Molecular barcoding of marine ornamental fish from the southern coast of West Java validates conventional identification

Agus Nuryanto1,*; Kusbiyanto Kusbiyanto1; Dian Bhagawati1

1Biology Faculty, Jenderal Soedirman University, Purwokerto 53122, Central Java, Indonesia

Abstract. Conventional identification of marine ornamental fish has faced difficulties due to similar color patterns of closed related species, or juvenile individuals have different color patterns from adult individuals. Molecular barcoding using the cytochrome c oxidase I (COI) gene provides a reliable tool for unmasking such difficulties. This study aimed to barcode marine ornamental fish from the southern coast of West Java. Fragment of the COI gene was sequenced from 54 morphotypes. In this study, we determined the taxonomic status of the samples based on a 5% genetic divergence, with the parameter including sequence percent identity, genetic distance, and length of monophyletic branch in a phylogenetic tree. The result showed that most samples had a high percentage of sequence identities, low genetic distances, and short chapters in monophyletic clades, but the remaining were not. Those data indicated that most samples could be identified at species-level without doubt and support conventional identification. Barcoding success is also depending on the availability of conspecific sequences in the databases. This study concluded that molecular barcoding could strengthen and validate traditional identification.

1 Introduction

Indonesian coral reef supports consumptive and non-consumptive fish species. Ornamental fish is a non-consumptive fish group that is utilized for recreation. This fish group is in high demand because of its beautiful colors and color pattern, both in juvenile and adult individuals. Ornamental fish has a broad market from national to international trading [1].

Trading of these wildlife commodities in Indonesia has been started since the 1990s either local or international trade. Many publications have reviewed the ornamental marine fish from Indonesia. However, mainly on trading values and data were collected from prominent exporters [1-3]. The study focused on species diversity of marine ornamental fish on particular sites where the commodities are collected relatively rare, especially on the southern coast of West Java. Data on marine ornamental fish production at the south coast of West Java were also not available.

* Corresponding author: agus.nuryanto@unsoed.ac.id

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
Two recent studies reported marine ornamental fish from the southern coast of West Java, which collected ornamental fish from Pangandaran [4] and Pelabuhan Ratu, Ujung Genteng, and Taman Manalusu [5]. In both studies, the researchers proved that the high species diversity of marine ornamental fish is involved in the aquarium trade on the southern coast of West Java.

In particular of ornamental fish groups, species identification mainly relied on morphological characters, such as color pattern faced difficulties and might lead to misidentification. On the one hand, closed related fish species might show only subtle morphological differences [6]. Different fish species might show similar colors and patterns in marine ornamental fish during the juvenile stage [7]. On the other hand, different life stages of ornamental fish offer different color and color patterns, such as *Pomacanthus semicircularis* [8].

In addition to morphological characters, this report utilized molecular characters for species identification of marine ornamental fish from the southern coast of West Java. This study used the cytochrome c oxidase 1 (COI) gene as a barcode marker. The COI gene has been a reliable technique for species-level identification [9, 10]. Some exceptions in some fish groups, COI barcodes could not differentiate closely related species [11]. Moreover, studies demonstrated that COI barcoding could reveal that cryptic species are also abundant [12, 13]. Other studies also proved that COI barcoding was strengthened and validated morphological identification [6].

The researchers utilized variable sequence homology values during species delimitation. A minimum sequence homology of 97% or 3% sequence divergences is used for species delimitation in Boldsystems [14]. A similar value was also used by previous studies [15, 16] Ward et al. (2009) and Amatya (2019). Other researchers used a minimum of 98% sequence homology as species threshold. However, low genetic homology (below 95%) was observed when the reference species came from different localities [17], while other studies used 99% homology for species determination [18]. At the same time, many studies also reported that intraspecific genetic distances in fish were wildly variable among species ranging from 0.0 to higher than 0.05 [19, 20, 21, 22, 23]. Higher genetic distance among species was reported when considering the geographic localities of the samples [24]. Another study said that an overlap genetic distance is observed between intra- and interspecific individuals [25].

This study aimed to identify marine ornamental fish collected on the southern coast of West Java based on cytochrome c oxidase one gene barcoding to validate morphological identification.

### 2 Materials and methods

#### 2.1 Sampling sites and times

A total of 367 ornamental fish samples were bought from the first collector in Pelabuhan Ratu and Ujung Genteng, Sukabumi Regency, Taman Manalusu Garut Regency, and Bojongsalawe Village, District of Parigi, Pangandaran Regency (Figure 1). Ornamental fish samples were collected during the field trips in 2018 and 2019.
Two recent studies reported marine ornamental fish from the southern coast of West Java, which collected ornamental fish from Panganda ran [4] and Pelabuhan Ratu, Ujung Genteng, and Taman Manalusu [5]. In both studies, the researchers proved that the high species diversity of marine ornamental fish is involved in the aquarium trade on the southern coast of West Java.

In particular of ornamental fish groups, species identification mainly relied on morphological characters, such as color pattern faced difficulties and might lead to misidentification. On the one hand, closed related fish species might show only subtle morphological differences [6]. Different fish species might show similar colors and patterns in marine ornamental fish during the juvenile stage [7]. On the other hand, different life stages of ornamental fish offer different color and color patterns, such as Pomacanthus semicircularis [8].

In addition to morphological characters, this report utilized molecular characters for species identification of marine ornamental fish from the southern coast of West Java. This study used the cytochrome c oxidase 1 (COI) gene as a barcode marker. The COI gene has been a reliable technique for species-level identification [9, 10]. Some exceptions in some fish groups, COI barcodes could not differentiate closely related species [11]. Moreover, studies demonstrated that COI barcoding could reveal that cryptic species are also abundant [12, 13]. Other studies also proved that COI barcoding was strengthened and validated morphological identification [6].

The researchers utilized variable sequence homology values during species delimitation. A minimum sequence homology of 97% or 3% sequence divergences is used for species delimitation in Boldsystems [14]. A similar value was also used by previous studies [15, 16] Ward et al. (2009) and Amatya (2019). Other researchers used a minimum of 98% sequence homology as species threshold. However, low genetic homology (below 95%) was observed when the reference species came from different localities [17], while other studies used 99% homology for species determination [18]. At the same time, many studies also reported that intraspecific genetic distances in fish were widely variable among species ranging from 0.0 to higher than 0.05 [19, 20, 21, 22, 23]. Higher genetic distance among species was reported when considering the geographic localities of the samples [24]. Another study said that an overlap genetic distance is observed between intra- and interspecific individuals [25].

This study aimed to identify marine ornamental fish collected on the southern coast of West Java based on cytochrome c oxidase one gene barcoding to validate morphological identification.

2.2 Marker amplification and sequencing

Molecular barcoding was carried out on 54 morphotypes identified morphologically. However, the results were questionable due to overlapping characters between closely related species. The genomic DNA of the samples was isolated from caudal fin clips using Chelex®100 methods [26] with slight modification [27]. The selected marker was amplified using primers FishF2 and FishR2 [28]. Reagent composition was as follow; 10X PCR buffer 5 µl, MgCl₂ (50 mM) 5 µl, 2 µl (0.01 mM) of each primer, 2 µl dNTPs (0.05 mM), 1 U Taq polymerase, and 4 µl of template DNA. Adjusted finale volumes to 50 µl were obtained by adding RNAse-DNAse free water.

The marker was multiplied using the following thermal cycles. Pre-denaturation was performed at 95°C for 5 minutes and continued by 35 cycles with the following conditions. The denaturation process was conducted at 94°C for 1 minute, annealing at a temperature range from 53°C to 55°C depending on the suspected species, and extension steps at 72°C for 1.5 minutes. We conducted the final extension for 5 minutes at 72°C.

Half of the fish samples were treated as follows to obtain sequences data. The genomic DNA was isolated using ZR Tissue and Insect DNA Miniprep Kit (Zymo Research, D6016) following the protocol from the manufacturer. The PCR amplification of the selected COI marker was performed using the MyTaq HS Red Mix (Bioline, BIO-25047),
while the sequencing of the COI gene was used in the bi-directional sequencing technique. All procedures of DNA analysis were conducted at Genetika Laboratory (PT. Genetika Science Indonesia).

2.3 Sequences editing and species determination

All the sequences were subjected to manual editing, and trimming using Bioedit 7.0 software packed [29]. With manual checking, pairwise multiple sequences alignment was conducted using ClustalW as applied in Bioedit 7.0 software packages [29]. The marker's confidence level as the actual COI sequence obtained from the translation process to the amino acid using the ORF Finder online version (https://www.ncbi.nlm.nih.gov/orffinder/). This study rechecked the translation results through the blast process with the formatting option search parameters plus the CDS feature. This process was carried out to ensure no stop codon in the middle of the COI gene base sequence is obtained.

We determined the taxonomic status of each morphotype based on the sequence identity or similarity value of 95%. The present study chose that value based on a consideration that species could have other sequences divergences within species [6] and geographic locality between the current samples and the references species [17] available in the barcode library (GenBank and Boldsystems). The Kimura 2-parameter genetic distance of 0.05 was selected as additional data for species-level identification. Support to those values was also obtained from the phylogenetic tree, reconstructed based on 527 base pair (bp) sequences. The tree was constructed using the Neighbor-joining method based on the Kimura 2-parameter substitution model. We obtained branching polarity from 1000 bootstraps pseudoreplication. Genetic distances calculation and tree reconstruction were performed in MEGAX [30]. Short branches in the monophyletic clade (maximum scale 0.05) were referred to as a single species. This study compared molecular barcoding with previous studies, which identified marine ornamental fish from the same sites but based on morphology [4, 5]. That step was conducted to check the validity of morphological identification.

3 Results and discussion

3.1 Results

3.1.1 Taxonomic status

Sequence identity test using essential local alignment search tool (BLAST) to the references species available in GenBank resulted in identity values ranging from 94.65% to 100%. This study also rechecked sequence similarities of the samples to conspecific references in Boldsystems. The current study obtained the lowest identity value of 94.65% for the sequence of WJM5. A detail data on sequence identity values and genetic distances between samples and their references species are presented in Table 1.
While the sequencing of the COI gene was used in the bi-directional sequencing technique. All procedures of DNA analysis were conducted at Genetika Laboratory (PT. Genetika Science Indonesia).

2.3 Sequences editing and species determination

All the sequences were subjected to manual editing and trimming using Bioedit 7.0 software packed [29]. With manual checking, pairwise multiple sequences alignment was conducted using ClustalW as applied in Bioedit 7.0 software packages [29]. The marker’s confidence level as the actual COI sequence obtained from the translation process to the amino acid using the ORF Finder online version (https://www.ncbi.nlm.nih.gov/orffinder/).

This study rechecked the translation results through the blast process with the formatting option search parameters plus the CDS feature. This process was carried out to ensure no stop codon in the middle of the COI gene base sequence is obtained.

We determined the taxonomic status of each morphotype based on the sequence identity or similarity value of 95%. The present study chose that value based on a consideration that species could have other sequences divergences within species [6] and geographic locality between the current samples and the references species [17] available in the barcode library (GenBank and Boldsystems). The Kimura 2-parameter genetic distance of 0.05 was selected as additional data for species-level identification. Support to those values was also obtained from the phylogenetic tree, reconstructed based on 527 base pair (bp) sequences. The tree was constructed using the Neighbor-joining method based on the Kimura 2-parameter substitution model. We obtained branching polarity from 1000 bootstraps pseudoreplication. Genetic distances calculation and tree reconstruction were performed in MEGAX [30].

Short branches in the monophyletic clad (maximum scale 0.05) were referred to as a single species. This study compared molecular barcoding with previous studies, which identified marine ornamental fish from the same sites but based on morphology [4, 5]. That step was conducted to check the validity of morphological identification.

3 Results and discussion

3.1 Results

3.1.1 Taxonomic status

Sequence identity test using essential local alignment search tool (BLAST) to the references species available in GenBank resulted in identity values ranging from 94.65% to 100%. This study also rechecked sequence similarities of the samples to conspecific references in Boldsystems. The current study obtained the lowest identity value of 94.65% for the sequence of WJM5. A detail data on sequence identity values and genetic distances between samples and their references species are presented in Table 1.

| Samples   | Accession number | Identity (%) | Genetic distance | Conspecific references                  |
|-----------|------------------|--------------|------------------|----------------------------------------|
| WJ01      | MK041042         | 99.19        | 0.009            | Myripristis hexagona QJ350118          |
| WJ02      | MK041043         | 98.72        | 0.014            | Plectorhinchus picus FJ583866          |
| WJ3Kc     | MK256660         | 95.67        | 0.044            | Chaetodon vagabundus KJ967962          |
| WJ04      | MK041044         | 100          | 0.000            | Hippocampus kuda GQ502154              |
| WJ05      | MK256661         | 96.41        | 0.037            | Arothron hispidus JQ431462             |
| WJ05GT    | MK246805         | 96.70        | 0.032            | Acentrogobius nebulosus MK962523       |
| WJ06      | MK041045         | 100          | 0.000            | Chaetodon kleinii MW034078              |
| WJ07      | MK041046         | 99.16        | 0.008            | Chaetodon auriga MF123777               |
| WJ08      | MK256662         | 99.52        | 0.005            | Dendrochirus zebra FJ583352             |
| WJ08KKT   | MK246806         | 98.02        | 0.018            | Chaetodon vagabundus JF434839           |
| WJ09_01   | MK041047         | 100          | 0.000            | Chelmon rostratus FJ583127              |
| WJ09      | MK256663         | 96.53        | 0.035            | Balistapus undulatus MN560967           |
| WJ10      | MK246812         | 97.59        | 0.021            | Centropyge eibli KT001113               |
| WJ10_1    | MK041048         | 99.35        | 0.006            | Dendrochirus zebra KP929813             |
| WJ11DA    | MK246807         | 96.73        | 0.033            | Dascyllus trimaculatus MF409512         |
| WJ13      | MK041049         | 99.84        | 0.001            | Pterois miles KU317873                  |
| WJ14_1    | MK041050         | 100          | 0.004            | Pomacanthus semicirculatus FJ583886     |
| WJ14      | MK246808         | 98.08        | 0.019            | Strophidon satheke MT318376             |
| WJ15      | MK041051         | 99.69        | 0.004            | Chaetodon collare KC626015              |
| WJ18      | MK256664         | 99.19        | 0.008            | Terapon jarbua FJ347886                 |
| WJM243    | MK246809         | 97.88        | 0.021            | Blenniella peripheralthalmus MF409604   |
| WJM342    | MK246810         | 96.52        | 0.035            | Chaetodon decussatus GU673801           |
| WJUG4     | MK246811         | 98.53        | 0.014            | Zanclus cornutus AP009162               |
| WJM1      | MK256665         | 95.65        | 0.045            | Neoglyphidodon boning FOAN677-11.COI-5P |
| WJM4      | MK256666         | 98.69        | 0.013            | Naso unicornis QJ350128                 |
| WJM5      | MK256667         | 94.65        | 0.055            | Lutjanus decussatus KF009608            |
| WJMG1     | MK256668         | 98.21        | 0.018            | Plectroglyphidodon lacrymatus KP194879   |
| WJMG2     | MK256669         | 98.87        | 0.011            | Chaetodon rafflesii FJ583077            |
| WJPR1     | MK256670         | 98.24        | 0.017            | Epinephelus merra MF185539              |
| WJPR2     | MK256671         | 99.38        | 0.006            | Siganus guttatus KJ013064               |
| Samples | Accession number | Identity (%) | Genetic distance | Conspecific references |
|---------|-----------------|--------------|------------------|------------------------|
| WJPR3  | MK256672        | 97.98        | 0.020            | Sufflamen bursa MK657647 |
| WJPR5  | MK256673        | 99.84        | 0.001            | Lutjanus fulvus KF009613 |
| WJUG1  | MK256674        | 98.71        | 0.013            | Stethojulis trilineata JN313092 |
| WJUG2  | MK256675        | 99.37        | 0.006            | Siganus spinus          |
| WJUG3  | MK256676        | 98.44        | 0.015            | Naso lituratus KC970406 |
| PGN013 | MT881550        | 99.36        | 0.006            | Centropyge eibli KT001113 |
| PGN014 | MT881551        | 100          | 0.000            | Acanthurus bariene KF009560 |
| PGN015 | MT881552        | 100          | 0.000            | Pseudobalistes flavimarginatus MW034195 and MH331840 |
| PGN021 | MT881553        | 99.84        | 0.001            | Chaetodon collare KX000917 |
| PGN024 | MT881554        | 100          | 0.000            | Thalassoma lunare KF715032 |
| PGN025 | MT881555        | 100          | 0.000            | Platax orbicularis MF123985 |
| PGN_028| MT881556        | 100          | 0.000            | Chaetodon lunula KF194718 |
| PGN_030| MT881557        | 99.84        | 0.001            | Ostracion cubicus JQ861019 |
| PGN_702| MT881558        | 100          | 0.000            | Sargocentron diadema MF409594 |
| PGN_705| MT881559        | 100          | 0.000            | Abudelfdu vaigiensis JF434721 |
| PGN_707| MT881560        | 99.84        | 0.001            | Chaetodon ephippium MN733557 |
| PGN_715| MT881561        | 100          | 0.000            | Balistoides viridescens KF025675 and JQ431476 |
| PGN_718| MT881562        | 99.84        | 0.001            | Sargocentron caudimaculatum HM034164 |
| PGN_719| MT881563        | 99.84        | 0.000            | Pterois miles KU317873 |
| PGN_728| MT881564        | 99.84        | 0.000            | Ostorhinchus novemfasciatus FJ459573 and FADLI017-17.COI-5P |
| PGN_729| MT881565        | 99.35        | 0.000            | Scorpaenodes guamensis KU893076 |
| PGN_819| MT881566        | 99.84        | 0.000            | Pterois miles KU317873 |
| PGN_828| MT881567        | 99.84        | 0.000            | Ostorhinchus novemfasciatus FJ459573 and FADLI017-17.COI-5P |
| PGN_924| MT881568        | 99.84        | 0.001            | Thalassoma lunare KF715032 |

The pairwise Kimura-2parameter (K2P) comparisons indicated that the samples had genetic distances between 0.000 and 0.055 (Table 1). We found the highest genetic distance of 0.055 between morphotype WJM5 and its references species Lutjanus decussatus.
The pairwise Kimura-2-parameter (K2P) comparisons indicated that the samples had genetic distances between 0.000 and 0.055 (Table 1). We found the highest genetic distance of 0.055 between morphotype WJM5 and its references species *Lutjanus decussatus*.

Fig. 1. Phylogenetic tree showing monophyly between samples and its references species.
Neighbor-joining (NJ) phylogenetic tree reconstructed based on the K2P model indicated that most samples formed a monophyletic clade with their references species. Most representatives formed a clade with a short branch length to the references species except between WJM5 and L. decussatus. Almost all clade had branch lengths lower than 0.05 scales, and only WJM5 and L. decussatus clade had branch lengths higher than the 0.05 scale. The phylogenetic tree is presented in Figure 2.

3.1.2 Molecular barcoding versus morphological identification

Comparison to previous studies [2, 5] proved that 51 out of 54 (94.44%) morphotypes resulted in similar taxonomic status between molecular barcoding and conventional identification based on morphological characters. The remaining three morphotypes (5.6%) were different between molecular and traditional identification. Complete data on the comparison between molecular barcoding and conventional identification is presented in Table 2.

Table 2. Molecular barcoding versus conventional identification

| Samples  | Molecular barcoding     | Conventional identification |
|----------|-------------------------|-----------------------------|
| WJ01     | Myripristis hexagona    | Myripristis hexagona        |
| WJ02     | Plectorhinchus picus     | Plectorhinchus picus        |
| WJ3Kc    | Chaetodon vagabundus     | Chaetodon vagabundus        |
| WJ04     | Hippocampus kuda         | Hippocampus kuda            |
| WJ05     | Arothron hispidus        | Arothron hispidus           |
| WJ5GT    | Acentrogobius nebulosus  | Acentrogobius nebulosus     |
| WJ06     | Chaetodon kleinii        | Chaetodon kleinii           |
| WJ07     | Chaetodon auriga         | Chaetodon auriga            |
| WJ08     | Dendrochirus zebra       | Dendrochirus zebra          |
| WJ08KKT  | Chaetodon vagabundus     | Chaetodon vagabundus        |
| WJ09     | Chelmon rostratus        | Chelmon rostratus           |
| WJ09_01  | Balistapus undulatus     | Balistapus undulatus        |
| WJ10     | Centropyge eibli         | Centropyge eibli            |
| WJ10_1   | Dendrochirus zebra       | Dendrochirus zebra          |
| WJ11DA   | Dascyllus trimaculatus   | Dascyllus trimaculatus      |
| WJ13     | Pterois miles            | Pterois miles               |
| WJ14_1   | Pomacanthus semicirculatus | Pomacanthus semicirculatus |
| WJ14     | Strophidon sathe         | Strophidon sathe            |
| WJ15     | Chaetodon collare        | Chaetodon collare           |
| WJ18     | Terapon jarbua           | Terapon jarbua              |
| WJM243   | Blenniella periphthalmus  | Blenniella periphthalmus    |
| WJM342   | Chaetodon decussatus     | Chaetodon decussatus        |
Neighbor-joining (NJ) phylogenetic tree reconstructed based on the K2P model indicated that most samples formed a monophyletic clade with their reference species. Most representatives formed a clade with a short branch length to the reference species except between WJM5 and L. decussatus. Almost all clades had branch lengths lower than 0.05 scales, and only WJM5 and L. decussatus clade had branch lengths higher than the 0.05 scale. The phylogenetic tree is presented in Figure 2.

3.1.2 Molecular barcoding versus morphological identification

Comparison to previous studies [2, 5] proved that 51 out of 54 (94.44%) morphotypes resulted in similar taxonomic status between molecular barcoding and conventional identification based on morphological characters. The remaining three morphotypes (5.6%) were different between molecular and traditional identification. Complete data on the comparison between molecular barcoding and conventional identification is presented in Table 2.

| Samples  | Molecular barcoding  | Conventional identification |
|----------|----------------------|-----------------------------|
| WJUG4    | Zanclus cornutus      | Zanclus cornutus            |
| WJM1     | Neoglyphidodon bonang| Neoglyphidodon bonang       |
| WJM4     | Naso unicornis       | Naso brevirostris           |
| WJM5     | Lutjanus decussatus  | Lutjanus decussatus         |
| WJMGG1   | Plectroglyphidodon lacrymatus | Plectroglyphidodon lacrymatus |
| WJMGG2   | Chaetodon rafflesii   | Chaetodon rafflesii         |
| WJPR1    | Epinephelus merra    | Epinephelus merra           |
| WJPR2    | Siganus guttatus     | Siganus guttatus            |
| WJPR3    | Sufflamen bursa      | Sufflamen bursa             |
| WJPR5    | Lutjanus fulvus      | Lutjanus bohar              |
| WJUG1    | Stethojulis trilineata| Stethojulis trilineata      |
| WJUG2    | Siganus spinus       | Siganus spinus              |
| WJUG3    | Naso lituratus       | Naso lituratus              |
| PGN013   | Centropyge eibli     | Centropyge eibli            |
| PGN014   | Acanthurus bariene   | Acanthurus maculiceps       |
| PGN015   | Pseudobalistes flavimarginatus | Pseudobalistes flavimarginatus |
| PGN021   | Chaetodon collare    | Chaetodon collare           |
| PGN024   | Thalassoma lunare    | Thalassoma lunare           |
| PGN025   | Platax orbicularis   | Platax orbicularis          |
| PGN_028  | Chaetodon lunula     | Chaetodon lunula            |
| PGN_030  | Ostracion cubicus    | Ostracion cubicus           |
| PGN_702  | Sargocentron diadema | Sargocentron diadema        |
| PGN_705  | Abudefduf vaigiensis | Abudefduf vaigiensis        |
| PGN_707  | Chaetodon ephippium  | Chaetodon ephippium         |
| PGN_715  | Balistoides viridescens | Pseudobalistes flavimarginatus |
| PGN_718  | Sargocentron caudimaculatum | Sargocentron caudimaculatum |
| PGN_719  | Pterois miles        | Pterois miles               |
| PGN_728  | Ostorhinchus novemfasciatus | Ostorhinchus novemfasciatus |
| PGN_729  | Scorpaenodes guamensis | Scorpaenodes guamensis      |
| PGN_819  | Pterois miles        | Pterois miles               |
| PGN_828  | Ostorhinchus novemfasciatus | Ostorhinchus novemfasciatus |
| PGN_924  | Thalassoma lunare    | Thalassoma lunare           |
3.2 Discussion

3.2.1 Taxonomic status

Fifty-three morphotypes had identity values above 95% to their conspecific references, with genetic distances below 0.05 (Table 1). Those morphotypes also formed monophyletic clades with branch lengths less than 0.05 to their conspecific references (Figure 2). Those three data (sequence identity, genetic distance, and branch length on monophyletic clade) proved that those 53 species could be assigned to species level. The assignment to the species level is defined according to the barcoding gap used in species determination is 5% genetic divergence, which means 95% genetic similarity between query samples with conspecific references. Several studies reported that 95% could be used for species-level barcoding [17, 22-23]. The use of 3% to 5% genetic divergences must be added by other data [31], including geographic localities [32]. This study utilized the geographic localities of the samples and reference species as additional considerations for species determination.

There were exciting findings that two morphotypes had high sequence identities to two different references species. The PGN015 has 100% to *Pseudobalistes flavimarginatus* and has 99.84% sequence identity to *Balistoides viridescens*. In contrast, PGN715 has an identity value of 100 to *B. viridescens* and 99.84% to *Pseudobalistes flavimarginatus*. In such a case, this study used the highest homology and the lowest genetic distances even though they formed a monophyletic clade with 0.00 branch length in the phylogenetic tree. Therefore, PGN015 and PGN715 were taxonomically referred to as *P. flavimarginatus* and *B. viridescens*, respectively. This situation was not surprising because previous studies also reported a similar condition in other fish groups [11]. They found a high homology value of *Mystus vittatus* sample to *M. vittatus* and *M. horai* in barcode databases (99% to each reference species, respectively). A similar high homology value was reported for *Bagarius bagarius* samples to *B. vagaries* and *B. yarrelli* in the databases, with homology values of 100% to both species, respectively [11].

Morphotype WJM5 had a sequence identity of 94.65% to 13 sequences of *L. decussatus* in GenBank and more than 50 sequences of *L. decussatus* in Boldsystems, genetic distance 0.055, branch length was longer than 0.05. The morphotype had sequence identity top hits to *L. decussatus*. However, because the used genetic gap was 95% sequences similarity and genetic gap 0.05, the morphotype WJM5 was referred to genus level *Lutjanus* and *Lutjanus* sp.

3.2.2 Molecular barcoding versus morphological identification

Based on the data in Table 2, the result of molecular identification was highly congruent (94.44%) to conventional identification [4, 5]. High Congruent between molecular and morphological identification was reported in a previous study with success between 90% and 99% [33]. Congruent between morphological and molecular identification was also reported in mosquitoes [34].

In the case of WJM5, although it has 0.055 (higher than 0.05) genetic distance and genetic identity lower than 95% (94.65%), the nearest relative in barcode libraries (GenBank, 13 individuals, and Boldsystems, > 50 individuals) were *L. decussatus*. The result was congruence with conventional identification. It is reasonable that WJM5 had higher genetic divergence to its nearest relative in barcode libraries because the researcher collected from different geographic regions or even different oceans. This study collected samples from the southern coast of West Java (East Indian Ocean). In contrast, the previous researcher collected conspecific reference *L. decussatus* (KF009608) previously published in GenBank from the Philippines (Pacific Ocean). Combining both molecular and
conventional identification for WJM5, we finally decided that WJM5 was referred to as *L. decussatus*. The argument was that samples of single species collected from different localities could have a relatively low genetic identity and high genetic distance to their conspecific references in barcode libraries [6, 17, 21-23, 31-32].

4 Conclusions

This study highlighted that under certain circumstances, molecular barcoding could strengthen and validate conventional identification. The success of species-level barcoding depends on the availability of conspecific sequences in databases.

5 Acknowledgments

We would like to thank the Directorate of Research and Community Service, Ministry of Research, Technology, and Higher Education for the funding with contract numbers: 1936/UN223.14/PN/2018 and P/1811/UN23/14/PN/2019. We appreciate the head of Research and Community Service Institute and Dean of Biology Faculty of Jenderal Soedirman University, who provide administrative support during the project. Thank students who help during laboratory assessment of the fish samples.

References

1. C. Wabnitz, M. Taylor, E. Green, T. Razak, From ocean to aquarium (UNEP-WMP, Cambridge, 2003)
2. C. C. Shuman, G. Hodgson, R.F. Ambrose, Environ. Conserv. 31, 339-348 (2004)
3. V. Nijman, Biodivers. Conserv. 19, 1101-1114 (2010)
4. A. Nuryanto, D. Bhagawati, Kusbiyanto, Biodivers. 21, 512-520 (2020)
5. A. Nuryanto, D. Bhagawati, Kusbiyanto, IOP Conf. Ser. Earth Environ. Sci. 746, 012021 (2021)
6. S. Yi, J. Zhong, S. Huang, S. Wang, W. Wang, Biochem. Syst. Ecol. 70, 50-59 (2017)
7. R. Froese, D. Pauly, Fishbase, www.fishbase.org (2021)
8. G. R. Allen, M.V. Erdmann, Reef fishes of the East Indies I-III (University of Hawaii Press, Hawaii, 2012)
9. N. Hubert, C.P. Meyer, H.J. Bruggeman, F. Guerin, R.J.L. Komen, B. Espiau, R. Caussee, J.T. Wiliams, S. Planes, PLoS One 7, 28987 (2012)
10. A. Karahan, J. Douek, G. Paz, N. Stern, A.E. Kideys, L. Shaish, M. Goren, B. Rinkevich, J. Nat. Conserv. 36, 1-9 (2017)
11. M. J. Bhattacharjee, B. A. Laskar, B. Dhar, S.K. Ghosh, PLoS One 7, 49950 (2012)
12. K. K. Dasmahapatra, M. Elias, R.I.Hill, J.I. Hoffman, J. Mallet, Mol. Ecol. Resour. 10, 264-273 (2010)
13. A. J. Guimaraes-Costa, F.S. Machado, R.R.S. Oliveira, V. Silva-Costa, M.C. Andrade, T. Giarrizzo, U. Saint-Paul, I. Sampaio, H. Schneider, Sci. Rep. 9, 7530 (2019)
14. R. Ratnasigham, P. D. N. Hebert, Mol. Ecol. Notes 7, 355-364 (2007)
15. R. D. Ward, R. Hanner, P.D.N. Hebert, J. Fish Biol. 74, 329-356 (2009)
16. B. Amatya, Int. J. Biol. 11, 88-100 (2019)
17. O. Abdalwahhab, A. Galal-Khallaf, S.A. El-Latif Saber, A.G.M. Osman, K. Mohammed-Geba, Aquat. Living Resour. 33, 10 (2020)
18. R. Hanner, S. Becker, N.V. Ivanova, D. Steinke, Mitochondrial DNA 22, 106-122 (2011)
19. L. H. G. Pereira, R. Hanner, F. Foresti, C. Oliveira, BMC Med. Genet. 14, 20 (2013)
20. M. Puckridge, N. Andreakis, S.A. Appleyard, R.D. Ward, Mol. Ecol. Resour. 13, 32-42 (2013)
21. J. Diaz, G.V. Villanova, F. Brancolini, F. del Pazo, V.M. Posner, A. Grimberg, S.A. Arranz, PLoS One 11, 0157419 (2016)
22. F. S. Ali, M. Ismail, W. Aly, Mol. Biol. Rep. 47, 5865-5877 (2020)
23. A. Sholihah, E. Delrieu-Trottin, T. Sukmono, H. Dahruddin, R. Risdawati, R. Elvyra, A. Wibowo, K. Kustiati, F. Busson, S. Sauri, U. Nurhaman, D. Dounias, M.S.A. Zein, Y. Fitriana, I.V. Utama, Z.A. Muchlisin, J-F Agnèse, R. Hanner, D. Wowor, D. Steinke, P. Keith, L. Rüber, N. Hubert, Sci. Rep. 10, 2818 (2020)
24. H. Asgharian, H.H. Sahafi, A.A. Ardalan, S. Shekarziz, E. Elahi, Mol. Ecol. Resour. 11, 461-472 (2011)
25. A. Bucklin, D. Steinke, L. Blanco-Bercial, Annu. Rev. Mar. Sci. 3, 471–508 (2011)
26. P. S. Walsh, D.A. Metzger, R. Higuchi, Biotechniques 10, 506-513 (1991)
27. A. Nuryanto, M. Kochzius, Coral Reefs 28, 607–619 (2009)
28. R. D. Ward, T.S. Zemlak, B.H. Innes, P.R. Last, P.D.N. Hebert, Phil. Trans. R. Soc. B 360, 1847–1857 (2005)
29. T. Hall, GERF Bull. Biosci 2, 60-61 (2011)
30. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, Mol. Biol. Evol. 35, 1547–1549 (2018)
31. N. W. Jefery, M. Elias-Guttierrez, S.J. Adamowicz, PLoS One 6, e18364 (2011).
32. R. Higashi, A. Tsukagoshi, H.Kimura, K. Kato, J. Crust. Biol. 31, 142-152 (2011)
33. R. A. Collins, K.F. Armstrong, R. Meier, Y. Yi, S.D.J. Brown, R.H. Cruickshank, S. Keeling, C. Johnston, PLoS One 7, 28381 (2012)
34. A. Chan, L-P. Chang, H.C. Hapuarachchi, C-H. Tan, S-C. Pang, R. Lee, K-S. Lee, L-C. Ng, S-G. Lam-Pua, Parasites Vectors 7, 569 (2014)