The DNA Barcoding of Several Commercial Fish from Simeulue Islands Coast of Aceh Using *Cytochrome Oxidase* Sub Unit I (COI) Gene Marker

Edwarsyah[1], Muhammad Nasir[2], Muhammad Banda Selamat[3]
[1]Department of Aquatic Resources, Faculty of Fisheries and Marine Science, Teuku Umar University, Indonesia
[2]Department of Economics, Faculty of Economics and Business Syiah Kuala University, Indonesia
[3]Department of Marine Science, Faculty of Marine Science and Fisheries,Hasanuddin University, Indonesia
edwarsyah@utu.ac.id, nasirmsi@unsyiah.ac.id mbandas@yahoo.com

Abstract : *Simeulue is a cluster of islands that rich of commercial fisheries in Aceh waters. Management of the fisheries products is highly related to species identification in order to ascertain the appropriate steps to manage the resources sustainably. Identification using DNA barcoding tools is the right answer for the problem that have not been able to be resolved even by morphological approach. Some of 11 individual samples were taken from 6 sampling points based on lobster's catchment and cultivation areas. All of the samples were identified using COI mitochodrial DNA resulting 6 species including Scomber scombus, Scomberomorus plurilineatus, Octopus cyanea, Taeniura Lymna, Sympterygia bonapartii and Panulirus versicolor. The DNA bases were aligned using MEGA application and Neigbour- Joining method, Kimura 2 parameters resulting a 690 base pairs nucleotides. Reconstruction of philogenetic tree shows that the species of Simeulue were conjoined into one clade with the sequences downloaded from genebank. It shows that those species were closely related genetically indicating by the bootstrap value of Scomber scombus (100), Scomberomorus plurilineatus (100), Octopus cyanea (100). The results of the study shows that the DNA barcoding tools can explicate not only the identification up to the species level but also the genetic relationship that can be seen from the interspecies bases composition.*
Keywords: *COI, DNA barcoding; Octopus cyanea, Panulirus versicolor; Simeulue island; Scomber scombus; Scomberomorus plurilineatus; Sympterygia bonapartii; Taeniura Lymna.*

I. Introduction

Simeleu Regency with Sinabang as the capital city is a cluster of islands consisting of 147 islets in surrounding. Located on the western side of Aceh province, the distance is around 105 nautical miles from Meulaboh regency of West Aceh or 76 nautical miles from Labuhan Haji, the South Aceh regency and at position of 2°15'-2°55' North Latitude and 95°40'-96°30' East Longitude (Earth Map, Scale 1:250.000 by Bakorsutanal) (Marine and Fisheries Service of Simeuleu Regency, 2014). The cluster islands of Simeulue are highly potential in marine tourism as it characterized by the archipelagic features that rich in coral reef, seagrass beds, and mangroves ecosystems.

Simeulue Islands, which contain abundant marine wealth, make the islands immensely rich in commercial creatures, which in turn lead to the management to become noteworthy. Commercial animals are constantly becoming target species by force of high consumers demand. Species identification is a major problem and very crucial to undertake in fisheries management. Throughout the world, animals are being commercialized using their local names which vary greatly from one region to another. Identification is an accurate process which can become a solution to provide sustainable fisheries management (Marko et. al. 2004). Morphological identification generating many doubts due to the similarity between the same type species, that even the taxonomists can experience such difficulties (Bangsal et. al 2009).

Molecular identification has rapidly developed and has become solution for identification since the process are prompt, accurate, and able to enrich the information of the species (Ardura et. al. 2010).

Genetic diversity is the lowest hierarchy in biodiversity level. Biodiversity comprises the diversity of habitat, community, population up to species. Genetic diversity is a reflection of diversity in a species named as subspecies (Suwardi, 2007). Dunham (2004) stated that genetic diversity is important for the long-term species survivality. Genetic diversity can maintain the fitness of a species or population by providing them the ability to adapt to environmental changes. The more diverse the genetic resources of the population, the higher capability of the population to survive in the long term, and the higher adaptability to the surrounding environmental changes. There are several factors that influence the level of genetic diversity of fish population, namely factors that can increase genetic diversity or add genes including mutation and migration and factors that can reduce it, namely inbreeding and selection (Soewardi, 2007).

Study of genetic diversity through molecular marking using Deoxyribonucleic acid (DNA) in the nucleus and mitochondrial DNA (mtDNA) can reveal the more detailed intraspecies and interspecies differences. It encompass the structure, composition and organization of genomes at the DNA level (Solihin 1994). The mitochondrial DNA constitutes a different structure from nucleus DNA and smaller size as well. The use of mtDNA as genetic marker has provided both qualititavely and quantitatively information. In addition the mtDNA is rapidly to change so that it can be used to trace the relatively recent events (Solihin 1994). One of the genes encoding mtDNA genome is Cytochrome oxidase subunit I (COI) gene.

The COI gene has many features for studying genetic characteristic, as it either very slightly encounters deletion and insertion in its sequences and has abundant conserved part (sustained) as well so that it can be used as DNA barcoding, namely an attribute of a species (Hebert et al 2003). The COI gene can be utilized to construct a philogenetic at the branch of evolution in species level (Palumbi 1996). Besides, the amino acid structure of protein encoded by COI gene rarely undergoes substitution thus the COI gene is quite stable and can be used as marker for philogeny analysis. However, the bases in the triplet codons are still changing and are silent, namely the changing of bases that create no changes in the amino acid types (Lynch & Jarrel 1993). The COI gene applied to study the genetic characteristics of inter species and between individual (Hebert et. al 2003). COI can be utilized as barcode as well. The literature progress on DNA barcode shows that a short fragment of COI can be used as variation marker that capable to accurately identify variety of animals to species level (Luo et al 2011). The molecular study of *Osteochilus* sp (serukan fish) in Indonesia is still very lacking. Based on the genetic database in GenBank, there is still no genome data of *Osteochilus* sp particularly originated from Aceh/Indonesia. The results of this study are expected to be used as a basic information on genetic characteristics of Serukan fish.

II. Methodology

2.1 Area of study

This research was conducted at clusters of island of Simeleu, Aceh Province. The location determination is based on lobster's catchment and cultivation areas in Simeulue islands. The representatives regions for the study in this research comprising Amaiteng Mulia, Lugu, Linggi and Sefoyan villages, at East Simeulue and Simeulue Cut districts.
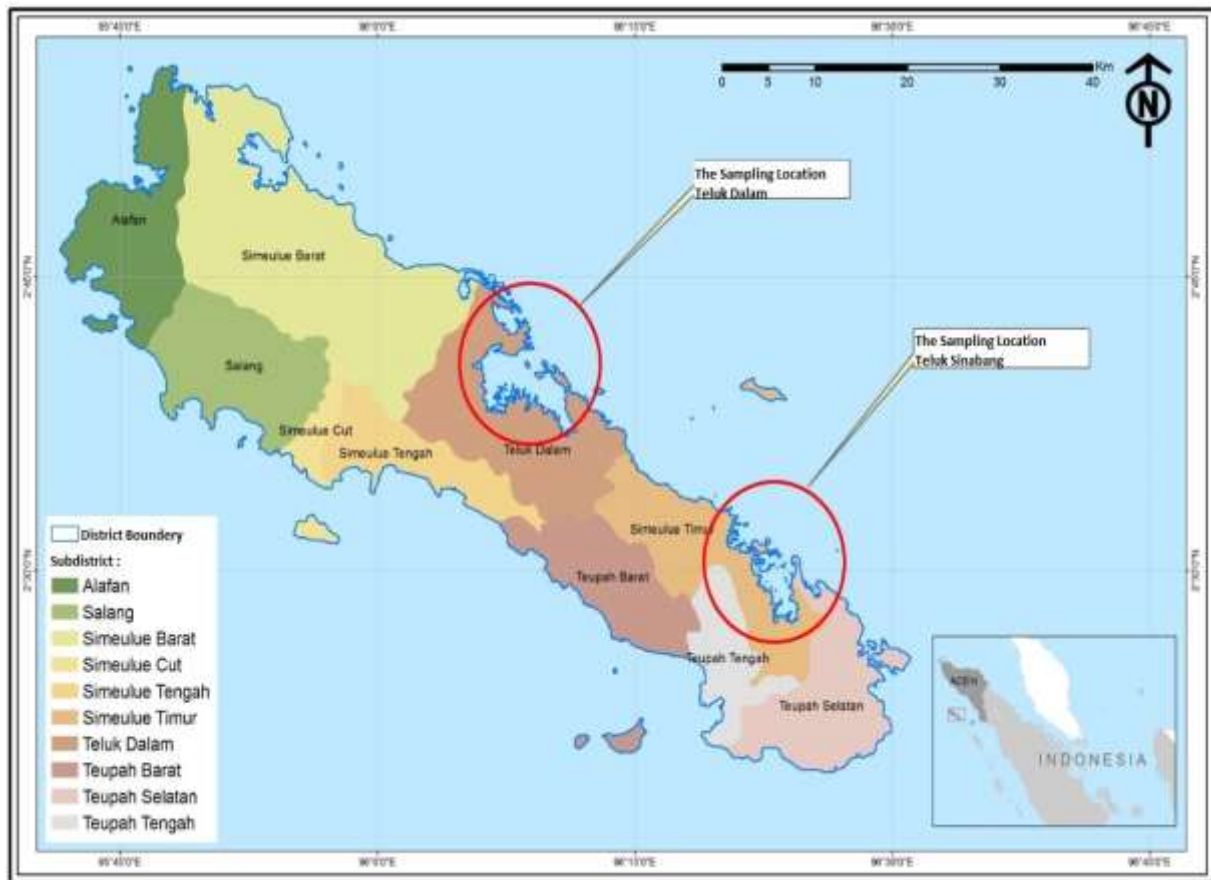
Figure 1. The sampling locations marked by red circles.

2.2 Collecting of fish for testing

The fish species for testing were obtained from local fishermen in Simeulue Islands, to the tune of 3 individual of each specimens (tuna tail fin, octopus leg muscles, stingrays fin, lobster abdomen), in fresh condition. The dorsal fin of each test fish was inserted to a collection tube containing absolute ethanol (EtOH) 70% for 1 hour until overnight. Next, the EtOH 70% was replaced by EtOH 96% as preservatives, labeled and stored in special box. The next step, the fish test were brought to the Reproduction and Genetics Laboratory of the Department of Aquaculture, Faculty of Fisheries and Marine Sciences, Bogor Agricultural University (IPB) to proceed in molecular analysis stage. The tissue (fin) was taken and inserted into 1,5 mL tube containing EtOH 96% dan the leftover fin were stored as voucher specimens at the Basic Fisheries Laboratory, Faculty of Fisheries and Marine Sciences, Teuku Umar University (UTU).

2.3 Molecular Analysis

Molecular analysis was conducted to obtain the DNA fragmen of the animals being tested. The molecular analysis conducted in the laboratory was consists of several stages, namely: DNA isolation and extraction, Polymerase Chain Reaction (PCR), Electrophoresis, and DNA sequencing.

2.3.1 DNA Isolation and extraction, Polymerase Chain Reaction (PCR), Electrophoresis and Sequencing of COI gene

DNA isolation and extraction were conducted using fin section of the test fish (tuna tail fin, octopus leg muscle, stingray fin, lobster abdomen) which were preserved in EtOH 96%. Approximately 30 mg of fin (about 0,3 mm^3) were taken from the storage, then soaked in distilled water for about 30 minutes. DNA isolation and extraction was proceeded using *Genomic DNA Mini Kit for Tissue* (*Geneaid*) following the manual instruction from the manufacturer. The result of DNA extraction then verified by 1,2% agarose gel electrophoresis. The PCR stage was carried out using *Kapa Extra Hot Start kit* and universal primer designed by Butet (2013, unpublished data). DNA amplifications were proceeded at predenaturation temperature of 94°C for 5 minutes, denaturation temperature of 94°C for 45 second, annealing temperature of 53°C for 1 minutes, elongation temperature of 72°C for 1 minutes, post elongation temperature of 72°C for 5 minutes, and storage temperature of 8°C for 5 minutes. The testing for DNA quality resulted from amplification was carried out by electrophoresis using 1,2% gel agarose in TAE 1x (40 mM Tris-acetate, 1 mM EDTA) buffer. Electrophoresis was run in gel containing 5 µL EtBr at 100 volt electric voltage for 30 minutes. The migration of DNA band on agarose gel then observed on transluminator (lightbox using UV as the light source).

Good quality of PCR product was shown by a single, bright and bold DNA band, which then being used as a template in PCR reaction for sequencing. Sequencing reaction following the Sanger method which was conducted through sequencing service company using dNTP big dye labelling.

2.4 Data Analysis

2.4.1 Validating nucleotide base sequences of COI gene

The nucleotide base sequences of individual fish test which were succesfully sequenced then validated by uploading the nucleotide sequences result of *Scomber scombrus, Scomberomorus plurilineatus, Octopus cyanea, Sympterygia bonapartii, Panulirus versicolor* in the *Basic Local Alignment Search Tool-nucleotide (BLASTn)* on the NCBI (*National Center for Biotechnology Information*) web. The results obtained were the percentage of similarity against the species found in the GenBank database. The value of similarity percentage then used as a reference to conduct species validation.

Figure 2. The NCBI web

2.4.2   Alignment of nucleotide sequence of COI gene

The nucleotide sequences which had been obtained, then edited manually based on chromatogram resulted from the sequence reader machine and assisted by the sequences of nucleotide which had been published (Table 1). In addition, the nucleotide editing was also carried out by translating it into the amino acids following the Table of genetic code for mitochondrial genome of invertebrates. Furthermore, the nucleotide bases of all sample species were aligned to homologous DNA segments which stored in GenBank using Clustal W program with the default alignment option embedded in the MEGA 5.2 program (Tamura *et al*. 2011).

2.4.3   Genetic Distance

The genetic distance of the sequences or nucleotide bases of COI intraspecies of fish test was analysed based on the pairwise distance method in the Mega 5.2 program (Tamura *et al*. 2011). The genetic distance presented in the form of data matrix which were used to analyze the tribe relationship among the species. The genetic distance values were also utilized to classify the species into intraspecies or interspecies categories.

### 2.4.4   Reconstructing the philogenetic tree

The phylogeny analysis of fish test genus was carried out based on the *bootstrapped Neighbour-Joinning* (NJ) method at 1000 repetitions which available in the MEGA 5.2 software (Tamura *et al*. 2011). The matrix data of genetic distance then used to reconstruct a philogenetic tree. The phylogenetic tree serves to identify the kinship relationship between the species.

2.4.5   Data Collection

The research data collection comprised of collecting fish for testing, molecular analysis, bioinformatics analysis. Molecular analysis consisted of DNA isolation and extraction, amplification of COI gene, quality assesment of amplification product, and sequencing. Furthermore, bioinformatics analysis was conducted through editing DNA chromatogram data result assisted by the MEGA 5.2 software.

III.   Results

The result of DNA testing was shown by DNA bands presenting in the figures below. The initial marker is a basepair marker followed by samples from Tn1 to Lb3 (Figure 3.1). The DNA bands seen in Figure 3.2 showed that they had been succesfully amplified, indicated by the emerging bar which is a positive result. One sample that was unable to amplify is the number 9 sample, that is the undetected Lb 2 sample.
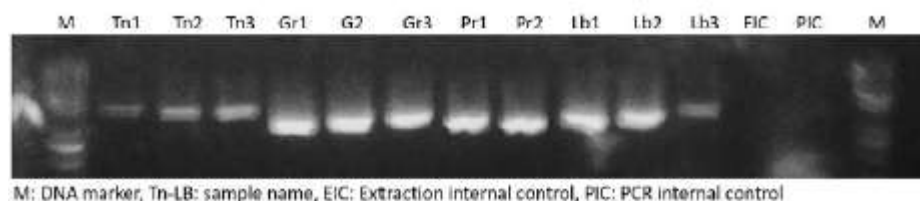
M: DNA marker, Tn-LB: sample name, EIC: Extraction internal control, PIC: PCR internal control

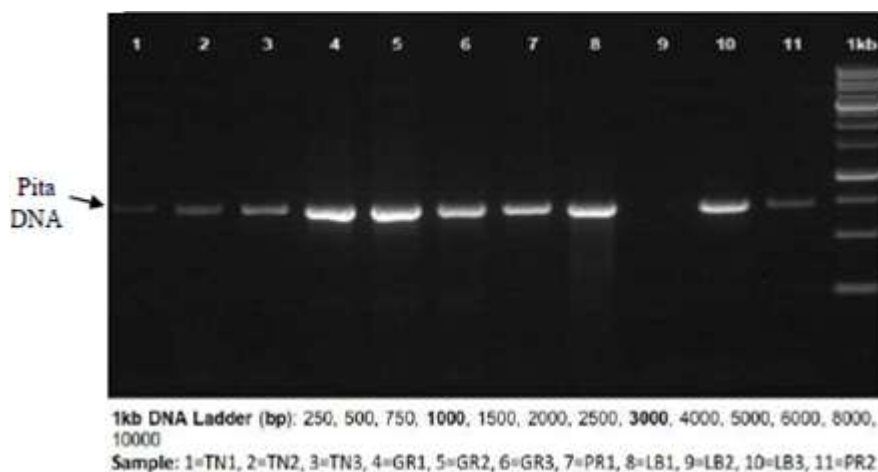Figure 3. DNA bands presented during extraction process

1kb DNA Ladder (bp): 250, 500, 750, **1000**, 1500, 2000, 2500, **3000**, 4000, 5000, 6000, 8000, 10000
Sample: 1=TN1, 2=TN2, 3=TN3, 4=GR1, 5=GR2, 6=GR3, 7=PR1, 8=LB1, 9=LB2, 10=LB3, 11=PR2

Figure 4. DNA bands presented during electrophoresis process

3.1. DNA Analysis

DNA analysis of all organisms were obtained from the sequencing result which then aligned to other organisms from reference data (GenBank). Based on the result of nucleotide alignments with the comparison species, a 690 bases nucleotide length were obtained.

Tabel. 1 The results of identification of fish species landed at Simeulue PPI with barcodes

Isolat	Identification of Bacode	Genebank Acession Number	Identities	Nucleotide Match	Query Cover
TN1*	Scomber scombrus	AB120717.1	72%	489/679	93%
TN2*	Scomberomorus plurilineatus	JF494457.1	73%	454/625	90%
TN3*	Scomber scombrus	AB120717.1	73%	482/660	98%
GR1	Octopus cyanea	AB191280.1	99%	656/657	97%
GR2	Octopus cyanea	AB191280.1	99%	653/657	94%
GR3	Octopus cyanea	AB191280.1	99%	656/657	95%
PR1	Taeniura lymma	KC250631.1	95%	619/652	96%
PR2	Sympterygia bonapartii	EU074616.1	74%	467/629	98%
LB1	Panulirus versicolor	KC107808.1	99%	687/691	99%
LB2	Not detected	-	-	-	-
LB3	Panulirus versicolor	KC107808.1	99%	679/685	100%

3.2 Phylogenetic

The sequences that had been downloaded from genebank (www.ncbi.nlm.nih.gov) then grouped with sequences that obtained directly from Simeulue waters and forming similar clade. It verify that the species in Simeulue waters are similar with the species types of which in the genebank. The phylogenetic tree exhibit that the tunas divided into two clades where *Scomber scombrus* and *Scomberommorus plurilineatus* tuna separated and each of which having its own

distinctive clade. Spesies of *Sympterygia bonaparti* and *Taeniura lymma* are two different clades as well. And so as the lobster and octopus, each of which having it's own clade.
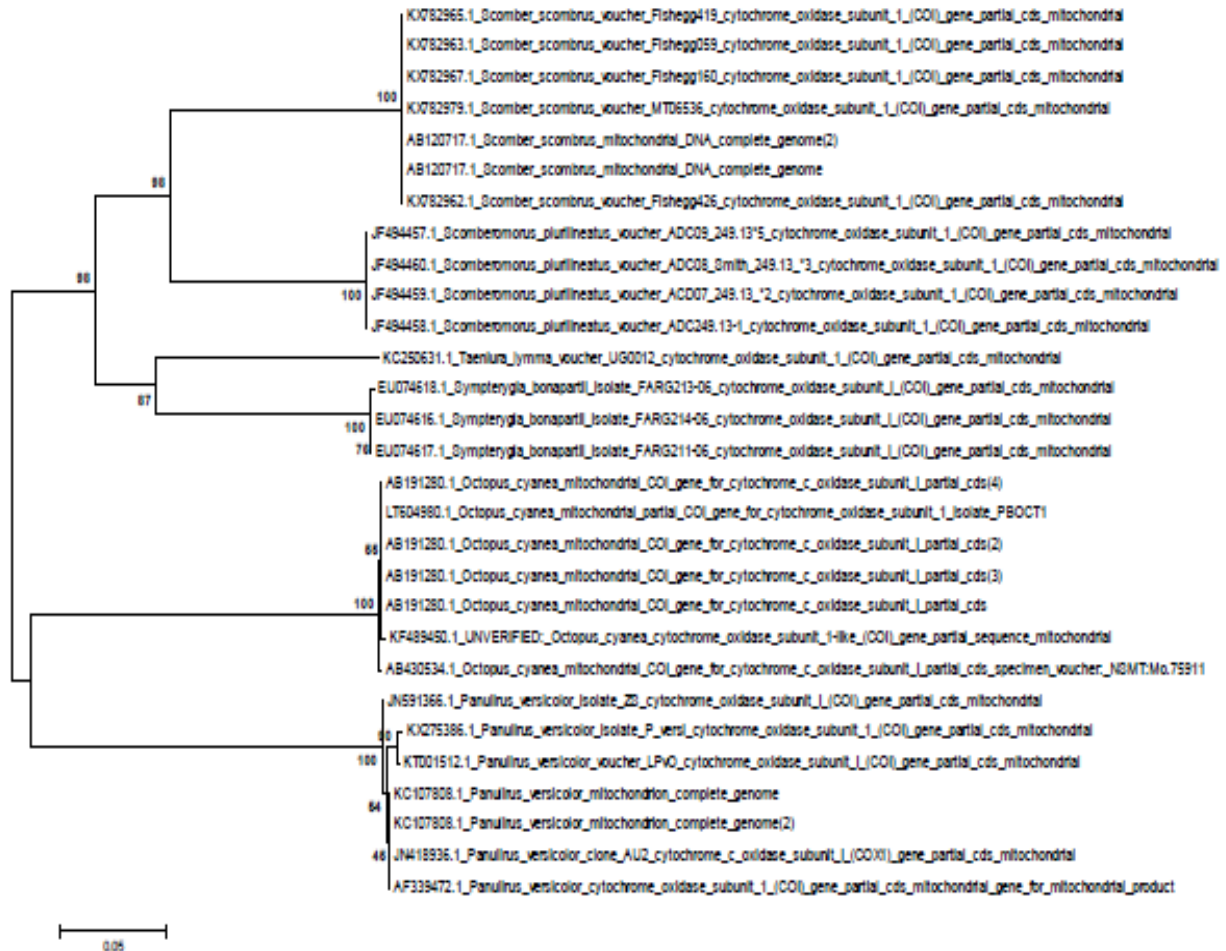

Figure 5. Filogenetik tree (some species found in Simeulue Island) Searched from Genebank

IV. Discussion

The DNA analysis is considerably succesful in addressing the problem of identification, not only as a tool of identification, but also it is capable to answer questions which haven't been able to be answered by morphological approach (Reeb dan Avise, 1990). The sequence analysis of all samples using the COI locus obtained an indent result of 72-99%. Total DNA fragments of 11 samples were then aligned until obtaining 690 bp nucleotides. Based on samples of 11 individuals taken from several sampling points, it was found that the individual composition of *Scomber scombus* is 2, *Scomberomorus plurilineatus* is 1, *Octopus cyanea* is 3, *Taeniura Lymna* is 1, *Sympterygia bonapartii* is 1, *Panulirus versicolor* is 2, and 1 individual sample was not able to identify. The species which unable to identify might have caused by the damaged tissue samples or errors in the DNA amplification process (Maulid, et. al 2016).

The phylogenetic tree reconstruction was intended to illustrate the genetic distance between all species by using the Neighbour-Joining method which evaluated by using the

bootstrap 1000x method (Tamura et. al 2013). The phylogenetic tree was made by inserting sequences from the genebank in order to observe the grouping and to prove that the species being identified are similar to those in the genebank. The clade of *Scomber scombrus* resulting bootstrap value of 100, observable as a perfect value which means that the tuna from Simeulue area are closely related to those whose information were downloaded from the genebank. This is similar to what appears in *Scomberomorus plurilineatus* and *Octopus cyanea,* both of which have 100 bootstrap value. *Panulirus versicolor* emerges as a separate clade and forms a subclade with bootstrap value of 45. The separation of clades determined by the number of bases which encountering changes (Ubaidillah and Sutrisno 2009), which means that the more frequent the mutation occur will lead to the farther distance result, and vice versa (Ramadhaniaty et al. 2018).

V. Conclusion

Identification of 11 samples using the COI locus resulting in 6 species, namely, *Scomber scombus*, *Scomberomorus plurilineatus*, *Octopus cyanea*, *Taeniura Lymna*, *Sympterygia bonapartii*, and *Panulirus versicolor.* The indent result of the eleven samples is 72-99 %, with the bootstrap value is 690 bp.

References

Aprilia FE, Aris S, Widodo, Abdul HAT. 2014. Amplifikasi Gen COI dan 16s rRNA dari Invertebrata Laut *Plakobranchus ocellatus*. *Biotropika*. 2(5): 276-278.

Ardura A, Linde AR, Moreira JC, Garcia-Vazquez E. 2010. DNA barcoding for conservation and management of Amazonian commercial fish. Biol Conserv 143:1438–1443

Dunham RA. 2004. Aquaculture and Fisheries Biotechnology: Genetic Approachs. CABI Publishing, UK. 372 pp.

Hebert PDN, Cywinska A, Ball SL, de Waard JR. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London.* 270: 313-322.

Luo A, Zhang A, Ho SYW, Xu W, Zhang Y, Shi W, Cameron SL and Zhu C. 2011. Potential efficacy of mitochondrial genes for animal DNA barcoding: a case study using eutherian mammals. *BMC Genomics* 12: 84-96. doi: 10.1186/1471-2164-12-84.

Marko PB, Lee SC, Rice AM, Gramling JM, Fitzhenry TM, McAlister JS, Harper GR, Moran AL. 2004. Mislabeling of a depleted reef fish. Nature 430:309–310.

Miya M and Sado T. 2013. Whole mitogenome of fishes. Natural History Museum & Institute, Chiba, Department of Zoology, 955-2 Aoba-cho, Chuo-ku, Japan.

Nelson JS. 1994. Fishes of the World. Third Edition. John Wiley & Sons. Inc, New York USA.

Ratnasingham S, Hebert PDN. 2007. BOLD: The Barcode of Life Data System.*Molecular Ecology Notes*. 7:355–364. doi: 10.1111/j.14718286.2006.016778.

Su LW, Liu ZZ, Tang WQ, Liu D, Wu CY, Yang JQ. 2013. Complete mitochondrial genome of *Osteochilus salsburyi* (*Cypriniformes, Cyprinidae*). *Mitochondrial DNA*, 24(3): 252–254. doi: 10.3109/19401736.2012.752482.

Soewardi K. 2007. Pengelolaan Keragaman Genetik Sumberdaya Perikanan dan Kelautan. Departemen Manajemen Sumberdaya Perairan, Fakultas Perikanan dan Ilmu Kelautan, Institut Pertanian Bogor.

Tamura K, Daniel P, Nicholas P, Glen S, Masatoshi N, Sudhir K. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evoluntary Distance, and Maximum Parsimony Methods. *Mol. Biol Evol*. 28(10): 2731-2739. doi:10.1093/molbev/msr121.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. Mega 5: Molecular evolution genetics analysis version 6.0. *Molecular Biology Evolution* 30: 2725-2729.

Ubaidillah R, Sutrisno H. 2009. *Pengantar Biosistematika : Teori dan Praktek*. LIPI Press. Jakarta

Yang L, Arunachalam M, Sado T, Levin BA, Golubtsov AS, Freyhof J, Friel JP, Chen WJ, Hirt MV, Manickam R, Agnew MK, Simons AM, Saitoh K, Miya M, Mayden RL, He S. 2012. Molecular phylogeny of the cyprinid tribe Labeonini (*Teleostei: Cypriniformes*). *Molecular Phylogenetics and Evolution*, 65(2) 362–379. doi: 10.1016/j.ympev.2012.06.007.

Zein MSA and Prawiradilaga DM. 2013. DNA Barcode Fauna Indonesia. Jakarta Kencana Prenadamedia Group.