Mole crab phylogenetics relationship analysis in Parangkusumo and Ketawang Beach Waters

M Krisanti¹, B F Ramadhan¹, A Mashar¹, N A Butet¹, A A Hakim¹ and Y Wardiatno¹
¹Department of Aquatic Resources Management, Faculty of Fisheries and Marine Science, IPB University (Bogor Agricultural University), Bogor, 16680, Indonesia

Abstract. Mole crabs are the high economic value of aquatic biota that has an important role in the trophic structure, but the continuous use of it can result in disruption of the population. Therefore, accurate information is needed as a reference for determining the right management base. This study aims to analyze the phylogenetic relationships of mole crabs in Yogyakarta and Purworejo coastal waters using COI gene markers. The study was conducted at Ketawang and Parangkusumo Beach. The analysis process consists of morphological and molecular identification. The results were sequenced and processed using MEGA 5. Mole crabs analysis results had a closeness value of 96.27%. The genetic distance of mole crabs from both locations is 0.000-0.005. Mole crabs obtained from the location were morphologically identified as Emerita emeritus and successfully validated using the COI gene with a value of 96%. Thus, mole crabs in this study have a very close kinship relationship.

1. Introduction
Mole crabs are benthic organisms from the Crustacean Hippidae family that lived in intertidal and subtidal regions [1]. Mole crabs have an ecological role in the food chain structure of coastal ecosystems as first-rate consumers. A study about molecular identification of mole crabs has not been done much. Molecular identification has been used extensively for the study of genetics population, taxonomy, and phylogenetic of various types of organisms [2]. The molecular identification technique commonly used is DNA barcoding. The results of DNA barcoding can be described in the form of phylogenetic trees [3].

Beaches in Yogyakarta and Purworejo were selected as research locations. The reason for choosing those locations is because they have regional characteristics that are suitable for mole crabs. The exploitation of mole crabs activities in this region was carried out throughout the year; this condition could have an impact on the mole crabs decreasing population [4], so that management actions are needed to keep the population sustainable. The management of resources requires accurate information about the resources to be managed. This study aims to analyze the phylogenetic relationship of mole crabs on the Parangkusumo and Ketawang beaches using COI gene markers. COI gene markers being used for this research because it can be amplified from an enormous range of species, it has always been acknowledged that primer binding sites within this protein-coding gene are highly conserved [5].
2. Methods

2.1. Sample collections
A total of 113 samples were collected from Ketawang and Parangkusumo Beaches. Sample collection was conducted by local fishermen. Mole crabs sample were preserved using 96% alcohol and then observed in Aquatic Molecular Biology Laboratory, Department of Aquatic Resources Management, Faculty of Fisheries and Marine Sciences, IPB University.

2.2. Morphological identification
Morphological identification of mole crabs refers to the [6] identification journal. Identification is conducted by measuring several parameters, such as carapace length, carapace width, and weights. Samples of mole crabs are measured using a caliper with an accuracy of 0.005 mm and weighed using an analytical balance with an accuracy of 0.05 mg.

2.3. Molecular identification
Molecular analysis refers to the standard commercial kit procedure (Gene Aid, Taiwan), and the sequencing process was carried out by the 1st base sequencing service company Pte Ltd in Singapore. The molecular analysis includes several stages.

2.3.1. Sample preparation. Sample preparation is the initial stage before DNA isolation and extraction. Samples were taken from part of the telson and legs of mole crabs. Since the samples were from the preserved one, samples must be washed using distilled water to remove alcohol before entering the isolation stage.

2.3.2. DNA isolation and extraction. Isolation procedures refer to the standard commercial kit procedure (Gene Aid, Taiwan). This procedure aims to separate DNA in cells from other substances. Mole crabs samples are isolated as much as 30 mg. The sample is added with a 200 µL buffer GT solution, which aims to homogenize the sample. After the homogeneous sample, then add Prot-K as much as 20 µL, which aims to lyse the sample. The mixed solution was incubated for 30 minutes at 60 °C and inverted every 5 minutes. After incubation, the sample was added with a solution of 200 µL of GBT buffer and then shaken for 5 minutes and incubated for 20 minutes at 60 °C with an invert every 5 minutes. After incubation is complete, add an absolute ETOH solution of 200 µL.

Sample of mole crabs is stored in the freezer for approximately 30 minutes, and at the same time, the EB solution is heated at 60°C. Sample of mole crabs was transferred to Spin GD Column and centrifuges at a speed of 14,000 g for 2 minutes using a centrifuge machine. The centrifugation process produces two types of solutions, namely natan and supernatant. The solution was taken while the supernatant located at the bottom of the GD Spin was discarded. The centrifugation process was carried out by adding different buffer solutions in sequence, namely W1 buffer 400 µL and W2 buffer 600 µL at a speed of 14,000 g for 30 seconds.

Supernatant solutions resulting from each centrifugation process are removed. The centrifugation process was carried out again for 3 minutes at a speed of 14,000 g without adding any buffer; then, the Spin GD Column was transferred to a new microtube. The elution is done twice. The first elution is done by adding 50 µL EB to the sample and wait for 10 minutes; then, centrifugation is carried out at a speed of 14,000 g for 30 seconds. The second elution is the same as the first elution but uses a new microtube. Both elution results are then stored in a freezer so that DNA is not damaged and can be used in the DNA quality testing process.

2.3.3. DNA quality testing. DNA quality test was carried out by the electrophoresis method. Electrophoresis is a way of checking the quality of DNA by separating DNA molecules in a medium with 100 volts of electricity for 25 minutes. The electrophoresis process begins with the manufacture
of 1.2% agarose gel. Agarose gel is made by mixing 1.2 g agarose powder with a 100 mL TAE buffer solution. The mixture is heated using a hotplate until it boils, then the hotplate is turned off, and the agarose solution is stirred using a stirrer. After mixing, 5 µL of fluorescence was added to make DNA visible when illuminated using ultraviolet light. Then the gel mold is prepared, agarose solution is transferred into the mold. The solution will harden to form a gel and put into an electrophoresis machine; then, a 1x TAE buffer solution is poured into the electrophoresis machine until the gel is completely submerged. The next step is the electrophoresis process. Loading dye, parafilm, and DNA isolation results were prepared. The loading dye is dripped sufficiently on the parafilm according to the number of samples to be tested. Then the sample is added to the loading dye using microtip. The loading dye functions as a weight (enhancer density) of the DNA sample to facilitate the placement of the sample in the wellbore and give color to the sample so that the sample can be easily observed during the electrophoresis process. After the process is complete, the agarose gel is transferred to an ultraviolet machine. Good quality DNA will produce clearly visible DNA bands.

2.3.4. DNA amplification and visualization. Good quality DNA will be selected for the amplification stage with the PCR (Polymerase Chain Reaction) technique using the My Taq HS Red Mix PCR kit consisting of aquadest, buffers, and primers. The primary used is the COI gene universal primer for several aquatic biotas designed by Butet [7]. The first stage was amplification process, namely pre-denaturation at a temperature of 94˚C for 5 minutes, continued with the second stage, namely the denaturation process at a temperature of 94˚C for 45 seconds, the third stage, namely the annealing process at 53˚C for 1 minute, the fourth stage, namely the elongation process at 72˚C for 1 minute, and finally the post PCR stage at 72˚C for 5 minutes. The second to the fourth stage was carried out as many as 35 cycles. After the amplification process is completed, the next step is to visualize the PCR product using 1.2% agarose gel on an ultraviolet machine using 100 V voltages for 25 minutes.

2.3.5. DNA sequencing. PCR products that have good quality will then go through the phase of gene sequencing. The gene sequencing stage is the process of reading the nucleic base sequences from the analyzed DNA using the Sanger and Nicklen method [8]. This stage is carried out by a sequencing service company in Singapore. The results of the amplification and sequencing will be sent back in the form of nucleotide base sequence data.

3. Data Analysis

3.1. Validation of nucleotide bases. Validation of nucleotide bases was carried out using the BLAST-n (Basic Local Alignment Search Tool - Nucleotide) program located on the National Center for Biotechnology Information (NCBI) website. Sequenced base nucleotide sequence data is uploaded to the NCBI website, and then the site will match the uploaded data with the stored database. Database with species that have the characteristics of the nucleotide bases closest to the nucleotide bases of sequencing results will be displayed.

3.2. Alignment of COI gene nucleotide sequences. The nucleotide sequences resulting from sequencing are aligned using the Clustal W method. This method is used because it has a high level of sensitivity, efficiency, and quick [9]. The principle of this method is all the nucleotide bases are sorted and compared with each other and the value or distance is calculated between each pair of nucleotide bases. The distance matrix between nucleotide base pairs is used to make dendrograms; then, dendrograms will be used as a basis to make various nucleotide base sequences by sorting by adjacent nucleotide base pairs [9] Alignment of nucleotide base sequences was carried out using MEGA 5 application.
3.3. Genetic distance determination. The method used to determine genetic distance is the Pairwise Distance method. This method analyzes all pairs of nucleotide bases and changes the difference between the two bases paired into a distance [10]. The results of the genetic distance calculation are presented in the form of a data matrix that can be used for the analysis of family-to-species relationships based on phylogenetic trees.

3.4. Phylogenetic analysis. Phylogenetic tree construction was carried out using the Neighbor-Joining method. The principle of this method is to find pairs of operational taxonomic units that minimize the total branches at each level of the operational taxonomic unit cluster starting from the star-shaped phylogenetic tree [11]. This method can be performed on the MEGA 5 application [12] with bootstrapped 1000 repetitions.

4. Results and discussion

4.1. Characteristic of habitat
In general, the sampling locations were beaches at South Java that facing the Indian Ocean. Beach profiles in both areas are dissipative beach types characterized by large waves, steep sloping, and fine sandy beach sediments (Figure 1). These characterized by a turbulent layer of water that washes up on the beach after an incoming wave has broken (swash zone). This is a suitable habitat of mole crabs that like sloping and sandy contours.

Parangkusumo and Ketawang beaches have different intensity of anthropogenic activities. Parangkusumo Beach is an area with a high intensity of domestic and international tourism area. However, tourism activities on Ketawang Beach are not too crowded, so the intensity of activities in this region is low.

![Figure 1. Beaches characteristics (A) Pantai Parangkusumo; (B) Pantai Ketawang.](image)

4.2. Morphological identification
Morphological species identification shows that the research sample is *Emerita emeritus* species (Figure 2) with taxonomic classification as follows:

- **Kingdom**: Animalia
- **Phylum**: Athropod
- **Subphylum**: Crustacea
- **Class**: Malacostraca
- **Order**: Decapod
- **Family**: Hippidae
- **Genus**: *Emerita*
- **Species**: *Emerita emeritus* Linnaeus, 1767
Figure 2. Mole crabs.

Morphological identification shows that species come from the family Hippidae included in the genus *Emerita* with the species *Emerita emeritus*. Species names are known through identification journals that specify in detail the morphological characteristics of this species [6]. *Emerita Emeritus* species are characterized by a simple dactylus of pereiopod I, antennal flagellum, which is the same size as the carapace length, and there are fine spines along the ventral. Morphological classification of mole crabs still allows for errors such as misclassification, so that molecular identification is needed so species certainty is obtained.

4.3. Total DNA
The results of the total DNA isolation sample of mole crabs that have been visualized using the electrophoresis method with 1.2% agarose media produce DNA with good quality (Figure 3).

Figure 3. Total DNA of the sample.

4.4. Amplification and visualization DNA fragments of the COI gene
Good amplification results will be sequenced to obtain good nucleotide sequences (Figure 4). The length of the COI gene nucleotide base sequence from the results of the amplification of the mole crabs is in the order of 500-600 bp.

4.5. Nucleotide sequence validation of mole crabs COI gene
Nucleotide sequences obtained from sequencing are then uploaded to the BLASTn (Basic Local Alignment Search Tool - nucleotide) located on the NCBI (National Center for Biotechnology Information) site to obtain certainty of species (certainty species) and find out their proximity to other species. An example in this study has a closeness with Emerita emeritus (Genbank: AF246159.1) of 96.27%. The results of this study are similar to research conducted by Mashar [4], which obtained the results of mole crabs validation of 96%, this could be due to differences in nucleotide base pairs
between the GenBank database and those in the sample. This condition causes not all nucleotide base pairs to be aligned. Besides, the species observed may be different from species found in the GenBank database.

4.6. Nucleotide sequence alignment of mole crab COI gene
Based on the results of COI gene sequencing, the mole crabs obtained nucleotide base length of 482 bp. The nucleotide base composition obtained was 32.6% thymine (T) base, 29.3% base Adenine (A), 21.0% cytosine base (C), and 17.2% Guanin base (G). The alignment of nucleotide sequences with species from the Hippidae and Albunidae families resulted in conserved, variable, and singleton values of 27.11% (135/498), 72.89% (363/498) and 3.82% (19/498). The nucleotides composition of mole crabs is dominated by A-T bases, which are weak bonds. This bond is more easily mutated. When A-T bases dominate the kinship of individual distance, this is inversely proportional to the results. A factor that causes close kinship is the similarity of environmental conditions between research sites.

4.7. The specific nucleotide of mole crab COI gene
The alignment of nucleotide sequences was carried out at the sample of two research sites on the nucleotide Emerita emeritus from the Genbank database. The results of the alignment note that there are 20 specific nucleotides (singleton) in samples from Parangkusumo and Ketawang Beaches. Specific sites that existed in this study sample shows that the sample has distinguishing characteristics between families. Mole crabs have two families, which are Hippidae and Albunidae. The appearance of this site shows the existence of specific evolution in the sample. The evolution of an individual involves genetic mutations and recombinant processes within species. Besides, the evolution of an organism can be identified through changes in its character [13].

4.8. Genetic distance
The genetic distance of COI gene fragments between mole crabs at the two study sites is worth 0.000-0.0042 while the genetic distance of mole crabs with comparative species of the hippidae and albunidae families ranges from 0.0356-0.6360 (Table 1). Comparison between examples in two locations (Parangkusumo Beach and Ketawang Beach) with Emerita emeritus from Genbank (Access code: AF246159.1), which is 0.0356. The greatest genetic distance is the distance between Parangkusumo and Ketawang mole crabs with Albunea paretti (Access code: MF490048.1) and Albunea gibbesii (Access code: MF490047.1), which is 0.6360. Differences in the genetic distance can be influenced by several factors, such as geographical conditions, coverage of migrations areas, environmental characteristics, and genetic structure [14].
### Table 1. Genetics distance of mole crabs sample between Hippidae and Albunidae Family.

| Sample Description                        | Distance  |
|-------------------------------------------|-----------|
| Undur-undur laut Parangkusumo 3          | 0.0021    |
| Undur-undur laut Parangkusumo 2          | 0.0021    |
| Undur-undur laut Parangkusumo 1          | 0.0000    |
| Undur-undur laut Ketawang 3              | 0.0000    |
| Undur-undur laut Ketawang 2              | 0.0021    |
| Undur-undur laut Ketawang 1              | 0.0000    |
| Undur-undur laut Kebumen 5*              | 0.0021    |
| Undur-undur laut Kebumen 4*              | 0.0000    |
| Undur-undur laut Kebumen 3*              | 0.0000    |
| Undur-undur laut Kebumen 2*              | 0.0000    |
| Undur-undur laut Kebumen 1*              | 0.0000    |
| Undur-undur laut Cicalap 3*              | 0.0000    |
| Undur-undur laut Cicalap 2*              | 0.0000    |
| Undur-undur laut Cicalap 1*              | 0.0000    |
| Hippa pacifica AF246161.1                | 0.1695    |
| Hippa ovalis KR047034.1                  | 0.5921    |
| Hippa admirabilis KR047031.1            | 0.6067    |
| Hippa adactyla KR047033.1               | 0.6130    |
| Emerita pilipes KT599463                 | 0.5941    |
| Emerita portenensis KP091520.1          | 0.5921    |
| Emerita ementes AF246159.1              | 0.0356    |
| Emerita brasiliensis KP091539            | 0.5858    |
| Emerita benedicta KP091525               | 0.5837    |
| Albunea parae MF490048.1                 | 0.6067    |
| Albunea gibbea MF490047.1                | 0.6339    |

Note: The table continues with more entries...
4.9. Phylogenetic analysis of mole crabs based on COI gene marker

Data on genetic distance values were used to construct phylogenetic trees of mole crabs from four research locations based on the pairwise distance method. The construction results show that the kinship between species is as shown in Figure 5.

![Phylogenetic tree of mole crabs sample between Hippidae and Albunidae Family.](image)

The results of phylogenetic tree construction show that there is a clear separation between Hippidae and Albunidae families. Phylogenetic trees, such as mole crabs (Parangkusumo, Ketawang, and secondary data from Kebumen and Cilacap), *E. emeritus*, *E. austroafricana*, *E. analoga*, *E. talpoida*, *E. benedicti*, *E. portoricensis*, *E. brasiliensis*, *H. adactyla*, *H. pacifica*, *H. ovalis*, *H. admirabilis*, *A. paretii*, and *A. gibbesii* are constructed based on pairwise distance genetic distance from the composition of COI nucleotide bases showing kinship relationships between species. Marine intraspecies grouped between study sites have a close kinship relationship. Phylogenetic tree branches for the Hippidae family have a close kinship when compared to the Albunidae family.
There is a clade formed based on the results of phylogenetic tree construction. This shows that the population of mole crabs is genetically mixed and spread in the Ketawang and Parangkusumo Beaches. According to Templeton [15], genetic mixing in a clade within a phylogenetic tree is caused by various factors such as environmental conditions, changes in haplotypes, random marriages, and connectivity between regions.

The phylogenetic tree construction of the mole crabs with the comparison of the Hippidae and Albunidae family shows that the sample of mole crabs has a closeness to *Emerita emeritus* in the GenBank database of 0.0356. This reinforces the results of morphological identification that has classified the research sample as *Emerita emeritus* species.

5. Conclusions
Mole crabs obtained from Ketawang and Parangkusumo Beach were morphologically identified as *Emerita emeritus* species and was successfully validated using COI gene markers with a value of 96.27%. The mole crabs in both locations have a very close phylogenetic relationship.

6. Acknowledgment
The authors are thankful to the support of Aquatic Molecular Biology Laboratory and Department of Aquatic Resources Management, Faculty of Fisheries and Marine Sciences, IPB University.

References
[1] Haye P A, Tam Y K and Kornfield I 2002 Molecular phylogenetics of mole crabs (Hippidae: *Emerita*) *Journal of Crustacean Biology* 22(4): 903-915
[2] Hillis D M, Moritz C and Mable B K 1996 *Molecular Systematics, 2nd edition* (Massachusetts: Sinauer Associates, Inc.)
[3] Ward R D, Zemlak T S, Innes B H, Last P R and Hebert P D N 2005 DNA barcoding Australias fish species *Phil. Trans. R. Soc. B.* 360 1847-1857
[4] Mashar A 2016 *Biologi Populasi Undur-Undur Laut (Crustacea: Hippidae) di Pantai Selatan Jawa Tengah* [PhD Thesis] (Bogor: Institut Pertanian Bogor)
[5] Deaegel B E, Jarman S N, Coissac E, Pompanon F and Taberlet P 2014 DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biol. Lett.* 10 1-4
[6] Boyko C B and Harvey A W 1999 Crustacea Decapoda: Albuneidae and Hippidae of the tropical Indo-West Pacific region *Résultats des Campagnes MUSORTOM* ed Crosnier A *Mémoires du Muséum national d’Histoire naturelle* 20 (180) 379-406
[7] Butet N A 2013 *Plastisitas Fenotip Kerang Darah Anadara granosa dalam Merespon Pencemaran Lingkungan: Studi Kasus di Perairan Pesisir Banten* [PhD Thesis] (Bogor: Institut Pertanian Bogor)
[8] Sanger F and S Nicklen 1977 DNA sequencing with chain terminating inhibitors *Proc. Natl. Acad. Sci. USA* 74 5463-5467
[9] Thompson J D, Desmond G H and Toby J G 1994 Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice *Nucleic Acids Research* 22(22) 4673-4680
[10] Tamura K, Dudley J, Nei M and Kumar S 2011 MEGA 5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods *Journal Molecular Biology Evolution* 28(10) 2731-2739
[11] Saitou N and Nei M 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees *Molecular Biology and Evolution* 4(4) 406-425
[12] Kumar S, Stecher G and Tamura K 2016 MEGA7: molecular evolutionary genetics analysis version 7. 0 for bigger datasets *Molecular Biological Evolution* 33(7) 1870-1874
[13] Schmidt H 2003 *Phylogenetic Trees from Large Datasets Inaugural* [PhD Thesis] (Dusseldorf: Dusseldorf University)
[14] Bermingham E and Moritz C 1998 Comparative phylogeography: concepts and applications *Molecular Ecology* 7(4) 367-369

[15] Templeton A R 2010 The diverse applications of cladistics analysis of molecular evolution, with special reference to nested clade analysis *International Journal of Molecular Sciences* 11 124-139