Integrative systematics unveils the controversial identity of Engraulidae fishing stocks in a Neotropical estuary, northeast Brazil

Correspondence:
Uedson Pereira Jacobina
uedson.jacobina@penedo.ufal.br

Fabiana Vicente[1], Marina V. Loeb[2], Andréa Carla Guimarães de Paiva[1], Claudio L. S. Sampaio[1], Leandro Araujo Argolo[3] and Uedson Pereira Jacobina[1]

In Brazil, the use and diversity of the common names of fish species, coupled with taxonomic uncertainties, hinder the reliability of fishing statistical data. In this scenario, there are the so-called pilombetas of the São Francisco River, an important fishing resource in region. Despite its importance, the real diversity of species identified in the area remains obscure. In order to properly identify and delimit the species popularly known as pilombetas, an integrative approach involving traditional taxonomy, geometric morphometrics and molecular systematics was applied. Results from geometric morphometrics and molecular analyses were consistent with the results of the traditional morphological analysis, also indicating the delimitation of six taxa belonging to Engraulidae in the lower São Francisco River. In addition, species delimitation methods revealed an intrapopulation genetic divergence of 1.7% for *Lycengraulis grossidens*. The results revealed that the currently known richness species of Engraulidae in the studied area has been underestimated. Thus, an updated taxonomic key is herein proposed for the Engraulidae species from the lower São Francisco River and estuary. The integrative analysis approach revealed to be effective to address taxonomic questions and help the management of stocks, ensuring the maintenance of local diversity of fishes in the Neotropical region.

Keywords: Artisanal Fishing, Criptic Diversity, DNA Barcode, Identification Key, São Francisco River.

1 Laboratório de Ictiologia e Conservação, Campus-Penedo/ Universidade Federal de Alagoas, Avenida Beira Rio s/n, 57200-000 Penedo, AL, Brazil. (FV) fabiana.vicente@outlook.com, (ACGP) andrea.guimaraes.paiva@gmail.com, (CLSS) claudio.sampaio@penedo.ufal.br; (UPJ) uedson.jacobina@penedo.ufal.br.

2 Museu de Zoologia da Universidade de São Paulo, Seção de Peixes, C. P. 42494, 04263-000 São Paulo, SP, Brazil. (ML) loeb.mv@gmail.com

3 Instituto de Biologia, Universidade Federal da Bahia, 40170-115 Salvador, BA, Brazil. (LAA) leo.argolo@msn.com.

No Brasil, o uso e a diversidade dos nomes comuns para espécies de peixes, associado a incertezas taxonômicas, dificultam a confiabilidade dos dados estatísticos da pesca. Nesse cenário, existem as chamadas pilombetas do rio São Francisco, um importante recurso pesqueiro da região. Apesar de sua importância, a real diversidade de espécies identificadas na área permanece obscura. Para identificar adequadamente as espécies conhecidas como pilombetas, uma abordagem integrativa envolvendo taxonomia tradicional, morfometria geométrica e sistemática molecular foi aplicada. Os resultados das análises moleculares e de morfometria geométrica foram consistentes com os resultados da análise morfológica tradicional, indicando também a delimitação de seis táxons pertencentes a Engraulidae no baixo São Francisco. Além disso, os métodos de delimitação de espécies revelaram divergência genética intrapopulacional de 1,7% em *Lycengraulis grossidens*. Nossos dados revelaram que a riqueza de espécies atualmente conhecida de Engraulidae na área estudada é subestimada. Assim, uma chave taxonômica atualizada é aqui proposta para as espécies de Engraulidae do baixo rio São Francisco e seu estuário. A abordagem de análise integrativa revelou ser efetiva para tratar de questões taxonômicas e ajudar no gerenciamento de estoques, garantindo a manutenção da diversidade local de peixes na região Neotropical.

Palavras-chave: Chave de Identificação, Diversidade Críptica, DNA Barcode, Pesca Artesanal, Rio São Francisco.

INTRODUCTION

In the Neotropical region, artisanal fishing is one of the oldest activities carried out by riverside communities, especially geared towards species of commercial importance (Begossi *et al.*, 2004; Lins, 2011). This region houses an expressive marine and estuarine fish diversity and richness, but little is known about how fish stocks are recruited throughout their distribution, hindering the management of relevant areas and species for fishing (Blaber, 2000; Polaz, Ribeiro, 2016).

In Brazil, fishing statistics were largely built considering the common names of species rather than the scientific ones, which lead to misidentifications considering either common names used to different taxa or different common names for the same taxon between different localities (Vasconcellos *et al.*, 2011; Previeiro *et al.*, 2013). This approach leads to the underestimation or overestimation of species richness in an area, as well as overexploitation fish stocks. Species of Engraulidae family, for example, are popularly known in Portuguese as apapás, mulatinhas, sardinhas boca torta, manjubas and pilombetas (Silva *et al.*, 2003; Previeiro *et al.*, 2013). Additionally, common names usually used to Engraulidae species, are also applied for fishes from different families, such as manjuba used for both Engraulidae in southeastern Brazil, and Atherinopsidae (*e.g.*, genus *Atherinella* Steindachner, 1875) in the Northeastern (Ribeiro, Molina, 2013). These oversimplifications can obscure the real conservation status of the stock of those fishes, which may already be in an overexploited state.

Engraulidae species are found in tropical and subtropical waters of the Americas, mainly in coastal regions and estuaries, but also in continental waters of South America (Whitehead *et al.*, 1988; Nelson, 2016). The representatives of the family are important fishing resources of great commercial value, catch volume, and wide geographical distribution, making up about 25% of the total production of fishes caught commercially in the world (FAO, 2010; Vasconcellos, Csirke, 2011; Froese, Pauly, 2019). Ecologically, the species play a functional role in aquatic environments as filter-feeding fishes, being great consumers of zooplanktons and serving as a link in the food chain for piscivorous fishes and sea birds (Anderson *et al.*, 1980; Baird, Ulanowicz, 1989; Nelson, 2016).

In the São Francisco River, artisanal fishing plays a crucial role in generating income and employment, with about 287 tons of Engraulidae caught in 2006 (Ibama, 2006). This main fishing resource has been showing a sharp decline in recent years, certainly related to the decreased river flow, overfishing effort, especially in breeding seasons (Da Mata-Oliveira, 2020, pers. comm.). In addition, the changes caused in the hydrographic basin by the decrease in flow in recent years have severely impacted several stocks, including pilombetas (Medeiros *et al.*, 2015). It is estimated that at least four species are "informally" known as pilombetas (Loeb, Figueiredo, 2014; Barbosa *et al.*, 2017).

In this region, a new species has recently been described, *Anchoviella sanfranciscana* Barbosa, Gomes da Silva, da Rocha Araújo & Carvalho, 2017 (Barbosa *et al.*, 2017). However, according to unpublished data (Agostinho, 2019) *A. sanfranciscana* is a junior synonym of *Anchoviella cayennensis* (Puyo, 1945). All these taxonomic issues are perpetuated, due to their great phenotypic similarity and overlapping morphological characters, obscuring the taxonomic identification, quantification of their real diversity and the understanding of their phylogenetic relationships.

In recent years, DNA barcode became the leading molecular tool for species identification, being often integrated with other data sources, enhancing the robustness of taxonomic studies (Ward, 2000; Durand *et al.*, 2017). Among these sources, traditional morphometric measurements used in taxonomy, and geometric morphometrics stand out, allowing the detection of subtle differences at the intra and interspecific level (Vasconcellos *et al.*, 2008; DiBattista *et al.*, 2014; Pérez-Quiñónez *et al.*, 2016; Nirchio *et al.*, 2018). Studies integrating morphology and DNA have been increasingly effective in detecting cryptic lineages (Padial *et al.*, 2010; Yeates *et al.*, 2011), facilitating the resolution of taxonomic issues in complex groups such as the representatives of the Engraulidae family. In this context, an integrated analysis (morphology and genetics) was used to infer the number of species of pilombetas in the lower São Francisco River that are commercially exploited. This study also aims to determine whether there are species of pilombetas not discriminated by traditional taxonomic methods.

MATERIAL AND METHODS

The study area covers the region between the city of Piaçabuçu, Alagoas State (AL) (10°24'20"S 36°26'2"W) and the mouth of the São Francisco River (10°31'50.7"S

36°23'53"W) between the states of Alagoas and Sergipe (SE), Brazil. This region houses the Piaçabuçu Environmental Protection Area (EPA), Federal Conservation Unit (Oliveira *et al.*, 2014), in the lower stretch of the São Francisco River floodplain (Fig. 1). The river mouth region has been one of the most affected in recent years, with reduced flow, introduction of exotic species and the transposition of its waters. In addition, it has been suffering from the advance of the Atlantic Ocean (Brito, Magalhães, 2017; Cavalcante *et al.*, 2020).

All specimens of pilombetas were retrieved from artisanal fisheries caught with a 12/13 mm mesh gillnet, from September 2017 to April 2018. The captured animals were placed in a thermal box containing ice and transported to the Ichthyology and Conservation Laboratory (LIC) at the Federal University of Alagoas. Voucher specimens for all species herein identified were cataloged in the Museu de História Natural da Bahia, Universidade Federal da Bahia-UFBA, *Anchovia* sp. (UFBA 9079), *Anchoviella cayennensis* (UFBA 9078), *Anchoviella lepidentostole* (Fowler, 1911) (UFBA 9075), *Anchoviella* sp. (MZUFBA 9076), *Cetengraulis edentulus* (Cuvier, 1829) (UFBA 9077), and *Lycengraulis grossidens* (Spix & Agassiz, 1829) (UFBA 9073, UFBA 9074).

Taxonomic identification. Diagnostic measurements and meristic data were examined for all 124 specimens of pilombetas according to Cervigón (1989), Loeb, Figueiredo (2014) and Vera-Alcaraz, Vuletich (2016). Data was gathered with an Opticam OPZTS triocular magnifying glass in the laboratory of the Universidade Federal de Alagoas (UFAL-Penedo). Based on the analyzed characters, an updated taxonomic key for Engraulidae species inhabiting the lower São Francisco estuary was developed.
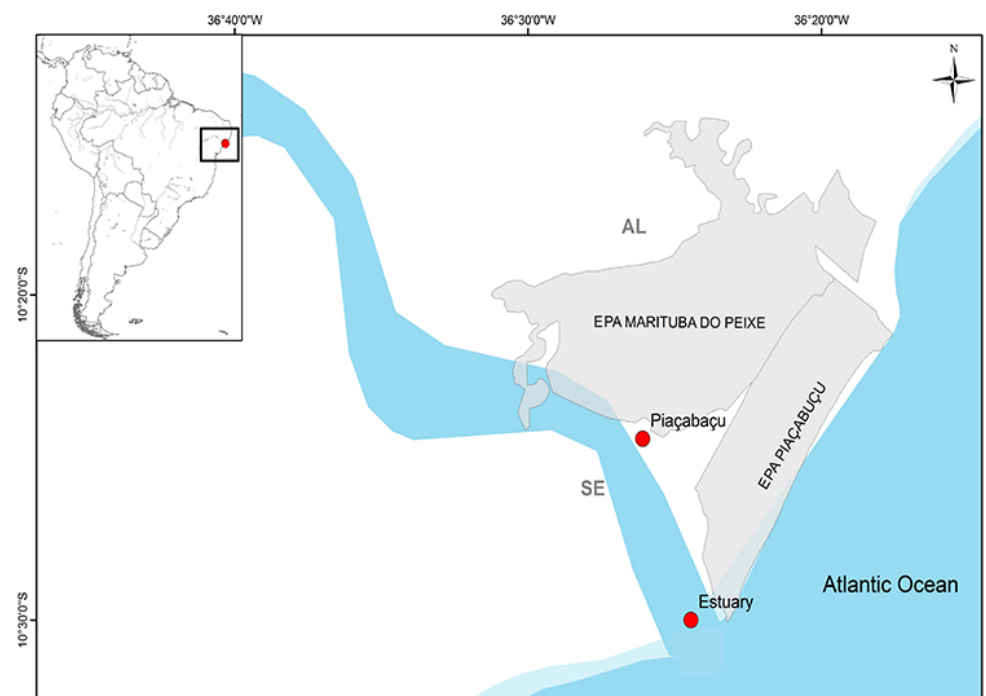
FIGURE 1 | Map of the sampling locality of pilombetas specimens in São Francisco estuary, northeastern Brazil. EPA = Environmental Protection Area.

Geometric morphometrics. All analyzed specimens were photographed from the left side with a 30 cm scale bar through a camera attached to a tripod. From the images, 10 body landmarks (Fig. 2) were chosen following standard procedures in studies of geometric morphometrics with fish and diagnostic characters of the group (Vasconcellos *et al.*, 2008; Kerschbaumer, Sturmbauer, 2011; Loeb, Figueiredo, 2014). The images were digitized and analyzed using the TpsUtil (Rohlf, 2010a) and TpsDig2 (Rohlf, 2010b) softwares. Shape data (*Procrustes coordinates*) were used for multivariate analyses through principal components (PCA) and canonical variates (CVA), aiming to identify morphological disparities between the species. Both tests were performed with MorphoJ 2.0 (Klingenberg, 2011).
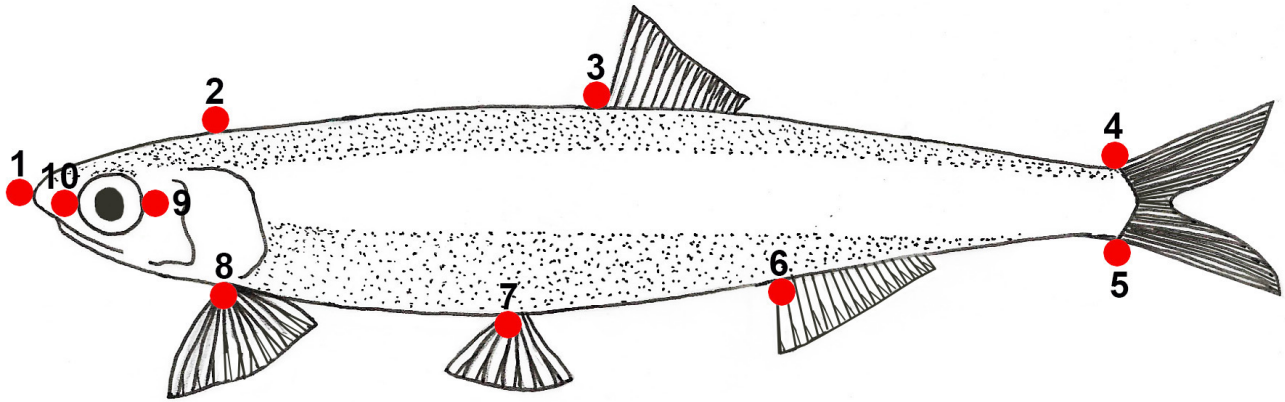
FIGURE 2 | Location of anatomical landmarks used for morphometric analyses: 1- distal point of the rostrum; 2- posterior end of the head; 3-anterior insertion of the dorsal fin; 4- insertion of the first upper radius of caudal fin; 5-insertion of the first lower radius of caudal fin; 6-anterior anal fin insertion; 7- insertion of the ventral fin; 8-insertion of the pectoral fin; 9- posterior end of the eye; 10- anterior end of the eye.

DNA extractions, amplification and sequencing. A total of 60 samples of specimens were selected and sequenced based on the preliminary identification, encompassing individuals of *Anchovia* sp. (n = 7), *Anchoviella cayennensis* (n = 10), *Anchoviella lepidentostole* (n = 7), *Anchoviella* sp. (n = 13), *Cetengraulis edentulus* (n = 9), and *Lycengraulis grossidens* (n = 14). Muscle tissue was obtained from the dorsal region of the fishes and stored in eppendorfs microtubes with absolute alcohol in the freezer at –20 °C. Subsequently, the genomic DNA was extracted using the Wizard genomic DNA purification kit (Promega) according to the manufacturer's instructions. The mitochondrial gene fragment cytochrome c oxidase subunit 1 (*cox I*) was amplified via polymerase chain reaction (PCR), using the universal primers BarcFish1 (5' TCAACCAACCACAAAGACATTGGCAC 3') and BarcFish22 (5' ACTTCAGGGTGACCGAAGAATCAGAA 3'), following the procedures of Ward *et al.* (2005). All PCR reactions were performed in a 25μl final volume containing 12μl of 2X Taq master mix (Vivantis), 2μl of 40ng / μl genomic DNA solution, 0.5μl of each primer (10MM), and 15μl of ultrapure water. The reactions consisted of an initial

denaturation of 2 minutes at 95 °C, 35 cycles of 94 °C for 30s, 57 °C for 30s, and 72 °C for 2 min, with a final extension of 72 °C for 7 min (modified from Ward *et al.*, 2005). Amplification products were purified with 20% polyethylene glycol (PEG). Sequencing was performed in two reads, forward and reverse, with the Big DyeTM Terminator v 3.1 Cycle Sequencing Ready Reaction kit (Applied Biosystems) with the additional primer M-13 and performed on a Model ABI 3130 sequencer (Applied Biosystems).

Data analysis. The electropherograms were analyzed and edited, and consensus sequences were generated in Geneious 9.1.8 (Kearse *et al.*, 2012), followed by a visual inspection for final adjustments. We use NCBI BLASTn and BOLD Identification System (IDS) tools under the Barcode of Life Data System (BOLD) Species Level Barcode Records option to compare and confirm the prior morphological identifications. These sequences were posteriorly aligned using the ClustalW method in Geneious (Kearse *et al.*, 2012). The final alignment was submitted to phylogenetic analysis under Bayesian Inference (BI) and Maximum Likelihood (ML) criteria, as well as the Neighbor Joining (NJ) distance method with the K2P evolutionary model, following Barcode of Life analytical procedures (www.boldsystems.org). To visualize putative differences between species, we calculated intra- and interspecific distances, using complete deletion with the evolutionary model K2P in MEGA6 software (Tamura *et al.*, 2013).

For the BI and ML analyses, the best fitting model of nucleotide substitution was HKY+I, according to the Akaike Information Criterion in PartitionFinder2 (Lanfear *et al.*, 2017). ML analysis was performed in RAxML v.8 and branch support was obtained with 1000 bootstrap pseudoreplicates (Stamatakis *et al.*, 2014). BI analysis was conducted through BEAST 1.8.4 software (Drummond, Rambaut, 2007) to generate ultrametric trees with the Yule Speciation prior model. The tree search was produced by 20 million generations, sampling at each 2 thousand 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. We applied a 10% burn-in and used the remaining trees to obtain a Maximum Clade Credibility tree. Branch support was based on the posterior probability values. The trees were visualized in FigTree v1.4.1 (Rambaut, 2009).

Species delimitation methods. Single locus species delimitation using two tree-based: Bayesian Poisson Tree Process (bPTP, Zhang *et al.*, 2013) and Generalized Mixed Yule-coalescent model GMYC, (Pons *et al.*, 2006); and one distance method: Automatic Barcode Gap Discovery (ABGD) (Puillandre *et al.*, 2012) were implemented to infer the number of species in our dataset. The bPTP was performed on the bPTP web server (http://species.h-its.org/ptp/) using a non-ultrametric tree obtained from ML analysis. The analysis was performed with 500 thousand generations sampling every 500 and using a 10% burn-in. GMYC, in turn, was implemented on the GMYC web server (http://species.h-its.org/gmyc/), assuming a single threshold for an uncalibrated tree, no outgroup, including one specimen per haplotype constructed using BEAST (Drummond *et al.*, 2012). Finally, ABGD was conducted on the online server (http://wwwabi.snv.jussieu.fr/public/abgd/) using an aperture width value of 1.0 for all available distances (p-distance, Kimura-2 -parameters and Jukes-Cantor) (Berbel-Filho *et al.*, 2018). The agreement between the delimitations of the OTUs was evaluated by comparing the composition of clusters between the different species delimitation methods.

RESULTS

Traditional taxonomy. Morphological analyses based on meristic and morphometric characters of the studied specimens of pilombetas revealed the presence of six taxa belonging to four genera in Engraulidae: *Anchoviella* Fowler, 1911, *Anchovia* Jordan & Evermann, 1895, *Cetengraulis* Günther, 1868 and *Lycengraulis* Günther, 1868. However, due to character overlap, only four taxa could be identified at species level, *Anchoviella lepidentostole* (n = 30, 83.6–128.7 mm SL), *Anchoviella cayennensis* (n = 20, 84.3–130.7 mm SL), *Cetengraulis edentulus* (n = 12, 100.4–129.6 mm SL) and *Lycengraulis grossidens* (n = 20, 85.7–131.8 mm SL). Two additional taxa were only identified at genus level, *Anchovia* sp. (n = 20, 109.9–165 mm SL) and *Anchoviella* sp. (n = 22, 73.5–80.6 mm SL). To facilitate the identifications of the species studied herein, a new taxonomic key using morphological features along with photographs (Fig. 3) for these taxa that occur in the lower São Francisco River estuary is presented.
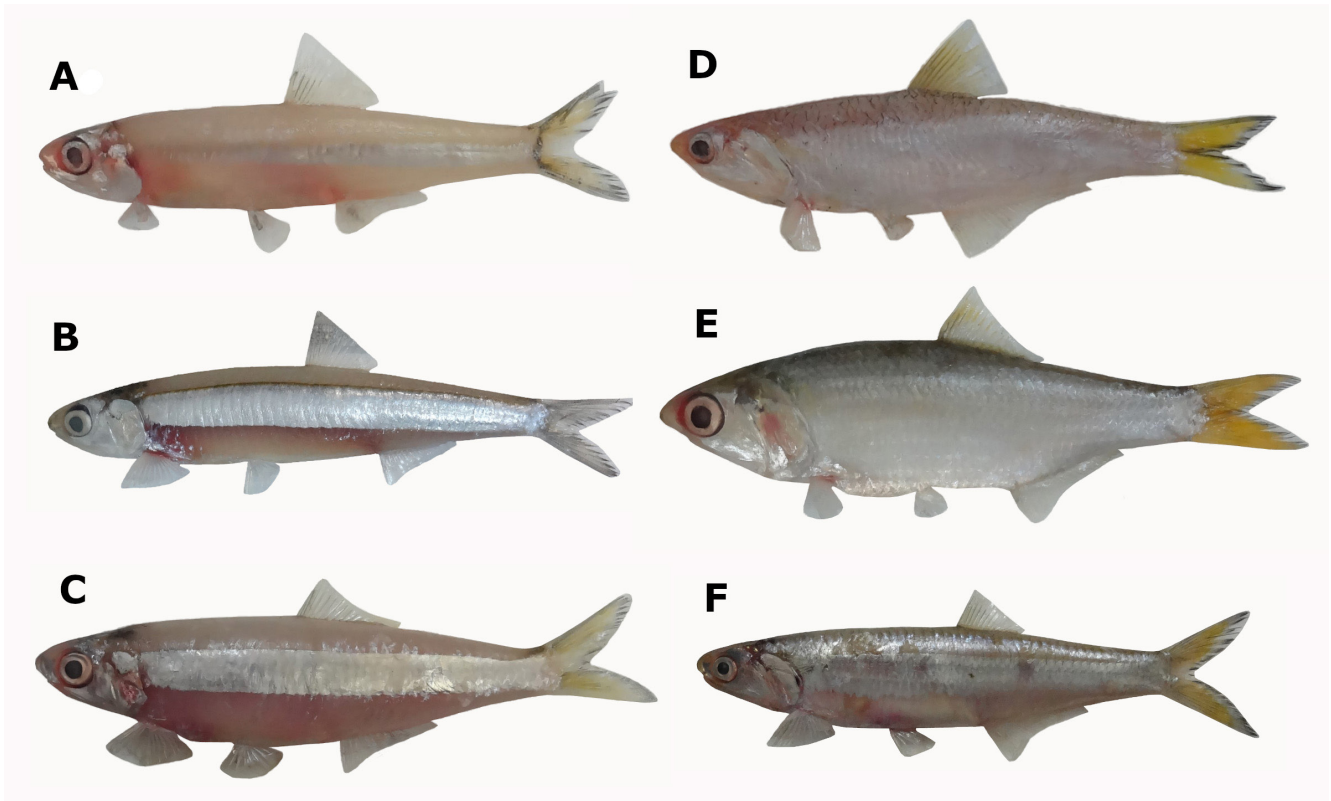
FIGURE 3 | Fresh specimens of the species identified in the study area. **A.** *Anchoviella brevirostris* (75.5 mm SL), **B.** *Anchoviella cayennensis* (90.3 mm SL), **C.** *Anchoviella lepidentostole* (83.2 mm SL), **D.** *Anchovia clupeoides* (120 mm SL), **E.** *Cetengraulis edentulus* (104.4 mm SL), **F.** *Lycengraulis grossidens* (98.7 mm SL). Photographs were taken by the first author.

Identification key of Engraulidae species from the lower San Francisco River

Engraulidae species are characterized by a fusiform small to moderate-sized body (up to 200 mm SL) laterally compressed; cycloid scales present on body and base of caudal and anal fins; body coloration brownish to pale, with a conspicuous silver longitudinal stripe present in most analyzed specimens; eyes laterally positioned in head; mouth subterminal, its posterior margin almost reaching posterior margin of opercle; teeth if present, canine or conical-shaped in a single row in the upper and lower maxilla; caudal fin forked.

- 1. Dentition in upper and lower maxilla absent or if present, with reduced conical teeth 2
- 1'. Dentition present, with conspicuous canine teeth in the upper and lower maxilla .... *Lycengraulis grossidens*

- 2. Number of gill rakers in the lower half of first branchial arch 19–35; 12–20 total anal-fin rays 3
- 2'. Number of gill rakers in the lower half of first branchial arch 40–116; 21–35 total anal-fin rays 5

- 3. Number of gill rakers in the lower half of first branchial arch 19–27; anal-fin origin positioned at vertical through dorsal-fin base 4
- 3'. Number of gill rakers in the lower half of first branchial arch 28–35; anal-fin origin at vertical through posterior margin of last dorsal-fin ray or posteriorly .... *Anchoviella cayennensis*
- 4. Upper maxilla short, not reaching the anterior margin of preopercle; 15–20 total anal-fin rays; snout rounded ..... *Anchoviella brevirostris*
- 4'. Upper maxilla long, reaching the anterior margin of preopercle; 12–14 total anal-fin rays; snout pointed .*Anchoviella lepidentostole*

- 5. Upper maxilla short, not reaching the anterior margin of preopercle; 21–27 total anal-fin rays; greatest body depth 3.5 times in SL .... *Cetengraulis edentulus*
- 5'. Upper maxilla long, reaching the anterior margin of preopercle; 28–35 anal-fin rays; greatest body depth 4.0 times in SL ....*Anchovia clupeoides*

Geometric morphometrics. The overall morphological variation in the Principal Component Analysis revealed 61% of the variance explained by the first two components (PC1 = 42% and PC2 = 19%). *Cetengraulis edentulus*, *Anchovia* sp. were unambiguously distinguished, while the other taxa *Anchoviella lepidentostole*, *Anchoviella* sp. and *Lycengraulis grossidens* showed partially overlapping distributions in the morphospace (Fig. 4). Procrustes distances were significant among all analyzed species (p-values <0.001; Tab. 1). In turn, the Canonical Variation Analysis (CVA) effectively discriminated the body shape of all representatives in six completely different taxa. The two axis CVA1 (47%) and CVA2 (27%) explained 74% of all morphological variation found (Fig. 5). The largest vector variations are related to body (landmarks 3, 6, 7, and 8; Fig. 2) and caudal peduncle depth (landmarks 4 and 5; Fig. 2).

Genetic analysis. The final alignment consisted of 596 bp with 423 conservative and 173 variable sites. The K2P distances revealed the highest genetic distance

TABLE 1 | Procrustes distance (diagonal bottom, black) and estimates of pairwise genetic divergence with the K2P evolutionary model (diagonal top, bold), among representatives of Engraulidae from the São Francisco River estuary, northeastern Brazil.

	A. cayennensis	A. clupeoides	A. lepidentostole	A. brevirostris	C. edentulus	L. grossidens
Anchoviella cayennensis	0	0.167	0.143	0.164	0.197	0.137
Anchovia clupeoides	0.1306	0	0.152	0.151	0.154	0.144
Anchoviella lepidentostole	0.0823	0.0658	0	0.150	0.162	0.147
Anchoviella brevirostris	0.0645	0.0804	0.0436	0	0.160	0.152
Cetengraulis edentulus	0.1262	0.1043	0.0832	0.1069	0	0.179
Lycengraulis grossidens	0.0605	0.092	0.0453	0.0577	0.1005	0
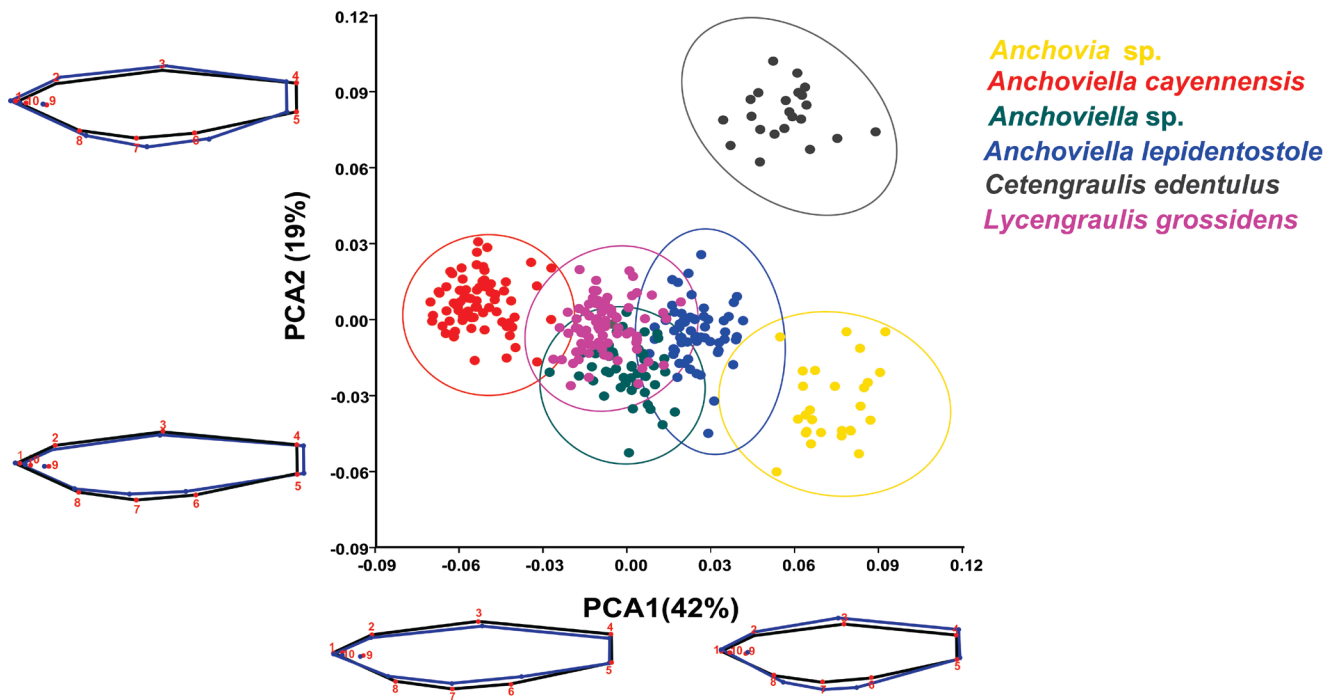
FIGURE 4 | Principal Components Analysis indicating the visual ordination of the six species of pilombetas in the morphospace. Estimated changes in dorsal and ventral view shape are shown as deformations from the mean shape along the first and second principal components.

between *L. grossidens* and *C. edentulus* (17.9%) and the lowest distance between *L. grossidens* and *A. cayennensis* (13.7%) (Tab. 1). The molecular identification methods 'Blastn and Species Level Barcode Records' identified *Anchoviella* sp. and *Anchovia* sp., previously unidentified at the species level by traditional morphology as *Anchoviella brevirostris* (Günther, 1868) and *Anchovia clupeoides* (Swainson, 1839), respectively. The phylogenetic analyses (Fig. 6) resulted in a well-supported tree (bootstrap = 100, and

posterior probabilities PP = 1) for most species clade, discriminating seven independent lineages among all specimens analyzed. The group called pilombetas is represented by two clades: clade I (*Anchoviella brevirostris* + *Anchovia clupeoides* and *C. edentulus*) and clade II (*A. lepidentostole*, *A. cayennensis* and *L. grossidens*). Moreover, these analyses also revealed the presence of two divergent lineages within the population of *Lycengraulis grossidens* from São Francisco River estuary (1.7% of genetic distance). These results were consistent among the three species delimitation methods (GMYC, bPTP, and ABGD; Fig. 6).
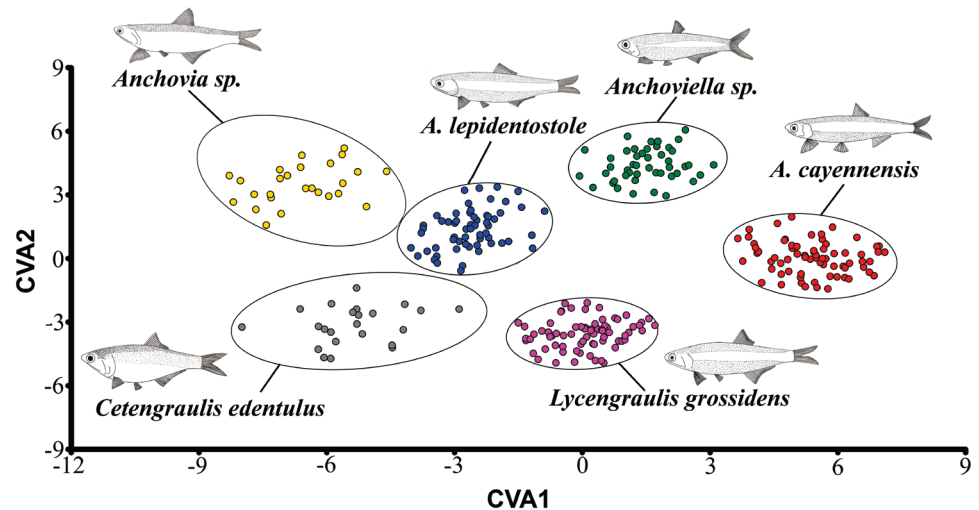

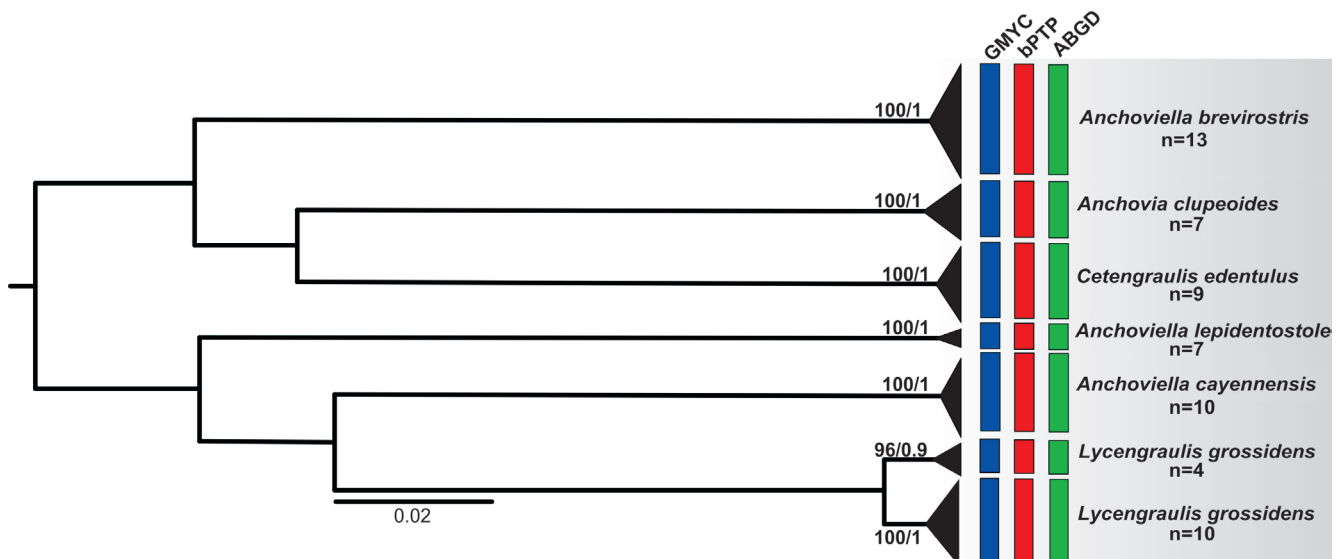**FIGURE 5 |** Canonical Variables Analysis discriminating for six pilombetas species.

FIGURE 6 | Bayesian topology and species delimitation using Generalized Mixed Yule-coalescent (GMYC), Bayesian Poisson Tree Process (bPTP) and Automatic Barcode Gap Discovery (ABGD) discriminating species denominated pilombetas.

DISCUSSION

Our integrative results revealed that seven lineages occurring in the lower São Francisco are currently being called pilombetas. The analysis of meristic characters along with the prior available taxonomic keys effectively discriminated four species (*A. lepidentostole*, *A. cayennensis*, *C. edentulus*, and *L. grossidens*). The diagnostic characters to identify these taxa were mostly related to the number of anal-fin rays and gill rakers on the upper and lower branches of the first gill arch. For the two taxa initially identified only at genus level, *Anchoviella* sp. and *Anchovia* sp., dorsal and pectoral-fin ray counts were useful to the identification, although the remaining characters overlapped, which hindered a specific identification of these two putative taxa.

Recently, Loeb, Figueiredo (2014), analyzed samples of *Anchoviella* from the São Francisco and Amazonas basins, and reported that the overlap of meristic characters for this family hampers the identification by traditional morphology, which requires great experience in studies with the group. In fact, the great morphological similarity between these commercially important species hinders their reliable identification based on external characteristics and, consequently, appropriate control and management actions for these taxa (Bakar *et al.*, 2018). The indiscriminate fishing of several taxa under the same common name may be unsustainably affecting the stocks of each of these species, which generates a worrying perspective, especially in estuaries, which are largely influenced by anthropic actions (Ferreira, Lacerda, 2016), and has a central role as cradles of diversity (Roque *et al.*, 2016).

Geometric morphometrics (GM) analyses have long been used as a tool to identify species and fish stocks (Cadrin, 2000; Silva *et al.*, 2003). In our GM analysis, the CVA was consistent with the traditional morphometry showing two more taxa, *Anchoviella* sp., and *Anchovia* sp., posteriorly identified by molecular barcode analysis as *Anchoviella brevirostris* and *Anchovia clupeoides* respectively, which substantially segregated in the morphospace visualization in comparison to the other species. Divergence by the amplitude of dorsal profile variation in axis 1 (CVA1) and caudal peduncle in axis 2 (CVA2) seems to follow the patterns commonly found in fish (Perazzo *et al.*, 2018). Duarte *et al.* (2017), using GM to evaluate the diversity of Carangidae of the coast of Rio de Janeiro in Brazil, were able to identify the occurrence of three species that were previously identified as a single taxon. Barreto *et al.* (2016), evaluating the species of *Nematocharax* Weitzman, Menezes & Britski, 1986 in seven locations in the basins of the Bahia and Minas Gerais states in Brazil, found variations in the shape and body size of specimens from allopatric populations, proposing a putative new species for the group, which was formally described (Barreto *et al.*, 2018), attesting the power of geometric morphometrics in distinguishing morphological differences at both intraspecific and interspecific levels. In the case of the present study, a fundamental difference in the fact that the questionable species studied are formally recognized as valid and even belong to different genera. However, the exploration of these groups by fishermen under the common name "pilombetas" reinforces the urgency of the need for effective alternative methodologies in identifying and controlling the use of fish stocks.

DNA barcoding, over the last decades, has been consolidated as the main tool to help classical systematics, especially in taxonomically controversial species that are difficult to identify by traditional methods (Carvalho *et al.*, 2011; Durand *et al.*, 2017).

The present results add to the growing number of studies that attest this effectiveness. In fact, all Engraulidae species identified based only in morphological analyses were confirmed by the molecular results of *cox* I. Additionally, records in the BOLD and NCBI databases revealed that species of *Anchoviella* and *Anchovia* that had not been identified at the species level are *Anchoviella brevirostris* and *Anchovia clupeoides*. Thus, it seems evident that building molecular inventories of regional ichthyofaunas (*e.g.*, Brandão *et al.*, 2016) and implementing DNA barcoding by control bodies can increase the effectiveness of fish surveillance.

Ardura *et al.* (2010) studying species of commercial importance in the northern region of Brazil, used the DNA barcoding to differentiate taxa popularly known and marketed as Acarás, which later revealed to be a complex composed of seven species of Cichlidae. In another study, Bakar *et al.* (2018) analyzed juvenile specimens morphologically discriminated as *Lutjanus lutjanus* Bloch, 1790 commercially captured in Malaysia and found two genetically distinct and well-supported monophyletic lineages, reaffirming the power of this technique to discriminate cryptic taxa not visualized by traditional methods where diversity may be underestimated. DNA barcoding for fish identification has been shown to be highly effective in distinguishing fish species from various locations (Ward *et al.*, 2005; Ardura *et al.*, 2010; Pereira *et al.*, 2013).

The species delimitation methods GMYC, bPTP and ABGD showed concordant results, suggesting that the Engraulidae of the lower São Francisco encompass seven lineages. Two distinct lineages within *L. grossidens* from the São Francisco River estuary indicate a cryptic species, which the divergence between them represents 1.7% of genetic distance, and they were not discriminated by morphological characters and morphometric analyses. Previous studies have shown the ecological plasticity and genetic diversity of this species in entering rivers and estuaries along the Brazilian coast (Mai *et al.*, 2016). In this context, future phylogeographic analysis will be able to assess in more detail the evolutionary routes of the *L. grossidens* populations. In fishing statistics, a crucial problem is the lack of precise identification of fish stocks, especially in Brazil, where common names are widely used. This is an important requirement to correctly estimate trends in productivity, biomass and population dynamics, which in turn helps to define appropriate management strategies. To confuse several species in an evaluation would be equivalent to ignoring the different population dynamics of each individual taxon and their different life histories (Karahan *et al.*, 2014). In the case of fish that are commercially exploited, it is important that they are correctly labeled in research, and mainly in the commercial market (Karahan *et al.*, 2014), as it should happen in the case of pilombetas, considering that they are an important resource of the fishing communities of the lower São Francisco River and in other regions of Brazil (Paiva-Filho *et al.*, 1986; Ibama, 2006; Sampaio, *et al.*, 2015).

The present results revealed that the diversity of species living in the lower São Francisco River called pilombetas has been underestimated. In addition, it demonstrated that combined analyses using morphological (traditional taxonomy and geometric morphometrics) and molecular data (DNA barcoding) can be used to assess species richness and genetic diversity, securing these finite resources for its long-term exploration. Moreover, a new taxonomic key that can accurately assist in the Engraulidae identification for the area is presented. We recommend the development of regional identification keys for other fish groups with similar taxonomic problems,

which can assist in this process of identification and improvement of fishing reports. This information is expected to contribute effectively to the proper management of these fish stocks in the lower São Francisco River, as well as in other regions along the Brazilian coast.

ACKNOWLEDGMENTS

REFERENCES

• **Agostinho LS.** Revisão taxonômica de *Anchoviella cayennensis* (Puyo, 1945) (Clupeiformes: Engraulidae), uma espécie de manjuba pouco conhecida do Atlântico ocidental. [Master Dissertation]. Macaé: Universidade Federal do Rio de Janeiro – Campus UFRJ Macaé; 2019.

• **Anderson DW, Gress F, Mais KF, Kelly PR.** Brown pelicans as anchovy stock indicators and their relationships to commercial fishing. CalCOFIs Rep. 1980; 21:54–61. Available from: https://www.calcofi.org/ccpublications/ccreports.html

• **Ardura A, Linde AR, Moreira JC, Garcia-Vazquez E**. DNA barcoding for conservation and management of Amazonian commercial fish. Biol Conserv. 2010; 143(6):1438–43. https://doi.org/10.1016/j.biocon.2010.03.019

• **Baird D, Ulanowicz RE.** The seasonal dynamics of the Chesapeake Bay. Ecol Monogr. 1989; 59(4):329–64. https://doi.org/10.2307/1943071

• **Bakar AA, Adamson EAS, Juliana LH, Nor Mohd SA, Wei-Jen C, Man A** *et al*. DNA barcoding of Malaysian commercial snapper reveals an unrecognized species of the yellow-lined *Lutjanus* (Pisces: Lutjanidae). PLoS ONE. 2018; 13(9):e0202945. https://doi.org/10.1371/journal.pone.0202945

• **Barbosa JM, Soares EC, Cintra IHA, Hermann M, Araújo AR.** Profile of the fish fauna of the São Francisco River basin. ActaFish. 2017; 5(1):70–90. https://doi.org/10.2312/Actafish.2017.5.1.70-90

• **Barreto SB, Nunes LA, Silva AT, Jucá-Chagas R, Diniz D, Sampaio I, Schneider H, Affonso PRAM**. Is *Nematocharax* (Actinopterygii, Characiformes) a monotypic fish genus? Genome. 2016; 59: 851–65. https://doi.org/10.1139/gen-2015-0166

• **Barreto SB, Silva AT, Batalha Filho H, Affonso PRAM, Zanata AM**. Integrative approach reveals a new species of *Nematocharax* (Teleostei: Characidae). J Fish Biol. 2018; 93(6):1151–62. https://doi.org/10.1111/jfb.13834

• **Begossi A, Silva AL, Seixas CS, Castro F, Pezzuti J, Hanazaki N, Peroni N, Silvano RAM.** Ecologia de Pescadores da Mata Atlântica e da Amazônia. São Paulo: Hucitec: Nepam/Unicamp: Nupaub/USP: Fapesp; 2004. Available from: https://fisheriesandfood.com/wp-content/uploads/2018/02/2004-Ecologia-de-Pescadores-da-Mata-Atlantica-e-da-Amazonia.pdf

• **Berbel Filho WM, Ramos TPA, Jacobina UP, Gama Maia D, Torres RA, Lima SMQ.** Updated checklist and DNA barcode-based species delimitations reveal taxonomic uncertainties among freshwater fishes from the mid-north-eastern Caatinga ecoregion, north-eastern Brazil. J Fish Biol. 2018; 93(2):311–23. https://doi.org/10.1111/jfb.13758

• **Blaber SJM.** Tropical estuarine fishes: ecology, exploitation and conservation. Oxford: Blackwell Science; 2000.

- **Brandao JHSG, Araújo Bitencourt J, Santos FB, Watanabe LA, Schneider H, Sampaio I, Affonso PRAM.** DNA barcoding of coastal ichthyofauna from Bahia, northeastern Brazil, South Atlantic: high efficiency for systematics and identification of cryptic diversity. Biochem Syst Ecol. 2016; 65:214–24. https://doi.org/10.1016/j.bse.2016.02.012

- **Brito MFG, Magalhães ALB.** Brazil's development turns river into sea. Science. 2017; 358(6360):179. https://doi.org/10.1126/science.aap9525

- **Cadrin SX.** Advances in morphometric identification of fishery stocks. Rev Fish Biol Fish. 2000; 10:91–112. https://doi.org/10.1023/A:1008939104413

- **Carvalho DC, Oliveira DA, Pompeu PS, Leal CG, Oliveira C, Hanner R.** Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. Mitochondrial DNA. 2011; 22(S1):80–86. http://dx.doi.org/10.3109/19401736.2011.588214

- **Cavalcante G, Vieira F, Campos E, Brandini N, Medeiros P. RP.** Temporal streamflow reduction and impact on the salt dynamics of the São Francisco River Estuary and adjacent coastal zone (NE/Brazil). Reg Stud Mar Sci. 2020; 38: 101363. https://doi.org/10.1016/j.rsma.2020.101363

- **Cervingon F.** Los peces marinos de Venezuela. Caracas: Editorial Arte; 1989.

- **DiBattista JD, Randall JE, Newman SJ, Bowen BW.** Round herring (genus *Etrumeus*) contain distinct evolutionary lineages coincident with a biogeographic barrier along Australia's southern temperate coastline. Mar Biol. 2014; 161(11):2465–77. https://doi.org/10.1007/s00227-014-2516-5

- **Drummond AJ, Rambaut A.** BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007; 7:214. https://doi.org/10.1186/1471-2148-7-214

- **Drummond AJ, Suchard MA, Xie D, Rambaut A.** Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012; 29(8):1969–73. https://doi.org/10.1093/molbev/mss075

- **Duarte MR, Tubino RA, Monteiro-Neto C, Martins RRM, Vieira FC, Andrade-Tubino MF, Silva EP.** Genetic and morphometric evidence that the jacks (Carangidae) fished off the coast of Rio de Janeiro (Brazil) comprise four different species. Biochem Syst Ecol. 2017; 71:78–86. https://doi.org/10.1016/j.bse.2017.01.013

- **Durand JD, Hubert N, Shen KN, Borsa P.** DNA barcoding grey mullets. Rev Fish Biol Fish. 2017; 27:233–43. https://doi.org/10.1007/s11160-016-9457-7

- **IBAMA.** Boletim Estatístico da Pesca Marítima e Estuarina (ESTATPESCA) do Nordeste do Brasil. Tamandaré:– CEPENE; 2006.

- **FAO.** FAO yearbook: fishery and aquaculture statistics 2008. Rome: Food and Agriculture Organization; 2010.

- **Ferreira AC, Lacerda LD.** Degradation and conservation of Brazilian mangroves, status and perspectives. Ocean Coast Manag. 2016; 125:38–46. https://doi.org/10.1016/j.ocecoaman.2016.03.011

- **Froese R, Pauly D.** FishBase: World Wide Web electronic publication [Internet]; 2019. Available from: www.fishbase.org

- **Karahan A, Borsa P, Gucu AC, Kandemir I, Ozkan E, Orek YA, Acan SC, Koban E, Togan I.** Geometric morphometrics, Fourier analysis of otolith shape, and nuclear-DNA markers distinguish two anchovy species (*Engraulis* spp.) in the Eastern Mediterranean Sea. Fish Res. 2014; 159(2014):45–55. http://dx.doi.org/10.1016/j.fishres.2014.05.009

- **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A.** Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012; 28(12):1647–49. https://doi.org/10.1093/bioinformatics/bts199

- **Kerschbaumer M, Sturmbauer C.** The utility of geometric morphometrics to elucidate pathways of cichlid fish evolution. Int J Evol Biol. 2011; 2011:290245. http://dx.doi.org/10.4061/2011/290245

- **Klingenberg CP.** MorphoJ: an integrated software package for geometric morphometrics. Mol Ecol Resour. 2011; 11(2):353–57. http://dx.doi.org/10.1111/j.1755-0998.2010.02924.x

- **Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B.** PartitionFinder 2: new methods for selecting partitioned models of evolution formolecular and morphological phylogenetic analyses. Mol Biol Evol. 2017; 34(3):772–73. http://dx.doi.org/10.1093/molbev/msw260

- **Lins PMO.** Técnico em Pesca e Aquicultura: Tecnologia Pesqueira. IFPA: UFRN; 2011.

- **Loeb MV, Figueiredo JL.** Redescription of the freshwater anchovy *Anchoviella vaillanti* (Steindachner, 1908) (Clupeiformes: Engraulidae) with notes on the distribution of estuarine congeners in the Rio São Francisco basin, Brazil. Arq Zool. 2014; 45:33–40. https://doi.org/10.11606/issn.2176-7793.v45iespp33-40

- **Mai ACG, Robe LJ, Marins LF, Vieira J**P. Genetic relationships between landlocked and coastal populations of *Lycengraulis grossidens* (Engraulidae) in Southeastern South-America: evidence for a continental colonization route with secondary transitions to the coastal region. Mar Environ Res. 2016; 68(2):342–51. https://doi.org/10.1071/MF15355

- **Medeiros PRP, Cavalcante Segundo GH, Magalhães EMM.** Comportamento da turbidez e material em suspensão, em um rio com vazão regularizada por sistema de barragens em cascata: Rio São Francisco (NE, Brasil). Geochim Bras. 2015; 29(1):35–44. https://doi.org/10.5327/Z0102-9800201500010004

- **Nelson JS, Grande TC, Wilson MVH.** Fishes of the World. Hoboken: John Wiley & Sons; 2016.

- **Nirchio M, Paim FG, Milana V, Rossi AR, Oliveira C.** Identification of a new mullet species complex based on an integrative molecular and cytogenetic Investigation of *Mugil hospes* (Mugilidae: Mugiliformes). Front Genet. 2018; 9:17. http://dx.doi.org/10.3389/fgene.2018.00017

- **Oliveira NS, Amorim CMF, Lemos RPL.** As riquezas das áreas protegidas no território Alagoano. Maceió: Instituto do Meio Ambiente do Estado de Alagoas: Mineração Vale Verde; 2014.

- **Padial JM, Miralles A, De La Riva I, Vences M.** The integrative future of taxonomy. Front Zool. 2010; 7:16. https://doi.org/10.1186/1742-9994-7-16

- **Paiva Filho AM, Zani-Teixeira ML, Kihara PK.** Contribuição ao conhecimento da biologia da manjuba *Anchoviella lepidentostole* (Fowler 1911) no Estuário de São Vicente (Osteichthyes, Engraulidae). Bol Inst Oceanogr. 1986; 34:71–77. Available from: https://www.scielo.br/pdf/bioce/v34/v34a06.pdf

- **Perazzo G, Correa F, Calviño P, Alonso F, Salzburger W, Gava A.** Shape and size variation of *Jenynsia lineata* (Jenyns 1842) (Cyprinodontiformes: Anablepidae) from different coastal environments. Hydrobiologia. 2018; 828:21–39. https://doi.org/10.1007/s10750-018-3794-6

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

- **Pérez-Quiñónez I, Quiñónez-Velázquez C, Vergara-Solana FJ, García-Rodríguez FJ.** Combining geometric morphometrics and genetic analysis to identify species of Opisthonema Gill, 1861 in the eastern Mexican Pacific. J Appl Ichthyol. 2016; 33(1): 84–92. https://doi.org/10.1111/jai.13051

- **Polaz CNM, Ribeiro KT.** Conservação de peixes continentais e manejo de unidades de conservação. BioBrasil. 2016; 7(1): 1–3.

- **Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP.** Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst Biol. 2006; 55:595–609. https://doi.org/10.1080/10635150600852011

- **Previero M, Minte-Vera CV, Moura RL.** Fisheries monitoring in Babel: fish ethnotaxonomy in a hotspot of common names. Neotrop Ichthyol. 2013; 11(2):467–76. https://doi.org/10.1590/S1679-62252013000200016

- **Puillandre N, Lambert A, Brouillet S, Achaz G**. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol Ecol. 2012; 21(8):1864–77. https://doi.org/10.1111/j.1365-294X.2011.05239.x

- **Rambaut A, Drummond AJ**. FigTree version 1.3. 1 [Internet]; 2009. 541 Available from: http://tree.bio.ed.ac.uk/software/figtree/

- **Ribeiro ED, Molina WF.** Marcada diferenciação cariotípica entre as "manjubas" *Atherinella blackburni* e *A. brasiliensis* (Atheriniformes). Biota Amaz. 2013; 3(2):40–52. https://doi.org/10.18561/2179-5746/biotaamazonia.v3n2p40-52

- **Rohlf FJ.** TpsUtil, tps file utility program, version 1.46 [Internet]. Stony Brook University; 2010a. Available from https://life.bio.sunysb.edu/morph/soft-utility.html

- **Rohlf FJ.** tpsDIG2, digitize landmarks & outlines from image files, scanner, or video, version 2.16 [Internet]. Stony Brook University; 2010b. Available from https://life.bio.sunysb.edu/morph/soft-utility.html

- **Sampaio CLS, Paiva ACG, Silva E, Soares EC.** Peixes, pesca e pescadores do baixo São Francisco, Nordeste do Brasil. In: Nogueira EMS, Sá MFP, organizers. A pesca artesanal no baixo São Francisco: atores, recursos, conflitos. Petrolina, PE: SABEH; 2015. p.105–148.

- **Silva MDA, Araújo FG, Azevedo CCD, Mendonça P.** Distribuição espacial e temporal de *Cetengraulis edentulus* (Cuvier) (Actinopterygii, Engaulidae) na Baía de Sepetiba, Rio de Janeiro, Brasil. Rev Bras Zool. 2003; 20(4):577–81. https://doi.org/10.1590/S0101-81752003000400003

- **Stamatakis A.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30(9):1312–13. https://doi.org/10.1093/bioinformatics/btu033

- **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30(12):2725–29. https://doi.org/10.1093/molbev/mst197

- **Vasconcellos AV, Vianna P, Paiva P, Schama R, Solé-Cava A.** Genetic and morphometric differences between yellowtail snapper (*Ocyurus chrysurus*, Lutjanidae) populations of the tropical West Atlantic. Genet Mol Biol. 2008; 31(1):308–16. https://dx.doi.org/10.1590/S1415-47572008000200026

- **Vasconcellos M, Diegues AC, Kalikoski DC**. Coastal fisheries of Brazil. In: Salas S, Chuenpagdee R, Charles A, Seijo JC, organizers. Coastal fisheries of Latin America and the Caribbean. Rome: FAO; 2011. p.73–116. Available from: http://www.fao.org/3/a-i1926e.pdf

- **Vasconcellos M, Csirke J.** Southwest Atlantic. In: Review of the state of world marine fishery resources.: FAO. 2011. p.93–106. Available from: http://www.fao.org/3/i2389e/i2389e.pdf

- **Vera-Alcaraz HS, Vuletich JDE.** Records of clupeiform fishes from Paraguay. ICP. 2016; 46:1–6.

- **Ward RD.** Genetics in fisheries management. Hydrobiologia. 2000; 420(1):191–201. https://doi.org/10.1023/A:1003928327503

- **Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN.** DNA barcoding Australia's fish species. Philos Trans R Soc Lond B Biol Sci. 2005; 360:1847–57. https://doi.org/10.1098/rstb.2005.1716

- **Whithead PJ, Nelson GJ, Wongrotana T.** FAO species catalogue. Rome: Food and Agriculture Organization of the United Nations; 1988. vol. 7, Clupeoid fishes of the world (suborder Clupeoidei): An annotated and illustrated catalogue of the herrings, pilchards, sprats, shads, anchovies and wolf-herrings, pt. 1, Chirocentridae, Clupeidae and Pristigasteridae. (FAO Fisheries Synopsis; no. 125).

- **YEATES DK, Seago A, Nelson L, Cameron SL, Joseph L, Trueman JWH.** Integrative taxonomy, or iterative taxonomy? Syst Entomol. 2011; 36(2):209–17. https://doi.org/10.1111/j.1365-3113.2010.00558.x

- **Zhang J, Pavlidis P, Stamatakis AA**. General species delimitation method with applications to phylogenetic placements. Bioinformatics. 2013; 29(22):2869–76. https://doi.org/10.1093/bioinformatics/btt499

AUTHOR'S CONTRIBUTION ●CRT

Fabiana Vicente: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing-original draft, Writing-review and editing.

Marina V. Loeb: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing-original draft, Writing-review and editing.

Andréa Carla Guimarães de Paiva: Data curation, Formal analysis, Investigation, Writing-original draft.

Claudio L. S. Sampaio: Conceptualization, Visualization, Writing-original draft.

Leandro Araujo Argolo: Conceptualization, Formal analysis, Writing-original draft, Writing-review and editing.

Uedson Pereira Jacobina: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Writing-original draft, Writing-review and editing.

ETHICAL STATEMENT

Samples were taken under Permanent license of Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) (SISBIO 13754- 4).

COMPETING INTERESTS

The authors declare no competing interests.

HOW TO CITE THIS ARTICLE

Neotropical Ichthyology

Official Journal of the
Sociedade Brasileira de Ictiologia