistent with Pliocene–Pleistocene population expansion (Avise, 2000) with additional signals of local haplotype radiations within black and white water river systems in particular. However, there are also several haplotypes shared between individuals from different water colours, further indicating that a simple model of allopatric divergence between ecotypes is not sufficient to explain the association of haplotype and water colour. Rather, our data probably represent a model of parapatric divergence, or divergence-with-gene-flow across water colours, in which an impermeable geographic barrier is not necessary to promote reproductive isolation. Instead, gene flow is reduced but not terminated across hydrochemical gradients between populations, a consequence of divergent natural selection that may eventuate in reproductive isolation via ecological speciation (Endler, 1977; Schluter, 2000; Coyne & Orr, 2004).

Despite the wide geographic distribution of P. squamosissimus (Casatti, 2005) our frequency-based analysis revealed substantial genetic structure associated with water colour. While we detected a significant association between genetic differentiation and geographic distance, we also found higher levels of genetic differentiation between selective environments than within them (Tables 3 & 4), corroborating our inferences made by genealogy-based analyses. In particular, using AMOVA we found that population differentiation was highly significant between white, black and clear water populations, but not significant between white water populations from different rivers (Amazon and Madeira rivers). This result was
particularly important as it ruled out the possibility of population structure generated by allopatric isolation within a tributary regardless of its water colour.

Our phylogeographic scenario for *P. squamosissimus* is likely to include the dimension of selection in addition to vicariant biogeographic history. Following the final establishment of the Amazon River, white water ecotypes of *P. squamosissimus* would have colonized the newly available habitat in an easterly direction, invading black and clear water habitats and exploiting new niches. Reduced competition in black and clear waters would have facilitated rapid adaptation and population expansion, driving evolution in parapatry with related white water populations. Indeed, it has been shown that ecologically based adaptive evolution can actually be reinforced related white water populations. Indeed, it has been shown that divergent natural selection would have been necessary to overcome the homogenizing effects of gene flow between river systems of different water colours.

Despite the fact that most notable examples of divergent natural selection involve freshwater fish (Schluter, 1995; Beheregaray & Sunnucks, 2001; Barluenga & Meyer, 2004; McKinnon et al., 2004; Bernatchez et al., 2010; Elmer et al., 2010), the process of adaptation driven by ecological processes has received little attention in studies of Amazonian speciation. Nonetheless, by identifying loci under selection using genome scans we have detected divergent natural selection between black, white and clear water rivers of the Amazon Basin across taxonomically diverse fishes (G.M.C., N.L.C. & L.B.B., unpublished). Furthermore, results for *P. squamosissimus* presented here are consistent with the overall pattern of reduced gene flow due to hydrochemical gradients observed in our other studies. In the light of this, we propose that the phylogeographic history of *P. squamosissimus* has been influenced by a combination of ecology and geomorphological history.

**CONCLUSIONS**

Here we have presented phylogenetic and biogeographic hypotheses consistent with geomorphological history and fossil data that account for the marine invasion and diversification of the sciaenid genus *Plagioscion* into freshwaters of South America. *Plagioscion* probably invaded the freshwaters of South America via a northern route during late Oligocene and early Miocene marine incursions, with the subsequent establishment of a southerly colonization route. However, diversification of this genus appears to have largely taken place following the cessation of the marine incursions during the late Miocene and the birth of the Amazon River system. Within *P. squamosissimus*, diversification and population structure also appear to be associated with water colour and ecological gradients. Unexpectedly, our work also revealed two new putative cryptic *Plagioscion* species, suggesting a taxonomic revision of the genus and adding to the ever-growing body of research detecting cryptic Amazonian fish species (e.g. Littmann et al., 2001; Hubert et al., 2007; Sistrom et al., 2009; Piggott et al., 2011). In conclusion, we show that divergent natural selection resulting in the adaptation of sciaenids to freshwaters has persisted through time and probably continues as a diversifying mechanism within contemporary populations of *P. squamosissimus* across freshwater hydrochemical gradients within the Amazon Basin. This result has implications for how we attribute mechanisms of diversification, such as geomorphological events and ecological settings, to the evolution of Amazonian fish diversity.

**ACKNOWLEDGEMENTS**

We thank Wei-Jen Chen for providing *Plagioscion tarnetzi*, *P. auratus*, *P. magdalenae* and *P. montei* samples. This study was funded by the Discovery Program of the Australian Research Council (ARC grant DP0556496 to L.B.B.) and by Macquarie University through a postgraduate travel grant and research award to G.M.C. Logistics and local arrangements were supported in part through the Brazilian National Council of Research and Technology CNPq-SEAP no. 408782/2006-4. Collection permit is under IBAMA no. 1920550, and ethical approval was received from Macquarie University, Approval number: 2007/033 under G.M.C.

**REFERENCES**


Proceedings of the National Academy of Sciences USA, 96, 13869–13873.

SUPPORTING INFORMATION
Additional Supporting Information may be found in the online version of this article:
Appendix S1 Maximum likelihood phylogenetic trees showing the genetic relationships inferred from (a) the mitochondrial ATPase 6 and 8 genes and (b) the nuclear RAG1 gene for Plagioscion and Pachyurus + Pachypops species within South America.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES
This study forms part of Georgina Cooke’s PhD research supervised by Luciano Beheregaray and Ning Chao. The authors share similar research interests in understanding the processes underpinning the evolution of biodiversity.

Editor: Luiz Rocha