Species identification among morphologically-similar *Caranx* species

Shenna Kate M. Torres¹², Brian S. Santos¹²*¹

¹ University of the Philippines, Natural Sciences Research Institute, Diliman, Quezon City 1101, Philippines
² University of the Philippines, Institute of Biology, Diliman, Quezon City 1101, Philippines

Abstract

Accurate identification of species is important in assessing biodiversity and in conservation and population management strategies. Species with similar morphology, however, could be difficult to identify correctly. Published papers have reported on *Caranx* species that have been incorrectly identified and careful analysis is needed when identifying these species. In this study, 21 specimens from the genus *Caranx* were identified using DNA barcoding. Specimens were collected from Batangas, Philippines, where freshwater *Caranx*, which are about 50% more expensive, are present. Using morphological characteristics, *C. ignobilis* and *C. sexfasciatus* were identified correctly while one *C. papuensis* specimen was identified as *C. sexfasciatus*. Barcodes for three species, namely *C. ignobilis*, *C. sexfasciatus*, and *C. papuensis*, were detected based on cytochrome c oxidase I (COI) sequences. With the addition of *Caranx* COI sequences from GenBank, the calculated average K2P distance among species (8.19%), was higher than the average K2P distance within species (0.30%). Analysis of all available COI sequences of *C. sexfasciatus* from GenBank showed multiple cases of likely misidentification in other studies. Geometric morphometric analyses revealed morphological differences between specimens barcoded as *C. papuensis* and *C. sexfasciatus* that could aid in identifying the species.

Introduction

Species identification is an integral step in monitoring biodiversity. This information can be used to determine the biology and ecology of the species, which then can be used for conservation strategies, aquaculture, and sustainable fishery management (Jordan, Okey, Bauer, & Libralato, 2008; Joshi et al., 2014). Misidentification of species may pose a threat not only to the species itself, but also to the ecosystem through inaccurate monitoring processes, inappropriate usage of resources for conservation efforts, and an unobserved decline in fish stock (Garcia-Vasquez, Schiaffino-Machado, Campo, & Juanes, 2012). Traditionally, morphometric and meristic characters are used to identify species (Cadrin, 2000) but are not always effective and reliable for correct identification (Murta, 2000).

Carangidae, which includes jacks, pompanos, and scads, is a family of teleost fish belonging to the order Perciformes. This diverse group of family is currently comprised of 4 subfamilies with 30-32 genera and 146 species which inhabit mostly marine tropical to subtropical waters and penetrates brackish waters (Damerau, Freese, & Hanel, 2017). These teleost fishes are characterized mainly by an elongated, slightly compressed and deep body, separate dorsal fins, two anterior spines on their anal fin, a narrow caudal peduncle, and a forked caudal fin (Honebrink, 2000; Chan, Talbot, & Sukhavaisidh, 1974). Carangidae are
regarded to be of high economic value (Wilette & Padin, 2014; Lin & Shao, 1999) as total production amounted to 1,641,791 metric tons in Southeast Asia (FAO, 2017). One of the genera belonging to this family is *Caranx* from the subfamily Caranginae (Damerau, Freese, & Hanel, 2017). The genus *Caranx* is characterized by having a moderately large to very large body size, deep body, 20-31 gill rakers on the first gill arch, 7 dorsal spines, and 1-2 anal spines. Currently, there are 18 species comprising the genus *Caranx*.

In the Philippines, six species from this family, namely, *C. ignobilis*, *C. lugubris*, *C. melampygus*, *C. papuensis*, *C. sexfasciatus*, and *C. tille* have been reported in the marine ecosystem (Smith-Vaniz, 1999). The Philippines was one of the top countries in terms of yield in *Caranx* species in 2007 and 2008, with a total of 130 and 150 tonnes, respectively. However, the yield has been declining throughout the years, dropping to 24 tonnes in 2015, and to 4 tonnes in 2016 (FAO, 2018). This indicates the need for better management and conservation strategies to maintain their populations.

*Caranx ignobilis* (bigeye trevally) and *C. sexfasciatus* (giant trevally) have freshwater forms in Taal Lake, Batangas (Aquilino et al., 2011; Papa & Mamaril, 2011). They travel through the Balayan Bay upstream to Pansipit River, which drains into the Taal Lake. They can traverse these environments with different saline conditions because they are euryhaline. These two species of *Caranx* are regarded to be an important source of the fishery in Batangas (Alaira & Rebancos, 2014). Freshwater forms of *Caranx* from Taal Lake are highly prized and cost about 50% more compared to their marine counterparts.

Past studies (Willett & Padin, 2014; Aquilino et al., 2011) have shown that misidentification of *Caranx* species is common because they undergo significant changes in morphology as they mature; juveniles of *C. ignobilis* possess a silvery-yellow to silvery-brown coloration, as well as five to six dark vertical bands at the lateral side of the fish (Lin & Shao, 1999). In contrast, the normal coloration for mature *C. ignobilis* is a silvery-olive to blue-green at the dorsal side and silvery white for the ventral side. In addition, *C. ignobilis* may change color depending on the environment, with the body displaying either a white or yellow coloration (Phuong, Anh, Phuong, & Linh, 2015). This incongruity between the morphology of juveniles and adults may lead to a general taxonomic confusion (Jaafar et al., 2012).

This study aims to identify and discriminate *Caranx* species found in Taal Lake and Balayan Bay through DNA barcoding and geometric morphometrics. DNA barcoding is an efficient technique for rapid identification of species (Bickford et al., 2006) and inferring phylogenetic relationships (Dahrudin, 2016; Santos & Quilang, 2011; Hebert, Cywinska, Ball, & deWaard, 2003). It utilizes a specific region in the genome that serves as "barcodes" (Zhang, 2011; Ward, Zemlack, Innes, Last, & Hebert, 2005). In animals the mitochondrial DNA cytochrome c oxidase subunit I (COI) is used because it shows a greater range of variability between species (Hebert, Cywinska, Ball, & deWaard, 2003) and has shown effectiveness in identification of different groups of animals (Hajibabaei, Singer, Hebert, & Hickey, 2007). Geometric morphometrics, on the other hand, visualizes and quantifies differences in shape based on landmark points assigned to each specimen (Santos & Quilang, 2012; Zelditch, Swiderski, Sheets, & Fink, 2004; Richtsmeier, Deleon, & Lele, 2002).

**Materials and Methods**

**Sample Collection**

Putative *Caranx* sp. specimens were obtained from Lemery (13°53′06″ N, 120°54′53″ E), San Nicholas (13°55′45″ N, 120°57′07″ E), and Talisay (14°05′18″ N, 121°01′15″ E) localities of Batangas. Lemery represents an area in Balayan Bay whereas San Nicholas and Talisay represent the southern and northern areas of Taal Lake, respectively. The collected specimens were subjected to preliminary morphometric analyses. A total of 21 specimens were collected across the three areas, specifically, 11 from Lemery, 6 from San Nicholas, and 4 from Talisay. The total weight and length measurements (total length, fork length, standard length, scutes length, and body depth) were determined using a weighing scale and ruler, respectively. In addition, the number of dorsal and anal spines were counted. Finally, the condition factor, which was calculated by dividing the total weight (in grams) by the cube of the total length (in centimeters) and multiplying the quotient by 100, was recorded to assess the well-being and degree of fatness of the fish (Zelditch et al., 2004).

**Morphological Identification**

The 21 putative *Caranx* sp. were subjected to initial identification using the taxonomic keys of Smith-Vaniz (1999) and Mansor et al. (1998) (Figure 1). *C. ignobilis* was described as having a deep body with a body depth 1/3 of the fork length, a steep head profile that strongly curves above the eye, greater than 15 cm fork length, and possesses a scaleless patch in the base of the pectoral fin and breast. *C. papuensis* was described as having a conspicuous pale spot behind the posterodorsal margin of the operculum and having a narrow white border on the posterior margin of the lower lobe of the caudal fin. *C. sexfasciatus* was described as having a body depth that is 2/7 of the fork length, with a black spot on upper edge of gill cover and breast that is fully scaled.

**Sample Processing and Tissue Extraction**

The left body side of each specimen was photographed using a Canon EOS 700D DSLR Camera.
The tissue of the specimens was obtained from the epaxial muscle behind the operculum on the right side of the fish and was submerged in absolute ethanol and stored at -20°C. 20-25 milligrams of muscle tissue was utilized for DNA extraction using Fujifilm QuickGene DNA Tissue Kit S (DT-S).

Polymerase Chain Reaction (PCR) Amplification and DNA Sequencing

The DNA extracts of all specimens were amplified for the cytochrome c oxidase I (COI) region. 25 µL of PCR mixture was used, consisting of 17.375 µL ultra-pure water, 2.50 µL PCR buffer (10x) containing 15mM of MgCl₂, 1.25 µL of forward primer (10 µM), 1.25 µL of reverse primer (10 µM), 0.50 µL dNTPs (10mM), 0.125 µL of Taq polymerase, and 2.00 µL of DNA. The forward primer used was Fish F2 -5' TCGACTAATCATAAAGATATCGGCAC – 3', while the reverse primer was Fish R1 5' – TGATTTCTTGCCACCGAAGTCTA -3' (Ward et al., 2005). All samples were subjected to the following PCR conditions: 95 °C for 2 minutes, followed by 35 cycles of 94 °C for 0.5 minutes, 54 °C for 0.5 minutes, and 72 °C for 1 minute. Afterward, samples were kept to 72 °C for 10 minutes and then held at 4 °C.

All PCR products were visualized using 1% agarose gel stained with ethidium bromide. Gel electrophoresis was ran for 35 minutes at 100 V. The bands with expected lengths of 650bp were excised and extracted using QIAGEN™ QIAquick Gel Extraction Kit (QIAGEN Inc., Valencia, CA). Purified PCR products were then sent to Macrogen, Inc. (Seoul, Korea) for bidirectional DNA sequencing using Applied Biosystems™ 3730xl DNA analyzer.

Figure 1. Morphological Characteristics of Caranx sp. (A) Caranx ignobilis, (B) Caranx papuensis, (C) Caranx sexfasciatus
Sequence Analysis

The sequences obtained were assembled using Staden Package v1.7 (Staden, Beal, & Bonfield, 2000). Sequence identity was verified using BLASTn (https://blast.ncbi.nlm.nih.gov). MEGA v7 (Kumar, Stecher, & Tamura, 2015) was used to align the consensus sequences. Additional COI sequences from Caranx species were obtained from GenBank for comparison. At most 5 sequences for each species were selected. Carangoides dinema which belong to the same subfamily as Caranx was selected as the outgroup taxon (Santini & Carnevale, 2015) A neighbor-joining tree (Saitou, 1987) was constructed with 1000 bootstrap replicates. Lastly, pairwise genetic distance was calculated using the Kimura-2-Parameter model (Kimura, 1980).

Image Digitization and Data Analyses

Eighteen (18) landmark points from the left body side (Figure 2) were plotted on each digital image using tpsDig2 software (Rohlf, 2010). Generalized Procrustes Analysis (GPA) was performed to superimpose landmark coordinates as shape variables using the MorphoJ software (Klingenberg, 2011) to ensure that differences in shape would be independent of size, position or orientation (Slice, 2007). Likewise, shape variations between species were analyzed using Multivariate Analysis of Variance (MANOVA) and were summarized using Canonical Variate Analysis (CVA) using MorphoJ software.

Results

Morphological and Molecular Identification

From the initial assessment of the specimens using taxonomic keys, 11 were identified as C. ignobilis (8 from Lemery and 3 from Talisay), 5 were identified as C. papuensis (2 from Lemery and 3 from San Nicholas), and 5 were identified as C. sexfasciatus (1 from Lemery, 2 San Nicholas, and 1 from Talisay). All identifications were confirmed as correct based on BLASTn results except for one specimen initially identified C. sexfasciatus but genetically matching C. papuensis.

Table 1 shows the morphometric values obtained from three species of Caranx. The weight of the fishes ranged from 29.00 g to 274.00 g, with an average of 148.33 g. The total length of the samples ranged from 13.50 cm to 28.40 cm, with an average length of 21.70 cm across species. The condition factor yielded an average of 1.45, 1.27 and 1.31 for C. ignobilis, C. papuensis, and C. sexfasciatus, respectively. For fork length, an average of 18.36 cm, 22.78 cm, and 15.43 cm was obtained from C. ignobilis, C. papuensis and C. sexfasciatus, respectively. The average values of 17.46 cm, 22.02 cm, 14.65 cm were obtained for the standard length of C. ignobilis, C. papuensis and C. sexfasciatus, respectively. Scutes length for C. ignobilis, C. papuensis, and C. sexfasciatus yielded an average length of 6.58 cm, 9.02 cm, and 6.05 cm, respectively. In terms of body depth, an average of 6.99 cm, 7.97 cm, and 5.25 cm was measured for C. ignobilis, C. papuensis and C. sexfasciatus, respectively. All samples had a dorsal spine count of 7, and an anal spine count of 1 to 2.
Sequence Analysis

A total of 51 COI sequences, 21 from this study and 30 from GenBank, were aligned. In the NJ tree, a distinct group clustered together according to their species designation. The groups that clustered together were designated as *C. ignobilis*, *C. tille*, *C. melampygus*, *C. papuensis*, *C. sexfasciatus*, and *Carangoides dinema* (Figure 3).

The pairwise K2P distances (Table 2) were calculated. A total of 244 comparisons were made within species of *Caranx*, with an average genetic distance of 0.30%. A total of 801 comparisons were made between species within *Caranx* genus, which show an average distance of 8.19%. A total of 230 comparisons were made between genera, with an average genetic distance of 15.21%.

A total of 59 COI sequences of *C. sexfasciatus* obtained from GenBank were aligned and utilized to construct an NJ tree (Figure 4). There were four groups on the tree, each with 72-99% bootstrap support. The formation of diverging clades in the *C. sexfasciatus* neighbor-joining tree (Figure 4) indicates the presence of multiple species. BLASTn analysis of representative sequences from each clade matched *C. sexfasciatus*, *C. tille*, *C. papuensis*, and *Selaroides leptolepis*, respectively (Table 2).

**Table 1.** Morphometrics values of different species of *Caranx*

<table>
<thead>
<tr>
<th></th>
<th><em>C. ignobilis</em> (n=11)</th>
<th><em>C. papuensis</em> (n=6)</th>
<th><em>C. sexfasciatus</em> (n=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Weight (g)</td>
<td>132.82 (87.00-168.00)</td>
<td>222 (153.00-274.00)</td>
<td>80.5 (29.00-148.00)</td>
</tr>
<tr>
<td>Total Length (cm)</td>
<td>20.9 (19.80-22.20)</td>
<td>25.92 (21.50-28.40)</td>
<td>17.55 (13.50-21.60)</td>
</tr>
<tr>
<td>Condition Factor</td>
<td>1.45 (1.04-1.62)</td>
<td>1.27 (1.11-1.54)</td>
<td>1.31 (1.18-1.47)</td>
</tr>
<tr>
<td>Fork Length (cm)</td>
<td>18.36 (17.00-19.30)</td>
<td>22.78 (19.40-25.10)</td>
<td>15.43 (12.10-18.80)</td>
</tr>
<tr>
<td>Standard Length (cm)</td>
<td>17.46 (16.10-18.50)</td>
<td>22.02 (18.50-24.40)</td>
<td>14.65 (11.20-17.90)</td>
</tr>
<tr>
<td>Scutes Length (cm)</td>
<td>6.58 (6.00-7.20)</td>
<td>9.02 (7.50-10.10)</td>
<td>6.05 (4.80-7.30)</td>
</tr>
<tr>
<td>Body Depth (cm)</td>
<td>6.99 (6.10-7.90)</td>
<td>7.97 (7.40-8.80)</td>
<td>5.25 (3.80-7.00)</td>
</tr>
</tbody>
</table>

**Geometric Morphometrics**

Canonical variate analysis (CVA) yielded two significant canonical variates (CVs), wherein the shape variations along the X-axis (CV1) and the shape differences along the Y-axis (CV2) accounted for 77.97% and 22.03% of the total variance, respectively (Figure 5). CV1 accounted for differences in the anterior insertion of the dorsal fin, wherein the more positive end is correlated with fishes with deeper body depth while the more negative end is correlated with fishes with shallow body depth. On the other hand, CV2 corresponded to the disparities in the anterior tip of the snout, wherein the more positive end is nearer the anterior extreme of the orbit while the more negative end is farther from the anterior extreme of the orbit.

**Discussion**

In this study, identification of *Caranx* species was not entirely successfully based on morphological characters alone. Only five out of the six *C. papuensis* specimens were correctly classified, with one initially classified as *C. sexfasciatus*. This misidentification was brought about by a black spot on upper edge of gill cover of the specimen, which is known to be a striking feature of *C. sexfasciatus*, as well as the lack of a well-defined

**Table 2.** Summary of percent genetic divergence within and between species

<table>
<thead>
<tr>
<th>Comparison</th>
<th>No. of Comparisons</th>
<th>Minimum</th>
<th>Distance Mean</th>
<th>Maximum</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within species</td>
<td>244</td>
<td>0.00</td>
<td>0.30</td>
<td>0.95</td>
<td>0.02</td>
</tr>
<tr>
<td>Between species within <em>Caranx</em> genus</td>
<td>801</td>
<td>1.93</td>
<td>8.19</td>
<td>11.00</td>
<td>0.08</td>
</tr>
<tr>
<td>Between <em>Caranx</em> and <em>Carangoides</em></td>
<td>230</td>
<td>12.75</td>
<td>15.21</td>
<td>18.33</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Figure 3. Neighbor-joining tree of 51 COI sequences from five species of *Caranx* and *Carangoides dinema* (outgroup taxon) based on K2P distances. Sequence IDs followed by asterisk (*) are from specimens used in this study.
Figure 4. Neighbor-joining tree of 59 COI sequences of *C. sexfasciatus* obtained from GenBank based on K2P distances. Bootstrap support using 1000 replicates are shown.
white border at the posterior portion of the lower lobe of the caudal fin, which is a characteristic feature of *C. papuensis*. The presence of the former trait and the absence of the latter trait contributed to the misidentification of the specimen using traditional morphological identification. In addition, there is a lack of measurement-based morphological identification techniques for *C. papuensis*, unlike for *C. ignobilis* and *C. sexfasciatus*. For *C. ignobilis*, it is classified as such if the specimen possesses a deep body with a body depth 1/3 of the fork length, whereas for *C. sexfasciatus*, the sample is identified as such if it has a body depth that is 2/7 of the fork length.

Based on the NJ tree, distinct clusters of *Caranx* were formed with high bootstrap value and separated the highly similar *C. sexfasciatus* and *C. papuensis*. The occasional exception of *C. sexfasciatus* (LCr17) and *C. sexfasciatus* (HQ654685) clustering in *C. papuensis* group is most likely a case of misidentification. An average K2P distance of 0.30% and 8.19% for within species and between species of *Caranx* genus were obtained, respectively. Hence, the average interspecific genetic distance was 27 times higher than the average intraspecific distances. The average intraspecific K2P distance in this study was within the standard screening threshold of 3.5% for fish species by Ward et al. (2009). This is also consistent with the findings of Hubert et al. (2008) in which the calculated intraspecific and interspecific distance for 190 Canadian freshwater fish species was 0.3% and 8.3%, respectively.

Most of the studies over the years identify different species of *Caranx* sp. as *C. sexfasciatus* (Table 3). Wilette and Padin (2014) also stated that samples previously identified as *C. sexfasciatus* were actually *C. papuensis*, which was determined through diagnostic coloration patterns and cytochrome *b* phylogenetic reconstruction. These findings emphasize the need for an integrative method for species identification, utilizing both morphological and molecular means to be certain of the identity of the samples.

Shape differences in the anterior insertion of the dorsal fin and anterior tip of the snout accounted for the morphological variation among species of *Caranx*. In the CVA plot, CV1 separated the population of *C. ignobilis* from *C. papuensis* and *C. sexfasciatus* samples. This separation implies that *C. ignobilis* has a deeper body depth compared to the other two species, validating the published morphological identification for this species, wherein the body depth of *C. ignobilis* is 1/3 of its fork length, while for *C. sexfasciatus*, its body depth is 2/7 of its fork length (Smith-Vaniz, 1999). Thus, the expected morphological appearance of *C. ignobilis* would have a body depth that is greater than that of *C. sexfasciatus*, which is consistent with the obtained geometric morphometric results. On the other hand, CV2 separated the population of *C. sexfasciatus* from the other two species. This implies that the anterior tip of the snout of *C. sexfasciatus* is nearer to the anterior extreme of the orbit of the eye compared to the other species. Conversely, this means that the anterior tip of the snout of *C. ignobilis* and *C. papuensis* is farther from the anterior extreme of the orbit of the eye of the fish. The steeper head profile of *C. ignobilis* contributed to the increased distance between the anterior tip of the

Figure 5. Canonical Variate Analysis of 21 specimens of *Caranx* sp. The CV1 and the CV2 axes account for 77.97% and 22.03% of the total variance, respectively. Deformation grids for CV1 and CV2 are shown.
Table 3. Possible misidentification of Caranx sexfasciatus

<table>
<thead>
<tr>
<th>Accession number from GenBank</th>
<th>Author</th>
<th>Place</th>
<th>Species match</th>
<th>Percent identity based on BLASTn</th>
</tr>
</thead>
<tbody>
<tr>
<td>KC970459.1</td>
<td>Yambot AV, Alcantara SG and Templonuevo RMT</td>
<td>Cuyo, Palawan, Philippines</td>
<td>Caranx papuensis      [KJ20138.1]</td>
<td>99%</td>
</tr>
<tr>
<td>HQ654681.1 - HQ654685.1</td>
<td>Aquilino SV, Tango JM, Fontanilla IK, Pagulayan RC, Basiao ZU, Ong PS and Quilang JP</td>
<td>Taal Lake, Philippines</td>
<td>Caranx papuensis [KP194683.1]</td>
<td>99%</td>
</tr>
<tr>
<td>KC508512.1</td>
<td>Priyanga K, Thangaraj M, Singh R, and Barathkumar TR</td>
<td>India</td>
<td>Selaroides leptolepis [KM079293.1]</td>
<td>100%</td>
</tr>
</tbody>
</table>

snout to the anterior orbit of the eye. These results indicate that the variation in the distance between the tip of the snout and the anterior orbit of the eye may be used to differentiate C. papuensis from C. sexfasciatus.

Information gathered from species identification can be used for conservation and protection strategies. The data presented in this paper identifies three species of morphologically-similar Caranx species in Batangas, Philippines. The utilization of both morphological and genetic analysis for species identification would be ideal for future studies to accurately identify species. In studies where misidentification of Caranx species was present, only one method of species identification was used; supporting the genetic analysis with identification of distinguishing anatomical traits would lead to a higher certainty. Another possible avenue of study for species as a window on diversity and conservation. Trends in Ecology and Evolution, 22 (3), 148-155. doi:10.1016/j.tree.2006.11.004


Acknowledgements

This study was funded by the Natural Science Research Institute of the University of the Philippines (Diliman, Quezon City) under Project Code Number BIO-18-1-01. Likewise, the authors would like to thank the Institute of Biology, College of Science, UP Diliman for the logistical support.

References


167–172. doi:10.1016/j.tig.2007.02.001


