International Journal of Zoology Studies
ISSN: 2455-7269
Impact Factor: RJIF 5.14
www.zoologyjournals.com
Volume 3; Issue 1; January 2018; Page No. 96-102

Molecular identification and validation of Indian mud crab genus *Scylla* (Decapoda: Portunidae) based on mitochondrial genes

[*1] Gatreddi Srinu, [2] G Lakshmana Rao, [3] SK Chand Basha, [4] B Pandayya
[1, 2, 3] Department of Zoology & Aquaculture, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India
[4] Department of Zoology, Adikavi Nannaya University, Rajamahendravaram, Andhra Pradesh, India

Abstract
Effective stock assessment and fisheries management require clearness in their taxonomy for any species. Indian mud crab (genus *Scylla* de Hann, 1883) is well known for its morphological plasticity and absence species specific taxonomic diagnostic characters due to which became much contentious in their identification. Although this fascinating genus *Scylla* studied using different morphological and molecular approaches, still there were ambiguities the identification of mud crab species in India. The present study was carried out to find out genetic divergence and to validate phylogenetic position of four species of the genus *Scylla* (*S. serrata, S. olivacea, S. tranquebarica, S. paramamosian*) by analyzing 127 partial sequences belongs to four mitochondrial (mt) genes *i.e.* Cytochrome Oxidase I (COI), Cytochrome b (Cyt b) complex, 12s rRNA and 16s rRNA genes. In the results, the mean intraspecific Kimura 2-parameter distances were consistently higher in all four species for COI gene followed by 12s rRNA and 16s rRNA genes whereas the least distances were observed in Cytochrome b gene. No overlap was found between intraspecific and interspecific distances, suggesting that the existence of distinct barcoding gap to delineate the species boundaries. The Neighbour Joining (N-J) phylogenetic tree for Cytb and 12s rRNA genes formed distinct clusters each containing same species. It can be concluded as Cytb and 12s rRNA sequences could used to delineate evolutionary relatedness with closely related species and this is an effective combinational approach for genetic studies and molecular taxonomic studies on genus *Scylla*.

Keywords: mtDNA, mud crab, genetic distance, *Scylla*, 16s gene, cytochrome oxidase I

Introduction
The genus Scylla is a commercially important crab in both natural fisheries and aquaculture practices. Its wide range of geographic distribution from south-eastern and eastern Africa to Southeast Asia and Indo-pacific regions [1] provides to cope up with different morphological characters according to diverse environment conditions. Their bigger size, high meat yield and delicious taste makes them as most preferable quality food item for both local and international markets [2]. The classification of *Scylla* has been controversial for a long time. Estampador (1949) [3] assigned *Scylla* to three species and one sub-species (four taxa) including *S. serrata, S. oceanica, S. tranquebarica,* and *S. serrata var. paramamosain* based on morphology and gametogenesis. Stephenson and Campbell (1960) [4] suggested that only one species is present for this genus (*S. serrata*). Serene (1952) divided the four *Scylla* species into two categories, recognized *S.oceanica* as the only species for the mankind, with its variety *S.oceanica tranquebarica* and *S. serrata* is the only valid species for the unmarked type, with its variety *S. serrata paramamosain* [5]. Subsequently, Keenan *et al*. (1998) [6] extensively revised the taxonomy of genus *Scylla* using integrative approaches and divided the this genus into four distinct species including *S. serrata, S. olivacea, S. tranquebarica,* and *S. paramamosain*. Afterwards, several researchers revised the taxonomic status of genus Scylla with morphological and molecular approaches following the Keenan *et al*. Majority of taxonomical works on genus *Scylla* in India referred only to *S. serrata*. Joel and Raj (1983) [7] reported the presence two species *S. serrata and S. tranquebarica* from Tamil Nadu near Pulicat lake. Many recently happened researches in India were based on Kathirvel and Srinivasagam (1992) [8], also reported 2 species in Indian mud crab from Kerala near Kochin back waters *S. tranquebarica* (a dark green morph) and *S. serrata* (a greenish brown morph).

The usage of mitochondrial DNA (mtDNA) in the molecular phylogeny and evolutionary studies was known for more than 30 years [9] particularly in Crustaceans at species and population levels [10 11]. The key characters making it powerful tool in the toolbox of evolutionary studies includes higher mutational rate lower effective population size compared to the nuclear DNA. This is the main reason which provoked to propose standardize DNA based species identification by analysing a single segment of mitochondrial genome. This scenario facilitated the workers across the globe in yielding the uniform sequence data. The interpretation of this data in conjunction with modern web analysis tools provides answers for many challenges in the field of taxonomy and evolution [12]. More recently, taxonomic studies on genus *Scylla* carried out by focusing both morphological [13, 14] and mitochondrial marker approaches [15, 16].

The limitations inherent in morphology based identification systems and the dwindling pool of taxonomists signal the need for a new approach to identify these closely related species. Moreover, phylogenetic studies were carried based on one or two mitochondrial genes but comprehensive study on genetic

distance and phylogenetic relationship among the species genus *Scylla* based on more number of genes were scarce. In this study, we collected 127 partial mitochondrial COI sequences belongs to Indian waters for genus *Scylla* to assess genetic divergence and tested phylogenetic signal for four mitochondrial (COI, Cyt b, 12s rRNA and 16s rRNA) genes in delineating its four species.

Materials and Methods

In total, 127 partial gene sequences obtained for four species with the range of sequences from 25 (*Scylla serrata*) to 1 (*Scylla tranquebarica* and *Scylla paramamosian*) (Table 1) as on 10th January, 2018. Maximum number of analysed partial sequences for the gene COI (47 sequences) followed by 16s rRNA gene (44 sequences), Cyt b gene (24 sequences) and 12s rRNA gene (12 sequences). Out of four species present in genus *Scylla*, including all four mitochondrial genes, the maximum sequences analysed for the species *Scylla serrata* (65 sequences) followed by *Scylla olivacea* (42 sequences), *Scylla tranquebarica* (15 sequences) and *Scylla paramamosian* (5 sequences). Among all sequences, maximum number of sequences were geographically belongs to the Tamil Nadu followed by Kerala and Maharashtra states. All sequences were assembled and end-trimmed to get homologous region to avoid errors during sequencing and those sequences subjected to aligned using ClustalW analysis tool [17]. Sequences with sufficient length only were considered with the view of bringing uniformity in analysis across all species. To ensure homology in heterogeneous sequences, some bases were trimmed. To bring this homogeneity in some sequences, missing sequence parts were adopted from most conserved regions of the sequences available in NCBI GenBank for the same species. Nucleotide composition (A, T, G, C, GC1, GC2 & GC3) calculated for homologous end-trimmed sequences using MEGA V.7.0 (Molecular Evolutionary Genetic Analysis) [18] software (Arizona). Inter and intra species evolutionary divergences in various hierarchical levels were analysed using Kimura 2 Parameter method [19]. The variation was estimated following the bootstrap method with 5000 bootstrap replicate values. The pair-wise deletion option was selected to treat the gaps or missing data between each compared specimen. Finally, the Neighbour-Joining (NJ) tree among species was created to give distance values using bootstrap method in Kimura 2 – parameter mode of analysis those values are in the units of the number of base substitutions per site [20]. To verify the robustness of the nodes of the N-J tree, bootstrap analysis was carried out using 1000 pseudo replicates [21]. Both transitions and transversions were cumulated and included as substitutions. Missing bases or gaps were treated by adopting pair wise deletion method employed in MEGA V 7.0.

Results and Discussion

All the retrieved sequences were verified thoroughly and no complexity or ambiguities were observed among them. Out of four analysed sequences, two of them (COI and Cyt b genes) were protein coding genes whereas remaining two (16s rRNA and 12s RNA genes) were non coding genes. The average length values of the aligned and end-trimmed sequences were presented in table 1.

Table 1: List of COI GenBank accession numbers along with their geographic source of sequences for four mitochondrial genes (N = No. of Sequences)

S. No.	Species (N)	GenBank accession no. & Source	Length
COI gene			
1	*Scylla serrata* [25]	KC760161-63 (Maharashtra), KC154079-83 (Kerala), AB861521 (Tamil Nadu), KY564434 (Kerala), KF717539-40 (Maharashtra), JN085428-36 (Tamil Nadu), KC200562 (Tamil Nadu), KC200564 (Tamil Nadu), KF612462 (Kerala), AB857346 (Tamil Nadu)	521
2	*Scylla olivacea* [12]	KC200563 (Tamil Nadu), KC200565 (Tamil Nadu), KC154069 (Kerala), KC154075-78 (Kerala), AB857347 (Tamil Nadu), AB861522 (Tamil Nadu), JN688965 (Tamil Nadu), KT921343 (Tamil Nadu), KT921346 (Tamil Nadu)	585
3	*Scylla tranquebarica* [8]	KC760164-68 (Tamil Nadu), JN688964 (Tamil Nadu), KT921348, KT921344	531
4	*Scylla paramamosian* [2]	KT921345 (Tamil Nadu), KT921347 (Tamil Nadu)	582
16s rRNA gene			
5	*Scylla serrata* [24]	JQ743345 (Tamil Nadu), KC154084-88 (Kerala), KF220527-36 (Tamil Nadu), JX446640-44 (Tamil Nadu), AB857343-45 (Tamil Nadu)	500
6	*Scylla olivacea* [17]	JQ743346 (Tamil Nadu), KC154070-74 (Kerala), JX446635-39 (Tamil Nadu), AB857337-42 (Tamil Nadu)	498
7	*Scylla tranquebarica* [1]	KF220537-45 (Tamil Nadu), KT921342 (Tamil Nadu)	526
8	*Scylla paramamosian* [2]	KT921340-41 (Tamil Nadu)	520
Cytochrome b gene			
9	*Scylla serrata* [9]	KJ607959-61 (Tamil Nadu), AB857348-53 (Tamil Nadu)	382
10	*Scylla olivacea* [13]	AB857354-59 (Tamil Nadu), AB861881-87 (Tamil Nadu)	382
11	*Scylla tranquebarica* [1]	KT921350 (Tamil Nadu)	382
12	*Scylla paramamosian* [1]	KT921349 (Tamil Nadu)	382
12s rRNA gene			
13	*Scylla serrata* [7]	KF246701-07 (Tamil Nadu)	393
14	*Scylla tranquebarica* [5]	KF246696-00 (Tamil Nadu)	354

Average genetic divergence values of COI gene for within species for *Scylla serrata*, *Scylla olivacea*, *Scylla tranquebarica* and *Scylla paramamosian* were 0.442%, 0.251%, 0.386% and 0.851% respectively whereas for 16s rRNA gene, the values were 0.055%, 0.017%, 0.040% and 0.002% respectively. Multiple genes to calculate genetic divergence within the species for Cyt b and 12s rRNA genes were available only for two species. The average genetic divergence values of Cyt b gene for within species for *Scylla serrata* and *Scylla olivacea* were 0.005% and 0.013 respectively. The average genetic divergence values of 12s rRNA gene for within species for *Scylla serrata* and *Scylla tranquebarica* were 0.091% and 0.031% respectively (Table 2). In almost all genes, the range of genetic distances were comprehensively higher for COI gene followed by 12s rRNA, 16s rRNA and Cyt b genes. (Figure 1). Hence, it is evident that COI gene undergone drastic changes in the process of evolution compared to other genes.

Table 2: Average and range of Kimura 2-parameter distance values within four *Scylla* species for 4 mitochondrial genes. (AD= Average distance, SE= Standard Error)

S. No.	Species	A.D. ± S.E.	Min	Max
COI gene				
1	*Scylla serrata*	0.442±0.43	0.00	0.976
2	*Scylla olivacea*	0.251±0.40	0.00	0.829
3	*Scylla tranquebarica*	0.386±0.43	0.002	0.913
4	*Scylla paramamosian*	0.851±0.00	0.851	0.851
16s rRNA gene				
5	*Scylla serrata*	0.055±0.14	0.00	0.123
6	*Scylla olivacea*	0.017±0.06	0.00	0.057
7	*Scylla tranquebarica*	0.040±0.08	0.00	0.107
8	*Scylla paramamosian*	0.002±0.00	0.002	0.002
Cytochrome b gene				
9	*Scylla serrata*	0.005±0.018	0.003	0.008
10	*Scylla olivacea*	0.013±0.046	0.003	0.035
12s rRNA gene				
11	*Scylla serrata*	0.091±0.173	0.016	0.229
12	*Scylla tranquebarica*	0.031±0.088	0.009	0.059

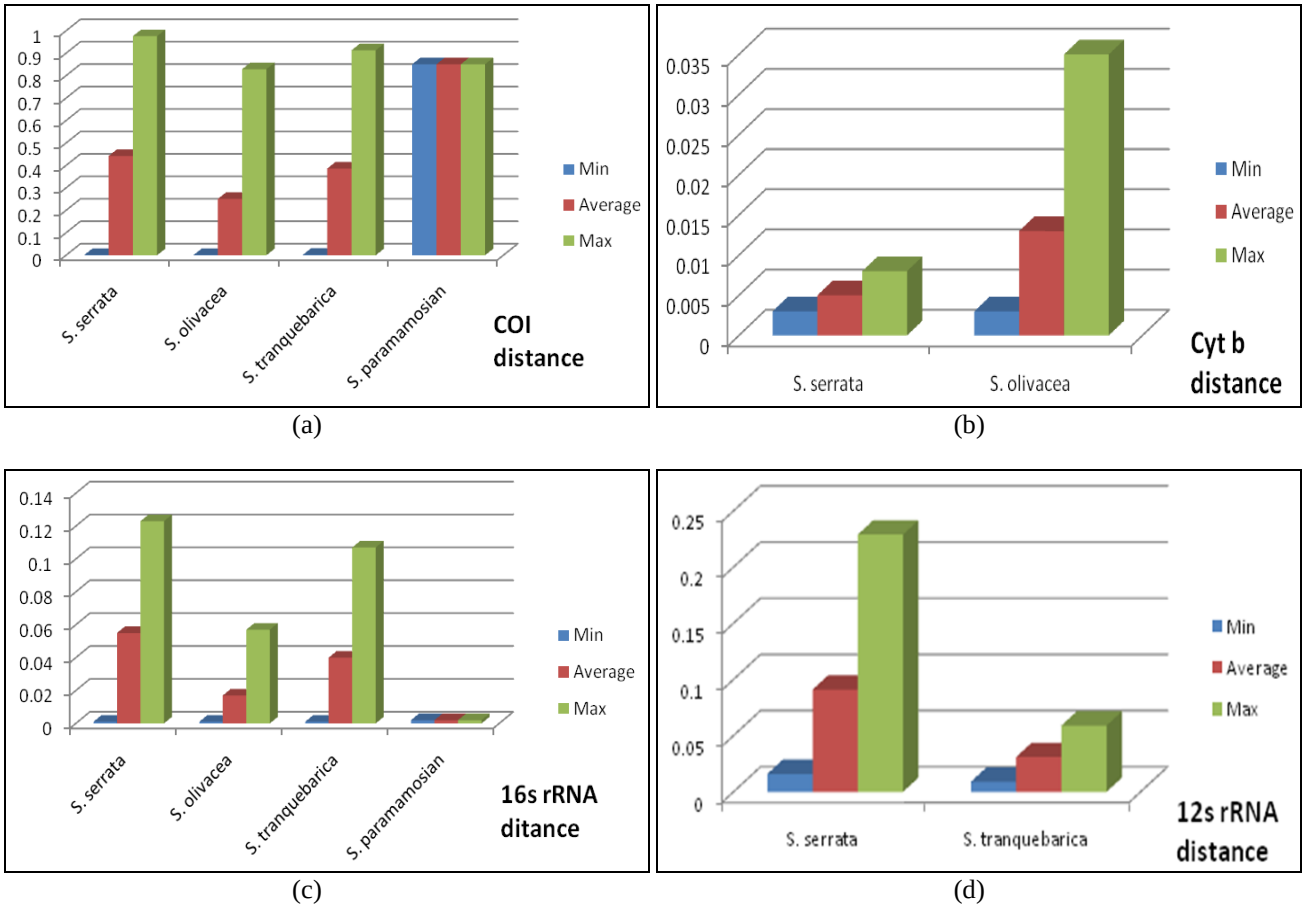

Fig 1: The average and range of intra genetic divergences for four mt- genes

Pair wise genetic distance matrix of *Scylla* species for four mt-genes sequences were represented in Table 3. The highest Kimura's 2-parameter (K-2P) genetic distance for COI gene was observed between *S. tranquebarica* and *S. olivacea* (0.795%) whereas the least value found between *S. tranquebarica* and *S. serrata* (0.459%)**.** The highest Kimura's 2-parameter genetic distance for 16s RNA gene was observed between *S. tranquebarica* and *S. olivacea* (0.059%) whereas the least value found between *S. tranquebarica* and *S. serrata* (0.459%)**.** The maximum Kimura's 2-parameter genetic distance for Cyt b gene was observed between *S. olivacea* and *S. serrata* (0.161%) whereas the least value found between *S. paramamosian* and *S. tranquebarica* (0.120%)**.** The highest intra K-2P genetic distances among all *Scylla* species were

noticed in COI gene followed by Cyt b gene and 16s rRNA gene.

Table 3: Inter species pair-wise genetic distances using K-2P model for 4 mt-genes

S. No.	Species	*S. serrata*	*S. olivacea*	*S. tranquebarica*
COI gene				
1	*S. serrata*	-	-	-
2	*S. olivacea*	0.763	-	-
3	*S. tranquebarica*	0.459	0.795	-
4	*S. paramamosian*	0.574	0.618	0.552
16s rRNA gene				
5	*S. serrata*	-	-	-
6	*S. olivacea*	0.070	-	-
7	*S. tranquebarica*	0.059	0.118	-
8	*S. paramamosian*	0.085	0.081	0.092
Cyt b gene				
9	*S. serrata*	-	-	-
10	*S. olivacea*	0.161	-	-
11	*S. tranquebarica*	0.134	0.154	-
12	*S. paramamosian*	0.150	0.148	0.120
12s rRNA gene				
13	*S. serrata*	-	-	0.140

The average nucleotide frequencies for four *Scylla* species were represented in Table 4. The average nucleotide percentage values for COI gene were A= 27.7%, T= 36.8%, G= 16.5%, C= 18.9%. The average nucleotide percentage values for Cyt b gene were A= 25.7%, T= 39.9%, G= 12.5%, C= 21.8%. A/T content is higher than the G/C content across four mitochondrial genes. Maximum value of A/T content observed in 12 s rRNA (72.0%) followed by 16s rRNA (69.75%), Cyt b gene (65.62%) and COI gene (64.5%). The comparative analysis for A/T and G/C content values for four mt-genes revealed that COI gene undergone most recent evolution followed by Cyt b, 16s rRNA and 12s rRNA genes. The average GC content in 3 codon positions (GC1, GC2 and GC3) of 2 protein coding genes (COI and Cyt b) for four *Scylla* species was depicted in figure 2. For COI gene, GC3 content is dominated followed by GC1 and GC2. This indicates that the synonymous mutations occur mostly at the GC3 followed by GC1 and GC2. This synonymous mutation doesn't effects resulted protein. The highest percentage of GC3 content for COI gene was noticed in *S. serrata* (47.7%) whereas the lowest value was observed in *Scylla tranquebarica* (34.6%). Contrarily, Cyt b sequences contain higher GC2 content followed by GC3 and GC1. Moreover, all the four species contains equal percentages of GC2 and GC3. Average GC content was higher than the COI gene GC content of Arthropod dataset suggested by Keskin & Atar (22). The average nucleotide percentage values for 16s rRNA gene were A= 36.1%, T= 33.6%, G= 19.0%, C= 11.17%. The average A, T, G, C nucleotide percentage values for 16s rRNA gene were 36.5%, 35.5%, 15.5% and 12.5% respectively (Table 5).

Table 4: Calculated nucleotide frequecies for coding genes along with their S.E. values

S. No.	Species	A	T	G	C	(GC)1	(GC)2	(GC)3
COI gene								
1	*Scylla serrata*	26.8±0.11	37.8±0.20	17.0±0.05	18.4±0.42	40.2±0.50	18.4±1.83	47.7±1.14
2	*Scylla olivacea*	28.4±0.14	34.5±0.13	16.4±0.07	20.7±0.10	42.5±0.52	27.2±1.20	41.8±0.63
3	*Scylla tranquebarica*	27.2±0.25	38.3±0.15	16.4±0.10	18.1±0.34	40.1±0.39	15.6±0.92	34.6±0.31
4	*Scylla paramamosian*	28.4±0.28	36.9±0.08	16.3±0.05	18.4±0.41	39.6±0.32	19.7±1.15	45.0±1.33
Cyt b gene								
5	*Scylla serrata*	26.9±0.03	38.4±0.04	13.1±0.04	21.5±0.05	24.3±0.14	44.3±0.05	35.4±0.00
6	*Scylla olivacea*	23.6±0.05	39.0±0.03	13.3±0.07	24.1±0.04	31.1±0.19	45.7±0.00	35.6±0.05
7	*Scylla tranquebarica*	26.2±0.00	42.1±0.00	11.8±0.00	19.9±0.00	13.3±0.00	46.5±0.00	35.4±0.00
8	*Scylla paramamosian*	26.2±0.00	40.1±0.00	12.0±0.00	21.7±0.00	18.8±0.00	47.2±0.00	35.4±0.00

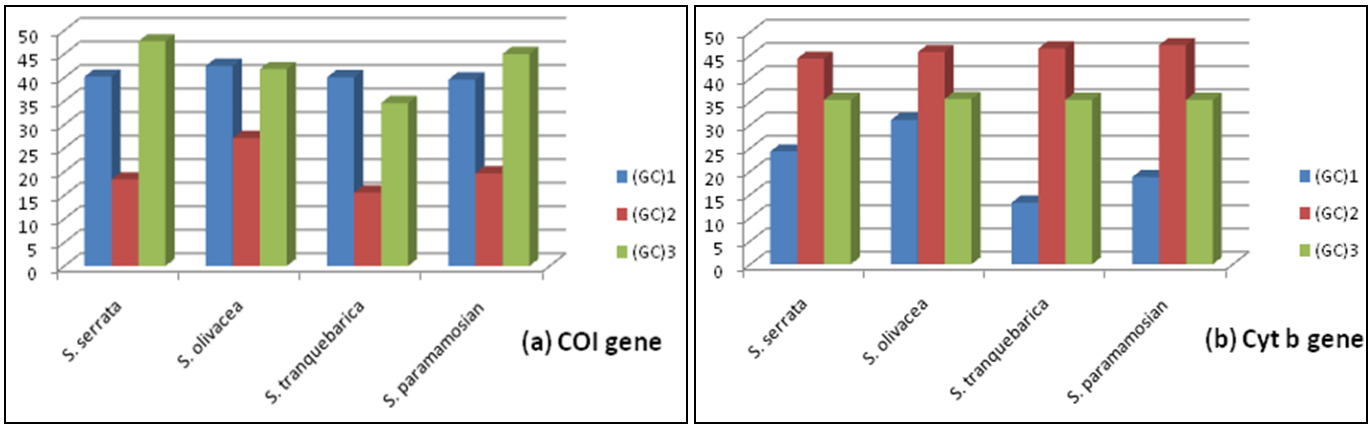

Fig 2: Variation in GC content (%) of (a) COI and (b) Cyt b genes for *Scylla* species

Table 5: Calculated nucleotide frequecies for non-coding genes along with their S.E. values

S. No.	Species	A	T	G	C
16s rRNA gene					
1	*Scylla serrata*	35.8±0.11	33.3±0.02	19.5±0.19	11.3±0.03
2	*Scylla olivacea*	36.4±0.04	33.4±0.05	18.6±0.07	11.5±0.04
3	*Scylla tranquebarica*	35.4±0.11	33.3±0.07	19.9±0.16	11.3±0.08
4	*Scylla paramamosian*	36.8±0.02	34.6±0.00	18.0±0.03	10.6±0.00
12s rRNA gene					
5	*Scylla serrata*	36.2±0.10	35.2±0.26	15.3±0.16	13.3±0.43
6	*Scylla olivacea*	36.8±0.14	35.8±0.13	15.7±0.07	11.7±0.41

The Neighbour-Joining trees for four mt genes were derived by using all 127 sequences of 4 species with MEGA 6.0 are shown in figure 3. Both the N-J trees resulted from COI sequences and 16s rRNA sequences were showed that all four species formed reciprocal monophyletic groups with unordered pattern where as Cytochrome b and 12s rRNA genes formed the clades with species belongs to same species under different clades with significant boot strap values. Based on clades pattern of COI and 16s rRNA sequences, it was evident that *S. serrata* is closely related to *S. tranquebarica* followed by *S. olivaceai*. Overall, Cytochrome b sequences better explained the phylogenitic relationship in coding genes whereas 12s rRNA exhibited proper phylogeny. So, it can be concluded that Cyt b and 12s rRNA sequences contains better phylogenetic signal over the other two genes.

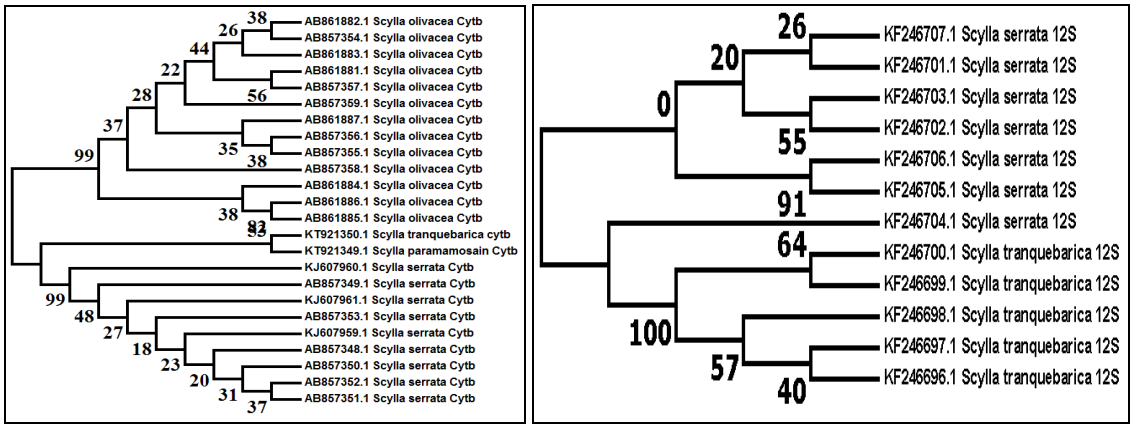

Fig 3: Traditional Kimura 2- Parameter distance Neighbour Joining tree constructed from partial mitochondrial gene sequences (a) COI (b) 16s rRNA (c) Cyt b and (d) 12s rRNA.

Conclusion

Therefore, it is confirmed that validation with mitochondrial markers was successful and helps in successful stock enhancement and breeding programs for domesticated stock that leads to conservation and sustainably exploitation India. This study also provided the efficacy of Cyt b and 12s rRNA genes in delineating the members of the evolutionary closely related organisms.

References

1. Fuseya R, Watanabe S. Genetic variability in the mud crab genus *Scylla* (Brachyura: Portunidae). Fisheries Sci. 1996; 62:705-709.
2. Knuckey IA. Maturity in male mud crabs, Scylla serrata, and the use of mating scars as a functional indicator. Journal of Crustacean Biology. 1996; 16(3):487-495.
3. Estampador EP. Studies on *Scylla* (Crustacea: Portunidae). Revision of the genus. Philipp. J. Sci. 1949; 78(1):95-108.
4. Stephenson W, Campbell B. The Australian Portunids (Crustacea: Portunidae) IV: remaining genera. Aust. J. Mar. Freshwater Res. 1960; 11:73-122.
5. Serene R. Les especes du genre *Scylla* a Nhatrang (Vietnam). Proceeding of Indo-Pacific Fisheries Council. 1952; 3(2):133-137.
6. Keenan CP, Davie PJF, Mann DL. A revision of the genus *Scylla* de Haan, 1833 (Crustacea: Decapoda: Brachyura: Portunidae). The Raffles Bulletin of Zoology. 1998; 46:217-245.
7. Joel DR, Raj PS. Taxonomic remarks on two species of the genus Scylla de Haan (Portunidae: Brachyura) from Pulicat Lake. Indian journal of fisheries. 1983; 30:13-26.
8. Kathirvel M, Srinivasagam S. Taxonomy of the mud crab, Scylla serrata (Forska°l), from India. In: Angell CA (ed) The mud crab, a report on the seminar convened in Surat Thani, Thailand. Bay of Bengal Programme, Madras, India. 1992; 5(8):172-132.
9. Brown WG. Rapid evolution of animal mitochondrial DNA. 1979.
10. Chu KH, Tong J, Chan TY. Mitochondrial cytochrome oxidase I sequence divergence in some Chinese species of *Charybdis* (Crustacea: Decapoda: Portunidae). Biochemical Systematics and Ecology. 1999; 27(5):461-468.
11. Baldwin JD, Bass AL, Bowen BW, Clark WH. Molecular Phylogeny and Biogeography of the Marine Shrimp Penaeus. Molecular phylogenetics and evolution. 1998; 10(3):399-407.
12. Padmavathi P, Srinu G. Emerging trends in DNA markers and their applications in Aquatic biodiversity with an emphasis on mitochondrial markers. International Journal of Zoology Studies. 2017; 2(6):213-219.
13. Trivedi JN, Vachhrajani KD. Taxonomic account of genus *Scylla* (de Haan, 1833) from Gujarat State, India with two new records of species. Arthropods. 2013; 2(4):159-171.
14. Devi PL, Joseph A. Taxonomy of mud crabs of the genus *Scylla* (De Haan) from Cochin Backwaters, Kerala, India. Indian Journal of Fisheries. 2015, 62(3).
15. Balasubramanian CP, Cubelio SS, Mohanlal DL, Ponniah AG, Kumar R, Bineesh KK *et al*. DNA sequence information resolves taxonomic ambiguity of the common mud crab species (Genus Scylla) in Indian waters. Mitochondrial DNA Part A. 2016; 27(1):270-275.
16. Mandal A, Varkey M, Sobhanan SP, Mani AK, Gopalakrishnan A, Kumaran G. Molecular markers reveal only two mud crab species of genus *Scylla* (Brachyura: Portunidae) in Indian coastal waters. Biochemical genetics. 2014; 52(7-8):338-354.
17. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic acids research. 1997; 25(24):4876-4882.
18. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular biology and evolution. 2016; 33(7):1870-1874.
19. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution. 1980; 16:111-120.
20. Saitou N, Nei M. The neighbor-joining method: A new

method for reconstructing phylogenetic trees. Molecular Biology and Evolution. 1987; 4:406-425.
21. Felsinstin J. PHYLIP (Phylogeny Inference Package) version 3.50 Distributed by the author. Department of Genetics, University of Washington, Seattle, USA. 1993.
22. Keskin E, Atar HH. DNA barcoding commercially important aquatic invertebrates of Turkey. Mitochondrial DNA. 2013; 24(4):440-450.