DNA barcoding and phylogenetic relationship of shrimps (Crustacea: Decapoda) of Bangladesh

Bioresearch Communications
Volume 7, Issue 1, January 2021

Md. Sagir Ahmed[1], Sumaiya Salam[1], Sayeda Sabrina Sarwar Rumana[1] and Anindita Barua[2]

[1]Department of Zoology, University of Dhaka, Dhaka 1000, Bangladesh.
[2]Department of Genetic Engineering and Biotechnology, University of Dhaka, Dhaka 1000, Bangladesh.

ABSTRACT: We adopted DNA barcoding technique using a 658-bp fragment of the mitochondrial cytochrome c oxidase I (COI) gene to identify shrimp species collected from the different areas of Bangladesh. A total of 24 sequences were generated belonging to 14 species including four new records- *Macrobrachium nipponense*, *Macrobrachium kistnense, Exopalaemon carinicauda* and *Alpheus malleator*. Genetic distance measured with Kimura 2 parameter showed that genetic divergence increased with higher taxonomic rank. The mean genetic divergence was evaluated and found to be 0.935%, 22.67% and 30.92% within species, genus and family, respectively. In addition to the barcode-based species identification system, phylogenetic relationships were established where individuals belonging to the same species were grouped under the same clade. Maximum likelihood (ML) was preferred as the statistical method and as expected, the phylogenetic tree complemented and ensured the conventional taxonomy. The present study evidently showed that DNA barcoding can be served as an effective tool to discriminate the shrimp species and this will enhance the understanding on evolution and conservation biology.

KEYWORDS: DNA barcoding, COI, Decapoda, Shrimp, Phylogeny

RECEIVED: 13 September 2020, ACCEPTED: 11 December 2020

TYPE: Original Research

CORRESPONDING AUTHOR: Dr. Md. Sagir Ahmed, Department of Zoology, University of Dhaka, Dhaka 1000, Bangladesh.
Email: sagir@du.ac.bd Tel: +880 1552 4422 621

Introduction

Bangladesh, due to its unique geographical location and a suitable climatic condition is exceptionally endowed with a rich variety of aquatic faunal diversity. Shrimps inhabits diverse regions ranging from low saline areas to inland hill streams and impounded water bodies. They have significant roles in the food chain and are an important food source for larger animals ranging from fish to whales. In Bangladesh, 64 taxa of shrimps were reported till now, and among them 27 were assessed as Least Concern and 37 as Data Deficient (IUCN Bangladesh, 2015). The shrimp and prawn fisheries plays an important role in the economy in terms of nutrition, income, employment and foreign exchange earnings (DoF, 2018). Nonetheless, there are some scattered publications that provide information on shrimps of Bangladesh and no in-depth taxonomic study particularly on molecular taxonomy on the decapod species of the country. Most of the commercially important shrimp stocks are either overexploited or under threat. We are exploiting many species without knowing their taxonomy and stock position.

Morphological characterization is not entirely robust. In shell fish especially, morphological plasticity between individuals of the same species is intrinsic. For instance, tones of body color and polymorphism shown by individuals of the same species may differ significantly depending on diet, habitat and season. (Khoo *et al*., 1997). Morphological identification of shrimps is very critical because this group has different larval stages, sexual dimorphism, plasticity, and also can be damaged by rough handling (Nicole *et al*., 2012). As shrimps are able to change color depending upon growth, background coloration and time of day due to chromo metaphors (CSIRO, 2013). For this it is likely to misidentify them and difficult to key to the genus or species level. In order to overcome these constraints, molecular identification may be an alternative approach.

DNA barcoding is a universal method designed to identify species authentically by using short, standardized gene regions of a 658 base pair mitochondrial cytochrome c oxidase subunit I (mtCOI) gene (Hebert *et al*., 2003). It is used as the mutation rate of the gene is often fast enough to distinguish closely related species and also its sequence is conserved among conspecifics (Hebert *et al*., 2003). The concept of using barcoding is to distinguish species and identify unrecognized specimens, such as larval stages, cuts of organs and materials undergoing processing, using fairly short gene sequences (Waugh, 2007). This method can be used as a rapid tool to investigate many uncertain species, composition of species and cryptic species (Spies *et al*., 2006).

Hence, this study aimed to determine the efficacy of mitochondrial cytochrome oxidase subunit I (mtCOI) gene to identify shrimps and also to evaluate the genetic variability and phylogenetic relationships. This baseline integrative approach would substantiate the further taxonomic research on shrimps from Bangladesh and other regions. The generated barcode data would enrich the global database, help to estimate the population structure of morphologically static species and also to detect commercial seafood fraud.

Materials and Methods

Sampling and morphological identification

The target specimens were collected mainly from the fish landing centers and different fish markets depending on their availability and coastal areas of Bangladesh especially, Chattagram and Cox's Bazar district, where most shrimps and prawns are landed. The shrimp specimens were immediately kept in a cool ice box after collection and transferred for further analysis to the Advanced Fisheries and DNA Barcoding Laboratory, Department of Zoology, University of Dhaka. Tissue samples from each of the specimens have also been collected and stored in 90% alcohol. Taxonomic key characters of each specimen were observed and later species level identification was confirmed by comparing with the described characters and taxonomic keys following the published literature by Ahmed *et al.* (2008) and Holthuis, (1980).

Genomic DNA isolation and PCR amplification

Genomic DNA was extracted from tissue samples following the Phenol Isoamyl alcohol extraction protocol (Sambrook, 2001) with minor modification. For amplification of mtCOI segment, the published universal Folmer Primer LCO1490 (forward) 5'- GGTCAACAAATCATAAAGATATTGG -3' and HCO2198 (reverse) 5'-TAAACTTCAGGGTGACCAAA AAATCA -3' (Folmer et al., 1994) was used. The 25ml PCR reaction mixture contains Taq polymerase 12.5 µL, Forward primer 1 µL, Reverse primer 1 µL, Extracted DNA 2 µL, Nuclease-free-water 8.5 µL. The PCR amplification was performed using a Thermal cycler (Applied Biosystem, Veriti 96 well thermal cycler), with the thermal profile set as initial denaturation at 95°C (5 min), followed by 41 cycles of denaturation at 95°C (30 s), annealing at 50-54°C (30 s), extension at 72°C (1 min), followed by a final extension at 72°C (7 min) and subsequent storage at -26 °C. The PCR products were visualized in a 1% agarose gel stained with 20µl Ethidium bromide to ensure that a fragment of the correct size had been amplified and visualized under the gel documentation system (AlphaImager HP). PCR products were purified using the GeneJET PCR Purification Kit (Thermo Scientific, Massachusetts, USA) and the purity and yield of the purified DNA were analyzed using Nanodrop spectrophotometry. The good quality purified PCR products of DNA concentration >5ng/l were sent to First BASE laboratories, Malaysia for sequencing. Sequencing was done by Sanger dideoxy sequencing technology using ABI PRISM 3730xl Genetic Analyzer exploiting the BigDye R Terminator v3.1 cycle sequencing kit chemistry.

Bioinformatic analyses

The raw sequences were viewed by CHROMAS software. The unwanted and noise bases were deleted to obtain a better-quality sequence. Then the sequences were transferred to the FASTA file format (.fas). BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was performed for finding homology. Query coverage, E-value, and identity were the three parameters that taken into the consideration for comparison with the pre-existing data of the National Center for Biodiversity Information database (NCBI) to determine the highest homology. Sequence alignment was done in order to find the genetic variation among the individuals of the same species and genus and to construct a phylogenetic tree. For this purpose, MUSCLE tool was used from the software Molecular Evolutionary Genetic Analysis (MEGA) X (Kumar *et al.,* 2018). The FASTA file (.fas) containing multiple sequences were used to identify genetic variation among different individual specimens that belong to the same species and Kimura-2-parameter (K2P) distance was calculated to identify genetic divergences. Phylogenetic trees were constructed using Maximum Likelihood (ML) based on the Kimura-2-parameter substitution model (Kimura *et al*., 1980). For bootstrapping analysis, a total of 1000 replicates were considered. The nucleotide composition and the %GC content was calculated by using Mega X software. The necessary statistical analysis was performed by using Rstudio (RStudio Team, 2015).

Results and Discussion

A total of 14 prawns and shrimps species under 3 Families (Class- Malacostraca, Order- Decapoda) were identified including four new records- *Macrobrachium nipponense*, *Macrobrachium kistnense*, *Exopalaemon carinicauda* and *Alpheus malleator* (**Table 1**) based on morphometric and molecular approach. The identified species were dominated by the two Families - Palaemonidae with 6 prawn species mainly of freshwater and estuarine origin and Penaeidae with 7 shrimp species of marine and coastal origin. Another family Alphidae has 1 prawn species *Alpheus malleator*. 24 sequences of the 14 shrimp species were generated, annotated and submitted to the global database NCBI GenBank. Their assigned accession numbers were summarized in **Table 1**.

Most of the generated sequences showed 98–100% similarities with the conspecific database sequences in GenBank confirmed by performing BLASTn. In present study the average nucleotide content were 27.13±1.230% A, 31.60±3.181% T, 18.70±1.174% G and 22.57±2.590% C **(Fig. 1)**. The average GC content 41.27±3.203% and AT content were 58.73±3.218%, found in the present study which was similar with the Japanese shrimps (Satoushi *et al.,* 2018). A total of 51 sequences, 27 NCBI and 24 DUZM sequences, were analyzed to perform divergence and evolutionary studies. The mean Kimura 2 parameter genetic divergence increased from lower taxa, species to the higher taxonomic rank, family with mean divergence 0.935±1.59%, 22.67±8.60%, 30.92±6.03% within species, genus and family, respectively **(Table 2)**.

Table-1. GenBank Accession number of the COI sequences of the shrimp species sequenced

SL No.	Family	Name of the Species	GB Accession numbers
1.	Palaemonidae	*Macrobrachium rosenbergii*	MH429358, MT895787-88
2.		*Macrobrachium nipponense*	MN200403
3.		*Macrobrachium kistnense*	MN200413
4.		*Macrobrachium villosimanus*	MT890526-27
5.		*Exopalaemon styliferus*	MN200417, MT890528-29
6.		*Exopalaemon carinicauda*	MN200393
7.		*Nematopalaemon tenuipes*	MT895789
8.	Penaeidae	*Penaeus monodon*	MN200401, MH429357
9.		*Fenneropenaeus indicus*	MN200396
10.		*Metapenaeus brevicornis*	MN200415, MH429352, MT890530
11.		*Metapenaeus ensis*	MN200399, MH429353
12.		*Parapenaeopsis sculptilis*	MN200398, MH429355
13.		*Parapenaepsis hardwickii*	MN200402
14.	Alpheidae	*Alpheus malleator*	MN200395

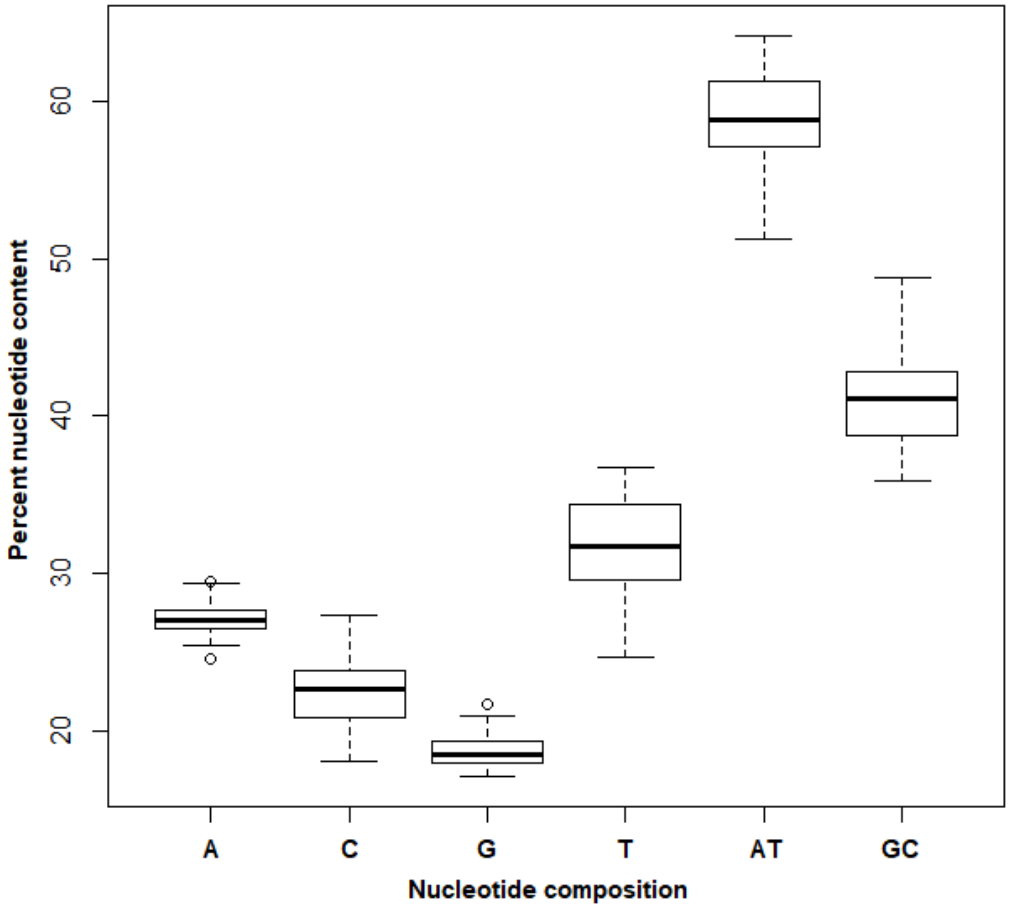

Figure-1. Percentage of nucleotide content of the 24 COI sequences analyzed

Table-2. The genetic divergence (K2P Distance %) within species, genus and family

Taxonomic rank	No. of comparisons	Min (%)	Max (%)	Mean (%)
Within species	72	0	6.522	0.935±1.59
Within genus	131	8.605	33.50	22.67±8.60
Within family	353	13.99	48.87	30.92±6.03

The average congeneric divergence was about 24 times higher than the conspecific. DNA barcoding gap was defined with the difference in K2P distance between highest conspecific and lowest congeneric genetic divergence. The gap of 2.1% and also the gradual increase in genetic variation with the increment of taxonomic level proves the efficiency of the DNA barcodes in species discrimination. Most importantly, this concluded that K2P distance of intraspecies being sharply less than the interspecies, supported the previous studies (Nursyahra *et al.*, 2018). The highest intragenus distance was 33.50%, found between *Macrobrachium rosenbergii* and *M. nipponense* and lowest was 8.605% between *M. rosenbergii* and *M. villosimanus*. Both the maximum and minimum genetic variation was found within family Penaeidae. The highest intrafamily divergence was 48.87% between *Penaeus monodon* and *Parapenaepsis hardwickii* and the lowest was 13.99% between *Metapenaeus ensis* and *Parapenaepsis hardwickii*.

Maximum likelihood (ML) phylogenetic tree was constructed and the lineage support was interpreted based on bootstrap percentage (BP) [BP: 100% maximal clade support, 95% to < 100% strong clade support, 75% to < 95% moderate clade support, 50% to < 75% weak clade support and < 50% negligible clade support]. It showed sequences of the same species formed monophyletic clade proving the effectiveness of the COI gene in species delimitation and also confirmed the authenticity of their recognition (**Fig. 2**).

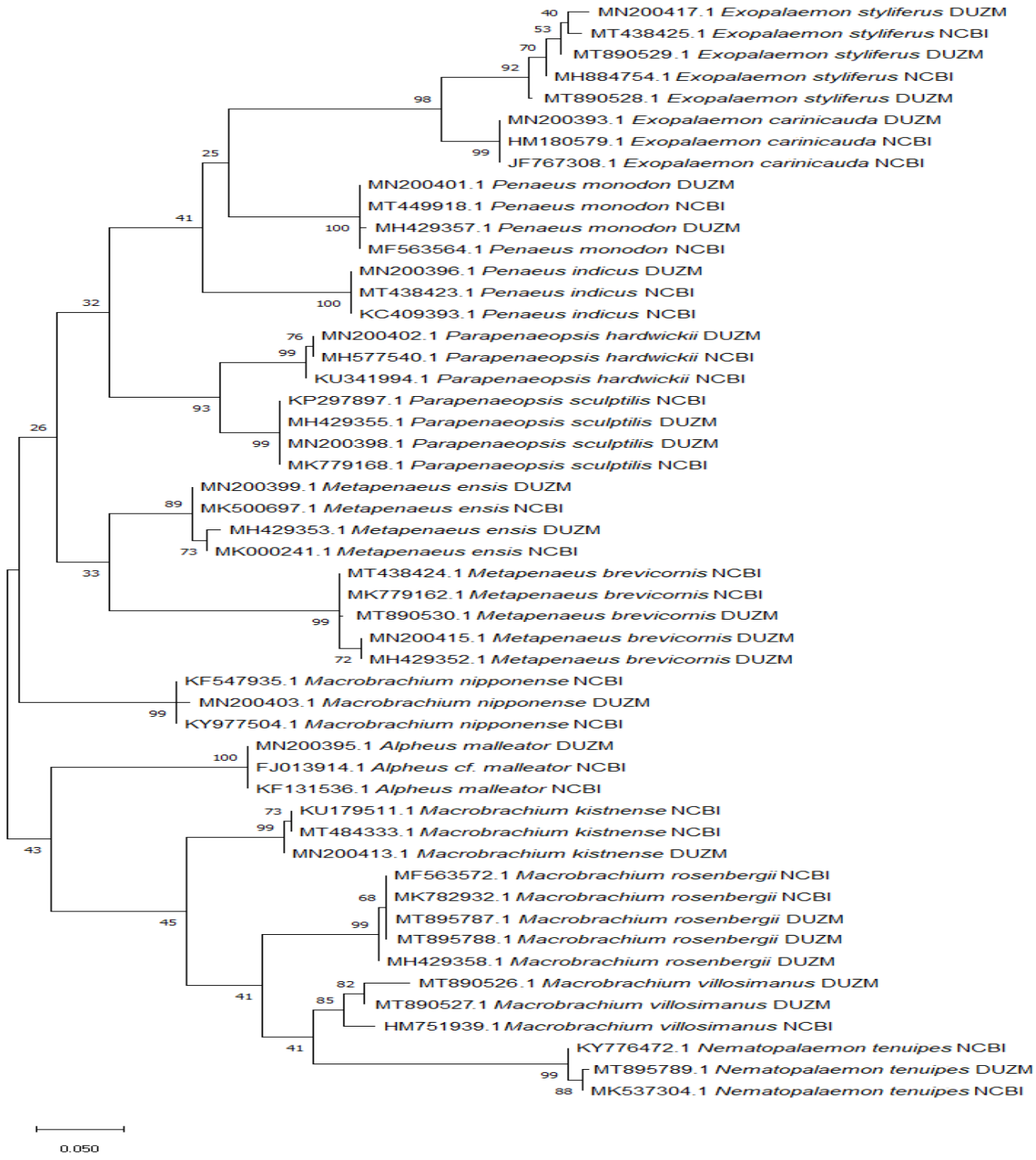

Figure-2. Maximum Likelihood tree showing the relationships among the shrimp species based on sequences analyzed in present study (designated as DUZM) with the pre-existing sequences of NCBI GenBank.

All the close association within species was with 99-100% bootstrap value support except for *E. styliferus* with 92%, *M. ensis* with 89% and *M. villosimanus* with 85% bootstrap value support. Also, the highest intraspecies genetic distance of 6.522% was found within *M. villosimanus* species, further supporting the reason of moderate BP. The congeneric species were also in close association except for *M. nipponense*, which was found to have evolved earlier and grouped with the common ancestor of *Metapenaeus*, *Penaeus* and *Parapenaeopsis* genus. The majority of the species in the tree are from the family Penaidae. This shrimp family has a strong association with coastal habitat features including mangrove systems especially during their juvenile phase (Mohan et al., 1995). The species within this family were in monophyly with however low BP 26%. The tree revealed strong terminal phylogenetic specificity in contrast to moderately or sometimes poorly supported deeper level relationships with DNA barcoding effectiveness in species identification but lower utility in elucidating relationships as all the Palaemonidae species are not in close association. Among two commercially important species *M. rosenbergii* show close association and lower genetic divergence with the *M. villosimanus* and *Penaeus monodon* was relatively closer to the *Exopaleomon* species. However, the bootstrap value support for both of them was less than 50%. In order to fully understand the rate of evolution and the time of their occurrence in a particular branch of the phylogenetic tree this study requires to expand and eventually complete an inventory of shrimps of Bangladesh.

Conclusions

Shrimp, commonly known as White Gold of Bangladesh, is the third most important source of high transnational value for the country. This study validates the efficacy of COI barcodes for identification of shrimp species. The use of molecular data was complementary to morphological analysis as such taxonomical identification of the particular species. The COI DNA barcoding also accelerates the pace of species discovery by allowing taxonomists to rapidly sort specimens and by highlighting divergent taxa that may represent new species. It is an initiative to extended the barcode to other shrimp species and prepare a complete DNA barcode reference library for shrimps and prawns of Bangladesh and also to further pave the way of shrimp culture, commercialization, management and conservation.

References

1. Ahmed, A. T. A., Kabir, S. M. H., Ahmad, M., Rahman, A. K. A., Haque, E. U., Ahmed, Z. U., Begum, Z. N. T., Hassan, M. A. and Khondker, M. (2008), *Encyclopedia of Flora and Fauna of Bangladesh, Vol.18. Part II. Arthropoda: Crustacea.* Asiatic society of Bangladesh, Dhaka pp. 1-149.
2. CSIRO (2013), "*Crustaceans Unique Colour Control System*", Location: Eco. Sciences Precinct - Dutton Park, 41 Boggo Road, Dutton Park Qld 4102, Australia.
3. DoF. (2018), *Yearbook of Fisheries Statistics of Bangladesh, 2017-18.* Fisheries Resources Survey System (FRSS), Department of Fisheries. Bangladesh: Ministry of Fisheries, 2018, Vol. 35, pp. 129.
4. Folmer, O.M., Black, M., Hoeh, R., Lutz, R., Vrijehoek, R. (1994), "DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates", *Mol. Mar. Biol. Biotech.*, Vol. 5, pp. 304–313.
5. Hebert, P. D. N., Cywinska, A., Ball, S. L. and deWaard, J. R. (2003), "Biological identifications through DNA barcodes", *Proc. R. Soc. Lond.* B, Vol. 270, pp. 313–322.
6. Holthuis, L. B. (1980), *FAO species catalogue. Vol. 1, Shrimps and prawns of the world: An annotated catalogue of species of interest to fisheries*, Rome: Food and Agriculture Organization of the United Nations.
7. IUCN Bangladesh. (2015), *Red List of Bangladesh Volume 6: Crustaceans*. IUCN, International Union for Conservation of Nature, Bangladesh Country Offce, Dhaka, Bangladesh, pp. xvi+256
8. Khoo, G., Loh, E. Y. F., Lim, T. M. and Phang, V. P. E. (1997), **"**Genetic variation in different varieties of Siamese fighting fish using isoelectric focusing of sarcoplasmic proteins", *Aquac. Int.*, Vol. *5*, pp. 537-549.
9. Kimura, M. (1980), "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences", *J Mol Evol.*, Vol. *16* No. 2, pp. 111-120.
10. Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018), "MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms", *Mol Biol Evol.*, Vol. 35, pp. 1547-1549.
11. Mohan R, Selvam V, Azariah J. (1995), "Temporal distribution and abundance of shrimp postlarvae and juveniles in the mangroves of Muthupet, Tamilnadu, India", *Hydrobiologia*, Vol. 295, pp. 183–191.
12. Nicole, S., Negrisolo, E., Eccher, G., Mantovani, R., Patarnello, T, Erickson, D. L., Kress, W. J. and Barcaccia, G. (2012), "DNA Barcoding as a reliable method for the authentication of commercial seafood products", *Food Technol Biotechnol*., Vol. *50*, pp. 387–398.
13. Nursyahra, Wahidi, I., Purnamasari, L. and Achmad Farajallah, A. (2018), "Clarification of freshwater shrimp species with DNA barcoding", *J Entomol Zool Stud.*, Vol. *6* No. 5, pp. 682-684.
14. RStudio Team. (2015), *RStudio: Integrated Development for R.* RStudio, Inc., Boston, MA. url: http://www.rstudio.com/
15. Sambrook, J. and Russell, D. W. (2001), "Molecular cloning: a laboratory manual (3-volume set)", *Immunology*, Vol. *49*, pp. 895-909.
16. Satoushi, K., Reiko, N., Hidehiro K. and Ikuo, H. (2018), "The complete mitochondrial genome sequence of the sakura shrimp, *Sergia lucens* (Crustacea, Decapoda,*Sergestidae*)", *Mitochondrial DNA Part B.*, Vol. 3.
17. Spies, I. B., Gaichas, S., Stevenson, D. E., Orr, J. W. and Canino, M. F. (2006), "DNA-based identification of Alaska skates (Amblyraja, Bathyraja and Raja: Rajidae) using cytochrome c oxidase subunit I (COI) variation", *J Fish Biol.*, Vol. *69* No. sb, pp. 283-292.
18. Waugh, J. (2007), "DNA barcoding in animal species: Progress, potential and pitfalls", *Bio Essays*, Vol. 29, pp. 188-197.