**Philippine Journal of Science**
150 (S1): 223-236, Special Issue on Biodiversity
ISSN 0031 - 7683
Date Received: 22 Jul 2020

# Genetic Diversity and Population Structure of Three *Caranx* Species from Batangas Province, Philippines

**Shenna Kate M. Torres**[1*] and **Brian S. Santos**[2]

[1]Natural Sciences and Research Institute; [2]Institute of Biology
College of Science, University of the Philippines Diliman
Quezon City 1101 Philippines

**Species belonging to the genus *Caranx*, locally known as "talakitok," belong to the economically important fish in the Philippines. The popularity of these species makes them prone to overexploitation, which may result to a decline in their population in the wild. Despite these circumstances, stock assessment and population genetic variation studies are scarce. In this study, three species from this genus (*C. ignobilis*, *C. papuensis*, *C. ignobilis*) from the Batangas region were subjected to genetic diversity and population structure analyses using the cytochrome b (cyt b) gene of the mitochondrial DNA (mtDNA) to address these information gaps. High haplotype and high nucleotide diversity were noted in *C. ignobilis*, which may indicate that the population is still large and stable. However, *C. papuensis* and *C. sexfasciatus* populations had a low nucleotide diversity and high haplotype diversity, which could mean a possible genetic bottleneck in the recent past. Likewise, each of these three species showed no genetic differentiation between marine and freshwater specimens, which can be attributed to their life history and biology. Analysis of molecular variance (AMOVA) revealed weak genetic structure, indicated by low percentage value between populations. Based on neutrality tests and mismatch distribution analysis, the three species may have possibly undergone demographic expansion. This study provides the genetic profile of *Caranx* species found in the Batangas region before the recent Taal Volcano eruption (January 2020), which can be used to investigate the effects of this eruption on the population of this species. Likewise, results obtained from this study served as preliminary data for the population genetics of *Caranx* spp. found in Batangas, Philippines.**

Keywords: aquaculture, Balayan Bay, "maliputo," Taal Lake, "talakitok"

## INTRODUCTION

Seafood – specifically fish – has become an integral part of our life, as it not only provides a source of food that is both tasteful and nutritious to humans but also contributes heavily to economic growth. Global fish production in 2018 provided a combined sale value of USD 401 billion from both the aquaculture and the capture fisheries sector (FAO 2020). However, despite increasing aquaculture production of fish, the growth of the human population continues to rise alongside it, which results in a relative scarcity of this resource. This has led to heavy exploitation of various aquatic resources for consumption – especially marine resources – subjecting them to stressors such as overfishing, rapid land use, and introduction of non-native and invasive species (Santos *et al*. 2010; Daly *et al.* 2019), which may have attributed to the observed decline in catch data. Proper management and conservation of marine resources for sustainability requires detailed knowledge on social, economic, ecological (*i.e.* species and environment), and pragmatic aspects (Thrush and

*Corresponding Authors: smtorres@up.edu.ph

Dayton 2010). However, relevant information regarding the biological aspects of many economically important fish remains elusive (Reiss *et al*. 2009).

Among the countries that have experienced an average decline in its fish resource from 2015–2018 production was the Philippines, with a noted decline of 3.07% (from 1,950,000 tons down to 1,890,000 tons) and 20% (from 200,000 tons to 160,000 tons) from its inland and marine capture production, respectively (FAO 2020). One species that has experienced an average decline in its catch data was the "talakitok" (*Caranx* spp.), an economically important fish in the Philippines. Data from the Bureau of Agricultural Statistics (2012, 2019) showed an average decline of 27.67% (from 32,709.43 metric tons down to 23,658.82 metric tons) from 2010–2018 for "talakitok." In the Philippines, seven species from the genus *Caranx* can be found in the marine environment; three of which thrive in freshwater environments – namely, *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* (Table 1) and are usually referred to as "talakitok" locally (Ganaden and Lavapie-Gonzales 1999). The *Caranx* spp. found in Batangas Province in the Philippines are highly valued since they are a source of income for the residents near the lake and are also utilized for aquaculture (Alaira and Rebancos 2014). Specifically, *C. ignobilis* is a high-priced food fish where a kilo usually amounts to PHP 350–600 (~ USD 7–12) and is regarded to be a specialty food in Batangas Province due to its tasty meat contributing to high demand for this species (Chavez *et al*. 2013; Alaira and Rebancos 2014). To capitalize on its economic significance and as a response to high demand for these fish species, the Bureau of Fisheries and Aquatic Resources – through the Freshwater Fisheries Research and Development Center in Butong, Taal, Batangas, Philippines (Mutia *et al.* 2015) – has started a project in 2006 that sought to breed *C. ignobilis* in captivity. Local efforts for breeding these species in captivity can also be seen, as evidenced by the practice of cage farming for *C. ignobilis* within the area of Taal Lake, Batangas, Philippines (Alaira and Rebancos 2014).

Batangas is a province located in the northwestern part of the Philippines near the West Philippine Sea, which opens to the South China Sea (www.batangas.gov.ph). Within this region, two bodies of water with known biodiversity – namely, Balayan Bay (marine environment) and Taal Lake (freshwater) – can be found. Taal Lake, formerly known as Bombon Lake, is the third-largest lake in the Philippines with a total surface area of 234.2 km$^2$ and an average depth of 60.1 m (Villadolid 1937). It was once continuous with Balayan Bay, which opens to the West Philippine Sea. A series of volcanic eruptions and other geologic processes altered the landscape and produced natural barriers, resulting in the formation of Taal Lake (Ramos 2010). The lake is almost completely devoid of connection with other bodies of water; there is a narrow river, called Pansipit River, that connects Taal Lake with Balayan Bay and is the sole drainage outlet of the lake (Willette and Padin 2014). These changing landscapes may have led to changes in the population structure of *Caranx* spp. found in Batangas, thus the need for studies regarding their population.

Population studies on the genus *Caranx* are limited. One possible reason would be their cryptic nature as a result of the significant changes they undergo as they grow and mature; their morphology also changes depending on its habitat (Honebrink 2000; Willette and Padin 2014), thus making them prone to misidentification (Willette and Padin 2014). One way to study the population structure of these species is by using genetic tools, such as population genetic analysis. In this method, mtDNA markers are commonly used to study population structures and to

**Table 1.** Species of *Caranx* found in Batangas Province, Philippines

| Species | Author and year | Local name | Reported location |
|---|---|---|---|
| *C. ignobilis* | D.V. Villadolid (1937) | "Maliputo" | Lake Bombon (Taal Lake), Batangas Province, Philippines<br>Pansipit River, Batangas Province, Philippines<br>Balayan Bay, Batangas Province, Philippines |
| | A.W. Herre (1953) | "Maliputo" and "lison" | Lake Bombon (Taal Lake), Batangas Province, Philippines |
| *C. papuensis* | D.A. Willette and J.I.M. Padin (2014) | – | Taal Lake, Batangas Province, Philippines |
| | S.K.M. Torres and B.S. Santos (2020) | – | Balayan Bay, Batangas Province, Philippines |
| *C. sexfasciatus* | A.W. Herre (1953) | "Muslo" and "lison" | Lake Bombon (Taal Lake), Batangas Province, Philippines<br>Pansipit River, Batangas Province, Philippines<br>Balayan Bay, Batangas Province, Philippines |

infer population histories since it is cost-effective and requires less laboratory work compared to other markers (Galtier *et al*. 2009).

This study aims to assess the genetic diversity and population structure of *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* found in Batangas, Philippines – both from marine (Balayan Bay) and freshwater (Taal Lake) environments – using the cyt b gene of the mtDNA as the molecular marker. In addition, the control region of the mtDNA was also used as a molecular marker for *C. ignobilis*. Specifically, the study seeks to determine the haplotype and nucleotide diversity of these three species, which may give insights as to whether their populations are stable or in decline. In addition, the extent of the genetic differentiation between populations located in marine and freshwater environments will also be determined. Based on genetic differentiation, it can be determined whether the population is continuous or if subpopulations have formed. Information regarding the population structure of these species shall provide an idea of demographic patterns, which may be useful for the identification of independent management units (MUs) for the formulation of policies for the sustainability of these species.

## MATERIALS AND METHODS

### Study Site and Specimen Collection

Three localities in Batangas were chosen as sampling sites: Lemery (13°53'06" N, 120°54'53" E), which represents Balayan Bay (marine environment), while San Nicholas (13°55'45" N, 120°57'07" E) and Talisay (14°05'18" N, 121°01'15" E) represent the southern and northern areas of Taal Lake (freshwater environment), respectively. A total of 87 putative *Caranx* spp. specimens were collected in 2018 – specifically, 30 of *C. ignobilis*, 30 of *C. papuensis*, and 27 of *C. sexfasciatus*. The specimens were obtained through the help of local fishermen in the area, which explicit instructions as to where the specimens should be caught. The collection of *C. ignobilis* specimens was also personally overseen by the researchers to ascertain that the specimens were obtained from the appropriate areas. The collected specimens were first analyzed using the taxonomic keys of Smith-Vaniz (1999) and Mansor *et al*. (1998). *C. ignobilis* specimens are identified through a steep head profile that strongly curves above the eye, a body depth that is 1/3 of the fork length, and a scaleless patch at the base of the pectoral fin and breast. *C. papuensis* is characterized by a conspicuous pale spot behind the posterodorsal margin of the operculum and a narrow white border on the posterior margin of the lower lobe of the caudal fin. Lastly, *C. sexfasciatus* normally has a body depth that is 2/7 of the fork length, with a black spot at the upper edge of the gill cover and breast that is fully scaled.

### Specimen Processing and DNA Extraction

A sample of the epaxial muscle tissue located behind the operculum of the right side of each specimen was excised and then stored in absolute ethanol at –20 °C prior to extraction. Then, ISOLATE II Genomic DNA Extraction Kit – which is manufactured by Meridian Bioscience® – was used for the DNA extraction of 20–25 mg of muscle tissue.

### PCR Amplification and DNA Sequencing

DNA extracts of all the specimens were further processed with the amplification of the cyt b gene of the mtDNA *via* polymerase chain reaction (PCR). A total volume of 25 µL was used for each PCR reaction, consisting of 17.375 µL ultra-pure water, 2.50 µL PCR buffer (10x) containing 15 mM of $MgCl_2$, 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 purified DNA. The forward primer used was 5′-AAC TGC AGC CCC TCA GAA TGA TAT TTG TCC TCA-3′, while the reverse primer was 5′- GTG ACT TGA AAA ACC ACC GTT G-3′ (Willette and Padin 2014). The PCR conditions were as follows: 95 °C for 7 min, followed by 35 cycles of 94 °C for 0.50 min, 58 °C for 0.58 min, and 72 °C for 2 min; afterward, the samples were kept at 74 °C for 7 min and then held at 4 °C.

The control region gene of the mtDNA of *C. ignobilis* was amplified *via* PCR and using the primers designed by the study. The forward primer used was 5'-TTC AGA CCA TTC AAT GTA TTA GCA A-3', while the reverse primer used was 5'-TTA GGG GCT TTC CTG TTT CC-3'. A total volume of 25 µL was used for each PCR reaction containing the same components and amount described above. The PCR conditions were as follows: 95 °C for 2 min, followed by 35 cycles of 94 °C for 0.45 min, 50 °C for 1 min, and 72 °C for 1 min; afterward, the samples were kept at 74 °C for 10 min and then held at 4 °C.

The PCR products were then visualized in 1% agarose gel using ethidium bromide. The PCR products were purified using a Promega Wizard® SV Gel and PCR Clean-Up System Kit. The purified PCR products were sent to Macrogen, Inc. (Seoul, Korea) for bidirectional DNA sequencing using the Applied Biosystems™ 3730xl DNA analyzer.

### Data Analyses

Staden Package was used to process the generated DNA sequences. The forward and reverse sequences were aligned using Pre-gap 4 version 1.6. A consensus sequence was generated using Gap version 4.11.2 software. The

sequences were then subjected to nucleotide BLAST for comparison of the generated sequences with publicly available sequences posted at GenBank database (https://www.ncbi.nlm.nih.gov/genbank/), thus individually validating the identity of the specimens whether they be *C. ignobilis* or *C. sexfasciatus* based on the percent identity match (98–100%) and query cover. For *C. papuensis*, due to lack of publicly available cyt b sequences at GenBank, validation was made through morphological identification and subjecting the specimens to DNA barcoding using the COI gene of the mtDNA (the data were published in the paper of Torres and Santos 2020).

The haplotype diversity (h), nucleotide diversity (π), and number of polymorphic sites of the specimens were calculated using DNA Sequencing Polymorphism version 6.12.03 software (Librado and Rozas 2009). These measures of genetic variation will help determine the status of the populations of the *Caranx* species.

On the other hand, to determine the level of genetic differentiation between sampling sites, pairwise $F_{ST}$ values for each specimen were calculated using Arlequin version 3.5.1.2 (Excoffier and Lischer 2010). To estimate the level of sequence divergence within and among populations, an AMOVA was performed using Arlequin version 3.5.1.2 software.

To evaluate the genetic relationship between specimens, a maximum likelihood tree and a haplotype network were generated. For the construction of maximum likelihood tree for cyt b sequences, 16 additional cyt b sequences of *Caranx* spp. were downloaded from GenBank for comparison – namely, two *C. ignobilis*, two *C. sexfasciatus*, two *C. tille*, two *C. crysos*, two *C. caballus*, one *C. latus*, one *C. vinctus*, one *C. caninus*, one *C. melampygus*, one *C. hippos*, and one *C. rhoncus*. Likewise, additional sequences of *Carangoides equula* (two sequences) were downloaded from GenBank to serve as an outgroup taxon. For the construction of maximum likelihood tree for control region sequences of *C. ignobilis*, five additional control region sequences of *Caranx* spp. were downloaded from GenBank for comparison – namely, two *C. ignobilis*, two *C. tille*, and one *C. melampygus*. Likewise, additional control region sequence of *Carangoides equuala* was added to the tree to serve as an outgroup taxon. For each molecular marker, appropriate model testing for base substitution frequencies using the jModelTest software (Darriba *et al*. 2012) was done. MEGA version 7 was used to construct the maximum likelihood tree with 1000 bootstrap replicates, while Population Analysis with Reticulate Trees software was used to construct the haplotype network (Leigh and Bryant 2015). The generated networks were used to determine how the unique haplotypes were connected, and to detect if separate populations have formed.

To assess the changes in demographic patterns and population growth of the three species, Tajima's D, Fu's FS, and mismatch distribution analysis were calculated using Arlequin version 3.5.1.2 (Excoffier and Lischer 2010). In addition, the construction of the mismatch distribution graph was done using Microsoft Excel® for Office 365. A statistically significant negative value for Tajima's D and Fu's FS may reveal the presence of rare alleles due to the effect of purifying selection or population expansion (Fu and Li 1993). For the mismatch distribution analysis, the calculated Harpending's raggedness index (Hri) and the sum of squared deviations (SSD) were used to examine whether the observed data fit the expected data for both demographic and spatial expansions models.

## RESULTS

The study yielded 30 cyt b sequences and 29 control region sequences for *C. ignobilis*, 30 cyt b sequences for *C. papuensis*, and 27 cyt b sequences for *C. sexfasciatus*. The sequence alignment consisted of 401 bp and 501 bp for cyt b and control region, respectively. For the *C. ignobilis* cyt b sequences, 10 sites were polymorphic (four singleton variable sites, six parsimony informative sites, and zero insertion/deletion), while the control region sequences for *C. ignobilis* generated 189 polymorphic sites (56 singleton variable sites, 133 parsimony informative sites, and zero insertion/polymorphic). For the haplotype analysis of *C. ignobilis* sequences, 10 and 29 unique haplotypes were generated from the cyt b and control region, respectively. On the other hand, *C. papuensis* sequences had six polymorphic sites (three singleton variable sites, three parsimony informative sites, and zero insertion/deletion) while *C. sexfasciatus* sequences had 10 polymorphic sites (five singleton variable sites, five parsimony informative sites, and zero insertion/deletion). Out of 30 cyt b sequences of *C. papuensis*, seven unique haplotypes were generated. For *C. sexfasciatus*, out of the 27 cyt b sequences, 10 unique haplotypes were generated.

Based on the haplotype analysis of cyt b and control region sequences of *C. ignobilis*, a high haplotype diversity (0.869 and 1.000, respectively) and high nucleotide diversity (0.504% and 7.87%, respectively) were observed (Table 2). On the other hand, AMOVA values calculated from the cyt b and control region revealed that most of the variation came from within the population (93.47% and 100%, respectively). This was further supported by the calculated $F_{ST}$ values that were not statistically significant for both cyt b and control region sequences, indicating the absence of genetic differentiation between *C. ignobilis* specimens from marine and freshwater environments ($F_{ST} = 0.0653$, $p = 0.0721$ and $F_{ST} = 0.0181$, $p = 0.4639$, respectively).

**Table 2.** Measures of genetic diversity in each population of *C. ignobilis* (legend: n – sample size; $n_H$ – number of haplotypes; h – haplotype diversity; s – number of segregating sites; k – average number of nucleotide differences; π – nucleotide diversity)

| Population | Cyt b | | | | | | Control region | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | $n_H$ | h | s | k | π (%) | n | $n_H$ | h | s | k | π (%) |
| Marine | 15 | 7 | 0.886 | 6 | 2.019 | 0.504 | 14 | 14 | 1.000 | 153 | 45.692 | 9.120 |
| Freshwater | 15 | 7 | 0.800 | 9 | 1.886 | 0.470 | 15 | 15 | 1.000 | 136 | 29.914 | 5.971 |
| All | 30 | 10 | 0.869 | 10 | 2.023 | 0.504 | 29 | 29 | 0.998 | 189 | 37.892 | 7.563 |

Calculation of the overall haplotype diversity in *C. papuensis* populations revealed a high value (h = 0.733); however, the overall nucleotide diversity was low (π = 0.362%) (Table 4). Similar results were found for *C. sexfasciatus* where the haplotype diversity was high (h = 0.772), and the nucleotide diversity was low (π = 0.384%) (Table 6). AMOVA for both *C. papuensis* and *C. sexfasciatus* sequences revealed that 100% of the genetic variation came from variation within the population. However, the calculated genetic differentiation between the marine and freshwater environment yielded non-statistically significant negative values ($F_{ST}$ = –0.0435, $p$ = 0.7478 and $F_{ST}$ = –0.0449, $p$ = 0.7297, respectively).

**Table 4.** Measures of genetic diversity of the cyt b gene in each population of *C. papuensis* (legend: n – sample size; $n_H$ – number of haplotypes; h – haplotype diversity; s – number of segregating sites; k – average number of nucleotide differences; **π** – nucleotide diversity)

| Population | n | $n_H$ | h | s | k | π (%) |
|---|---|---|---|---|---|---|
| Marine | 21 | 5 | 0.700 | 3 | 1.190 | 0.297 |
| Freshwater | 9 | 5 | 0.861 | 6 | 2.194 | 0.547 |
| All | 30 | 7 | 0.733 | 6 | 1.451 | 0.362 |

**Table 6.** Measures of genetic diversity of the cyt b gene in each population of *C. sexfasciatus* (legend: n – sample size; $n_H$ – number of haplotypes; h – haplotype diversity; s – number of segregating sites; k – average number of nucleotide differences; **π** – nucleotide diversity)

| Population | n | $n_H$ | h | s | k | π (%) |
|---|---|---|---|---|---|---|
| Marine | 5 | 4 | 0.900 | 3 | 1.200 | 0.299 |
| Freshwater | 22 | 10 | 0.788 | 11 | 1.662 | 0.415 |
| All | 27 | 10 | 0.772 | 10 | 1.538 | 0.384 |

The relationship among the 27 unique haplotypes of *Caranx* spp. generated from this study and the additional 18 haplotypes downloaded from GenBank were inferred using a maximum likelihood tree (Figure 1). The optimal model selected was the Hasegawa Kishino Yano with invariables sites (GTR + I). The maximum likelihood tree showed eight clades with 100% bootstrap support value. Specimens of *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* formed a separated clade from each other. For the *C. ignobilis* clade, the two haplotypes of *C. ignobilis* from Hawaii did not form a separate subclade. For the *C. sexfasciatus* clade, the two haplotypes of *C. sexfasciatus* from the North Pacific Ocean and Taiwan did not also form a separate subclade. Likewise, other species of *Caranx* downloaded from GenBank joined the *C. sexfasciatus* clade – namely, *C. caninus*, *C. tille*, and *C. melampygus*. On the other hand, the *C. papuensis* specimens formed a separate clade. Other separate clades that formed in the tree include the clade for *C. crysos*, *C. caballus*, *C. rhoncus*, and the outgroup taxa – namely, *Carangoides equula*. In addition, the subclades formed in the clade of *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* did not discriminate between marine and freshwater specimens of these species.

For the maximum likelihood of control region sequences of *C. ignobilis*, Tamura 3 Parameter with invariable sites were used as an optimal model (Figure 2). The 29 unique haplotypes of *C. ignobilis* sequences from the control region generated from the study formed a separate clade from other sequences of *Caranx* downloaded from GenBank. The two additional control region sequences of *C. ignobilis* from Hawaii, which were downloaded from GenBank, did not form a separate subclade from *C. ignobilis* haplotypes from the Philippines. Likewise, *C. ignobilis* specimens from marine and freshwater also did not form a separate subclade.

The result of the maximum likelihood tree for cyt b sequences was further supported by the generated haplotype network (Figure 3). Three groups were formed in the network, which separated the three species. However, this division was not based on the source of the population (marine *vs.* freshwater). The network had a star-burst appearance with multiple branches (Ferreri *et al*. 2011). For the *C. ignobilis* network, one central haplotype (haplotype 5) with a frequency of nine sequences generated from the study can be observed. Likewise, the two *C. ignobilis* sequences from Hawaii joined the most common haplotype (haplotype 5). On the other hand, the *C. papuensis* network had also one central haplotype (haplotype 3) with a frequency of 13 sequences generated

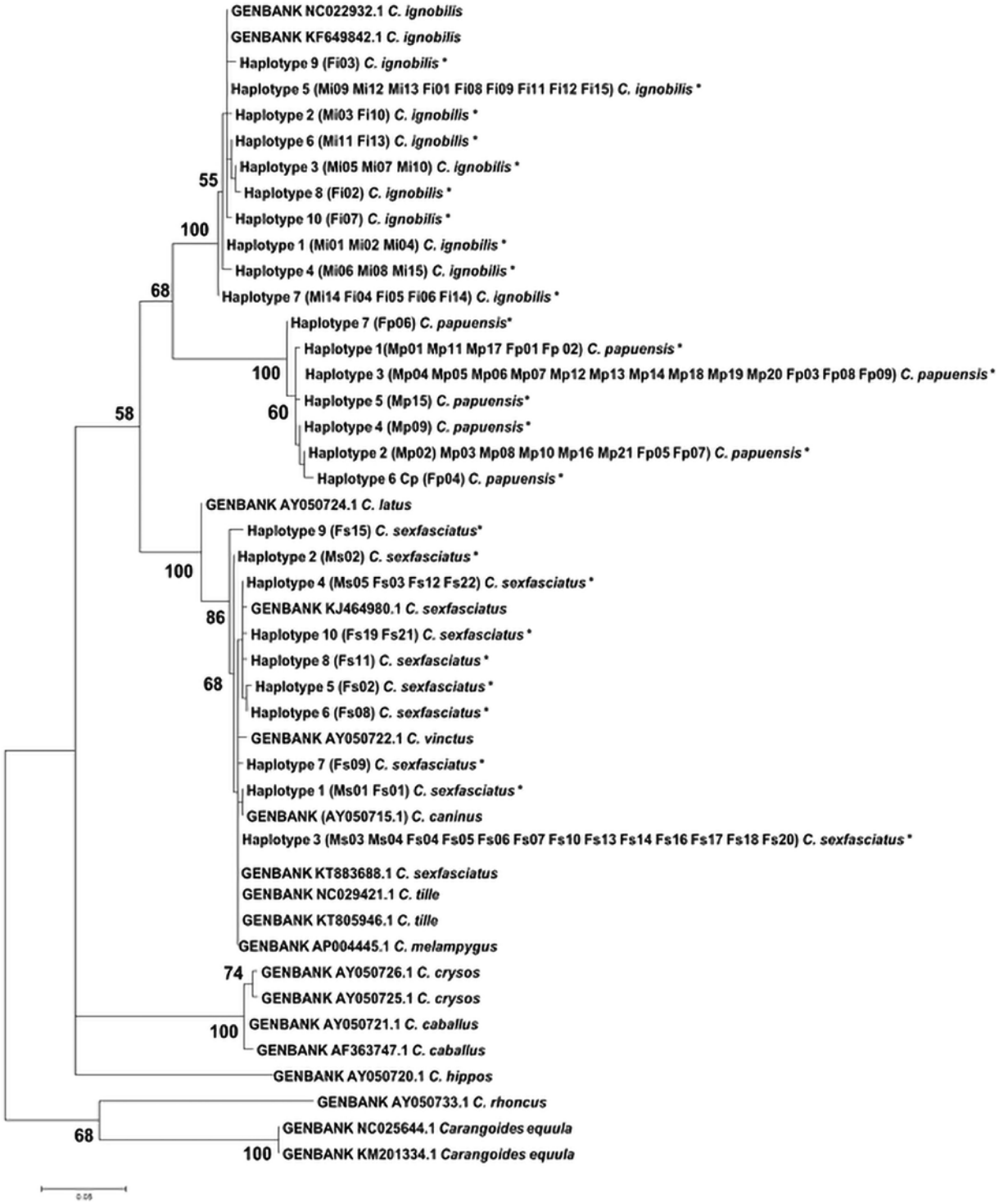


Figure 1. Maximum likelihood tree of cyt b sequences of three species of *Caranx* from Batangas, Philippines based on HKY + I model. The values at nodes represent the bootstrap support (1000 replicates). Bootstrap support value of less than 50% are not shown in the tree. The tree includes 16 cyt b sequences of different species of *Caranx* from other countries for comparison. Likewise, two cytochrome b sequences of *Carangoides equula* were included serving as outgroup taxon. Sequences followed by an asterisk (*) are generated from the study (legend: Mi – *C. ignobilis* from marine; Fi – *C. ignobilis* from freshwater; Mp – *C. papuensis* from marine; Fp – *C. papuensis* from freshwater; Ms – *C. sexfasciatus* from marine; Fs – *C. sexfasciatus* from freshwater)

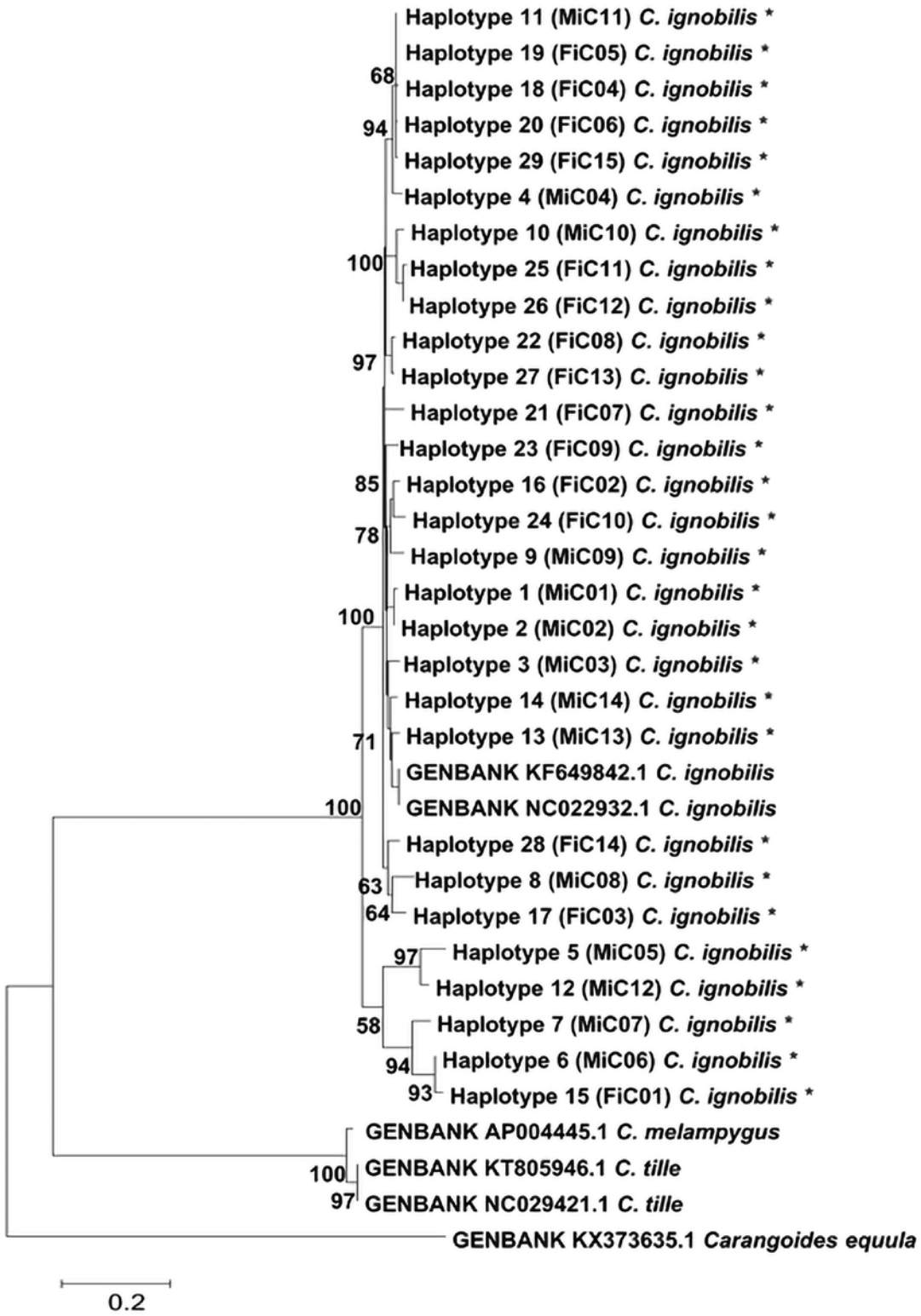


Figure 2. Maximum likelihood tree of control region sequences of *C. ignobilis* from Batangas, Philippines based on T92 + G model. The values at nodes represent the bootstrap support (1000 replicates). Bootstrap support value of less than 50% are not shown in the tree. The tree includes five control region sequences of different species of *Caranx* from other countries for comparison. Likewise, one control region sequence of *Carangoides equula* was included serving as outgroup taxon. Sequences followed by an asterisk (*) are generated from the study (legend: Mi – *C. ignobilis* from marine; Fi – *C. ignobilis* from freshwater; Mp – *C. papuensis* from marine; Fp – *C. papuensis* from freshwater; Ms – *C. sexfasciatus* from marine; Fs – *C. sexfasciatus* from freshwater).

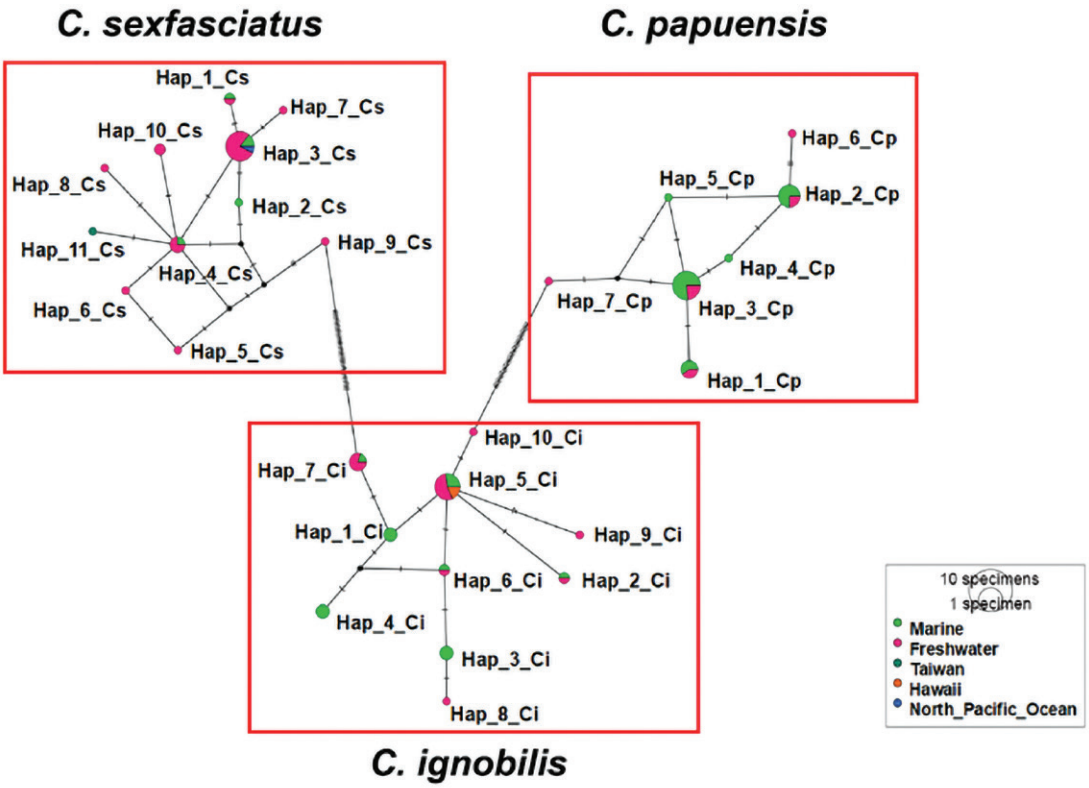


**Figure 3.** Median-joining haplotype network for cyt b sequences. Each circle represents a haplotype. The size of the circle is proportional to the number of haplotypes. The color(s) refers to the locality(ies) where the haplotype was collected. The lines on the branches represent the number of nucleotide changes while the black dot represents the theoretical median vectors introduced by the software (legend: Ci – *C. ignobilis*; Cp – *C. papuensis*; Cs – *C. sexfasciatus*).

from the study. Lastly, the *C. sexfasciatus* haplotype had also one central haplotype (haplotype 3) with a frequency of 13 sequences generated from the study. It can also be observed that the sequence of *C. sexfasciatus* from the North Pacific Ocean joined the central haplotype (haplotype 3), while the sequence of *C. sexfasciatus* from Taiwan had a separate haplotype (haplotype 11). On the other hand, the generated haplotype network for control region sequences of *C. ignobilis* did not present a star-like pattern, showing an ancestral haplotype as to where the other haplotypes could have branched off (Figure 4).

The values obtained for the neutrality tests, namely Tajima's D and Fu's FS, were used to infer the demographic history of each population. For the *C. ignobilis* population, the overall Tajima's D and Fu's FS values, using both cyt b and control region (Table 3) sequences, were negative; however, it was not statistically significant. On the other hand, the *C. papuensis* population had a non-statistically positive value for Tajima's D and a non-statistically negative value for Fu's FS test (Table 5). Lastly, the *C. sexfasciatus* population had a non-statistically negative value for Tajima's D, while a statistically negative value was obtained for Fu's FS test (Table 7).

The graph generated from the mismatch distribution analysis revealed that *Caranx* spp. from Batangas followed a unimodal distribution pattern (Figure 5). The calculated mismatch distribution parameters (SSD and Hri) were not statistically different from the expected values for populations experiencing demographic expansion or spatial expansion (Table 8).

The values obtained for the neutrality tests, namely Tajima's D and Fu's FS, were used to infer the demographic history of each population. For the *C. ignobilis* population, the overall Tajima's D and Fu's FS values, using both cyt b and control region (Table 3) sequences, were negative; however, it was not statistically significant. On the other hand, the *C. papuensis* population had a non-statistically positive value for Tajima's D and a non-statistically negative value for Fu's FS test (Table 5). Lastly, the *C. sexfasciatus* population had a non-statistically negative value for Tajima's D, while a statistically negative value was obtained for Fu's FS test (Table 7).

The graph generated from the mismatch distribution analysis revealed that *Caranx* spp. from Batangas followed a unimodal distribution pattern (Figure 5). The

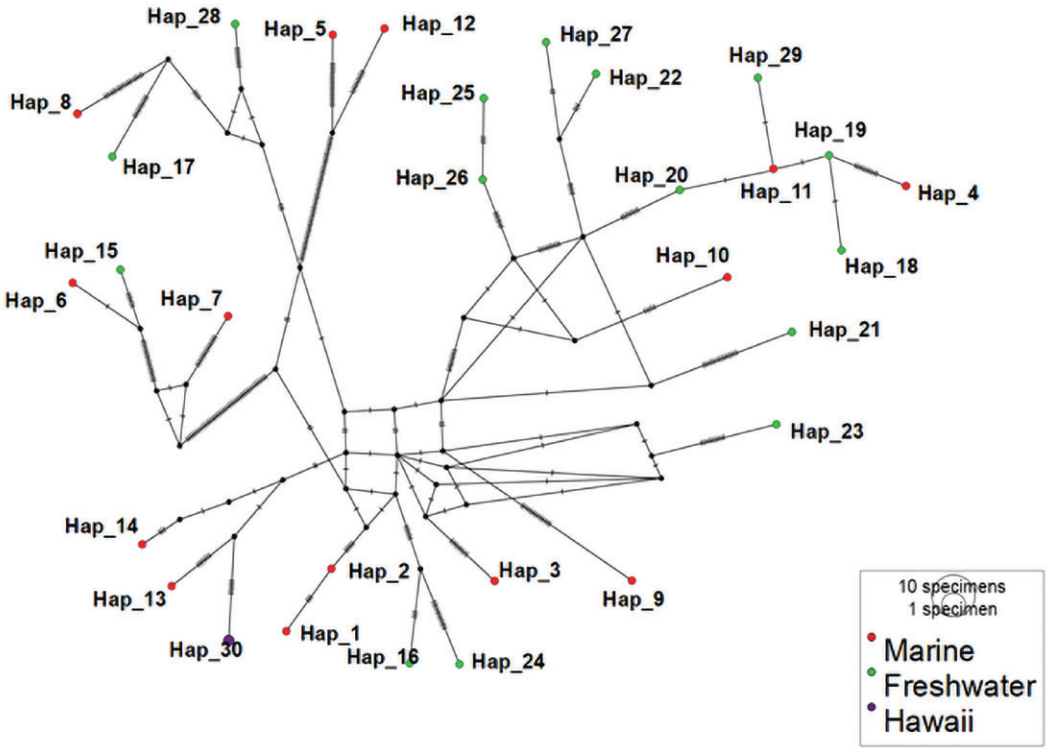


**Figure 4.** Median-joining haplotype network for control region sequences of *C. ignobilis*. Each circle represents a haplotype. The size of the circle is proportional to the number of haplotypes. The color(s) refers to the locality(ies) where the haplotype was collected. The lines on the branches represent the number of nucleotide changes while the black dot represents the theoretical median vectors introduced by the software.

**Table 3.** Neutrality test results of sequences of *C. ignobilis*. *P*-values < 0.05 are in bold print.

| Population | Cyt b | | | | Control region | | | |
|---|---|---|---|---|---|---|---|---|
| | Tajima's D | | Fu's FS | | Tajima's D | | Fu's FS | |
| | D | *P*-value | D | *P*-value | D | *P*-value | D | *P*-value |
| Marine | 0.3297 | 0.6580 | -1.8533 | 0.0940 | -0.2257 | 0.4240 | -1.6433 | 0.1270 |
| Freshwater | -1.1951 | 0.1280 | -2.0655 | 0.0640 | -1.2542 | 0.0920 | -3.0014 | 0.0620 |
| Mean | -0.4327 | 0.3930 | -1.9594 | 0.0790 | -0.7399 | 0.2580 | -2.3224 | 0.0945 |

**Table 5.** Neutrality test results of cyt b gene sequences of *C. papuensis*. *P*-values < 0.05 are in bold print.

| Population | Tajima's D | | Fu's FS | |
|---|---|---|---|---|
| | D | *P*-value | D | *P*-value |
| Marine | 1.1234 | 0.8760 | –0.6052 | 0.3700 |
| Freshwater | –0.0260 | 0.5110 | –0.6414 | 0.2670 |
| Mean | 0.5487 | 0.6935 | –0.6233 | 0.3185 |

**Table 7.** Neutrality test results of cyt b gene sequences of *C. sexfasciatus*. *P*-values < 0.05 are in bold print.

| Population | Tajima's D | | Fu's FS | |
|---|---|---|---|---|
| | D | *P*-value | D | *P*-value |
| Marine | –1.0485 | 0.1060 | –1.9379 | 0.0190 |
| Freshwater | **–1.5540** | **0.0460** | **–5.1559** | **0.0000** |
| Mean | –1.3013 | 0.0760 | **–3.5469** | **0.0095** |

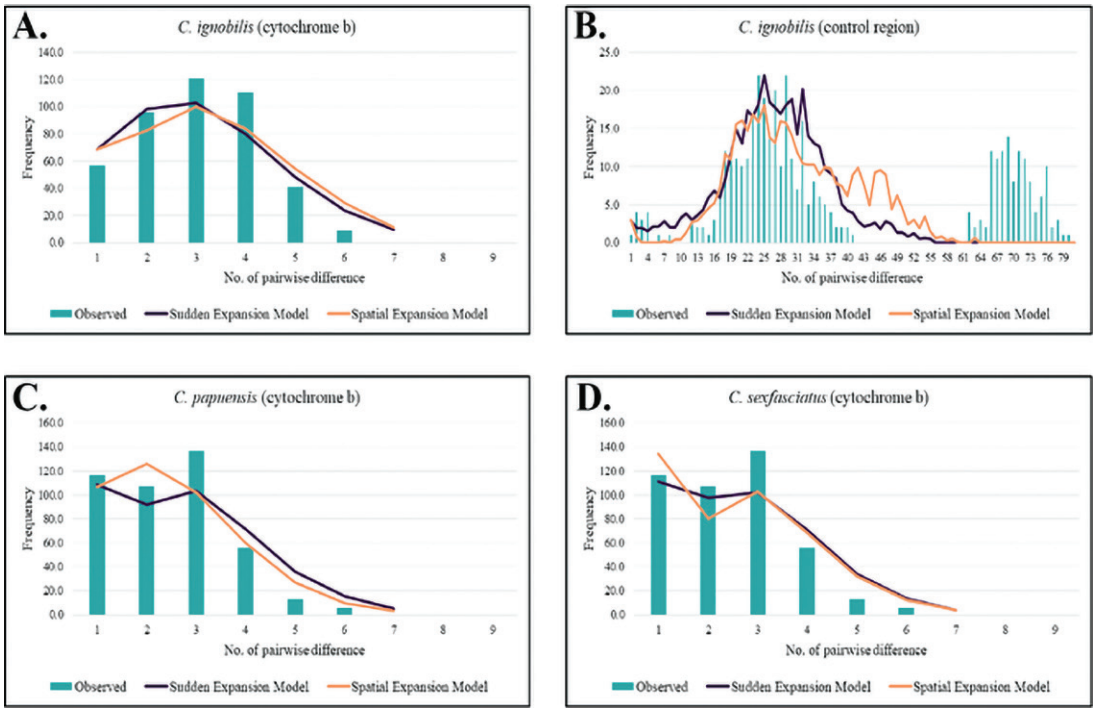


**Figure 5**. Mismatch distribution analysis plots for three species of *Caranx* spp. [legend: A – *C. ignobilis* (using cyt b); B – *C. ignobilis* (using control region); C – *C. papuensis* (using cyt b); D – *C. sexfasciatus* (using cyt b)].

**Table 8.** Mismatch distribution of sequences of *C. ignobilis, C. papuensis, and C. sexfasciatus. P*-values < 0.05 are in bold print (legend: SSD – sum of the squared differences; Hri – raggedness index)

| Population | Demographic expansion model | | | | Spatial expansion model | | | |
|---|---|---|---|---|---|---|---|---|
| | SSD | P (SSD) | Hri | P (Hri) | SSD | P (SSD) | Hri | P (Hri) |
| *C. ignobilis* (cyt b) | 0.0053 | 0.1000 | 0.0436 | 0.0700 | 0.0050 | 0.2300 | 0.0436 | 0.1900 |
| *C. ignobilis* (control region) | 0.0105 | 0.2000 | 0.0068 | 0.5000 | 0.0107 | 0.2000 | 0.0068 | 0.6000 |
| *C. papuensis* (cyt b) | 0.0100 | 0.2800 | 0.0500 | 0.3500 | 0.0238 | 0.2700 | 0.0500 | 0.6800 |
| *C. sexfasciatus* (cyt b) | 0.0100 | 0.4000 | 0.0500 | 0.4300 | 0.0077 | 0.4500 | 0.0500 | 0.6000 |

calculated mismatch distribution parameters (SSD and Hri) were not statistically different from the expected values for populations experiencing demographic expansion or spatial expansion (Table 8).

## DISCUSSION

### Genetic Diversity

Species from the genus *Caranx* are regarded to have significant economic value due to it being commonly used as a food fish, yet detailed information regarding their genetic diversity and population structure is limited (Lédée *et al*. 2015). The analysis done for *C. ignobilis* found in Batangas revealed that it has a high haplotype and nucleotide diversity. These values for the molecular diversity indices fall under the fourth category described by Grant and Bowen (1998) for observed haplotype and nucleotide diversity in marine fishes. The fourth category indicates that the population may have arisen either from a large and stable population with a long evolutionary history or from secondary contact between previously differentiated allopatric lineages. The high genetic diversity of *C. ignobilis* observed from this study was similar to the results obtained by Santos *et al*. (2010) for

*C. ignobilis* found in High Hawaii Islands, which have a high haplotype diversity (h = 0.697) and high nucleotide diversity ($\pi$ = 0.600%).

The high genetic diversity observed for *C. ignobilis* may also be explained by their biology (Martinez *et al*. 2018). *C. ignobilis* is regarded to be a migratory fish, with juveniles usually inhabiting estuaries whereas adults tend to go back to marine environments (Honebrink 2000; Daly *et al.* 2019). In the case of *C. ignobilis* found in Batangas, juveniles usually tend to travel upstream from Balayan Bay (marine environment) to Taal Lake (freshwater environment) through the Pansipit River connecting the two environments. This life-history strategy may have enabled them to introduce novel alleles into different habitats, which may be a possible mechanism for the maintenance of high genetic diversity (Martinez *et al*. 2018). Another possible mechanism would be their capacity to travel large distances (Daly *et al*. 2019), which allows them to reach more resource-filled environments that may allow them to have a higher carrying capacity, thus reducing the chance of genetic drift, which in turn can potentially increase their genetic diversity (Martinez *et al*. 2018).

On the other hand, results of the molecular diversity indices for *C. papuensis* and *C. sexfasciatus* revealed a high haplotype diversity but low nucleotide diversity. This pattern of molecular diversity values falls under the second category for observed haplotype and nucleotide diversity classifications described by Grant and Bowen (1998). This category indicates that the population may have undergone a bottleneck effect and a subsequent rapid increase in population size. For these two species, low nucleotide diversity was observed from marine specimens. The bottleneck effect is usually caused by genetic drift and targeted selection (Martinez *et al*. 2018). The observed bottleneck may be attributed to the evolution of the cyt b gene of the mtDNA, which is the molecular marker used in the study. The possible event of purifying selection that might happen in the mtDNA of these species could have caused a constraint in the evolution of their cyt b gene, affecting its nucleotide diversity (Santos *et al.* 2019). These selective pressures could be distinct to the environment where the species resides and favors a certain trait, which could imply how they could have survived in that particular environment (Consuegra *et al*. 2015). However, still, the usage of other molecular markers and an additional number of specimens should be done to further validate the results of the study.

**Population Structure**

Based on the calculated pairwise $F_{ST}$ values, the genetic differentiation between the marine (from Balayan Bay) and freshwater (from Taal Lake) specimens of *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* found in Batangas was not statistically significant. These results were further supported by the calculated molecular variance, wherein most of the variation came from within the population. The findings of the study revealed that genetic differentiation between the two different environments (marine *vs.* freshwater) was absent.

One of the possible explanations for the absence of genetic differentiation between marine and freshwater specimens of each of the three *Caranx* species used in the study would be their biology. Carangids are migratory fish that are widely distributed across the tropical and subtropical waters of the Atlantic, Indian, and Pacific regions (Honebrink 2000). Their primary motivations for migration are for breeding, as well as predation, of which they are a top predator in the aquatic system (Honebrink 2000). *C. ignobilis* adults can maintain a sustainable movement distance of up to 633 km (Daly *et al.* 2019). Similarly, *C. sexfasciatus* adults can also travel long distances up to 200 km. For *C. papuensis*, the only known range of movement for this species is based on their minimum marine reserve movement which is about 10 km (Green *et al*. 2014). In addition, carangids juveniles can tolerate a wide range of salinity, enabling them to enter estuaries (Honebrink 2000). It is possible that the Pansipit River, which connects Balayan Bay to Taal Lake and which serves as a nursery site for carangids juveniles, was not completely obstructed. Likewise, this river serves as a migratory route for these species to travel upstream towards the Taal Lake (Willette and Padin 2014). Thus, the migratory behavior and movement pattern of carangids can be a possible mechanism that promotes the mixing among the populations of *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* between marine and freshwater environments. Further studies regarding the movement and migratory pattern of carangids – especially for *C. papuensis*, which could be used to determine the population connectivity of these species – may be conducted to support the results of the study.

The homogenization of the population of *C. ignobilis*, as well as for the populations of *C. papuensis* and *C. sexfasciatus*, was further supported by the maximum likelihood tree and haplotype network. The networks do not discriminate against the marine specimens from freshwater specimens. In addition, the carangid sequences obtained from other countries did not form a separate clade, which was evident in the sharing of haplogroups obtained from the study and from the other countries in the maximum likelihood tree, which may indicate a shared ancestral haplotype. This finding might also be again attributed to the migratory behavior of carangids facilitating their expansion across the Indo-Pacific Region (Honebrink 2000). Another possible mechanism may be historical events that occurred during the Pleistocene era (~ 2.6 million yr ago) (Voris 2000). During this period, there was a simultaneous movement of glaciers

and tectonic plates, which caused significant changes to the sea levels as well as to mass configurations of land (Voris 2000). It is possible that during low sea levels, haplogroups were formed. Later on, when sea levels had risen, these haplogroups came together (Chanthran *et al*. 2020), which may be a possible mechanism for population homogenization and expansion. Next, the Tajima's D and Fu's FS values obtained for *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* were negative values, although only *C. sexfasciatus* had a statistically significant negative value. The results may indicate that there is an occurrence of an excess number of rare alleles probably due to recent population expansion (Fu and Li 1993), which further supports the results of the maximum likelihood tree and haplotype network. Lastly, the mismatch distribution analysis further supports the claim of population expansion due to the observed unimodal graph, which indicates that demographic expansion happened with the population (Ferreri *et al*. 2011).

Overall, this study was the first to report the genetic variation of *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus* populations found in Batangas. The results of the study suggest the absence of genetic differentiation between marine and freshwater specimens of *C. ignobilis*, *C. papuensis*, and *C. sexfasciatus*. Likewise, high genetic diversity was observed for *C. ignobilis*, which may indicate that the population is large and stable. In contrast, *C. papuensis* and *C. sexfasciatus* had a low nucleotide diversity despite having a high haplotype diversity, which may indicate that a population bottleneck with rapid expansion had occurred. The results of the study can serve as preliminary data for the population genetics of the three species of *Caranx* in the region, specifically before the recent 2020 Taal Volcano Eruption (12 Jan 2020) (https://www.rappler.com/newsbreak/iq/timeline-taal-volcano-eruption-2020). Although the results of the study need further validation, the data regarding the genetic pattern of species of *Caranx* found in Batangas Province generated from this study can serve as a preliminary guide for the identification and establishment of independent MUs in Balayan Bay and Taal Lake. The identification of putative MUs will help managers and conservationists formulate effective policy strategies, such as implementing quotas and size regulation during catches or establish closure of fishing during spawning periods (Palsbøll *et al*. 2006).

For future studies, it could be a good avenue of study to compare the genetic patterns of these three species of *Caranx* before and after the 2020 Taal Volcano eruption to investigate whether this volcanic event caused a significant change in the biodiversity found in Batangas Province. Likewise, it should be noted that the use of additional markers – not only from the mtDNA as used in the study but also nuclear DNA such as microsatellites – should be done to further strengthen the genetic patterns observed in this study.

## ACKNOWLEDGMENTS

This study was funded by the Natural Science Research Institute of the University of the Philippines (UP) in Diliman, Quezon City under Project Code Number BIO-19-1-07. Likewise, the authors would like to thank the Institute of Biology, College of Science, UP Diliman for the logistical support.

## STATEMENT ON CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

## REFERENCES

ALAIRA S, REBANCOS C. 2014. Maliputo (*Caranx ignobilis* Foorskal) fish cage farming practices among selected operators in Taal Lake, Batangas, Philippines. JNS 13(2): 25–40.

[BAS] Bureau of Agricultural Statistics. 2012. Fisheries Statistics of the Philippines 2009–2011. Department of Agriculture, Quezon City, Philippines. 421p.

[BAS] Bureau of Agricultural Statistics. 2019. Fisheries Statistics of the Philippines 2016–2018. Department of Agriculture, Quezon City, Philippines. 321p.

CHANTHRAN SSD, LIM P-E, LI Y, LIAO T-Y, POONG S-W, DU J, HUSSEIN MAS, SADE A, RUMPET R, LOH K-H. 2020. Genetic diversity and population structure of *Terapon jarbua* (Forskål, 1775) (Teleostei, Terapontidae) in Malaysian waters. ZooKeys 911: 139–160. doi:10.3897/zookeys.911.39222

CHAVEZ CJM, DE-ACOB JEB, DUEÑAS BJD, MACALALAD MMC. 2013. Native specialty foods of selected restaurants in Batangas City [BS Thesis]. Lyceum of the Philippines, Batangas City, Philippines. 7p. Retrieved from https://research.lpubatangas.edu.ph/wp-content/uploads/2014/08/JTHR-Native-Specialty-Foods.pdf

CONSUEGRA S, JOHN E, VERSPOOR E, DE-LEANIZ CG. 2015. Patterns of natural selection acting on the mitochondrial genome of a locally adapted fish species. Genet Sel Evol 47(1). doi:10.1186/s12711-015-0138-0

DALY R, FILMALTER JD, DALY CAK, BENNETT RH, PEREIRA MAM, MANN BQ, DUNLOP SW, COWLEY PD. 2019. Acoustic telemetry reveals multi-seasonal spatiotemporal dynamics of a giant trevally *Caranx ignobilis* aggregation. Mar Ecol Prog Ser 621: 185–197. doi:10.3354/meps12975

DARRIBA D, TABOADA GL, DOALLO R, POSADA D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9(8): 772. doi:10.1038/nmeth.2109

EXCOFFIER L, LISCHER HE. 2010. Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10(3): 564–567. doi:10.1111/j.1755-0998.2010.02847.x

[FAO] Food and Agriculture Organization. 2020. The state of world fisheries and aquaculture – sustainability in action. Rome. 20p. doi:https://doi.org/10.4060/ca9229en

FERRERI M, QU W, HAN B. 2011. Phylogenetic networks: a tool to display character conflict and demographic history. Afr J Biotechnol 10(6): 12799–12803. doi:10.5897/AJB11.010

FU YX, LI WH. 1993. Statistical tests of neutrality of mutations. Genetics 133(3): 693–709.

GALTIER N, NABHOLZ B, GLEMIN S, HURST GDD. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Mol Ecol 18(22): 4541–4550. doi:10.1111/j.1365-294x.2009.04380.x

GANADEN SR, LAVAPIE-GONZALES F. 1999. Common and local names of marine fishes of the Philippines. Bureau of Fisheries and Aquatic Resources, Quezon City, Philippines. p. 1–385.

GRANT W, BOWEN BW. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. J Hered 89(5): 415–426. doi:10.1093/jhered/89.5.415

GREEN AL, MAYPA AP, ALMANY GR, RHODES KL, WEEKS R, ABESAMIS RA, GLEASON MG, MUMBY PJ, WHITE AT. 2014. Larval dispersal and movement patterns of coral reef fishes, and implications for marine reserve network design. Biol Rev. p. 1–34. doi:10.1111/brv.12155

HERRE AW. 1953. Checklist of Philippine Fishes. Research Report: US Fish and Wildlife Service. 271p.

HONEBRINK RR. 2000. A Review of the Biology of the Family Carangidae, with Emphasis on Species Found in Hawaiian Waters. Technical Report 20–01 Honolulu, Hawaii: Department of Land and Natural Resources. 43p.

LÉDÉE EJ, HEUPEL MR, TOBIN AJ, SIMPFENDORFER CA. 2015. Movements and space use of giant trevally in coral reef habitats and the importance of environmental drivers. Anim Biotelemetry 3(6): 1–14. doi:10.1186/s40317-015-0024-0

LEIGH JW, BRYANT D. 2015. PopART: full-feature software for haplotype network construction. Methods Ecol Evol 6(9): 1110–1116. doi:10.1111/2041-210X.12410

LIBRADO P, ROZAS J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452. doi:10.1093/bioinformatics/btp187

MANSOR MI, KOHNO H, IDA H, NAKAMURA HT, AZNAN Z, ABDULLAH S. 1998. Field Guide to Important Commercial Marine Fishes of the South China Sea. Malaysia: Marine Fishery Resources Development and Management Department, Southeast Asian Fisheries Development Center. 287p.

MARTINEZ AS, WILLOUGHBY JR, CHRISTIE MR. 2018. Genetic diversity in fishes is influenced by habitat type and life-history variation. Ecol Evol. p. 1–10. doi:10.1002/ece3.4661

MUTIA MTM, MUYOT FB, MAGISTRADO ML. 2015. Induced breeding of giant trevally, maliputo (*Caranx ignobilis*). In: Proceedings of the International Workshop on Resource Enhancement and Sustainable Aquaculture Practices in Southeast Asia 2014 (RESA). Romana-Eguia MRR, Parado-Estepa FD, Salayo ND, Lebata-Ramos MJH eds. Aquaculture Department, Southeast Asian Fisheries Development Center, Tigbauan, Iloilo, Philippines. 311p.

PALSBØLL, PJ, BERUBE M, ALLENDORF FW. 2007. Identification of management units using population genetic data. Trends Ecol Evol 22(1): 11–16. doi:10.1016/j.tree.2006.09.003

RAMOS E. 2010. Origin and geologic features of Taal Lake, Philippines. Aquat Ecosyst Health Manag 2: 155–162. doi:10.1080/14634980290031794

REISS H, HOARAU G, DICKEY-COLLAS M, WOLFF WJ. 2009. Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish Fish 10: 361–395. doi:10.1111/j.1467-2979.2008.00324.x

SANTOS SR, XIANG Y, TAGAWA, AW. 2010. Population structure and comparative phylogeography of jack species (*Caranx ignobilis* and *C. melampygus*) in the High Hawaiian Islands. J Hered 102(1): 47–54. doi:10.1093/jhered/esq101

SANTOS BS, BASIAO ZU, QUILANG JP. 2019. Genetic diversity and patterns of demographic expansion in natural populations of milkfish, *Chanos chanos* (Forsskål, 1775), in the Philippines. Mitochondr DNA. p. 1–13. doi:10.1080/24701394.2018.1504931

SMITH-VANIZ WF. 1999. Carangidae. Jacks and scads (also trevallies, queenfishes, runners, amberjacks, pilotfishes, pampanos, *etc.*). In: FAO Species Identification Guide for Fishery Purposes: The Living Marine Resources of The Western Central Pacific, Vol. 4. Carpenter KE, Niem VH eds. FAO, Rome. p. 2659–2756.

THRUSH SF, DAYTON PK. 2010. What can ecology contribute to ecosystem-based management? Annu Rev Mar Sci 2: 419–441. doi:10.1146/annurev-marine-120308-081129

TORRES SKM, SANTOS BS. 2020. Species identification among morphologically-similar *Caranx* species. Turk J Fish Aquat Sc 20(2): 159–169. doi:10.4194/1303-2712-v20_2_08

VILLADOLID DV. 1937. The fisheries of Lake Taal, Pansipit River, and Balayan Bay, Batangas Province, Luzon. Philipp J Science 63(2): 191–228.

VORIS HK. 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. J Biogeogr 27(5): 1153–1167. doi:10.1046/j.1365-2699.2000.00489.x

WILLETTE DA, PADIN JIM. 2014. Identifying the biodiversity of marine jacks (Carangidae) in the freshwater Taal Lake, Philippines using phenotypic features and mitochondrial DNA. J Appl Ichthyol 30(3): 490–495. doi:10.1111/jai.12411