DNA barcoding reveals cryptic diversity and peculiar phylogeographic patterns in mojarras (Perciformes: Gerreidae) from the Caribbean and South-western Atlantic

Uedson Pereira Jacobina, Rodrigo Augusto Torres, Paulo Roberto Antunes de Mello Affonso, Ewerton Vieira dos Santos, Leonardo Luiz Calado and Jamille de Araújo Bitencourt

Abstract

The mojarras (Eucinostomus) are a widespread group of coastal fishes of controversial taxonomy because of similarities in their external morphology. In the present study, we assessed the genetic diversity of species and populations of Eucinostomus using DNA barcodes using a systematic and phylogeographic context. In total, 416 COI sequences of all valid Eucinostomus representatives were analysed based on public databases and collected specimens from the north-eastern coast of Brazil (Western South Atlantic). Several cases of misidentification were detected in the barcode dataset (E. argenteus, E. harengulus, E. gula, E. dowii and E. jonesii) that could account for the taxonomic issues in this genus. In contrast, we identified four molecular operational taxonomic units (MOTUs), with divergence above 2% in the Western Atlantic, that correspond to cryptic forms within E. argenteus, E. harengulus, E. gula and E. melanopterus. These data suggest that Plio-Pleistocene events (rise of the Panama isthmus, Amazonas outflow and sea-level fluctuations) played a major role in the diversification of mojarras. While subtle morphological differences have been used as proxies to discriminate Eucinostomus species, the genetic data proved to be efficient in differentiating them and revealing potentially undescribed taxa. Therefore, we recommend that further taxonomic studies in mojarras should incorporate DNA-based evidence.

Introduction

The apparent weakness of geographic barriers over large distances in marine environments hinders the identification of cladogenesis and several cases of cryptic lineages might remain overlooked (Rocha, 2003; Luiz et al., 2012; Da Silva et al., 2016). Moreover, considering that speciation processes might take place without major changes in morphology, the actual number of species might be underestimated when external body features are analysed in isolation (Winker, 2005). Failure of species diagnosis has a particularly negative impact on economically important species in fisheries where distinct morphotypes putatively related to a single species might encompass multiple evolutionary units (Mayden, 1997; Winker, 2005; Da Silva et al., 2018).

The mojarras (genus Eucinostomus, family Gerreidae) comprise 11 nominal taxa (Froese & Pauly, 2019) widespread in the Western Atlantic (WA; seven species) and Eastern Pacific (EP; four species) that have commercial relevance to local fisheries. These fish are found in coastal areas, including estuaries and hypersaline lakes because of their high osmoregulation adaptability (Nelson et al., 2014). Morphologically, they are characterized by dorsal compression, protractile mouth, smooth gill arches, cycloid scales on the head and ctenoid scales over the body (Nelson et al., 2016). From a systematic viewpoint, the genus Eucinostomus is considered one of the most problematic genera of coastal fishes from the New World since their putative interspecific morphological traits are subtle and usually overlapped, while their evolutionary interrelationships remain poorly known (Matheson & McEachran, 1984; De La Cruz-Agüero & Galvan-Magana, 1993; De La Cruz-Agüero, 2013). Because of their controversial taxonomy, some species in this group, such as Ulaema lefroyi (E. lefroyi), are regarded as inquierendae (Nelson et al., 2016). In addition, other taxonomic methods such as cytotaxonomic analyses have been inefficient in discriminating mojarras species (Calado et al., 2014).

The difficulties in recognizing Eucinostomus species based on traditional taxonomy were recently reported in a wide study of DNA barcoding in fish species from north-eastern Brazil, WA (Brandão et al., 2016). These authors verified that 82 specimens of mojarras...
morphologically identified as *Eucinostomus melanopterus* or only to the genus level (*Eucinostomus* sp.) encompassed five molecular units with deep genetic divergence, representing *E. melanopterus* (N = 5), *E. lefroyi* (N = 7), *E. gula* (N = 2), *E. harengulus* (N = 64) and *E. jonesii* (N = 4). Moreover, this work represented the first report about the occurrence of *E. harengulus* and *E. jonesii* in north-eastern Brazil.

The accurate diagnosis of taxa is a key step to biodiversity conservation (Bickford et al., 2007). The utilization of mitochondrial markers such as the Cytochrome c oxidase subunit I (COI) as DNA barcodes has been particularly useful in the identification of cryptic species and in the resolution of taxonomic uncertainties, with several examples in fishes (Ferreira et al., 2014; Hyde et al., 2014; Winterbottom et al., 2014; Barreira et al., 2016; Nirchio et al., 2018). This approach has also been improved over the last decade by the incorporation of distinct algorithms to test the hypothesis of evolutionary independent lineages and to avoid synonyms in the dataset (Brown et al., 2012; Puillandre et al., 2012; Zhang et al., 2013; Luo et al., 2018).

Considering the problematic systematics, the wide range of mojarras and the efficiency of DNA-based identification, we assessed the genetic diversity of all *Eucinostomus* representatives with available COI sequences throughout most of their geographic distribution. Based on distinct analytical methods, including species delimitation algorithms, we reviewed their systematic relationships within a phylogeographic context, providing evidence of overlooked phylogenetic diversity and variation in the range of evolutionary units, with emphasis in the WA.

### Materials and methods

#### Sampling

A total of 400 COI sequences of *Eucinostomus*, comprising 11 recognized taxa were analysed: *E. argenteus* (N = 90), *E. cf. gula* (N = 6), *E. curranii* (N = 7), *E. dawii* (N = 6), *E. entomelas* (N = 10), *E. gracilis* (N = 2), *E. gula* (N = 49), *E. havana* (N = 2), *E. harengulus* (N = 88), *E. jonesii* (N = 86), *E. lefroyi* (N = 11) and *E. melanopterus* (N = 20), as well as other sequences named as *Eucinostomus* sp. (N = 23). These sequences were downloaded from the public datasets, for example Bold Systems and GenBank (NCBI) (supplementary material 1). We also added 16 specimens of mojarras, *E. argenteus* (N = 1), *E. gula* (N = 8), *E. jonesii* (N = 2), *E. lefroyi* (N = 5), collected in the state of Rio Grande do Norte (05°05′03″S 36°16′03″W), north-eastern Brazilian coast, that were identified according to Woodland (2006). Fish were collected by gillnets and euthanasia was accomplished by immersion in cold water for 10–15 min. Afterwards, muscle tissues were removed from each individual and stored in ethanol at −20°C (Blessing et al., 2010).

#### DNA isolation, amplification and sequencing

Total DNA was isolated from stored muscle tissues, using the DNeasy kit (QIAGEN). A fragment of 645 base pairs (bp) of the COI gene was amplified via PCR using the primers VF2_t1 (5′-TGGTAAACGCGCCGATCAACCAACAAAGACAT-TGGCAG3′) and Fish2_r1 (5′-CAGGAGAAGCTGATGACATTCAGGGTGACCGAAGATCAGA3′) as described by Ward et al. (2005). The reactions encompassed 12 μl of 2× Taq master mix (Vivantis), 2 μl of template DNA solution at 40 ng/μl, 0.5 μl of each primer (10 μM) and ultrapure water to a final volume of 25 μl. The PCR steps (adapted from Ward et al., 2005) included a first denaturation at 95°C for 2 min, 35 cycles at 94°C (30 s), 57°C (30 s) and 72°C (2 min), plus a final extension at 72°C for 7 min. The PCR products were purified with ExoSap IT enzymatic system (Affimetrix). The sequencing of COI fragments was carried out using the BigDye™ Terminator v 3.1 Cycle Sequencing Ready Reaction kit (Applied Biosystems) using the M-13 initiator followed by reading in ABI 3130 automatic sequencer (Applied Biosystems).

The electropherograms were checked and edited in the software Geneious (Kearse et al., 2012), followed by visual inspection of consensus sequences for final edition adjustments. In addition, the COI sequences were uploaded and deposited in the BOLD platform under the project ‘Assessing the genetic diversity of species *Eucinostomus* – EUCl’, being automatically assigned to a Barcode Index Number – BIN (group of sequences that should correspond to a single taxon), following the analytical procedures of Ratnasingham & Hebert (2013). Afterwards, these sequences were aligned with those available in Bold Systems and GenBank by using the ClustalW method.

#### Phylogenetic and distance analyses

Phylogenetic reconstructions were carried out based on Maximum likelihood (ML) and Bayesian inference (BI). The best evolutionary model for both ML and BI trees was HKY + I as indicated in the software PartitionFinder (Lanfear et al., 2012). Branch support in ML analysis was based on 1000 bootstrap replicates using RAxML (Stamatakis et al., 2012). In the case of BI, ultrametric trees following the Yule Speciation prior model were generated in the software BEAST 1.8.4 (Drummond & Rambout, 2007), based on 20 million generations with sampling at every 2000 generations. The convergence of Markov chains was inspected in Tracer 1.6 (Drummond et al., 2012). All values of Effective Sample Size (ESS) were above 200. Based on 10% of burn-in, the remaining trees were used to obtain a consensus tree and the branch support was based on the posterior probability values.

Based on BI and ML results, we calculated the genetic distance of each cluster with support values higher than 1 of probability or 95% of bootstrap to build a Neighbour-joining (NJ) tree in the software MEGA6 (Tamura et al., 2013) using the Kimura 2-parameter (K2P) evolutionary model as identified by the Barcode of Life initiative (www.boldsystems.org). All trees were visualized using FigTree v. 1.4.1 (Rambout & Drummond, 2009).

#### Species delimitation methods

To establish a potential threshold among *Eucinostomus* species, we built a distance matrix based on the K2P model using the function spDistMatrix available in the R package SPIDER v1.3–0 (Brown et al., 2012). Based on this matrix, we created a density object with the minimum local function which disregards any previous knowledge about the species identity to indicate potential thresholds to infer intra- and interspecific variation levels (Brown et al., 2012). We also used three widely used algorithms for species delimitation from molecular data: the Bayesian Poisson Tree Process (bPTP) (Zhang et al., 2013), the generalized mixed Yule-coalescent model (GMYC) (Pons et al., 2006) and the Automatic Barcode Gap Discovery (ABGD) (Puillandre et al., 2012).

The bPTP method was carried out in the web server <http://species.h-its.org/ptp/> using as input a non-ultrametric tree based on ML inference. This analysis was performed with 500,000 generations with sampling at every 500 generations and 10% of burn-in. The GMYC was implemented in the web server <http://species.h-its.org/gmyc/> based on an ultrametric tree obtained in BEAST (Drummond et al., 2012). The ABGD was carried out based on the pairwise genetic distances (based on p-distance, K2P and Jukes–Cantor models). The
analysis was done using a gap width value of 1.0 for all distances using the software available in <http://wwwabi.snv.jussieu.fr/public/abgd/>. The congruence among the delimitation of MOTUs was evaluated by comparing the clusters inferred from each algorithm.

**Population analysis**

Haplotype networks were built in the software Popart (Leigh & Bryant, 2015) using the Median-joining network algorithm to elucidate the genealogical relationships of *E. argenteus* and *E. gula*, because both species encompassed a high number of sequences from distinct localities throughout most of their range.

**Results**

The final alignment of COI sequences comprised 593 bp. Insertions, deletions or stop codons were absent, indicating that all sequences correspond to functional COI genes and not to putative pseudogenes or nuclear mitochondrial DNA segments (numts).

Phylogenetic analyses combined with distance methods and tools of molecular identification were helpful in identifying 34 potentially misidentified sequences in the analysed dataset of *Eucinostomus*, related to *E. argenteus*, *E. harengulus*, *E. jonesii*, *E. dowii* and *E. gula* (see supplementary material 2). The genetic variation in these samples ranged from 0.2–1.7%. Therefore, these sequences were reallocated to their respective groups and analysed according to their actual taxonomic classification to avoid biased evolutionary inferences. Sequences with divergence above 2%, which did not fit into any of the described taxonomic species of *Eucinostomus* were included with the acronym ‘cf.’. They are: *E. cf. argenteus*, *E. cf. melanopterus*, *E. cf. harengulus* and *E. cf. gula*.

The BI and ML analyses recovered 15 MOTUs distributed into highly supported clades 1/1=05, as follows: *E. jonesii*, *E. dowii*, *E. cf. argenteus*, *E. harengulus*, *E. cf. harengulus*, *E. gracilis*, *E. gula*, *E. cf. gula*, *E. entomelas*, *E. argenteus*, *E. lefroyi*, *E. melanopterus*, *E. cf. melanopterus*, *E. currani* and *E. havana* (Figure 1).

The species delimitation algorithms (GMYC, bPTP, BINs and ABDG) also separated these clusters, besides revealing four additional MOTUS within *E. gula*, *E. argenteus*, *E. melanopterus* and *E. harengulus* (Figure 1). However, only the barcode index (BINs) was not found for *E. cf. harengulus*, since these sequences are only available in NCBI.

According to the distance method based on the K2P evolutionary model, the highest genetic distance was observed between *E. havana* and *E. argenteus* (21.8%), while the lowest value was detected between *E. cf. melanopterus* and *E. melanopterus* (2.9%). Two sequences erroneously assigned to *E. argenteus* were closely related to *E. harengulus* with 3.3% genetic divergence, being thus referred to as *E. cf. harengulus*. The highest values of intraspecific genetic variation were observed in *E. gula* (1.6%) and *E. cf. gula* (1.2%), while values close to 0% were observed in most of the *Eucinostomus* representatives (Table 1). The threshold potential from intra- to interspecific variation in mojarras as inferred by the minimal local function using Spider was established as 2.7% (Figure 1). This value was in agreement with the species delimitation algorithms where sequences with genetic distances above 2.7% were recovered as distinct MOTUs.

Haplotype networks based on the most representative taxa with information about their geographic origin (*E. argenteus* and *E. gula*) revealed two haplogroups within each. The haplogroups composed of *E. argenteus* + *E. cf. argenteus* and *E. gula* + *E. cf. gula* were separated by 103 mutation steps (genetic distance = 20.7%) and 23 mutation steps (genetic distance = 4.4%), respectively (Figure 2). Particularly within *E. argenteus*/*E. cf. argenteus*, the range of each haplogroup corresponds to the Caribbean and Brazilian provinces.

**Discussion**

Besides identifying cryptic diversity, DNA barcoding analyses have been particularly effective in detecting synonyms and misidentifications among fish (Ferreira et al., 2014; Hyde et al., 2014). Accordingly, the pairwise divergence in COI sequences of *Eucinostomus* associated with comparative tools (BLASTn and Species Level Barcode Records) and phylogenetic inferences detected several cases of misidentification in the sequences available from Bold Systems and GenBank, as observed in *E. argenteus*, which has four BINs representing different species with the same morphological identification. It is possible to observe that two of these BINs include the species *E. jonesii*, *E. dowii*, *E. harengulus* and *Eucinostomus* sp. as well as a possible cryptic lineage for the coast of Brazil. Two clusters were also found for *E. gula*, *E. melanopterus* and *E. harengulus* species. For the latter sequences are only available on GenBank (see supplementary material and Figure 1). Furthermore, some sequences were
Table 1. Matrix of genetic distance based on K2P model among *Eucinostomus* species

<table>
<thead>
<tr>
<th>Species</th>
<th>Distance within species</th>
<th>HAR</th>
<th>HAR1</th>
<th>JON</th>
<th>GUL</th>
<th>GUL1</th>
<th>ARG</th>
<th>ARG1</th>
<th>ENT</th>
<th>DOW</th>
<th>HAV</th>
<th>CUR</th>
<th>LEF</th>
<th>MEL</th>
<th>MEL1</th>
<th>GRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAR</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAR1</td>
<td>0.007</td>
<td>0.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JON</td>
<td>0.003</td>
<td>0.116</td>
<td>0.120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUL</td>
<td>0.016</td>
<td>0.144</td>
<td>0.134</td>
<td>0.138</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUL1</td>
<td>0.012</td>
<td>0.151</td>
<td>0.141</td>
<td>0.143</td>
<td>0.044</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARG</td>
<td>0.004</td>
<td>0.197</td>
<td>0.182</td>
<td>0.181</td>
<td>0.167</td>
<td>0.161</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARG1</td>
<td>0.001</td>
<td>0.139</td>
<td>0.121</td>
<td>0.127</td>
<td>0.171</td>
<td>0.173</td>
<td>0.207</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENT</td>
<td>0.001</td>
<td>0.165</td>
<td>0.164</td>
<td>0.150</td>
<td>0.175</td>
<td>0.165</td>
<td>0.103</td>
<td>0.181</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOW</td>
<td>0.001</td>
<td>0.128</td>
<td>0.117</td>
<td>0.110</td>
<td>0.159</td>
<td>0.158</td>
<td>0.185</td>
<td>0.117</td>
<td>0.157</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAV</td>
<td>n/c</td>
<td>0.203</td>
<td>0.183</td>
<td>0.185</td>
<td>0.211</td>
<td>0.201</td>
<td>0.218</td>
<td>0.207</td>
<td>0.190</td>
<td>0.190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUR</td>
<td>0.001</td>
<td>0.177</td>
<td>0.168</td>
<td>0.203</td>
<td>0.182</td>
<td>0.180</td>
<td>0.163</td>
<td>0.178</td>
<td>0.206</td>
<td>0.206</td>
<td>0.215</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEF</td>
<td>0.004</td>
<td>0.152</td>
<td>0.145</td>
<td>0.146</td>
<td>0.162</td>
<td>0.175</td>
<td>0.165</td>
<td>0.155</td>
<td>0.166</td>
<td>0.207</td>
<td>0.176</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEL</td>
<td>0.005</td>
<td>0.178</td>
<td>0.167</td>
<td>0.194</td>
<td>0.209</td>
<td>0.203</td>
<td>0.193</td>
<td>0.200</td>
<td>0.193</td>
<td>0.202</td>
<td>0.135</td>
<td>0.175</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEL1</td>
<td>0.001</td>
<td>0.132</td>
<td>0.169</td>
<td>0.191</td>
<td>0.185</td>
<td>0.187</td>
<td>0.204</td>
<td>0.198</td>
<td>0.198</td>
<td>0.198</td>
<td>0.202</td>
<td>0.178</td>
<td>0.178</td>
<td>0.029</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRA</td>
<td>0.000</td>
<td>0.170</td>
<td>0.127</td>
<td>0.135</td>
<td>0.129</td>
<td>0.142</td>
<td>0.182</td>
<td>0.168</td>
<td>0.132</td>
<td>0.132</td>
<td>0.192</td>
<td>0.151</td>
<td>0.151</td>
<td>0.182</td>
<td>0.174</td>
<td></td>
</tr>
</tbody>
</table>

identified only at genus level. In general, these sequences are derived from reports about larvae and juveniles of Caribbean fish without any previous information about COI sequences for comparative analysis, thus hindering their precise identification (Valdez-Moreno et al., 2010; Weigt et al., 2012). Indeed, even the morphological identification of adult stages in Eucinostomus species is still under debate (Matheson & McEachran, 1984; De La Cruz-Agüero & Galvan-Magaña, 1993).

Another issue commonly observed in studies using DNA barcodes is the restricted range of sampling from putative widespread taxa (Ward et al., 2005; de Ribeiro et al., 2012; Brandão et al., 2016). This issue restrains further biogeographic inferences, including the potential detection of phylopatric evolutionary lineages (Neves et al., 2016). In this sense, the phylogenetic and species delimitation analyses carried out in the present study were particularly informative in assessing the actual diversity in Eucinostomus, since 15 distinct lineages (MOTUs) were clearly identified. Four MOTUs were related to cryptic forms in the for-

Eucinostomus species delimitation analyses carried out in the present study lineages (Neves et al., 2016). This issue restrains further biogeographic inferences, including the potential detection of phylopatric evolutionary lineages (Neves et al., 2016). In this sense, the phylogenetic and species delimitation analyses carried out in the present study were particularly informative in assessing the actual diversity in Eucinostomus, since 15 distinct lineages (MOTUs) were clearly identified. Four MOTUs were related to cryptic forms in the for-

mal taxa E. argenteus, E. gula, E. melanopterus and E. harengulus from the Western Atlantic (Caribbean and Brazilian provinces). Moreover, the tree topologies suggest distinct evolutionary histories for these lineages (Figure 1), as also supported by their high genetic diversity (20.7% for E. argenteus, 4.4% for E. gula, 3.3% for E. harengulus and 2.9% for E. melanopterus).

Similarly, molecular studies in distinct animal groups have shown the remarkable isolation of Caribbean lineages when compared with other regions (De Biasse et al., 2016; Fields et al., 2016; Hurtado et al., 2017). Furthermore, recent reports have revealed that the biogeographic patterns in fishes from the Brazilian Province are quite heterogeneous, with growing evidence for high levels of endemism according to body size, dispersal routes and environmental features (Argolo et al., 2018; Pinheiro et al., 2018). Therefore, some of these vicariant effects could be responsible for the distinction among the lineages within the complex E. argenteus. In the Caribbean province, the great biodiversity reflects ancient and recent patterns of circulation and topography from the Miocene, when the Americas were not yet connected. Climatic changes during this period conditioned the opening and closing of currents at different depths that modified the circulation pattern, and influenced global climate changes during the Pliocene (Williams & Duda, 2008; Williams et al., 2013; Thacker, 2017). In addition, change in the flow of the Amazon on the north coast of South America between 10–6 mya and closure of the Isthmus of Panama between 5.5–3 mya might have, although temporarily, affected other taxa, such as E. gula, E. melanopterus and E. harengulus. Therefore, major diversification events during the Miocene–Pliocene could have become more and more accentuated during Pleistocene sea level fluctuations, leading to population isolation and divergence of cryptic MOTUs in E. cf. argenteus, E. cf. melanopterus, E. cf. gula and E. cf. harengulus (Lambeck & Chappell, 2011; Müller et al., 2008).

It should also be pointed out that the divergence between the cryptic lineages in E. argenteus was seven times higher than that established as the threshold from intra- to interspecific variation (2.7%). Similar cases of deep variation in COI sequences have also been reported in marine fish such as Scorpaena nornia (18%; Landi et al., 2014), lantern fishes (17–25%; Pappalardo et al., 2015), Gonostomatidae, Sternoptychidae and Myctophidae (16–23%; Kenchington et al., 2017). In fact, E. argenteus is a controversial taxon within Eucinostomus, being regarded as a putative species complex (Matheson & McEachran, 1984). Moreover, the present data diagnosed that most of misidentified sequences in public data-

sets are related to E. argenteus, bringing more noise to their taxo-
nomic status that remains to be resolved. Besides these aspects, the close relationship between the cryptic lineages of E. argenteus from the Atlantic in relation to E. entomelas and E. downi from the Pacific (∼10–12% of divergence) is also intriguing. This pattern reinforces the role of the Panama isthmus in cladogenetic events of the ichthyofauna from the Eastern Pacific and Western Atlantic (Bacon et al., 2015; O’Dea et al., 2016; Thacker, 2017).

In Eucinostomus gula a different scenario was observed. Even though the sampled region of E. gula overlaps with the cryptic lineages of E. argenteus, the genetic difference between the lineages in the former is 50% higher than the threshold of 2.7% for interspecific distinction, suggesting a more recent evolutionary history for the cryptic lineages of E. gula. The close relationship between both lineages in this formal taxon indicates a recent expansion of their range in the Brazilian Province.

Fig. 2. Map and haplotype network based on COI sequences in populations of E. gula, E. cf. gula, E. argenteus and E. cf. argenteus throughout their range in the Western Atlantic. The bars over lines indicate the number of mutations between the haplotypes.
The lack of shared haplotypes in the complex *E. gula* provides strong evidence for phylogeographic disjunction caused by allopatry. Therefore, *E. gula* is likely to encompass two Evolutionary Significant Units (ESUs), one of them representing an undescribed taxon since a divergence of 4.4% was observed in COI sequences in relation to the other lineages. Putatively, a Brazilian endemic lineage evolved in allopatry as a result of major past vicariant events previously mentioned for *E. argenteus*. A similar pattern was also reported in *Chromis multilineata* (Rocha et al., 2008), a reef fish species of similar distribution to *E. gula*.

The hypothesis of allopatric evolution of distinct lineages in *E. gula* could be weakened by the lack of geographic isolation of both lineages. However, the evidence combined with reciprocal monophyly, geographic coexistence and lack of a shared haplotype reinforces that both ESUs of *E. gula* in the tropical Atlantic evolved in allopatry followed by dispersal that determined their secondary contact. This evidence combined with reciprocal monophyly, geographic coexistence and lack of a shared haplotype reinforces that both ESUs of *E. gula* in the tropical Atlantic evolved in allopatry followed by dispersal that determined their secondary contact. Unfortunately, the lack of georeferenced data in the other cryptic lineages of *E. melanopterus* and *E. harengulus* restrains further inferences about their diversification processes. Therefore, further phylogeographic studies focusing on both formal taxa are highly recommended.

**Final remarks**

Our phylogenetic analyses combining species molecular delimitation in a biogeographic context revealed new operational taxonomic molecular units (MOTUs) in *Eucinostomus* species from the Atlantic. Based on extensive barcode datasets we were able to recover the intraspecific diversity for this genus, estimating a potential threshold of 2.7%. Our data also corroborated the efficiency of COI markers in detecting cryptic lineages of mojarra that could be useful in resolving the taxonomic uncertainties in this fish group by providing a database to further support integrative approaches involving morphological, ecological features and other molecular markers. Subtle and overlapped morphological differences have served as proxies for species discrimination within *Eucinostomus*, however, in future systematic studies are strongly encouraged to consider genetic evidence as well.

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S0025315419001206.

**Financial support.** UPJ thanks Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) for research funding (BCT-0125-2.04/15), Conselho Nacional de Desenvolvimento Científico e Tecnológico CNPq (425080/2016-1) and Fundação de Amparo do Estado de Alagoas, FAPEAL (APQ – 1026/2016).

**References**


Da Silva WC, Marcenikut AP, Sales JBL and Araripe J (2016) Early Pleistocene lineages of Bagre bagre (Linnaeus, 1766) (Siluriformes: Ariidae), from the Atlantic coast of South America, with insights into the demography and biogeography of the species. Neotropical Ichthyology 14, e150184.


Hyde JR, Underkoffler KE and Sundberg MA (2014) DNA barcoding provides support for a cryptic species complex within the globally distributed and fishery important opah (*Lampris guttatus*). Molecular Ecology Resources 14, 1239–1247.


